Wykład 4.

Matematyka 2, semestr letni 2010/2011

Po nauczeniu się solidnej porcji algebry liniowej możemy w końcu przejść do analizy funkcji wielu zmiennych rzeczywistych. Będziemy pracować także z odwzorowaniami z $\mathbb{R}^n$ do $\mathbb{R}^m$, tzn odwzorowaniami „składającymi się” z m funkcji n zmiennych:

$$F : \mathbb{R}^n \ni (x^1, x^2, \ldots, x^n) \longmapsto (\ F^1(x^1, \ldots, x^n),\ F^2(x^1, \ldots, x^n), \ldots, F^m(x^1, \ldots, x^n)) \in \mathbb{R}^m$$

Zanim zajmiemy się różniczkowaniem musimy omówić problemy związane z ciągłością i pojęciem granicy odwzorowania w punkcie. Przypomnijmy sobie definicję ciągłości funkcji jednej zmiennej w punkcie. Niech I będzie odcinkiem otwartym w $\mathbb{R}$ (bez końców). Mówimy, że funkcja $f : I \to \mathbb{R}$ jest ciągła w punkcie $x_0 \in I$ jeśli wartości funkcji w punktach bliskich x_0 są bliskie $f(x_0)$. Precyzyjnie warunek ciągłości zapisujemy

$$\forall \varepsilon > 0 \ \ \exists \delta > 0 : \quad |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon. \tag{1}$$

Powyższą definicję (zwaną *definicją ciągłości Cauchy'ego*) można uogólnić na przypadek odwzorowań z $\mathbb{R}^n$ do $\mathbb{R}^m$ posługując się pojęciem odległości. Załóżmy, że potrafimy mierzyć odległość między punktami w $\mathbb{R}^n$ i w $\mathbb{R}^m$. Niech $x = (x^1, x^2, \ldots, x^n)$, $y = (y^1, y^2, \ldots, y^n)$ będą elementami $\mathbb{R}^n$, wtedy odległość między nimi oznaczymy $d(x, y)$. W przestrzeni wartości, czyli w $\mathbb{R}^m$ użyjemy symbolu ρ, tzn jeśli $a = (a^1, a^2, \ldots, a^m)$, $b = (b^1, b^2, \ldots, b^m)$ to odległość między punktami a i b oznaczymy $\rho(a, b)$. Teraz możemy napisać warunek ciągłości w punkcie $x_0 \in \mathbb{R}^n$: Mówimy, że odwzorowanie

$$F : \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

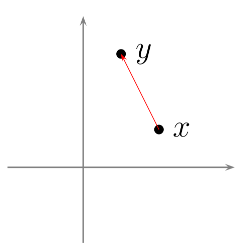
jest ciągłe w punkcie x_0, jeśli spełniony jest warunek

$$\forall \varepsilon > 0 \ \ \exists \delta > 0 : \quad d(x, x_0) < \delta \Rightarrow \rho(F(x), F(x_0)) < \varepsilon. \tag{2}$$

Warunek (2) jest bardzo podobny do (1) - wartość bezwzględną różnicy liczb należy zamienić na odległość między punktami odpowiednio w $\mathbb{R}^n$ i $\mathbb{R}^m$. Wygląda więc na to, że ciągłość odwzorowania zależy od sposobu mierzenia odległości w dziedzinie i zbiorze wartości. W ogólności rzeczywiście tak jest. Oto przykład:

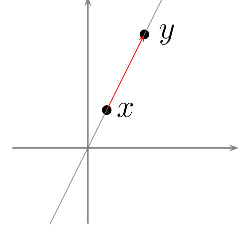
Przykład 1. Rozważmy dwa sposoby mierzenia odległości między punktami na płaszczyźnie ($\mathbb{R}^2$). Pierwszy sposób - euklidesowy:

$$x = (x^1, x^2), \quad y = (y^1, y^2), \qquad d_2(x, y) = \sqrt{(x^1 - y^1)^2 - (x^2 - y^2)^2},$$



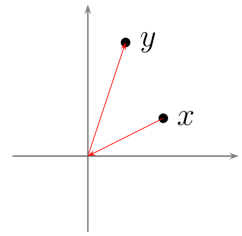
a drugi klolejowy (stacja węzłowa i tory): jeśli punkty x i y leżą na jednej prostej przechodzącej przez punkt $(0, 0)$ wtedy odległość mierzymy jak zwykle:

$$d_k(x, y) = \sqrt{(x^1 - y^1)^2 - (x^2 - y^2)^2},$$



w przeciwnym przypadku jedziemy najpierw do stacji węzłowej:

$$d_k(x,y) = \sqrt{(x^1)^2 + (x^2)^2} + \sqrt{(y^1)^2 + (y^2)^2}.$$



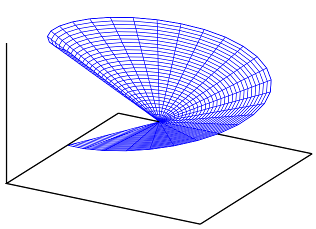
Rozważmy teraz funkcję, która przypisuje punktowi $x = (x^1, x^2)$ wartość

$$F(x^1, x^2) = r(x)\varphi(x),$$

gdzie $r(x)$ jest odległością punktu x od $(0,0)$ w sensie euklidesowym a $\varphi(x) \in [0, 2\pi[$ jest kątem jak we współrzędnych biegunowych. Rozważmy punkt $x_0 = (1,0)$ Funkcja ta jest nieciągła w x_0 względem metryki euklidesowej. Istotnie, Odległość euklidesowa od punktu x_0 do punktu $x = (1, -a)$ dla $a > 0$ jest równa a. Różnica w wartości zaś jest bliska 2π i dla a dążącego do 0 rośnie do 2π a nie zmniejsza się. W metryce kolejowej punt $(1, -a)$ nie jest dobrym punktem do konstruowania kontrprzykładu na ciągłość, gdyż odległość kolejowa

$$d_k(x_0, x) = 1 + \sqrt{1 + a^2} > 2.$$

Naprawdę blisko punktu x_0 leżą jedynie punkty położone na osi $0X$. Dla tych punktów wartość funkcji jest stała i równa zero. Funkcja F jest więc ciągła względem odległości kolejowej. Oto wykres funkcji F:



♣

Korzystając z pojęcia odległości możemy również mówić o ciągach zbieżnych w $\mathbb{R}^n$. Niech $(x_k)_{k=1}^{\infty}$ będzie ciągiem punktów w $\mathbb{R}^n$ (tzn. każde x_k ma n współrzędnych: $(x_k^1, x_k^2, \ldots, x_k^n)$). Mówimy, że *ciąg* (x_k) *jest zbieżny do* $x_0 = (x_0^1, x_0^2, \ldots, x_0^n)$ jeśli spełniony jest warunek

$$\forall \varepsilon > 0 \;\; \exists N \in \mathbb{N} : \;\; \forall k > N d(x_k, x_0) \leqslant \varepsilon. \tag{3}$$

Pojęcie ciągłości w $\mathbb{R}^n$ ma też swoją ciągową wersję, podobną do definicji Heinego:

Fakt 1. *Odwzorowanie* $F : \mathbb{R}^n \to \mathbb{R}^m$ *jest ciągłe w* $x_0 \in \mathbb{R}^n$ *wtedy i tylko wtedy, gdy dla każdego ciągu* (x_k) *zbieżnego do* x_0*, ciąg* $F(x_k)$ *jest zbieżny do* $F(x_0)$*.*

Nie będziemy dowodzić tego faktu, gdyż dowód nie różni się od dowodu dla jednego wymiaru. Ciągowa wersja definicji ciągłości nadaje się dobrze do wykazywania, że jakaś funkcja jest nieciągła. Wystarczy wtedy podać jeden ciąg (x_k) stanowiący kontrprzykład, tzn taki, który jest zbieżny do x_0, ale ciąg wartości $F(x_k)$ nie jest zbieżny do $F(x_0)$. Definicja Cauchy'ego nadaje się lepiej do dowodzenia ciągłości.

W definicji ciągłości Cauchy'ego badamy wartości funkcji F na punktach $\mathbb{R}^n$ znajdujących się w odległości mniejszej niż δ od x_0. Zbiór tych punktów nazywać będziemy kulą otwartą. Dokładniej, *kulą otwartą* w $\mathbb{R}^n$ o środku w x_0 i promieniu δ względem odległości d nazywamy zbiór

$$K_d(x_0, \delta) = \{\, x \in \mathbb{R}^n : \ d(x, x_0) < \delta \,\}.$$

Warunek ciągłości można przeformułować używając kul:

$$\forall \varepsilon > 0 \ \exists \delta > 0 : \ \ x \in K_d(x_0, \delta) \Rightarrow F(x) \in K_\rho(F(x_0), \varepsilon).$$

W pojęciu ciągłości ważny jest więc kształt kul względem danej odległosci. Zdarza się, że różne sposoby mierzenia odległości prowadzą do takiego samego zbioru funkcji ciągłych. Zanim omówimy to szczegółowo sformułujmy precyzyjnie pojęcie *sposobu mierzenia odległości*:

Sposób mierzenia odległości wyrażamy matematycznie za pomocą funkcji zwanej metryką. Mówimy, że $d : X \times X \to \mathbb{R}$ jest *metryką* na zbiorze x gdy spełnione są warunki

$$\begin{aligned}
&(1) \quad \forall x, y \in X \ \ d(x, y) \geqslant 0 \\
&(2) \quad \forall x, y \in X \ \ d(x, y) = d(y, x) \\
&(3) \quad d(x, y) = 0 \ \Leftrightarrow \ x = y \\
&(4) \quad \forall x, y, z \in X \ \ d(x, z) \leqslant d(x, y) + d(y, z)
\end{aligned}$$

Warunek (2) oznacza, że metryka jest symetryczna, warunek (3) wyraża niezdegenerowanie, warunek (4) nosi nazwę nierówności trójkąta. Oprócz całego mnóstwa dziwnych metryk na $\mathbb{R}^n$ istnieją trzy bardzo zwyczajne, których będziemy używać: Pierwsza metryka to *metryka euklidesowa*, zapewne najbardziej dla państwa naturalna, ale i najtrudniejsza w użyciu ze względu na nieporęczny pierwiastek:

$$d_2(x, y) = \sqrt{(x^1 - y^1)^+ \cdots + (x^n - y^n)^2}.$$

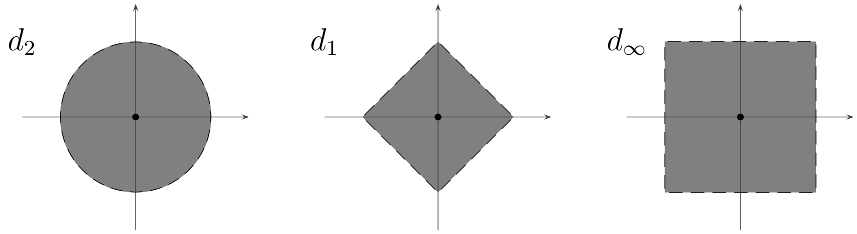
Druga metryka nazywana jest *metryką miejską*, gdyż odległość mierzona jest wzdłuż prostopadłych ulic:

$$d_1(x, y) = |x^1 - y^1| + \cdot + |x^n - y^n|.$$

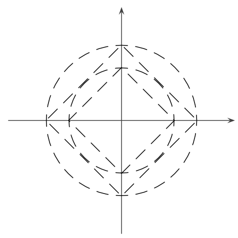
Ostatnia metryka, to tak zwana *metryka maksimum*:

$$d_\infty(x, y) = \max\{|x^1 - y^1|, \ldots + |x^n - y^n|\}.$$

Wspomnieliśmy wcześniej, że dla pojęcia ciągłości istotny jest kształt kuli względem metryki. Żeby wyrobić sobie jakieś wyobrażenie o trzech używanych przez nas metrykach narysujmy kule o środku w $(0, 0)$ i promieniu 1 w $\mathbb{R}^2$ względem każdej z nich:

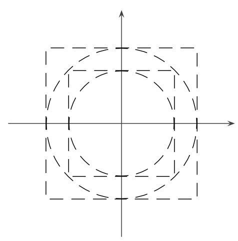


Zauważmy także, że dobierając odpowiednio promienie możemy zawrzeć jedną kulę w drugiej: np dla d_2 i d_1:



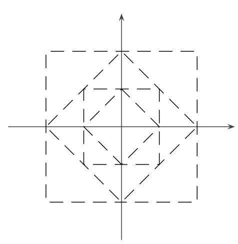
$$K_1((0,0),\frac{\sqrt{2}}{2}) \subset K_2((0,0),\frac{\sqrt{2}}{2}) \subset K_1((0,0),1) \subset K_2((0,0),1)$$

Podobnie dla pary d_2, d_∞:



$$K_2((0,0),\frac{\sqrt{2}}{2}) \subset K_\infty((0,0),\frac{\sqrt{2}}{2}) \subset K_2((0,0),1) \subset K_\infty((0,0),1)$$

To samo dla pary d_1, d_∞:



$$K_1((0,0),\frac{1}{2}) \subset K_\infty((0,0),\frac{1}{2}) \subset K_1((0,0),1) \subset K_\infty((0,0),1)$$

Jeśli więc funkcja $f : \mathbb{R}^n \to \mathbb{R}$ jest ciągła względem d_2, tzn dla każdego ε znajdziemy δ taką, że warunek ciągłości (2) zachodzi dla metryki d_2, to będzie zachodził także dla $\delta' = \frac{\sqrt{2}}{2}\delta$ dla metryki d_∞ (bo kula względem d_∞ i promieniu δ' zawiera się w kuli o promieniu δ względem metryki d_2). Warunek ten będzie zachodził także dla d_1 jeśli weźmiemy promień $\delta'' = \delta$. Podobnie, jeśli zapiszemy warunek dla metryki d_1, to możemy dobrać takie promienie względem d_2 i d_∞, żeby odpowiednie kule się zawierały w kuli względem d_1 i warunek zachodził. To samo, jeżeli zaczniemy od d_∞. Widać więc, że metrykom d_1, d_2, d_∞ odpowiadają takie same zbiory funkcji ciągłych. Dopuszczenie wielu wymiarów w przestrzeni wartości nie zmienia sytuacji: także te trzy metryki mają te same zbiory odwzorowań ciągłych. Będziemy więc we wszelkich dowodach ciągłości i zbieżności używać tych trzech metryk wymiennie - w zależności od wygody rachunkowej. Podobnie jest ze zbieżnością ciągów w $\mathbb{R}^n$. Wszystkie trzy metryki definiują takie same zbiory ciągów zbieżnych. Niech (x_k) będzie ciągiem zbieżnym do x_0 względem którejkolwiek z trzech metryk. Wtedy jest zbieżny także względem metryki maksimum. Oznacza to, że

$$\forall \varepsilon > 0 \ \exists N \in \mathbb{N} \ \forall n > N \quad \max\{|x_k^1 - x_0^1|, \ldots |x_k^n - x_0^n|\} < \varepsilon$$

Skoro największa z różnic jest mniejsza niż ε to także każda z nich jest mniejsza. To prowadzi do wniosku, że każda ze współrzędnych ciągu (x_k) zbiega oddzielnie do odpowiedniej współrzędnej punktu granicznego x_0. Odwrotnie, jeśli każda ze współrzędnych ciągu (x_k) zbiega do odpowiedniej współrzędnej punktu (x_0), to znaczy, że dla każdego indeksu i możemy znaleźć N_i takie, że dla $n > N_i$ zachodzi nierówność $|x_k^i - x_0^i| < \varepsilon$ Największe z N_i zapewnia spełnienie warunku zbieżności ciągu x_k do x_0. Zbieżność w $\mathbb{R}^n$ oznacza więc zbieżność po współrzędnych.

Oto dwa przykłady niecodziennych sytuacji, które możemy napotkać badając ciągłość funkcji zależnych od wielu zmiennych:

Przykład 2. Rozważmy funkcję $f : \mathbb{R}^2 \to \mathbb{R}$ w otoczeniu punktu $(0, 0)$

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

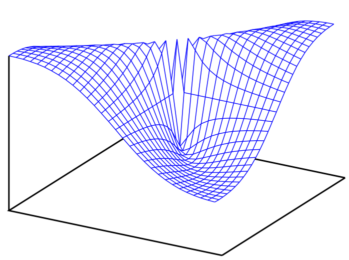
Na osiach współrzędnych, czyli na zbiorach $\{(0, y)\}$ i $\{(x, 0)\}$ funkcja f przyjmuje wartość 0. Gdy jednak zmierzamy do zera wzdłuż innej prostej, np $y = ax$ okazuje się, że f jest nieciągła w $(0, 0)$. Niech $\alpha_k = (\frac{1}{n}, \frac{a}{n})$. Jest oczywiste, że

$$\lim_{k\to\infty} \alpha_k = (0,0)$$

Ciąg wartości jest ciągiem stałym

$$f(\alpha_k) = f(\frac{1}{n}, \frac{a}{n}) = \frac{\frac{a}{n^2}}{\frac{1}{n^2} + \frac{a}{n^2}} = \frac{a}{1+a^2}$$

z granicą $\frac{a}{a+a^2} \neq 0$.



♣

Przykład 3. Rozważmy funkcję $g : \mathbb{R}^2 \to \mathbb{R}$ w otoczeniu punktu $(0,0)$

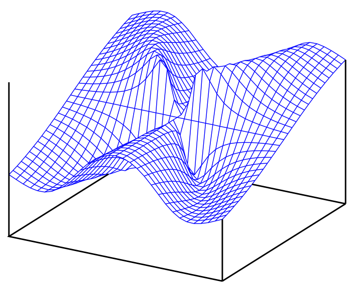
$$\begin{cases} g(x,y) = \dfrac{xy^2}{x^2+y^4} & (x,y) \neq (0,0) \\ 0 & (x,y) \neq (0,0) \end{cases}$$

Na osiach współrzędnych, czyli na zbiorach $\{(0,y)\}$ i $\{(x,0)\}$ funkcja g przyjmuje wartość 0. Na dowolnej prostej przechodzącej przez zero ($y = ax$) otrzymujemy

$$g(x,ax) = \frac{a^2x^3}{x^2+a^4x^4} = \frac{a^2}{1+a^4x^2} \longrightarrow_{(x\to 0)} 0$$

Gdy jednak zmierzamy do zera wzdłuż innej krzywej, np $x = ay^2$ okazuje się, że g jest nieciągła w $(0,0)$.

$$g(ay^2,y) = \frac{ay^4}{a^2y^4+y^4} = \frac{a}{1+a^2} \neq 0 \quad \text{dla} \quad a \neq 0.$$



♣

Różniczkowanie funkcji i odwzorowań zależnych od wielu zmienych rzeczywistych. Oznaczmy przez F pewne odwzorowanie z $\mathbb{R}^n$ do $\mathbb{R}^m$. Oznacza to, że F składa się z m funkcji n zmiennych:

$$F : \mathbb{R}^n \ni (x^1,\dots,x^n) \longmapsto \Big(F^1(x^1,\dots,x^n),\dots,F^m(x^1,\dots,x^n) \Big) \in \mathbb{R}^m.$$

Mówimy, że F jest *różniczkowalne w punkcie* $x_0 = (x_0^1,\dots,x_0^n)$ jeśli istnieje odwzorowanie liniowe $A \in L(\mathbb{R}^n,\mathbb{R}^n)$ takie, że

$$F(x_0+h) = F(x_0) + A(h) + r(x_0,h),$$

gdzie reszta $r(x_0,h)$ spełnia warunek

$$\lim_{h\to 0} \frac{\|r(x_0,h)\|}{\|h\|} = 0$$

W powyższym wzorze $\|h\|$ oznacza długość wektora h. Odwzorowanie A nazywamy pochodną F w punkcie x_0. Oznaczamy je także $F'(x_0)$.

Powyższa definicja nie jest niestety konstruktywna, tzn. nie precyzuje postaci odwzorowania A. Spróbujmy ją odgadnąć. Ustalmy pewne odwzorowanie F i załóżmy, że jest ono różniczkowalne. $A = F'(x_0)$ jest odwzorowaniem liniowym z $\mathbb{R}^n$ do $\mathbb{R}^m$, czyli macierzą mającą m-wierszy i n-kolumn. Wyrazy macierzowe znajdziemy (jak zawsze dla odwzorowań liniowych) obliczając wartość odwzorowania na wektorach bazowych. Użyć należy oczywiście bazy standardowej.

Niech $e = (e_1, \ldots, e_n)$ oznacza bazę standardową w $\mathbb{R}^n$ a $f = (f_1, \ldots, f_m)$ bazę standardową w $\mathbb{R}^m$. Wyraz macierzowy $A^i{}_j$ jest i-tą współrzędną wektora $A(e_j)$ w bazie f, tzn:

$$A(e_j) = A^1{}_j f_1 + \cdots + A^m{}_j f_m.$$

Znajdźmy $A(e_j)$. Z definicji wiemy, że dla przyrostu $h = te_j$, $t \in \mathbb{R}$, mamy

$$F(x_0 + te_j) = F(x_0) + A(te_j) + r(x_0, te_j) = F(x_0) + tA(e_j) + r(x_0, te_j).$$

Wyznaczamy $A(e_j)$:

$$tA(e_j) = F(x_0 + te_j) - F(x_0) - r(x_0, te_j), \quad \text{tzn.} \quad A(e_j) = \frac{F(x_0 + te_j) - F(x_0)}{t} - \frac{r(x_0, te_j)}{t}.$$

W granicy, dla $t \to 0$ wyrażenie z resztą znika (co wynika z definicji różniczkowalności) i otrzymujemy wzór:

$$A(e_j) = \lim_{t \to 0} \frac{F(x_0 + te_j) - F(x_0)}{t}$$

bardzo przypominający definicję pochodnej dla funkcji jednej zmiennej. Pamiętać jednak należy, że w liczniku jest różnica dwóch punktów z przestrzeni $\mathbb{R}^m$, a więc także element $\mathbb{R}^m$. Jeśli skorzystamy z bazy standardowej f otrzymamy

$$A(e_j) = \left(\lim_{t \to 0} \frac{F^1(x_0 + te_j) - F^1(x_0)}{t} \right) f_1 + \cdots + \left(\lim_{t \to 0} \frac{F^m(x_0 + te_j) - F^m(x_0)}{t} \right) f_m.$$

Wyraz macierzowy $A^i{}_j$ ma więc postać

$$A^i{}_j = \lim_{t \to 0} \frac{F^i(x_0 + te_j) - F^i(x_0)}{t}.$$

Wypiszmy jeszcze explicite argumenty funkcji F^i:

$$A^i{}_j = \lim_{t \to 0} \frac{F^i(x_0^1, x_0^2, \ldots, x_0^{j-1}, x_0^j + t, x_0^{j+1}, \ldots, x_0^n) - F^i(x_0^1, x_0^2, \ldots, x_0^{j-1}, x_0^j, x_0^{j+1}, \ldots, x_0^n)}{t}.$$

Powyższe wyrażenie nazywamy *pochodną cząstkową funkcji* F^i *w punkcie* x_0 *po współrzędnej* x^j i oznaczamy

$$\frac{\partial F^i}{\partial x^j}(x_0).$$

Praktycznie pochodną cząstkową obliczamy różniczkując funkcję F^i względem x^j, zgodnie z zasadami obowiązującymi dla funkcji jednej zmiennej, traktując pozostałe zmienne jako ustalone. Okazało się, że jeśli F jest różniczkowalne, to jego pochodna w punkcie x_0 ma postać:

$$F'(x_0) = \begin{bmatrix} \frac{\partial F^1}{\partial x^1}(x_0) & \frac{\partial F^1}{\partial x^2}(x_0) & \cdots & \frac{\partial F^1}{\partial x^n}(x_0) \\ \frac{\partial F^2}{\partial x^1}(x_0) & \frac{\partial F^2}{\partial x^2}(x_0) & \cdots & \frac{\partial F^2}{\partial x^n}(x_0) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial F^m}{\partial x^1}(x_0) & \frac{\partial F^m}{\partial x^2}(x_0) & \cdots & \frac{\partial F^m}{\partial x^n}(x_0) \end{bmatrix}$$

Przykład 4. Jako przykład rozważmy odwzorowanie $\Phi : \mathbb{R}^2 \to \mathbb{R}^2$ $(m = 2,\ n = 2)$, dane wzorem

$$\mathbb{R}^2 \ni (r, \varphi) \longmapsto (x(r, \varphi), y(r, \varphi)) = (r \cos \varphi,\ r \sin \varphi) \in \mathbb{R}^2,$$

które obcięte do odpowiedniego podzbioru w dziedzinie opisuje współrzędne biegunowe na płaszczyźnie. Sprawdzimy, czy odwzorowanie to jest różniczkowalne w punkcie (r_0, φ_0). Naturalnym kandydatem na pochodną jest macierz złożona z pochodnych cząstkowych funkcji składowych:

$$A = \begin{bmatrix} \dfrac{\partial x}{\partial r} & \dfrac{\partial x}{\partial \varphi} \\ \dfrac{\partial y}{\partial r} & \dfrac{\partial y}{\partial \varphi} \end{bmatrix} = \begin{bmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{bmatrix}.$$

Oznaczając przyrosty względem r i φ przez h_r i h_φ otrzymujemy

$$\Phi(r_0 + h_r, \varphi_0 + h_\varphi) = \Phi(r_0, \varphi_0) + A \begin{bmatrix} h_r \\ h_\varphi \end{bmatrix} + r((r_0, \varphi_0), (h_r, h_\varphi))$$

Badanie różniczkowalności polega na analizowniu postaci reszty. Wyznaczamy resztę z powyższego równania

$$r((r_0, \varphi_0), (h_r, h_\varphi)) = \Phi(r_0 + h_r, \varphi_0 + h_\varphi) - \Phi(r_0, \varphi_0) - A \begin{bmatrix} h_r \\ h_\varphi \end{bmatrix} =$$

Wstawiamy wzory określające Φ i A:

$$= ((r_0 + h_r) \cos(\varphi_0 + h_\varphi),\ (r_0 + h_r) \sin(\varphi_0 + h_\varphi)) -$$
$$(r_0 \cos \varphi_0,\ r_0 \sin \varphi_0) - \begin{bmatrix} \cos \varphi_0 & -r_0 \sin \varphi_0 \\ \sin \varphi_0 & r_0 \cos \varphi_0 \end{bmatrix} \begin{bmatrix} h_r \\ h_\varphi \end{bmatrix} =$$

i wykonujemy wskazane działania:

$$= ((r_0 + h_r) \cos(\varphi_0 + h_\varphi) - r_0 \cos \varphi_0 - \cos \varphi_0 h_r + r_0 \sin \varphi_0 h_\varphi,$$
$$(r_0 + h_r) \sin(\varphi_0 + h_\varphi) - r_0 \sin \varphi_0 - \sin \varphi_0 h_r - r_0 \cos \varphi_0 h_\varphi)$$

Każdą ze współrzędnych, niebieską i czerwoną, omówimy oddzielnie. Zaczynamy od niebieskiej, stosujemy wzór

$$\cos(\varphi_0 + h_\varphi) = \cos \varphi_0 \cos h_\varphi - \sin \varphi_0 \sin h_\varphi$$

$$(r_0 + h_r) \cos(\varphi_0 + h_\varphi) - r_0 \cos \varphi_0 - \cos \varphi_0 h_r + r_0 \sin \varphi_0 h_\varphi =$$
$$(r_0 + h_r)(\cos \varphi_0 \cos h_\varphi - \sin \varphi_0 \sin h_\varphi) - r_0 \cos \varphi_0 - \cos \varphi_0 h_r + r_0 \sin \varphi_0 h_\varphi =$$
$$r_0 \cos \varphi_0 \cos h_\varphi - r_0 \sin \varphi_0 \sin h_\varphi + h_r \cos \varphi_0 \cos h_\varphi -$$
$$h_r \sin \varphi_0 \sin h_\varphi - r_0 \cos \varphi_0 - \cos \varphi_0 h_r + r_0 \sin \varphi_0 h_\varphi =$$

W ostatnich dwóch linjkach wzoru grupujemy wyrazy: pierwszy z piątym, drugi z ostatnim i trzeci z przedostatnim:

$$= r_0 \cos \varphi_0 (\cos h_\varphi - 1) + r_0 \sin \varphi_0 (h_\varphi - \sin h_\varphi) + h_r \cos \varphi_0 (\cos h_\varphi - 1) - h_r \sin \varphi_0 \sin h_\varphi =$$

Wypisujemy przybliżone wartości pomijając wyrazy z potęgami przyrostów h_r i h_φ wyższymi niż trzecia (korzystamy z odpowiednich rozwinięć w szereg):

$$\approx r_0 \cos \varphi_0 \left(-\frac{h_\varphi^2}{2} \right) + r_0 \sin \varphi_0 \left(\frac{h_\varphi^3}{6} \right) + h_r \cos \varphi_0 \left(-\frac{h_\varphi^2}{2} \right) + \sin \varphi_0 h_r \left(h_\varphi - \frac{h_\varphi^3}{6} \right).$$

Okazuje się, że w wyrażeniu na pierwszą współrzędną nie ma składników rzędu niższego niż dwa. To samo robimy z czerwoną współrzędną. Rachunki pomijamy (są bardzo podobne) i otrzymujemy

$$(r_0 + h_r)\sin(\varphi_0 + h_\varphi) - r_0 \sin\varphi_0 - \sin\varphi_0 h_r - r_0 \cos\varphi_0 h_\varphi \approx \\ r_0 \sin\varphi_0 \left(-\frac{h_\varphi^2}{2}\right) + r_0 \cos\varphi_0 \left(-\frac{h_\varphi^3}{6}\right) + \sin\varphi_0 h_r \left(-\frac{h_\varphi^2}{2}\right) + cos\varphi_0 h_r \left(h_\varphi - \frac{h_\varphi^3}{6}\right).$$

Warunek na resztę jest postaci:

$$\lim_{h\to 0} \frac{\|r((r_0,\varphi_0),(h_r,h_\varphi))\|}{\|h\|} = 0.$$

Wygodnie zamiast metryki euklidesowej użyć metryki maksimum. Wówczas

$$\|h\| = \max\{|h_r|, |h_\varphi|\},$$

metryki maksimum używamy także dla reszty. Jeśli wykażemy, że każda ze współrzędnych reszty podzielona przez $\|h\|$ znika, gdy h dąży do 0, będziemy wiedzieć, że większa z tych współrzędnych będąca długością reszty znika. Zacznijmy od niebieskiej współrzędnej:

$$\frac{r_0 \cos\varphi_0 \left(-\frac{h_\varphi^2}{2}\right) + r_0 \sin\varphi_0 \left(\frac{h_\varphi^3}{6}\right) + h_r \cos\varphi_0 \left(-\frac{h_\varphi^2}{2}\right) + \sin\varphi_0 h_r \left(h_\varphi - \frac{h_\varphi^3}{6}\right)}{\|h\|}$$

Najniższa potęga przyrostu w liczniku jest dwa (także wyrażenie $h_r h_\varphi$ traktujemy jak kwadratowe, bo oba czynniki dążą do zera). Mianownik jest rzędu jeden (jest równy $|h_r|$ lub $|h_\varphi|$). Oznacza to, że cały ułamek jest rzędu jeden w przyroście. Gdy przyrost dąży do zera, cały ułamek także. Identycznie rzecz się ma z czerwoną współrzędną. W ten sposób pokazaliśmy, że warunek na resztę w definicji różniczkowalności jest spełniony. Odwzorowanie Φ jest różniczkowalne w (r_0, φ_0) i jego pochodna jest równa

$$\Phi'(r_0,\varphi_0) = \begin{bmatrix} \cos\varphi_0 & -r_0\sin\varphi_0 \\ \sin\varphi_0 & r_0\cos\varphi_0 \end{bmatrix}$$

♣

Powyższy przykład i wcześniejsze rozumowanie pokazują, że różniczkowanie funkcji i odwzorowań zależnych od wielu zmiennych polega na obliczaniu odpowiednich pochodnych cząstkowych i zapisywaniu ich w macierzach. Badanie różniczkowalności w sposób przedstawiony w przykładzie (4) jest nużące – na szczęście mamy do dyspozycji twierdzenie:

Twierdzenie 1. *Jeśli dla odwzorowania $F : \mathbb{R}^n \to \mathbb{R}^m$ istnieją wszystkie pochodne cząstkowe $\frac{\partial F^i}{\partial x^j}$ w punkcie (x_0) i są ciągłe w x_0 jako funkcje określone na $\mathbb{R}^n$, to odwzorowanie F jest różniczkowalne w x_0 i jego pochodna traktowana jako odwzorowanie*

$$\mathbb{R}^n \ni x \longmapsto F'(x) \in L(\mathbb{R}^n, \mathbb{R}^m).$$

jest ciągła w x_0. Mówimy w takim przypadku, że F jest różniczkowalna w x_0 w sposób ciągły.

Przestrzeni odwzorowań liniowych, która jest oczywiście przestrzenią wektorową można liczyć długość wektorów, jeśli jest określona długość w dziedzinie i obrazie. Jeśli $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ to

$$\|A\| = \sup_{\|v\|_n = 1} \|A(v)\|_m, \tag{4}$$

gdzie $\|\cdot\|_n$ jest długością wektora w dziedzinie a $\|\cdot\|_m$ długością wektora w przestrzeni wartości odwzorowania A. W powyższym twierdzeniu ciągłość odwzorowania o wartościach w $L(\mathbb{R}^n, \mathbb{R}^m)$ określana jest względem tej właśnie długości. Warto także wiedzieć, że ponieważ $\mathbb{R}^n$ i $\mathbb{R}^m$ są przestrzeniami skończenie-wymiarowymi, to także $L(\mathbb{R}^n, \mathbb{R}^m)$ jest przestrzenią skończenie-wymiarową (wymiaru nm), ponadto może być ona w sposób naturalny utożsamiona z $\mathbb{R}^{nm}$. Oznacza to, że mamy w niej trzy metryki d_1, d_2, d_∞ względem których można definiować ciągłość. Okazuje się, że zbiór odwzorowań ciągłych względem krórejkolwiek z metryk i zbiór odwzorowań ciągłych względem długości zdefiniowanej w (4) są identyczne.

Dla funkcji jednej zmiennej obowiązują prawa ułatwiające różniczkowanie skomplikowanych funkcji: reguła Leibniza różniczkowania iloczynu, wzór na pochodną funkcji złożonej, wzór na pochodną ilorazu itd. W świecie wielu zmiennych także są podobne prawa:

Twierdzenie 2 (Formalne prawa różniczkowania)**.** *Dla funkcji i odwzorowań określonych na $\mathbb{R}^n$ obowiązują następujące reguły:*

(1) *Jeśli odwzorowania $F, G : \mathbb{R}^n \to \mathbb{R}^m$ są różniczkowalne w x_0 to także*

$$\alpha F + \beta G, \qquad (\alpha F + \beta G)(x) = \alpha F(x) + \beta G(x)$$

jest różniczkowalne i

$$(\alpha F + \beta G)'(x_0) = \alpha F'(x_0) + \beta G'(x_0).$$

(2) *Jeśli odwzorowanie $F : \mathbb{R}^n \to \mathbb{R}^m$ jest różniczkowalne w x_0 oraz funkcja $f : \mathbb{R}^n \to \mathbb{R}$ jest różniczkowalna w x_0 to odwzorowanie*

$$fF : \mathbb{R}^n \ni x \longmapsto f(x)F(x) \in \mathbb{R}^m$$

także jest różniczkowalne w x_0 i pochodna ma postać

$$(fF)'(x_0) = f'(x_0)F(x_0) + f(x)F'(x_0),$$

tzn.

$$(fF)'(x_0)h = [f'(x_0)h]F(x_0) + f(x)[F'(x_0)h].$$

(3) *Jeśli $F : \mathbb{R}^n \to \mathbb{R}^m$ i $G : \mathbb{R}^m \to \mathbb{R}^k$ oraz F jest różniczkowalne w x_0 a G jest różniczkowalne w $F(x_0)$ to złożenie*

$$G \circ F : \mathbb{R}^n \to \mathbb{R}^k, \qquad G \circ F(x) = G(F(x))$$

także jest odwzorowaniem różniczkowalnym w x_0. Jego pochodna wyraża się wzorem

$$(G \circ F)'(x_0) = G'(F(x_0)) \cdot F'(x_0),$$

gdzie kropka po prawej stronie równania oznacza mnożenie macierzy.

Powyższego twierdzenia nie będziemy dowodzić. Zajmiemy się jedynie analizą przykładów. Punkt (1) wyraża zasadę liniowości pochodnej i nie wymaga komentowania. Punkt (2) jest wielozmiennową wersją reguły Leibniza. Zwróćmy uwagę na znaczenie wzoru na pochodną:

$$(fF)'(x_0) = f'(x_0)F(x_0) + f(x)F'(x_0) \quad \text{tzn.} \quad (fF)'(x_0)h = [f'(x_0)h]F(x_0) + f(x)[F'(x_0)h].$$

Pochodna $f'(x_0)$ jest odwzorowaniem liniowym z $\mathbb{R}^n$ do $\mathbb{R}$. Działając na przyrost daje liczbę przez którą można mnożyć wartość odwzorowania F. Pochodna $F'(x_0)$ jest odwzorowaniem liniowym z $\mathbb{R}^n$ do $\mathbb{R}^m$. Działając na przyrost h daje element $\mathbb{R}^m$, który następnie może być pomnożony (każda jego współrzędna) przez wartość $f(x_0)$.

Przykład 5. Niech Φ będzie odwzorowaniem związanym z biegunowym układem współrzędnych. Pomnóżmy Φ przez funkcję $f(r,\varphi) = r\varphi$ i obliczmy pochodną. Zrobimy to na dwa sposoby: wprost i korzystając ze wzoru: Najpierw wprost:

$$f(r,\varphi)\Phi(r,\varphi) = r\varphi(r\cos\varphi, r\sin\varphi) = (r^2\varphi\cos\varphi, r^2\varphi\sin\varphi)$$

Pochodna ma zatem postać

$$\begin{bmatrix} 2r\varphi\cos\varphi & r^2(\cos\varphi - \varphi\sin\varphi) \\ 2r\varphi\sin\varphi & r^2(\sin\varphi + \varphi\cos\varphi) \end{bmatrix}.$$

Używając wzoru z punktu (2) twierdzenia musimy zauważyć, że pierwszy człon tzn $f'(x_0)F(x_0)$ należy zapisać w odwrotnej kolejności jeśli chce się korzystać z rachunku macierzowego. Najpierw bowiem działamy $f'(x_0)$ na przyrost a dopiero potem mnożymy przez $F(x_0)$. W naszym przykładzie mamy

$$f(r,\varphi) = r\varphi, \quad f'(r,\varphi) = [\varphi \;\; r].$$

pochodną mnożymy z lewej strony przez $\Phi(r,\varphi)$ zapisane jako wektor w kolumnie:

$$\begin{bmatrix} r\cos\varphi \\ r\sin\varphi \end{bmatrix} [\varphi \;\; r] = \begin{bmatrix} \varphi r\cos\varphi & r^2\cos\varphi \\ \varphi r\sin\varphi & r^2\sin\varphi \end{bmatrix}$$

Dodajemy drugi składnik sumy, tzn $f(x_0)F'(x_0)$:

$$\begin{bmatrix} \varphi r\cos\varphi & r^2\cos\varphi \\ \varphi r\sin\varphi & r^2\sin\varphi \end{bmatrix} + r\varphi \begin{bmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{bmatrix} = \begin{bmatrix} 2\varphi r\cos\varphi & r^2\cos\varphi - r^2\varphi\sin\varphi \\ 2\varphi r\sin\varphi & r^2\sin\varphi + r^2\varphi\cos\varphi \end{bmatrix}$$

Obie metody doprowadziły do tego samego wyniku. ♣

punkt (3) twierdzenia o formalnych prawach rożniczkowania mówi o różniczkowaniu złożenia.

Przykład 6. Rozważmy funkcję

$$f : \mathbb{R}^2 \to \mathbb{R}, \qquad f(x,y) = ax^2 + by^2, \quad a, b > 0$$

Jej pochodną w biegunowym układzie współrzędnych możemy obliczyć na dwa sposoby: wprost i korzystając ze wzoru na pochodną złożenia. Funkcja f w biegunowym układzie współrzędnych jest złożeniem:

$$f(\Phi(r,\varphi)) = f(r\cos\varphi, r\sin\varphi) = ar^2\cos^2\varphi + br^2\sin^2\varphi = r^2(a\cos^2\varphi + b\sin^2\varphi).$$

Otrzymaną funkcję różniczkujemy:

$$[2r(a\cos^2\varphi + b\sin^2\varphi), \;\; r^2(-2a\cos\varphi\sin\varphi + 2b\cos\varphi\sin\varphi)] =$$
$$[2r(a\cos^2\varphi + b\sin^2\varphi), \;\; r^2\sin(2\varphi)(b-a)].$$

Używając wzoru na pochodną złożenia obliczymy

$$\Phi'(r,\varphi) = \begin{bmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{bmatrix}, \qquad f'(x,y) = [2ax, \;\; 2by], \qquad f'(\Phi(r,\varphi) = [2ar\cos\varphi, \;\; 2br\sin\varphi],$$

i pomnożymy:

$$(f\circ\Phi)'(r,\varphi) = f'(\Phi(r,\varphi))\Phi'(r,\varphi) = [2ar\cos\varphi, \;\; 2br\sin\varphi] \cdot \begin{bmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{bmatrix} =$$
$$[2ar\cos^2\varphi + 2br\sin^2\varphi, \;\; -2ar^2\cos\varphi\sin\varphi + 2br^2\cos\varphi\sin\varphi] =$$
$$[2r(a\cos^2\varphi + b\sin^2\varphi), \;\; r^2\sin(2\varphi)(b-a)]$$

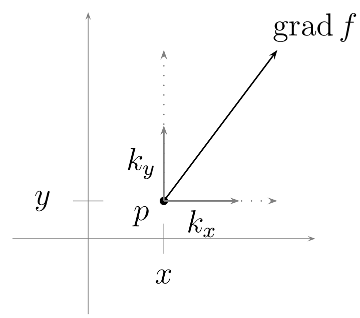
Obie metody doprowadziły do tego samego wyniku. ♣

Przykład 7 (Pojęcie gradientu funkcji)**.** Z całą pewnością byli już państwo zmuszeni na ćwiczeniach z fizyki używać gradientu funkcji określonej na $\mathbb{R}^2$, lub $\mathbb{R}^3$, zapisanego w biegunowym (w $\mathbb{R}^2$), walcowym i sferycznym (w $\mathbb{R}^3$) układzie współrzędnych. Pamiętam z moich czasów studenckich, że rachunki tego rodzaju sprawiały mi kłopoty. Później okazało się, że źródłem tych kłopotów był brak zrozumienia struktur geometrycznych, których używa się do definicji gradientu. Być może chodziło o to, że tak naprawdę pojęcie to zaliczane jest do dziedziny matematyki nazywanej *geometrią różniczkową*. My obracamy się raczej w ramach klasycznej analizy. Opisywanie układów fizycznych wymaga jednak używania wielu narzędzi matematycznych już na początkowym etapie nauki - nie możemy czekać aż odbędą państwo kurs geometrii różniczkowej.

Zastanówmy się więc najpierw co to jest gradient funkcji. Dla ustalenia uwagi będziemy pracować na $\mathbb{R}^2$. Przestrzeń tę potraktujemy jak przestrzeń afiniczną na której określimy funkcję $f : \mathbb{R}^2 \to \mathbb{R}$. Gradient $\operatorname{grad} f$ w punkcie $p = (x, y)$ jest wektorem zaczepionym w p o współrzędnych

$$\operatorname{grad} f(p) = \begin{bmatrix} \frac{\partial f}{\partial x}(x,y) \\ \frac{\partial f}{\partial y}(x,y) \end{bmatrix}$$

Przestrzeń wektorowa wektorów zaczepionych w punkcie p będzie oznaczana V. Jest to oczywiście także $\mathbb{R}^2$. Wprowadzamy jednak abstrakcyjne oznaczenie, gdyż w przestrzeni tej będziemy używać różnych baz. Oto stosowny rysunek:



$$\operatorname{grad} f = \frac{\partial f}{\partial x}(p) k_x + \frac{\partial f}{\partial y}(p) k_y$$

Składowe tego wektora względem bazy $k = (k_x, k_y)$ związanej z kartezjańskim układem współrzędnych są jednocześnie składowymi pochodnej funkcji f. Czasami myśli się więc o gradiencie jako o pochodnej. Jest to jednak błąd, który prowadzi do trudności w rachunkach. Gradient jest bowiem wektorem (elementem przestrzeni wektorowej) a pochodna jest odwzorowaniem liniowym na przestrzeni wektorowej. W zapisie widzimy także różnicę: gradient (wektor) jest „pionowy” a pochodna (odwzorowanie liniowe) jest „poziome”.

Przestrzeń wektorowa $\mathbb{R}^2$ (i każda $\mathbb{R}^n$) wyposażona jest w kanoniczny iloczyn skalarny. Zadaje on izomorfizm między przestrzenią wektorową $V = \mathbb{R}^2$ wektorów zaczepionych w p a przestrzenią funkcji liniowych na V, do której należy pochodna funkcji f. Funkcja liniowa odpowiadająca wektorowi $v \in V$ to funkcja

$$V \ni w \longmapsto (v \,|\, w) \in \mathbb{R}.$$

Można ją oznaczyć $(v \,|\, \cdot)$, gdzie kropka oznacza miejsce „czekające na argument”. Odwzorowanie między V a $L(V, \mathbb{R})$ jest izomorfizmem, zatem jest odwracalne: można mówić o funkcjach liniowych odpowiadających wektorom ale także o wektorach odpowiadających funkcjom liniowym. Gradient jest właśnie takim „wektorem odpowiadającym pochodnej”. Abstrakcyjny wzór wyglądałby tak:

$$f'(p) = (\operatorname{grad} f(p) \,|\, \cdot).$$

Jeśli formę dwuliniową będącą iloczynem skalarnym oznaczymy dodatkowo przez G (potrzebujemy jakąś literę do zapisywania macierzy) to w bazie k mamy:

$$[f'(p)]_k = ([\operatorname{grad} f(p)]^k)^T [G]_k,$$

a w innej, dowolnej bazie v:

$$[f'(p)]_v = ([\operatorname{grad} f(p)]^v)^T [G]_v.$$

Skupmy się na równaniu dla bazy v. Mamy wyznaczyć $[\operatorname{grad} f(p)]^v$. Transponujmy więc obie stony równania:

$$([f'(p)]_v)^T = ([G]_v)^T [\operatorname{grad} f(p)]^v = [G]_v [\operatorname{grad} f(p)]^v$$

Skorzystaliśmy także z faktu, że $[G]_v$ jest macierzą symetryczną, więc transpozycja nie wpływa na postać macierzy. Macierz $[G]_v$ jest odwracalna, więc

$$[\operatorname{grad} f(p)]^v = ([G]_v)^{-1} ([f'(p)]_v)^T. \tag{5}$$

Powyższy wzór może posłużyć do wyznaczenia postaci gradientu w dowolnym wybranym przez nas układzie współrzędnych.

Zajmijmy się teraz praktyczną stroną zapisywania gradientu we współrzednych biegunowych. Istnieją co najmniej dwie metody z których możemy skorzystać. Pierwsza z nich, którą nazywać będę *metodą zamiany zmiennych* jest następująca: Nowe zmienne, biegunowe, związane są z kartezjańskimi odwzorowaniem

$$\Phi : U \longrightarrow \mathbb{R}^2, \quad \mathbb{R}^2 \supset U = \{(r, \varphi) : \; r > 0, \varphi \in]0, 2\pi[\}.$$

$$\Phi(r, \varphi) = (r \cos\varphi, r \sin\varphi).$$

Różniczkując $f \circ \Phi$ jak funkcję złożoną dostajemy równania opisujące $\frac{\partial f}{\partial r}$, $\frac{\partial f}{\partial \varphi}$. Przyjęło się oznaczać wyjściową funkcję i tę złożoną z Φ tym samym symbolem, dla podkreślenia, że chodzi o ten sam obiekt, tylko zapisany w różnych układach współrzędnych.

$$[\frac{\partial f}{\partial r}, \;\; \frac{\partial f}{\partial \varphi}] = [(f \circ \Phi)'(r, \varphi)] =$$

$$[\frac{\partial f}{\partial x}(\Phi(r, \varphi)), \;\; \frac{\partial f}{\partial y}(\Phi(r, \varphi))] \begin{bmatrix} \cos\varphi & -r \sin\varphi \\ \sin\varphi & r \cos_0 \varphi \end{bmatrix} =$$

$$[\cos\varphi \frac{\partial f}{\partial x} + \sin\varphi \frac{\partial f}{\partial y}, \;\; -r \sin\varphi \frac{\partial f}{\partial x} + r \cos\varphi \frac{\partial f}{\partial y}].$$

Żeby wyznaczyć pochodne po x i y w zależności od pochodnych po r i φ należy rozwiązać układ równań:

$$\begin{cases} \dfrac{\partial f}{\partial r} = \cos\varphi \dfrac{\partial f}{\partial x} + \sin\varphi \dfrac{\partial f}{\partial y} \\ \dfrac{\partial f}{\partial \varphi} = -r\sin\varphi \dfrac{\partial f}{\partial x} + r\cos\varphi \dfrac{\partial f}{\partial y} \end{cases}$$

Otrzymujemy:

$$\begin{cases} \dfrac{\partial f}{\partial x} = \cos\varphi \dfrac{\partial f}{\partial r} - \dfrac{1}{r}\sin\varphi \dfrac{\partial f}{\partial \varphi} \\ \dfrac{\partial f}{\partial y} = \sin\varphi \dfrac{\partial f}{\partial r} + \dfrac{1}{r}\cos\varphi \dfrac{\partial f}{\partial \varphi} \end{cases}$$

Wstawiając to wszystko do wzoru na gradient otrzymujemy

$$\operatorname{grad} f \quad = \quad \frac{\partial f}{\partial x}k_x + \frac{\partial f}{\partial y}k_y \quad = \quad \left(\cos\varphi \frac{\partial f}{\partial r} - \frac{1}{r}\sin\varphi \frac{\partial f}{\partial \varphi}\right)k_x + \left(\sin\varphi \frac{\partial f}{\partial r} + \frac{1}{r}\cos\varphi \frac{\partial f}{\partial \varphi}\right)k_y$$

Dalej jeszcze musimy zamienić bazę na bazę ortonormalną związaną z biegunowym układem współrzędnych:

$$e_r = \cos\varphi k_x + \sin\varphi k_y \tag{6}$$

$$e_\varphi = -\sin\varphi k_x + \cos\varphi k_y \tag{7}$$

O tym, skąd wzięły się powyższe zależności będzie później. Wyznaczamy teraz k_x i k_y:

$$k_x = \cos\varphi e_r - \sin\varphi e_\varphi$$
$$k_y = \sin\varphi e_r + \cos\varphi e_\varphi$$

Wstawiamy wyniki do wzoru na gradient i wykonujemy działania:

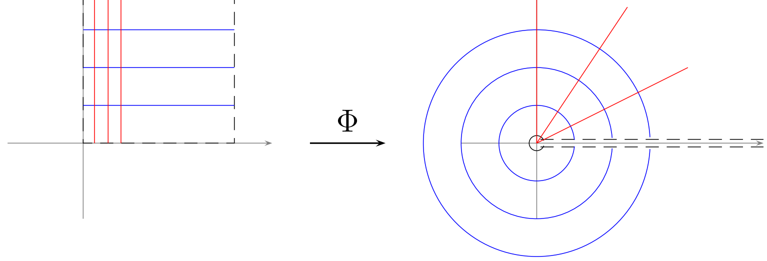
$$\operatorname{grad} f = \left(\cos\varphi \frac{\partial f}{\partial r} - \frac{1}{r}\sin\varphi \frac{\partial f}{\partial \varphi}\right)k_x + \left(\sin\varphi \frac{\partial f}{\partial r} + \frac{1}{r}\cos\varphi \frac{\partial f}{\partial \varphi}\right)k_y =$$

$$\left(\cos\varphi \frac{\partial f}{\partial r} - \frac{1}{r}\sin\varphi \frac{\partial f}{\partial \varphi}\right)(\cos\varphi e_r - \sin\varphi e_\varphi) + \left(\sin\varphi \frac{\partial f}{\partial r} + \frac{1}{r}\cos\varphi \frac{\partial f}{\partial \varphi}\right)(\sin\varphi e_r + \cos\varphi e_\varphi)$$

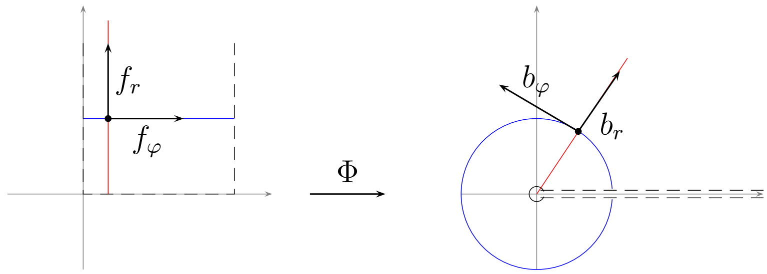
$$= \frac{\partial f}{\partial r}e_r + \frac{1}{r}\frac{\partial f}{\partial \varphi}e_\varphi.$$

Druga metoda polega na *skorzystaniu z definicji gradientu*, czyli użyciu wzoru (4) W przestrzeni A wprowadzamy współrzędne na dwa różne sposoby: kartezjańskie zadające w przestrzeni V bazę $k = (k_x, k_y)$ i biegunowe, z którymi związana jest baza $b = (b_r, b_\varphi)$. Przestrzeń V wyposażona jest w iloczyn skalarny, który w bazie k związanej ze współrzędnymi kartezjańskimi ma macierz równą macierzy jednostkowej.

Przyjrzyjmy się obrazkom ilustrującym sytuację: Biegunowy układ współrzędnych związany jest z odwzorowaniem Φ obciętym do zbioru $\{(r, \varphi) : r > 0, \varphi \in]0, 2\pi[\}$:



Odwzorowanie $\Phi'(r,\varphi)$ jest izomorfizmem między przestrzeniami wektorów zaczepionych w (r,φ) i w $(x,y)=\Phi(r,\varphi)$.



Baza b otrzymana jest z działania pochodnej $\Phi(r,\varphi)$ na wektory bazy standardowej $f=(f_r,f_\varphi)$ w przestrzeni wektorów zaczepionych w (r,φ) w dziedzinie odwzorowania Φ. Baza b składa się z wektorów prostopadłych do siebie, ale wektor b_φ nie jest długości 1. Sprawdźmy:

$$b_\varphi=\Phi'(r,\varphi)f_\varphi=\begin{bmatrix}\cos\varphi & -r\sin\varphi\\ \sin\varphi & r\cos\varphi\end{bmatrix}\begin{bmatrix}0\\1\end{bmatrix}=\begin{bmatrix}-r\sin\varphi\\ r\cos\varphi\end{bmatrix}=-r\sin\varphi e_x+r\cos\varphi e_y \tag{8}$$

Długość wektora b_φ liczymy korzystając z iloczynu skalarnego:

$$\|b_\varphi\|^2=(b_\varphi\,|\,b_\varphi)=(-r\sin\varphi e_x+r\cos\varphi e_y\,|\,-r\sin\varphi e_x+r\cos\varphi e_y)=r^2$$

Wektor b_φ jest długości r. Wektor b_r jest długości 1 co można sprawdzić zapisując go w bazie e:

$$b_r=\Phi'(r_0,\varphi_0)f_r=\begin{bmatrix}\cos\varphi & -r\sin\varphi\\ \sin\varphi & r\cos\varphi\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix}=\begin{bmatrix}\cos\varphi\\ \sin\varphi\end{bmatrix}=\cos\varphi e_x+\sin\varphi e_y \tag{9}$$

$$(b_r|b_r)=(\cos\varphi e_x+\sin\varphi e_y|\cos\varphi e_x+\sin\varphi e_y)=1$$

Równania (8) i (9) wraz z informacją, że $\|b_\varphi\|=r$ stanowią uzasadnienie dla użytych wcześniej wzorów (6), (7).

Pochodna funkcji f zapisana w bazie k ma postać

$$[f'(p)]_k=[\frac{\partial f}{\partial x},\ \frac{\partial f}{\partial y}]$$

zaś zapisana w bazie b związanej z biegunowym układem współrzędnych ma postać:

$$[f'(p)]_b=[\frac{\partial f}{\partial r},\ \frac{\partial f}{\partial \varphi}].$$

Wiedząc, że

$$[G]_b=\begin{bmatrix}1 & 0\\ 0 & r^2\end{bmatrix},\quad \text{więc}\quad [G]_b^{-1}=\begin{bmatrix}1 & 0\\ 0 & \frac{1}{r^2}\end{bmatrix}$$

wyznaczamy
$$[\text{grad}\, f]^b = \begin{bmatrix} 1 & 0 \\ 0 & \dfrac{1}{r^2} \end{bmatrix} \begin{bmatrix} \dfrac{\partial f}{\partial r} \\ \dfrac{\partial f}{\partial \varphi} \end{bmatrix} = \begin{bmatrix} \dfrac{\partial f}{\partial r} \\ \dfrac{1}{r^2}\dfrac{\partial f}{\partial \varphi} \end{bmatrix},$$
zatem
$$\text{grad}\, f = \frac{\partial f}{\partial r} b_r + \frac{1}{r^2}\frac{\partial f}{\partial \varphi} b_\varphi.$$
Zazwyczaj chcemy, aby baza, z którą pracujemy była nie tylko ortogonalna ale i ortonormalna. Zamiast bazy b musimy więc wziąć bazę $e' = (e_r, e_\varphi)$, gdzie $e_r = b_r$, $e_\varphi = \frac{1}{r} b_\varphi$:
$$\text{grad}\, f = \frac{\partial f}{\partial r} b_r + \frac{1}{r}\frac{\partial f}{\partial \varphi}\left(\frac{1}{r} b_\varphi\right) = \frac{\partial f}{\partial r} e_r + \frac{1}{r}\frac{\partial f}{\partial \varphi} e_\varphi.$$

Metoda korzystania z definicji wymagała od nas trochę rozważań abstrakcyjnych. Jednak teraz możemy bez dalszych rachunków zapisać gradient w jakichkolwiek współrzędnych. Potrzebujemy do tego jedynie macierzy iloczynu skalarnego zapisanej w bazie związanej z tym układem współrzędnych. Jeśli bazę tę oznaczymy v gradient uzyskujemy według wzoru (4):
$$[\text{grad}\, f(p)]^v = ([G]_v)^{-1}([f'(p)]_v)^T,$$
gdzie wektor $([f'(p)]_v)^T$ jest transpozycją pochodnej, składającej się z pochodnych cząstkowych funkcji f względem nowych współrzędnych. ♣