9.2 Pochodna kowariantna

W poprzednim podrozdziale zajmowaliśmy się różniczkowaniem pól tensorowych wzdłuż pola wektorowego, czyli pochodną Liego $\mathcal{L}_X$. Wartość pochodnej Liego zależy w sposób bardzo istotny od pola wektorowego wzdłuż którego różniczkujemy. Zależność ta jest poważniejsza niż tylko zależność od wartości pola w danym punkcie q - jest to zależność od tego jakie jest pole w otoczeniu q. Oczywiście pochodna Liego różniczkuje swój argument, to znaczy jeśli argument pomnożymy przez funkcję, wynik będzie zależał od pochodnych tej funkcji: Dla pola wektorowego otrzymamy

$$\mathcal{L}_X(fY) = [X, fY] = f[X,Y] + (Xf)Y = f\mathcal{L}_X Y + (\mathcal{L}_X f)Y \tag{31}$$

a dla formy α

$$\mathcal{L}_X(f\alpha) = (\mathcal{L}_X f)\alpha + f\mathcal{L}_X\alpha \tag{32}$$

Wzory (31) i (32) wyrażają *regułę Leibniza* dla pochodnej Liego. Okazuje się, że podobnie jest jeśli pomnożymy przez funkcję pole wzdłuż którego różniczkujemy. Porównajmy $\mathcal{L}_X$ z $\mathcal{L}_{fX}$ dla gładkiej funkcji f. Najpierw pochodna Liego pola wektorowego

$$\mathcal{L}_{fX}Y = [fX, Y] = f[X,Y] - (Yf)X = f\mathcal{L}_X Y - \langle \mathsf{d}f, Y\rangle X \tag{33}$$

potem pochodna formy

$$\begin{aligned} \mathcal{L}_{fX}\alpha &= \imath(fX)\mathsf{d}\alpha + \mathsf{d}(\imath(fX)\alpha) = f\imath(X)\mathsf{d}\alpha + \mathsf{d}(f\imath(X)\alpha) = \\ &= f\imath(X)\mathsf{d}\alpha + \mathsf{d}f\wedge\imath(X)\alpha + f\mathsf{d}\imath(X)\alpha = f\mathcal{L}_X\alpha + \mathsf{d}f\wedge\imath(X)\alpha \end{aligned} \tag{34}$$

W obu wzorach (33) i (34) pojawia się różniczka funkcji f, co wskazuje na zależność od wartości pola w otoczeniu q, a nie jedynie w punkcie q. My jednak cały czas poszukujemy odpowiednika różniczkowania po współrzędnych, które jest możliwe na $\mathbb{R}^n$, lub ogólniej, na przestrzeni afinicznej. We współrzędnych (x^a) związanych z punktem q_0 i bazą (e_a) w przestrzeni modelowej V różniczkowanie to wygląda następująco

$$D_v(X) = v^a(\partial_a X^b)e_b, \tag{35}$$

gdzie $v = v^a e_a$ oraz $X = X^b e_b$. Wzór jest poprawny, tzn jest niezmienniczy ze względu zmianę bazy w V i zmianę punktu początkowego a_0. Wzór (35) jest poprawny dlatego, że wiązka styczna do przestrzeni afinicznej jest trywialna: przestrzeń styczna w każdym punkcie jest kanonicznie izomorficzna z V. Jeszcze inaczej można powiedzieć, że wzór (35) jest poprawny ponieważ wiemy co to znaczy przesuwać wektor styczny wzdłuż krzywej w A. Rozważając powierzchnię $M \subset A$ nie możemy użyć wzoru (35), bo na powierzchni nie mamy wyróżnionej klasy układów współrzędnych. Jeśli zaś użyjemy dowolnych współrzędnych wzór przestanie być niezmienniczy, czyli nie będzie definiował żadnego obiektu geometrycznego. Pole wektorowe $X : M \to \mathsf{T}M$ możemy oczywiście zapisać w bazie (e_a) otrzymując $X = X^a e_a$ (zatem na oko podobnie jak poprzednio), z tą jednak różnicą, że funkcje X^a określone są jedynie w punktach należących do M. Możemy także, korzystając z zanurzenia M w A wyliczyć $D_v X$ jeśli tylko $v \in \mathsf{T}M$. Problem w tym, że otrzymany wektor najprawdopodobniej nie jest styczny do M. Z tego da się wybrnąć jeśli w V mamy wyróżniony iloczyn skalarny. (Tak jest oczywiście jeśli $A = \mathbb{R}^n$.)

Iloczyn skalarny pozwala rozłożyć przestrzeń $\mathsf{T}_qA = V$ styczną do A w punkcie q na sumę prostą przestrzeni $\mathsf{T}_qM = W_q$ i $W_q^\perp$, tzn

$$V = W_q \oplus W_q^\perp \tag{36}$$

Możemy teraz napisać definicję

Definicja 28 Niech M będzie powierzchnią zanurzoną w przestrzeni afinicznej A z wyróżnionym iloczynem skalarnym. *Pochodną kowariantną* pola wektorowego X na M w kierunku wektora stycznego $v \in \mathsf{T}_qM$ nazywamy wektor

$$\nabla_v X = (D_v X)^\parallel, \tag{37}$$

gdzie $(\cdot)^\parallel$ oznacza rzut na W_q związany z rozkładem (36).

W powyższej definicji współrzędne nie są używane, nie ma więc wątpliwości, że (37) definiuje pewien wektor z W_q. Wyrażenie na współrzędnych też będzie nam potrzebne. Wprowadźmy w tym celu w otoczeniu q w M układ współrzędnych (φ^i), $i = 1 \dots k$. Wektory ∂_i tworzą więc bazę W_q. Potrzebujemy jeszcze $n-k$ wektorów rozpinających $W_q^\perp$. Oznaczymy je n_α, $\alpha = 1 \dots n-k$. Każdy z wektorów e_a zapisać możemy w bazie złożonej z ∂_i oraz n_α otrzymując

$$e_a = A^i_a \partial_i + A^\alpha_a n_\alpha$$

Macierz A, której pierwsze k wierszy to A^i_a a dalsze $n-k$ wierszy to A^α_a jest macierzą przejścia z bazy (e_a) do bazy (∂_i, n_α). Wyrazy macierzowe A^i_a i A^α_a są funkcjami na M, można je więc wyrazić jako funkcje współrzędnych (φ^i). Pole wektorowe X zapisujemy w bazie e_a: $X = X^a(\varphi^i)e_a$, wektor v zapisujemy w bazie (∂_i) i wyznaczamy

$$D_v X = v^i(\partial_i X^a)e_a.$$

Żeby wziąć część styczną do M używamy macierzy A:

$$D_v X = v^i(\partial_i X^a)(A^j_a \partial_j + A^\alpha_a n_\alpha).$$

Widać więc, że

$$\nabla_v X = v^i(\partial_i X^a)A^j_a \partial_j.$$

W powyższym wzorze nie podoba mi się jszcze to, że pole X cały czas mamy wyrażone w bazie e_a a nie w bazie ∂_i, która jest dla niego naturalniejsza.

$$X^a = (A^{-1})^a_i X^i,$$

zatem

$$\begin{aligned}\nabla_v X = v^i\partial_i((A^{-1})^a_k X^k)A^j_a\partial_j = v^i\partial_i((A^{-1})^a_k)X^k A^j_a\partial_j + v^i(A^{-1})^a_k\partial_i(X^k)A^j_a\partial_j = \\ \partial_i(X^j)v^i\partial_j + [\partial_i((A^{-1})^a_k)A^j_a]v^i X^k \partial_j\end{aligned} \tag{38}$$

Zauważmy, że fragment wzoru zapisany na niebiesko nie zależy wcale od pola X ani od kierunku w którym różniczkujemy. Jest on związany z powierzchnią na której o wszystko się dzieje. Niebieskie funkcje oznaczamy Γ^j_{ik}, tzn

$$\Gamma^j_{ik} = \partial_i((A^{-1})^a_k)A^j_a \tag{39}$$

i nazywamy *symbolami Christoffela* (Erwin Bruno Christoffel (1829-1900), matematyk niemiecki). Z użyciem symboli Christoffela pochodną kowariantną zapisujemy następująco

$$\nabla_v X = \partial_i(X^j)v^i + \Gamma^j_{ik} v^i X^k. \tag{40}$$

Przykład 18 Policzmy (niektóre) symbole Christoffela na sferze $S^2 \subset \mathbb{R}^3$. Mamy $n = 3$, $k = 2$, współrzedne x^a to (x, y, z), współrzędne (φ^i) to (φ, θ) pochodzące ze sferycznego układu współrzędnych. Baza e_a to $(\partial_x, \partial_y, \partial_z)$, baza (∂_i) to $(\partial_\varphi, \partial_\theta)$, jako wektor normalny możemy wziąć ∂_r. Posługując się znanymi wzorami

$$\begin{aligned} x &= r\cos\varphi\sin\theta \\ y &= r\sin\varphi\sin\theta \\ z &= r\cos\theta \end{aligned}$$

i przyjmując $r = 1$ najdujemy wyrazy macierzowe macierzy A^{-1}

$$\begin{aligned} \partial_\varphi &= -\sin\varphi\sin\theta\partial_x + \cos\varphi\sin\theta\partial_y \\ \partial_\theta &= \cos\varphi\cos\theta\partial_x + \sin\varphi\cos\theta\partial_y - \sin\theta\partial_z \\ \partial_r &= \cos\varphi\sin\theta\partial_x + \sin\varphi\sin\theta\partial_y + \cos\theta\partial_z \end{aligned}$$

Macierz A^{-1} ma postać

$$A^{-1} = \begin{bmatrix} -\sin\varphi\sin\theta & \cos\varphi\cos\theta & \cos\varphi\sin\theta \\ \cos\varphi\sin\theta & \sin\varphi\cos\theta & \sin\varphi\sin\theta \\ 0 & -\sin\theta & \cos\theta \end{bmatrix},$$

zatem

$$A = \begin{bmatrix} -\dfrac{\sin\varphi}{\sin\theta} & \dfrac{\cos\varphi}{\sin\theta} & 0 \\ \cos\varphi\cos\theta & \sin\varphi\cos\theta & -\sin\theta \\ \cos\varphi\sin\theta & \sin\varphi\sin\theta & \cos\theta \end{bmatrix}.$$

Skorzystamy ze wzoru (39) zauważając, że można go zapisać na dwa sposoby:

$$\Gamma^j_{ik} = \partial_i((A^{-1})^a_k)A^j_a = -\partial_i(A^j_a)(A^{-1})^a_k.$$

Przykładowe rachunki:

$$\Gamma^\varphi_{\varphi\varphi} = -\partial_\varphi\left(-\frac{\sin\varphi}{\sin\theta}\right)(-\sin\varphi\sin\theta) - \partial_\varphi\left(\frac{\cos\varphi}{\sin\theta}\right)(-\cos\varphi\sin\theta) = \\ -\cos\varphi\sin\varphi + \cos\varphi\sin\varphi = 0$$

$$\Gamma^\varphi_{\theta\varphi} = -\partial_\varphi\left(-\frac{\sin\varphi}{\sin\theta}\right)(\cos\varphi\cos\theta) - \partial_\varphi\left(\frac{\cos\varphi}{\sin\theta}\right)(\sin\varphi\cos\theta) = \\ \cos^2\varphi\cot\theta + \sin^2\varphi\cot\theta = \cot\theta$$

Pozostałe współczynniki Christoffela na sferze proszę wyznaczyć samodzielnie. Żeby oszczędzić sobie rachunków można skorzystać z faktu (który udowodnimy wkrótce), że

Fakt 17 *Współczynniki Christoffela na powierzchni zanurzonej są symetryczne, tzn* $\Gamma^i_{jk} = \Gamma^i_{kj}$.

Dowód: Symetria wpółczynników Christoffela dla pochodnej kowariantnej pochodzącej od zanurzenia powierzchni w przestrzeń afiniczną z iloczynem skalarnym wynika z symetrii drugich pochodnych cząstkowych. Istotnie, wyrazy macierzowe występujące we wzorze (39) pochodzą od zamiany zmiennych, tzn

$$(A^{-1})^a_k = \frac{\partial x^a}{\partial \varphi^k}.$$

Wstawiając powyższą postać wyrazów macierzowych $(A^{-1})^a_k$ do (39) otrzymujemy

$$\Gamma^j_{ik} = \partial_i \left(\frac{\partial x^a}{\partial \varphi^k} \right) A^j_a = \frac{\partial^2 x^a}{\partial \varphi^i \partial \varphi^k} A^j_a = \partial_k \left(\frac{\partial x^a}{\partial \varphi^i} \right) A^j_a = \Gamma^j_{ki} \quad \square$$

Fakt 19 *Pochodna kowariantna i przesunięcie równoległe na powierzchni zanurzonej mają charakter wewnętrzny, tzn zależą jedynie od iloczynu skalarnego (metryki) indukowanej na powierzchni.*

Dowód: Zaczniemy od wyliczenia $\nabla_v g$, gdzie v jest dowolnym wektorem stycznym do M a g jest formą dwuliniową symetryczną reprezentującą iloczyn skalarny, czyli inaczej tensorem metrycznym na M. Tensor g pochodzi z obcięcia iloczynu skalarnego z $\mathsf{T}A$. W bazie (e_a) macierz g ma postać $g_{ab} = (e_a|e_b)$. Wyrazy macierzowe g_{ab} są stałe, tzn. nie zależą od punktu na powierzchni. Obcięcie iloczynu skalarnego można zapisać także w bazie związanej z układem współrzędnych na powierzchni. Wtedy $g_{ij} = (\partial_i|\partial_j)$. Wyrazy macierzowe g_{ij} są funkcjami współrzędnych (φ^i). W notacji tensorowej iloczyn skalarny zapiszemy jako

$$g = g_{ij} \mathsf{d}\varphi^i \otimes \mathsf{d}\varphi^j.$$

Do tej pory liczyliśmy pochodną kowariantną jedynie pól wektorowych. Można ją jednak rozszerzyć (poprzez regułę Leibniza) na dowolne pola tensorowe. Jak pochodna kowariantna działa na formy (pola kowektorowe)? Niech $\alpha = \alpha_i \mathsf{d}\varphi^i$ będzie polem kowektorowym a $X = X^i \partial_i$ polem wektorowym. Wtedy $\langle \alpha, X \rangle$ jest funkcją na M. Pochodna kowariantna funkcji jest pochodną zwykłą, czyli działaniem wektora stycznego na funkcję

$$\nabla_v \langle \alpha, X \rangle = v^i \partial_i (\alpha_j X^j) = v^i (\partial_i \alpha_j) X^j + v^i \alpha_j (\partial_i X^j). \tag{46}$$

Z drugiej strony, zgodnie z regułą Leibniza powinno być

$$\nabla_v \langle \alpha, X \rangle = \langle \nabla_v \alpha, X \rangle + \langle \alpha, \nabla_v X \rangle = v^i (\nabla_i \alpha)_j X^j + v^i \alpha_j (\partial_i X^j + \Gamma^j_{ik} X^k) \tag{47}$$

Z porównania (46) i (47) wynika, że

$$v^i (\partial_i \alpha_j) X^j = v^i (\nabla_i \alpha)_j X^j + v^i \alpha_j \Gamma^j_{ik} X^k,$$

czyli

$$v^i (\nabla_i \alpha)_k X^k = v^i (\partial_i \alpha_k) X^k - v^i \alpha_j \Gamma^j_{ik} X^k.$$

Ostatecznie

$$(\nabla_i \alpha) = \left(\partial_i \alpha_k - \alpha_j \Gamma^j_{ik} \right) \mathsf{d}\varphi^k. \tag{48}$$

Dalej, korzystając cały czas z reguły Leibniza, otrzymujemy

$$\nabla_i (g_{jk} \mathsf{d}\varphi^j \otimes \mathsf{d}\varphi^k) = (\partial_i g_{jk}) \mathsf{d}\varphi^j \otimes \mathsf{d}\varphi^k + g_{jk} \mathsf{d}(\nabla_i \mathsf{d}\varphi^j) \otimes \mathsf{d}\varphi^k + g_{jk} \mathsf{d}\varphi^j \otimes (\nabla_i \mathsf{d}\varphi^k).$$

Wynika z tego, że

$$(\nabla_i g)_{jk} = \partial_i g_{jk} - \Gamma^l_{ij} g_{lk} - \Gamma^l_{ik} g_{jl}. \tag{49}$$