8 Twierdzenie Schönhage

8.1 Iloczyn kartezjański wielu przestrzeni liniowych

Iloczyn kartezjański dwóch przestrzeni został zdefiniowany w rozdziale 3.3. Teraz będziemy rozważać iloczyny kartezjańskie dowolnej (skończonej) rodziny przestrzeni liniowych. Jeżeli $(U_t)_{t\in T}$ jest rodziną przestrzeni liniowych indeksowanych elementami zbioru T, to iloczyn (produkt) kartezjański tej rodziny będziemy oznaczać symbolem $\prod_{t\in T} U_t$. Elementy tej przestrzeni są funkcjami określonymi w zbiorze T, które indeksowi $t \in T$ przyporządkowują element ze zbioru U_t. Funkcje takie dodajemy, dodając ich wartości dla poszczególnych indeksów, i mnożymy przez c, mnożąc przez c wszystkie ich wartości. Funkcje należące do $\prod_{t\in T} U_t$ często przedstawiamy jako ciągi podając ich wartości w ustalonej kolejności. W takim przypadku wartość funkcji f dla argumentu t będziemy współrzędną, wyrazem lub pozycją (ciągu f) dla indeksu t.

Aby utworzyć bazę przestrzeni $\prod_{t\in T} U_t$, bierzemy bazy wszystkich przestrzeni U_t, i dla każdego $t \in T$ tworzymy elementy produktu, których współrzędne dla indeksu t są dowolnymi elementami bazy U_t, a pozostałe współrzędne są równe 0. Zbiór takich elementów jest bazą iloczynu $\prod_{t\in T} U_t$. Tak więc wymiar produktu $\prod_{t\in T} U_t$ jest sumą wymiarów przestrzeni U_t dla $t \in T$.

Najbardziej znanym przykładem produktu jest przestrzeń $\prod_{i=1}^{n} k = k^n$ (produkt n przestrzeni liniowych, z których każda jest ciałem k). Ciało k jest przestrzenią wymiaru 1. Bazę standardową w ciele k tworzy pojedyńczy element k równy 1. Bazę produktu $\prod_{i=1}^{n} k$ tworzy biorąc ciągi, które na jednej pozycji mają element bazy z ciała k (a więc 1), a na pozostałych pozycjach – 0.

Przypuśćmy, że mamy dwie przestrzenie liniowe U_1 i U_2 i liczbę $n \in N$. Będziemy zajmować się produktami postaci

$$\mathcal{U}_n = \prod_{\varphi:\{1,\ldots,n\}\to\{1,2\}} U_{\varphi(1)} \otimes U_{\varphi(2)} \otimes \ldots \otimes U_{\varphi(n)}$$

Są to produkty indeksowane funkcjami $\varphi : \{1, \ldots, n\} \to \{1, 2\}$ Aby skonstruować ten produkt, n-krotnie wybieramy jedną z dwóch danych przestrzeni (funkcje f zawierają informacje o tym, jak wybieraliśmy), tworzymy produkt tensorowy wybranych przestrzeni, a następnie tworzymy iloczyn kartezjański wszystkich otrzymanych w ten sposób przestrzeni.

Produkt $\mathcal{U}_3$ został schematycznie przedstawiony na rysunku 4. Na tym rysunku, w poszczególnych wierszach są wymienione wszystkie „produktowane” przestrzenie. Np. w wierszu czwartym znajduje się przestrzeń $U_1 \otimes U_2 \otimes U_2$, przestrzeń ta odpowiada indeksowi (funkcji) φ takiemu, że $\varphi(1) = 1$, $\varphi(2) = 2$ i $\varphi(3) = 2$.

8.2 Potęga tensorowa iloczynu kartezjańskiego

Dla danych przestrzeni U_1 i U_2, i liczby n pokażemy, że produkt

$$\mathcal{U}_n = \prod_{\varphi:\{1,\ldots,n\}\to\{1,2\}} U_{\varphi(1)} \otimes U_{\varphi(2)} \otimes \ldots \otimes U_{\varphi(n)}$$

jest iloczynem tensorowym n egzemplarzy przestrzeni $U_1 \times U_2$ (jest n-tą potęgą tensorową przestrzeni $U_1 \times U_2$). Aby to zrobić, zdefiniujemy przekształcenie n-liniowe mające razem z podaną przestrzenią własność jednoznacznej uniwersalnej faktoryzacji.

Dowód tego faktu zostanie podany w przypadku $n = 3$. W tym przypadku mamy zdefiniować przekształcenie

$$\otimes : (U_1 \times U_2) \times (U_1 \times U_2) \times (U_1 \times U_2) \to \prod_{\varphi:\{1,\ldots,n\}\to\{1,2\}} U_{\varphi(1)} \otimes U_{\varphi(2)} \otimes U_{\varphi(3)}$$

$$
\boxed{\begin{array}{llllll}
U_1 & & \otimes U_1 & & \otimes U_1 & \\
U_1 & & \otimes U_1 & & \otimes & U_2 \\
U_1 & & \otimes & U_2 \otimes U_1 & & \\
U_1 & & \otimes & U_2 \otimes & & U_2 \\
 & U_2 \otimes U_1 & & & \otimes U_1 & \\
 & U_2 \otimes U_1 & & & \otimes & U_2 \\
 & U_2 \otimes & & U_2 \otimes U_1 & & \\
 & U_2 \otimes & & U_2 \otimes & & U_2
\end{array}}
$$

Rysunek 4: Pewien produkt iloczynów tensorowych

Na rysunku 5 jest przedstawiona dziedzina i zbiór wartości iloczynu tensorowego, który mamy zdefiniować. Na nim jest także przedstawiona definicja przekształcenia $\otimes$. Aby zobaczyć tę definicję, wystarczy przyjąć, że U_1 i U_2 nie są przestrzeniami liniowymi, a są elementami tych przestrzeni. Tak więc, aby dla $u_{1,1}, u_{1,2}, u_{1,3} \in U_1$ oraz $u_{2,1}, u_{2,2}, u_{2,3} \in U_2$ wyliczyć wartość

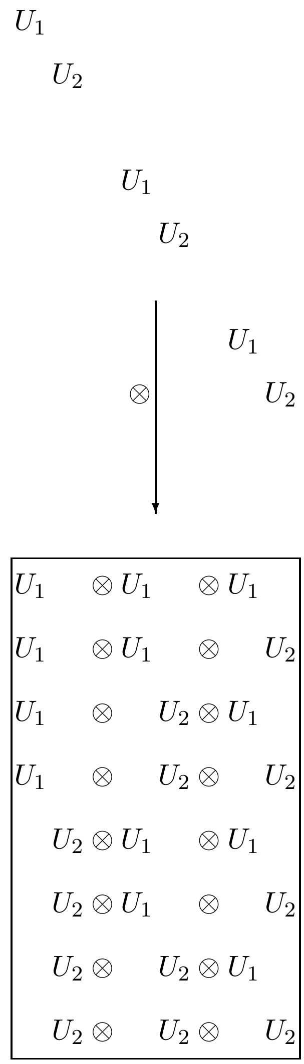
$$(u_{1,1}, u_{2,1}) \otimes (u_{1,2}, u_{2,2}) \otimes (u_{1,3}, u_{2,3}),$$

należy na wszystkie możliwe sposoby wybrać z trzech podanych par po jednej współrzędnej, wyliczyć iloczyn tensorowy wybranych współrzędnych i wynik umieścić w odpowiednim miejscu liczonej wartości. Jeżeli wybraliśmy wartości wskazane przez funkcję φ, to obliczamy $u_{\varphi(1),1} \otimes u_{\varphi(2),2} \otimes u_{\varphi(3),3}$ i tę wartość umieszczamy na pozycji odpowiadającej indeksowi φ. Zauważmy, że znak $\otimes$ we wzorze $u_{\varphi(1),1} \otimes u_{\varphi(2),2} \otimes u_{\varphi(3),3}$ oznacza mnożenie tensorowe przyjmujące wartości w przestrzeni $U_{\varphi(1)} \otimes U_{\varphi(2)} \otimes U_{\varphi(3)}$, a nie definiowane mnożnie tensorowe.

Jeżeli jeden z argumentów przekształcenia wieloliniowego jest równy 0, to wartość przekształcenia jest równa 0. Tak więc

$$0 \otimes u_{\varphi(2),2} \otimes u_{\varphi(3),3} = u_{\varphi(1),1} \otimes 0 \otimes u_{\varphi(3),3} = u_{\varphi(1),1} \otimes u_{\varphi(2),2} \otimes 0 = 0.$$

Weźmy teraz trzy pary $(u_{1,1}, u_{2,1}), (u_{1,2}, u_{2,2}), (u_{1,3}, u_{2,3})$ z bazy standardowej przestrzeni $U_1 \times U_2$. W każdej parze dokładnie jeden z elementów jest równy 0. Przyjmijmy, że $u_{\psi(i),i} \neq 0$ dla $i = 1, 2, 3$. Nietrudno zauważyć, że jeżeli $\varphi \neq \psi$, to przynajmnij jeden z czynników w wyrażeniu $u_{\varphi(1),1} \otimes u_{\varphi(2),2} \otimes u_{\varphi(3),3}$ jest równy 0 i całe wyrażenie też jest równe 0. W przeciwnym razie mamy oczywiście, że $u_{\varphi(1),1} \otimes u_{\varphi(2),2} \otimes u_{\varphi(3),3} = u_{\psi(1),1} \otimes u_{\psi(2),2} \otimes u_{\psi(3),3}$. Wobec tego, w ciągu $(u_{1,1}, u_{2,1}) \otimes (u_{1,2}, u_{2,2}) \otimes (u_{1,3}, u_{2,3})$, wszystkie współrzędne z wyjątkiem wyznaczonej przez ψ są równe 0, a ta ostatnia jest elementem bazy standardowej przestrzeni $U_{\psi(1)} \otimes U_{\psi(2)} \otimes U_{\psi(3)}$. Tak więc zdefiniowane przekształcenie przeprowadza trójki wektorów bazowych na elementy bazy. Jest też oczywiste, że w ten sposób można otrzymać wszystkie układy bazowe z przestrzeni $\mathcal{U}_3$. Porównując wymiary tego produktu i przestrzeni $U_1 \times U_2$ stwierdzamy, że zdefiniowane przekształcenie $\otimes$ jest różnowartościowe na trójkach wektorów bazowych. Teraz własność jednoznacznej uniwersalnej faktoryzacji wynika z lematu 3.10.



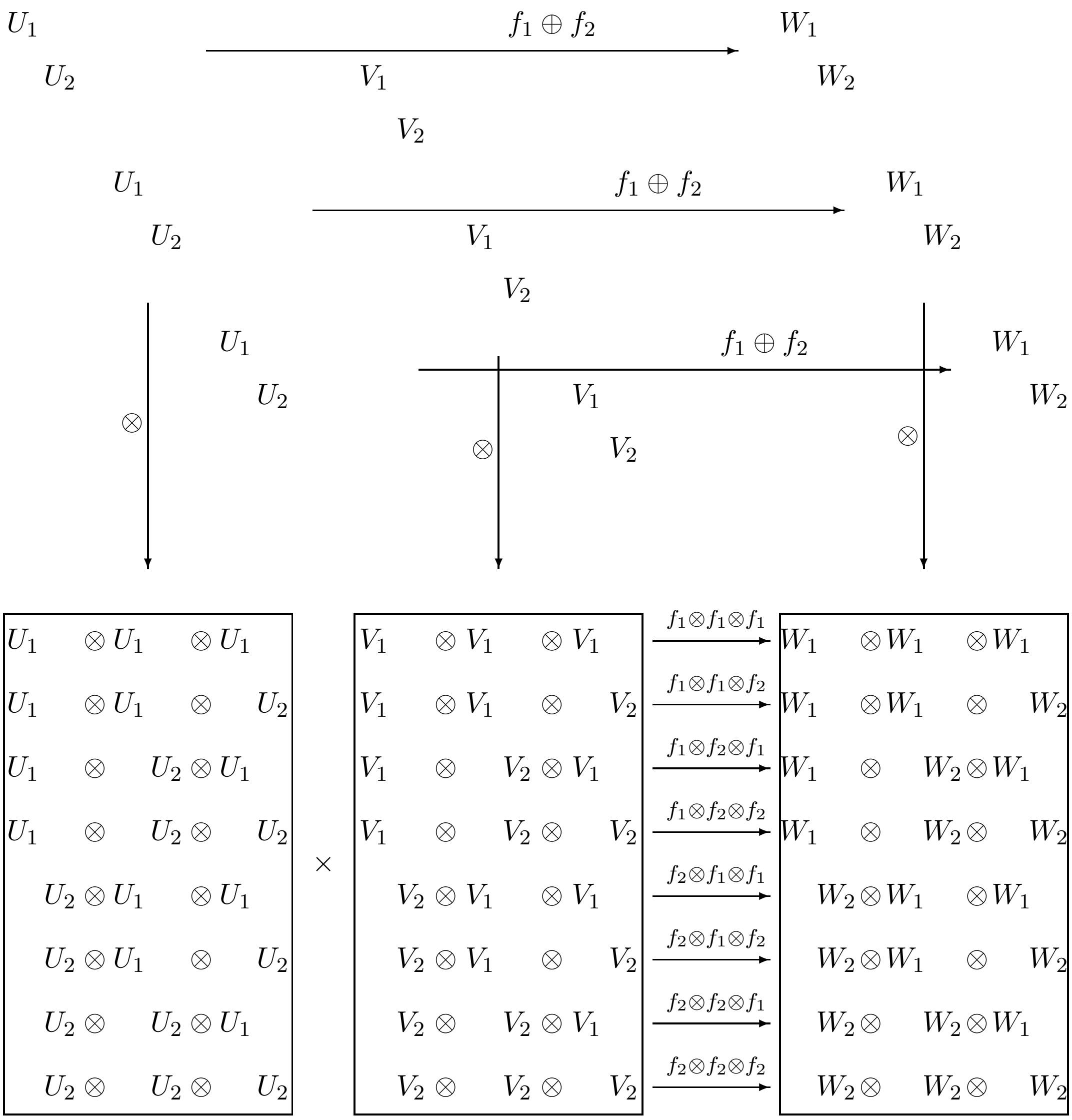
Rysunek 5: Potęga tensorowa iloczynu kartezjańskiego

8.3 Potęga tensorowa sumy prostej przekształceń dwuliniowych

Przypuśćmy, że mamy dwa przekształcenia dwuliniowe $f_1 : U_1 \times V_1 \to W_1$ oraz $f_2 : U_2 \times V_2 \to W_2$. Suma prosta tych przekształceń $f_1 \oplus f_2$ jest funkcją przekształcającą $f_1 \oplus f_2 : (U_1 \times U_2) \times (V_1 \times V_2) \to W_1 \times W_2$. Będziemy obliczać n-tą potęgę tensorową $(f_1 \oplus f_2) \otimes \ldots \otimes (f_1 \oplus f_2)$. Obliczenia przeprowadzimy przy założeniu $n = 3$. Będzie więc interesować nas funkcja $(f_1 \oplus f_2) \otimes (f_1 \oplus f_2) \otimes (f_1 \oplus f_2)$. W dolnej części rysunku 6 została zdefiniowana funkcja $\Psi : \mathcal{U}_3 \times \mathcal{V}_3 \to \mathcal{W}_3$. Funkcja ta jest pewną sumą prostą:

$$\Psi = \bigoplus_{\varphi:\{1,\ldots,3\}\to\{1,2\}} f_{\varphi(1)} \otimes f_{\varphi(2)} \otimes f_{\varphi(3)}.$$

Inaczej, aby wyliczyć wartość $\Psi(X, Y) \in \mathcal{W}_3$, wystarczy wyliczyć wszystkie współrzędne $\Psi(X, Y)$; aby wyliczyć współrzędną o indeksie φ tej wartości, należy wziąć współrzędne X_φ i Y_φ o indeksie φ



Rysunek 6: Potęga tensorowa sum prostych przekształceń

argumentów X i Y, i wyliczyć

$$(\Psi(X,Y))_\varphi = f_{\varphi(1)} \otimes f_{\varphi(2)} \otimes f_{\varphi(3)}(X_\varphi, Y_\varphi).$$

Pokażemy, że diagram z rysunku 6 jest przemienny. W ten sposób dowiedziemy, że $\Psi = (f_1 \oplus f_2) \otimes \ldots \otimes (f_1 \oplus f_2)$ (patrz rozdział 3.2.3).

Przemienność diagramu wystarczy sprawdzić dla układów wektorów bazowych. Weźmy wiec dowolny układ trzech wektorów bazowych $(u_{1,1}, u_{2,1}), (u_{1,2}, u_{2,2}), (u_{1,3}, u_{2,3})$ z przestrzeni $U_1 \times U_2$ i podobny układ $(v_{1,1}, v_{2,1}), (v_{1,2}, v_{2,2}), (v_{1,3}, v_{2,3})$ z przestrzeni $V_1 \times V_2$. W każdej wybranej parze dokładnie jeden z wektorów jest zerowy. Przyjmijmy, że $u_{\gamma(i),i} \neq 0$ i $v_{\eta(i),i} \neq 0$ dla $i = 1, 2, 3$. Dalej będziemy „przechodzić" przez diagram na dwa sposoby.

Iloczyn tensorowy $X = (u_{1,1}, u_{2,1}) \otimes (u_{1,2}, u_{2,2}) \otimes (u_{1,3}, u_{2,3})$ jest ciągiem, którego prawie wszystkie wyrazy są równe 0. Jedynie wyraz o indeksie γ jest równy $u_{\gamma(1),1} \otimes u_{\gamma(2),2} \otimes u_{\gamma(3),3}$ (patrz rozdział 8.2). Podobnie jest w przypadku iloczynu $Y = (v_{1,1}, v_{2,1}) \otimes (v_{1,2}, v_{2,2}) \otimes (v_{1,3}, v_{2,3})$, którego jedyny różny od 0 wyraz (o indeksie η) jest równy $v_{\eta(1),1} \otimes v_{\eta(2),2} \otimes v_{\eta(3),3}$.

Obliczając $\Psi(X,Y)$ ustalamy wyrazy X_φ i Y_φ o indeksie φ ciągów X i Y, a następnie obliczamy $f_{\varphi(1)} \otimes f_{\varphi(2)} \otimes f_{\varphi(3)}(X_\varphi, Y_\varphi)$. Obliczamy wartość pewnego przekształcenia dwuliniowego, a więc jeżeli $X_\varphi = 0$ lub $Y_\varphi = 0$, to także $f_{\varphi(1)} \otimes f_{\varphi(2)} \otimes f_{\varphi(3)}(X_\varphi, Y_\varphi) = 0$. Wobec tego, prawie wszystkie współrzędne $\Psi(X,Y)$ są równe 0. W szczególności, jeżeli $\gamma \neq \eta$, to $\Psi(X,Y)$ jest wektorem zerowym, gdyż niezerowe wyrazy ciągów X i Y w tym przypadku znajdują się na różnych pozycjach. Jeżeli $\gamma = \eta$, to niezerowe współrzędne X i Y mają indeks γ i wszystkie wyrazy $\Psi(X,Y)$ są równe 0 z wyjątkiem wyrazu o indeksie γ, który jest równy

$$f_{\gamma(1)} \otimes f_{\gamma(2)} \otimes f_{\gamma(3)}(u_{\gamma(1),1} \otimes u_{\gamma(2),2} \otimes u_{\gamma(3),3}, v_{\gamma(1),1} \otimes v_{\gamma(2),2} \otimes v_{\gamma(3),3}). \tag{32}$$

Drugi sposób przechodzenia diagramu wymaga najpierw obliczenia

$$X_i = (f_1 \oplus f_2)((u_{1,i}, u_{2,i}), (v_{1,i}, v_{2,i})).$$

Jeżeli $\gamma(i) = 1$, to

$$X_i = (f_1(u_{1,i}, v_{1,i}), f_2(u_{2,i}, v_{2,i})) = (f_1(u_{1,i}, v_{1,i}), f_2(0, v_{2,i})) = (f_1(u_{1,i}, v_{1,i}), 0).$$

W przeciwnym razie

$$X_i = (0, f_2(u_{2,i}, v_{2,i})).$$

Tak więc jedna współrzędna pary X_i jest równa 0, a jeżeli $X_i \neq 0$, to $\gamma(i)$ jest numerem niezerowej współrzędnej tej pary. Z tych samych powodów, tylko $\eta(i)$ może być numerem niezerowej współrzędnej X_i. Ostatecznie otrzymujemy, że jeżeli $\gamma(i) \neq \eta(i)$, to $X_i = (0,0)$, a w przeciwnym razie para X_i ma współrzędna o numerze $\gamma(i)(= \eta(i))$ równą $f_{\gamma(i)}(u_{\gamma(i),i}, v_{\gamma(i),i})$ i druga współrzędną równa 0.

Teraz już łatwo wyliczyć iloczyn $X_1 \otimes X_2 \otimes X_3$. Jeżeli $\gamma \neq \eta$, to jeden z czynników tego iloczynu jest równy 0 i cały iloczyn też jest równy 0. Tak więc w tym przypadku diagram z rysunku 6 jest przemienny.

W przeciwnym razie (gdy $\gamma \neq \eta$), obliczamy $X_1 \otimes X_2 \otimes X_3$ tak, jak w rozdziale 8.2. Wszystkie współrzędne tego iloczynu są równe 0 z wyjątkiem współrzędnej o indeksie γ równej

$$f_{\gamma(1)}(u_{\gamma(1),1}, v_{\gamma(1),1}) \otimes f_{\gamma(2)}(u_{\gamma(2),2}, v_{\gamma(2),2}) \otimes f_{\gamma(3)}(u_{\gamma(3),3}, v_{\gamma(3),3}). \tag{33}$$

Wyrażenia z wzorów (32) i (33) są równe na podstawie definicji iloczynu tensorowego $f_{\gamma(1)} \otimes f_{\gamma(2)} \otimes f_{\gamma(3)}$. To dowodzi przemienności diagramu z rysunku 6 także w przypadku $\gamma \neq \eta$.

8.4 Pewna własność potrzebna w dowodzie twierdzenia Schoenhage

Przypuśćmy, że $R(\langle m,n,k\rangle) \leq r$. Wtedy dla odpowiednio dobranych funkcjonałów liniowych $\alpha_t : k^{m\times n} \to k$ i $\beta_t : k^{n\times k} \to k$ oraz macierzy $w_t \in k^{m\times k}$, dla wszystkich $A_1 \in k^{m\times n}$ oraz $A_2 \in k^{n\times k}$ zachodzi wzór

$$XY = \langle m,n,k\rangle(A_1,A_2) = \sum_{t=1}^{r} \alpha_t(A_1)\beta_t(A_2)w_t.$$

Pokażemy, że podobny wzór jest słuszny dla mnożenia macierzy o wymiarach $mm'\times nn'$ i $nn'\times kk'$. W tym rozdziale symbol $\otimes$ będzie stale oznaczał iloczyn Kroneckera zdefiniownany w rozdziale 3.2.2.

Weźmy funkcjonał liniowy $\alpha : k^{m\times n} \to k$ i zdefiniujmy przekształcenie $\alpha' : k^{m\times n} \times k^{m'\times n'} \to k^{m'\times n'}$ przyjmując, że

$$\alpha'(A,B) = \alpha(A)B.$$

Oczywiście, przekształcenie α' jest dwuliniowe. Korzystając z faktu, że iloczyn Kroneckera jest szczególnym iloczynem tensorowym i ma własność jednoznacznej uniwersalnej faktoryzacji, znajdujemy przekształcenie liniowe $\tilde{\alpha} : k^{mm'\times nn'} \to k^{m'\times n'}$ takie, że

$$\tilde{\alpha}(A\otimes B) = \alpha(A)B$$

dla dowolnych $A \in k^{m\times n}$ oraz $B \in k^{m'\times n'}$. Analogicznie, dla funkcjonałów liniowych $\beta : k^{n\times k} \to k$ bierzemy przekształcenia liniowe $\tilde{\beta} : k^{nn'\times kk'} \to k^{n'\times k'}$ takie, że

$$\tilde{\beta}(A\otimes B) = \beta(A)B$$

dla dowolnych $A \in k^{n\times k}$ oraz $B \in k^{n'\times k'}$.

Niech f oznacza funkcję dwóch zmiennych określoną w zbiorze $k^{mm'\times nn'} \times k^{nn'\times kk'}$, przyjmującą wartości ze zbioru $k^{mm'\times kk'}$ i zdefiniowaną wzorem

$$f(X,Y) = \sum_{t=1}^{r} w_i \otimes (\tilde{\alpha}_t(X)\cdot\tilde{\beta}_t(Y)).$$

Bez trudu sprawdzamy, że f jest przekształceniem dwuliniowym.

Pokażemy, że f jest odpowiednim mnożeniem macierzy, a więc $f = \langle mm', nn', kk'\rangle$. Aby to zrobić, wystarczy pokazać, że obie funkcje przyjmują te same wartości dla macierzy bazowych. Pokażemy, że tak jest dla wszystkich argumentów postaci $A\otimes B$, a więc dla argumentów, wśród których są macierze z baz standardowych. Zauważmy, że

$$\begin{aligned}
f(A_1\otimes B_1, A_2\otimes B_2) &= \sum_{t=1}^{r} w_i \otimes (\tilde{\alpha}_t(A_1\otimes B_1)\cdot\tilde{\beta}_t(A_2\otimes B_2)) = \\
&= \sum_{t=1}^{r} w_i \otimes ((\alpha_t(A_1)B_1)\cdot(\beta_t(A_2)B_2)) = \left(\sum_{t=1}^{r} \alpha_t(A_1)\beta_t(A_2)w_i\right)\otimes(B_1B_2) = \\
&= (A_1A_2)\otimes(B_1B_2) = (A_1\otimes B_1)(A_2\otimes B_2).
\end{aligned}$$

Ostatnia z powyższych równości jest prawdziwa na mocy wzoru (20).

W ten sposób dowiedliśmy, że wzór

$$XY = \langle mm', nn', kk'\rangle(X,Y) = \sum_{t=1}^{r} w_t \otimes (\tilde{\alpha}_t(X)\cdot\tilde{\beta}_t(Y)) \tag{34}$$

zachodzi dla wszystkich $X \in k^{mm'\times nn'}$ oraz $Y \in k^{nn'\times kk'}$.

8.5 Asymptotyczna nierówność dla sum prostych

Twierdzenie 8.1 (Schönhage) *Jeżeli* $\tilde{R}(\langle m_1, n_1, k_1\rangle \oplus \langle m_2, n_2, k_2\rangle) \leq r$*, to*

$$(m_1 n_1 k_1)^{\omega/3} + (m_2 n_2 k_2)^{\omega/3} \leq r.$$

Dowód. W tym dowodzie można posługiwać się dowolnym iloczynem tensorowym, ale wygodnie zakładać, że jest to iloczyn Kroneckera. Tak więc dalej symbol $\otimes$ będzie oznaczać odpowiedni iloczyn Kroneckera.

Przypuśćmy, że

$$\tilde{R}_q(\langle m_1, n_1, k_1\rangle \oplus \langle m_2, n_2, k_2\rangle) \leq r.$$

Z lematu 7.5 wynika, że

$$\tilde{R}_{N(q-1)+1}((\langle m_1, n_1, k_1\rangle \oplus \langle m_2, n_2, k_2\rangle)^N) \leq r^N.$$

Jeżeli teraz skorzystamy z lematu 7.2, to otrzymamy nierówność

$$R((\langle m_1, n_1, k_1\rangle \oplus \langle m_2, n_2, k_2\rangle)^N) \leq (Nq)^2 r^N.$$

Weźmy liczbę $\mu \leq N$ i obetnijmy funkcję $(\langle m_1, n_1, k_1\rangle \oplus \langle m_2, n_2, k_2\rangle)^N$ (opisaną w rozdziale 8.3) do współrzędnych wyznaczonych przez funkcje φ przyjmujące wartość 1 dokładnie μ razy. Innymi słowy, zamiast przekształceniem

$$\Psi = \bigoplus_{\varphi:\{1,\ldots,N\}\to\{1,2\}} f_{\varphi(1)} \otimes \ldots \otimes f_{\varphi(N)},$$

dla $f_1 = \langle m_1, n_1, k_1\rangle$ i $f_2 = \langle m_2, n_2, k_2\rangle$ będziemy zajmować się przekształceniem

$$\Phi = \bigoplus_{\varphi \text{ przyjmuje wartość 1 dla } \mu \text{ argumentów}} f_{\varphi(1)} \otimes \ldots \otimes f_{\varphi(N)},$$

która dwóm układom $\binom{N}{\mu}$ macierzy przyporządkowuje $\binom{N}{\mu}$ iloczynów tych macierzy obliczonych zgodnie z wzorem

$$(\Phi(X, Y))_\varphi = f_{\varphi(1)} \otimes \ldots \otimes f_{\varphi(N)}(X_\varphi, Y_\varphi).$$

Nietrudno zauważyć posługując się wzorem (21), że każda ze współrzędnych $f_{\varphi(1)} \otimes \ldots \otimes f_{\varphi(N)}$ przekształcenia Φ jest równa $\langle m_1^\mu m_2^{N-\mu}, n_1^\mu n_2^{N-\mu}, k_1^\mu k_2^{N-\mu}\rangle$. Wobec tego,

$$\Phi = \bigoplus_{i=1,\ldots,\binom{N}{\mu}} \langle m_1^\mu m_2^{N-\mu}, n_1^\mu n_2^{N-\mu}, k_1^\mu k_2^{N-\mu}\rangle.$$

Możemy wyobrażać sobie, że obliczenie wartości funkcji Φ polega na „równoległym” obliczeniu tych samych iloczynów dla $\binom{N}{\mu}$ par (być może różnych) macierzy. Co więcej, funkcja Φ jest obcięciem potęgi $(\langle m_1, n_1, k_1\rangle \oplus \langle m_2, n_2, k_2\rangle)^N$ w sensie opisanym w rozdziale 6.2 (funkcje α i β zgodnie z oznaczeniami z rozdziału 6.2 dopisują do danych układów macierzy odpowiednią liczbę macierzy zerowych, funkcja γ usuwa z obliczonego układu macierze znajdujące się na niepotrzebnych współrzędnych). Wobec tego, zgodnie z lematem 6.1 mamy oszacowanie

$$R(\Phi) \leq (Nq)^2 r^N. \tag{35}$$

Weźmy teraz dowolne $\varepsilon > 0$. Posługując się definicją ω znajdujemy $c_\varepsilon \in N$ takie, że

$$R(\langle n, n, n\rangle) \leq c_\varepsilon n^{\omega+\varepsilon}$$

dla wszystkich liczb naturalnych n. Jeżeli nierówność ta jest zgodnie z definicją prawdziwa tylko dla prawie wszystkich liczb naturalnych, to powiększając stosownie stałą c_ε spowodujemy, że będzie zachodzić dla wszystkich n. Podstawiając zamiast n w tej nierówności $M = \left\lfloor \binom{N}{\mu}^{1/(\omega+\varepsilon)} \right\rfloor$ otrzymamy, że

$$R(\langle M, M, M\rangle) \leq \rho = c_\varepsilon \binom{N}{\mu}.$$

Stąd i z nierówności (35) wyprowadzimy, że

$$R(\langle Mm_1^\mu m_2^{N-\mu}, Mn_1^\mu n_2^{N-\mu}, Mk_1^\mu k_2^{N-\mu}\rangle) \leq c_\varepsilon (Nq)^2 r^N. \tag{36}$$

Po pierwsze weźmy sumę prostą c_ε egzemplarzy funkcji Φ:

$$c_\varepsilon \Phi = \bigoplus_{i=1,\ldots,c_\varepsilon} \Phi.$$

Intuicyjnie, funkcja $c_\varepsilon \Phi$ „równolegle” oblicza $c_\varepsilon \binom{N}{\mu}$ iloczynów macierzy. Zgodnie z lematem 7.1 mamy

$$R(c_\varepsilon \Phi) \leq c_\varepsilon (Nq)^2 r^N.$$

Aby dowieść nierówność (36), wystarczy teraz zauważyć, że interesujący nas iloczyn $I = \langle Mm_1^\mu m_2^{N-\mu}, Mn_1^\mu n_2^{N-}$ jest obcięciem (w sensie rozdziału 6.2) przekształcenia $c_\varepsilon \Phi$. Wynika to z wzoru (34) dla $m = n = k = M$, $m' = m_1^\mu m_2^{N-\mu}$, $n' = n_1^\mu n_2^{N-\mu}$, $k' = k_1^\mu k_2^{N-\mu}$ oraz $r = \rho = c_\varepsilon \binom{N}{\mu}$. Zgodnie z tym wzorem, aby obliczyć iloczyn $I(X, Y)$ macierzy X i Y należy obliczyć układy $(\tilde{\alpha}_1(X), \ldots, \tilde{\alpha}_\rho(X))$ oraz $(\tilde{\beta}_1(Y), \ldots, \tilde{\beta}_\rho(Y))$ (obliczenie tych układów odpowiada obliczeniu wartości funkcji α i β z rozdziału 6.2), następnie dla obliczonych układów znaleźć wartość przekształcenia $c_\varepsilon \Phi$ (w ten sposób obliczymy układ $(\tilde{\alpha}_1(X)\tilde{\beta}_1(Y), \ldots, \tilde{\alpha}_r(X)\tilde{\beta}_r(Y))$), a na koniec dla tego układu obliczyć wartość funkcji γ zdefiniowanej wzorem

$$\gamma(C_1, \ldots, C_\rho) = \sum_{t=1}^{\rho} w_t \otimes C_t$$

(jest to odpowiednik funkcji γ z rozdziału 6.2). Zgodnie z wzorem (34), obliczymy w ten sposób iloczyn I, a z lematu 6.1 otrzymamy nierówność $R(I) \leq c_\varepsilon (Nq)^2 r^N$, czyli nierówność (36).

Z nierówności (36) i wniosku 6.13 mamy, że

$$(M^3 m_1^\mu m_2^{N-\mu} n_1^\mu n_2^{N-\mu} k_1^\mu k_2^{N-\mu})^{\omega/3} = M^\omega ((m_1 n_1 k_1)^\mu (m_2 n_2 k_2)^{N-\mu})^{\omega/3} \leq c_\varepsilon (Nq)^2 r^N.$$

Korzystając z nierówności $a/2 \leq \lfloor a \rfloor$ słusznej dla $a \geq 1$ pokazujemy, że

$$M^\omega = \left(\left\lfloor \binom{N}{\mu}^{1/(\omega+\varepsilon)} \right\rfloor \right)^\omega \geq \frac{1}{2^\omega} \binom{N}{\mu}^{\omega/(\omega+\varepsilon)} = \frac{1}{2^\omega} \binom{N}{\mu}^{1-\varepsilon/(\omega+\varepsilon)}.$$

Łącząc dwie ostatnie nierówności dostajemy dla liczby μ nierówność

$$\binom{N}{\mu} ((m_1 n_1 k_1)^{\omega/3})^\mu ((m_2 n_2 k_2)^{\omega/3})^{N-\mu} \leq 2^\omega \binom{N}{\mu}^{\varepsilon/(\omega+\varepsilon)} c_\varepsilon (Nq)^2 r^N \leq 2^{\omega + N\varepsilon/(\omega+\varepsilon)} c_\varepsilon (Nq)^2 r^N.$$

Jeżeli zsumujemy powyższe nierówności dla $\mu = 0, \ldots, N$, to otrzymamy

$$((m_1n_1k_1)^{\omega/3} + (m_2n_2k_2)^{\omega/3})^N \leq (N+1)2^{\omega+N\varepsilon/(\omega+\varepsilon)}c_\varepsilon(Nq)^2r^N.$$

Po spierwiastkowaniu stronami dostajemy

$$(m_1n_1k_1)^{\omega/3} + (m_2n_2k_2)^{\omega/3} \leq \sqrt[N]{(N+1)2^\omega c_\varepsilon(Nq)^2} \cdot 2^{\varepsilon/(\omega+\varepsilon)}r,$$

a po przejściu z N do nieskończoności –

$$(m_1n_1k_1)^{\omega/3} + (m_2n_2k_2)^{\omega/3} \leq \lim_{N\to\infty} \sqrt[N]{(N+1)2^\omega c_\varepsilon(Nq)^2} \cdot 2^{\varepsilon/(\omega+\varepsilon)}r = 2^{\varepsilon/(\omega+\varepsilon)}r.$$

Jeżeli teraz przejdziemy z ε do 0, to otrzymamy nierówność

$$(m_1n_1k_1)^{\omega/3} + (m_2n_2k_2)^{\omega/3} \leq \lim_{\varepsilon\to 0} 2^{\varepsilon/(\omega+\varepsilon)}r = r. \ \square$$

8.6 Dolne oszacowanie rzędu przekształcenia dwuliniowego

Lemat 8.2 *Przypuśćmy, że przekształcenie dwuliniowe $f : U \times V \to W$ jest definiowane wzorem*

$$f(u,v) = \sum_{i=1}^{r} f_i(u)g_i(v)w_i$$

dla funkcjonałów liniowych $f_i \in U^*$ i $g_i \in V^*$ oraz wektorów $w_i \in W$. Jeżeli 0 jest jedynym wektorem u takim, że $f(u,v) = 0$ dla wszystkich $v \in V$, to r jest większe od wymiaru U lub jest jemu równe.

Dowód. Rozważmy funkcje $\varphi : U \to k^r$ i $\psi : k^r \to Lin(V,W)$ takie, że

$$\varphi(u) = (f_1(u), \ldots, f_r(u))$$

oraz

$$\psi(a_1, \ldots, a_r)(v) = \sum_{i=1}^{r} a_ig_i(v)w_i.$$

Oczywiście, funkcje φ i ψ są liniowe oraz $\psi(\varphi(u))(v) = f(u,v)$. Pokażemy, że złożenie $\psi\varphi$ jest różnowartościowe.

Przypuśćmy, że $\psi(\varphi(u_1)) = \psi(\varphi(u_2))$. Wtedy

$$f(u_1 - u_2, v) = \psi(\varphi(u_1 - u_2))(v) = \psi(\varphi(u_1))(v) - \psi(\varphi(u_1))(v) = 0$$

dla dowolnego $v \in V$. Z założonej własności f otrzymujemy, że $u_1 = u_2$.

Jeżeli złożenie $\psi\varphi$ jest różnowartościowe, to także przekształcenie φ jest różnowartościowe. Z istnienia różnowartościowego przekształcenia liniowego z U w k^r wynika, że wymiar U jest $\leq r$. $\square$

Lemat 8.3 *Przypuśćmy, że przekształcenie dwuliniowe $f : U \times V \to W$ jest definiowane wzorem*

$$f(u,v) = \sum_{i=1}^{r} f_i(u)g_i(v)w_i$$

dla funkcjonałów liniowych $f_i \in U^*$ i $g_i \in V^*$ oraz wektorów $w_i \in W$. Jeżeli 0 jest jedynym wektorem v takim, że $f(u,v) = 0$ dla wszystkich $u \in U$, to r jest większe od wymiaru V lub jest jemu równe. $\square$

Lemat 8.4 *Przypuśćmy, że przekształcenie dwuliniowe $f : U \times V \to W$ jest definiowane wzorem*

$$f(u,v) = \sum_{i=1}^{r} f_i(u) g_i(v) w_i$$

dla funkcjonałów liniowych $f_i \in U^*$ i $g_i \in V^*$ oraz wektorów $w_i \in W$. Jeżeli zbiór wartości f generuje W, to r jest większe od wymiaru W lub jest jemu równe.

Dowód. Jest oczywiste, że jeżeli zbiór watości przekształcenia f generuje W, to także wektory $w_1, \ldots, w_r$ generują W. Stąd przestrzeń W, jako generowana przez r wektorów, ma wymiar $\geq r$. □

Przekształcenie dwuliniowe $f : U \times V \to W$ nazywamy zwięzłym, jeżeli

1) 0 jest jedynym wektorem u takim, że $f(u,v) = 0$ dla wszystkich $v \in V$,

2) 0 jest jedynym wektorem v takim, że $f(u,v) = 0$ dla wszystkich $u \in U$,

3) zbiór wartości przekształcenia f generuje przestrzeń W.

Wniosek 8.5 *Jeżeli przekształcenie dwuliniowe $f : U \times V \to W$ jest zwięzłe, to rząd f jest większy lub równy od wymiaru każdej z przestrzeni U, V i W.* □

Lemat 8.6 *Wszystkie mnożenia macierzy są zwięzłe.*

Dowód. Będziemy dowodzić, że przekształcenie $\langle m, n, k \rangle$ jest zwięzłe.

Weźmy macierz $X \in k^{m \times n}$ o wyrazach $x_{i,j}$ i pomnóżmy ją przez $E_{j,1}^{n \times k}$:

$$XE_{j,1}^{n\times k} = \begin{pmatrix} x_{1,1} & \ldots & x_{1,n} \\ & \ddots & \\ x_{m,1} & \ldots & x_{m,n} \end{pmatrix} \begin{pmatrix} 0 & \ldots & 0 \\ & \ddots & \\ 1 & \ldots & 0 \\ & \ddots & \\ 0 & \ldots & 0 \end{pmatrix} = \begin{pmatrix} x_{1,j} & 0 & \ldots & 0 \\ & \ddots & & \\ x_{m,j} & 0 & \ldots & 0 \end{pmatrix}.$$

Jeżeli dla wszystkich Y macierze XY są zerowe, to także wszystkie macierze $XE_{j,1}^{n\times k}$, a stąd, na podstawie podanego wzoru, wszystkie wyrazy macierzy X są równe 0.

Drugi warunek z definicji zwięzłości dowodzimy w podobny sposób.

Z wzoru $E_{i,j}^{m\times n} E_{j,l}^{n\times k} = E_{j,l}^{m\times k}$ i fakty, że macierze $E_{j,l}^{m\times k}$ generują przestrzeń $k^{m\times k}$, wynika warunek 3) z definicji zwięzłości. □

Wniosek 8.7 *Iloczyny $\langle m, n, 1 \rangle$, $\langle m, 1, n \rangle$ i $\langle 1, m, n \rangle$ mają rząd mn.* □

8.7 Przykład Schönhage

Dla liczb m i k definiujemy przekształcenie

$$\mathcal{F} : (k^{m\times 1} \times k^{1\times(m-1)(k-1)}) \times (k^{1\times k} \times k^{(m-1)(k-1)\times 1}) \to k^{m\times k} \times k^{1\times 1}$$

przyjmując, że

$$\mathcal{F}((A,B),(C,D)) = (\langle m,1,k\rangle(A,C), \langle 1,(m-1)(k-1),1\rangle(B,D)) = (AC, BD).$$

Podany wzór świadczy o tym, że funkcja $\mathcal{F}$ jest sumą prostą iloczynów $\langle m,1,k\rangle$ i $\langle 1,(m-1)(k-1),1\rangle$, czyli

$$\mathcal{F} = \langle m,1,k\rangle \oplus \langle 1,(m-1)(k-1),1\rangle.$$

Oczywiście,

$$\mathcal{F} \in Bil(k^{m\times 1} \times k^{1\times(m-1)(k-1)}, k^{1\times k} \times k^{(m-1)(k-1)\times 1}, k^{m\times k} \times k^{1\times 1}).$$

Najpierw przedstawimy $\mathcal{F}$ jako kombinację liniową przekształceń bazowych. Wymaga to wprowadzenia odpowiednich oznaczeń. Symbole $E^1_{i,j}$ i E^2 będą oznaczały odpowiednio pary $(E^{m\times k}_{i,j},0)$ oraz $(0,E^{1\times 1}_{1,1})$ należące do przestrzeni $k^{m\times k}\times k^{1\times 1}$. Oczywiście, zbiór tych par jest bazą tej przestrzeni.

Przyjmijmy, że $\delta_{i,j}$ oznacza funkcję, która dowolnej macierzy X przyporządkowuje wyraz macierzy X znajdujący się i-tym wierszu i j-tej kolumnie. Zdefiniujemy funkcjonały liniowe

$$\delta^1_i : k^{m\times 1}\times k^{1\times(m-1)(k-1)} \to k \ \text{ oraz } \ \delta^2_j : k^{1\times k}\times k^{(m-1)(k-1)\times 1} \to k,$$

a także

$$\Delta^1_i : k^{m\times 1}\times k^{1\times(m-1)(k-1)} \to k \ \text{ oraz } \ \Delta^2_j : k^{1\times k}\times k^{(m-1)(k-1)\times 1} \to k.$$

Funcjonały te są definiowane równościami

$$\delta^1_i(A,B) = \delta_{i,1}(A) \ \text{ i } \ \delta^2_j(A,B) = \delta_{1,j}(B)$$

oraz

$$\Delta^1_j(C,D) = \delta_{1,j}(C) \ \text{ i } \ \Delta^2_i(C,D) = \delta_{i,1}(D).$$

Zauważmy, że

$$\begin{aligned}
&\mathcal{F}((A,B),(C,D)) = \\
&= \mathcal{F}(((a_{1,1},\dots,a_{m,1})^T,(b_{1,1},\dots,b_{1,(m-1)(k-1)})),(((c_{1,1},\dots,c_{1,k}),(d_{1,1},\dots,d_{(m-1)(k-1),1})^T)) = \\
&= \left(\begin{pmatrix} a_{1,1}c_{1,1} & \dots & a_{1,1}c_{1,k} \\ & \ddots & \\ a_{m,1}c_{1,1} & \dots & a_{m,1}c_{1,k}\end{pmatrix}, \sum_{l=1}^{(m-1)(k-1)} b_{1,l}d_{l,1}\right) = \\
&= \sum_{i,j=1}^{m,k} a_{i,1}c_{1,j}E^1_{i,j} + \sum_{l=1}^{(m-1)(k-1)} c_{1,l}d_{l,1}E^2 = \\
&= \sum_{i,j=1}^{m,k} \delta^1_i(A,B)\Delta^1_j(C,D)E^1_{i,j} + \sum_{l=1}^{(m-1)(k-1)} \delta^2_l(A,B)\Delta^2_l(C,D)E^2.
\end{aligned}$$

Ostatecznie otrzymujemy, że

$$\mathcal{F} = \sum_{i,j=1}^{m,k}\delta^1_i\Delta^1_jE^1_{i,j} + \sum_{l=1}^{(m-1)(k-1)}\delta^2_l\Delta^2_lE^2.$$

Ze względów technicznych przestrzenie $k^{1\times(m-1)(k-1)}$ i $k^{(m-1)(k-1)\times 1}$ będziemy utożsamiać z przestrzenią $k^{(m-1)\times(k-1)}$. Spowoduje to zmianę numeracji funkcjonałów δ^2_j i Δ^2_j. Dla $i=1,\dots,m-1$ oraz $j=1,\dots,k-1$ przyjmijmy, że $\delta^2_{i,j} = \delta^2_{(i-1)(k-1)+j}$ oraz $\Delta^2_{i,j} = \Delta^2_{(i-1)(k-1)+j}$. Posługując się tymi oznaczeniami otrzymujemy wzór

$$\mathcal{F} = \sum_{i,j=1}^{m,k}\delta^1_i\Delta^1_jE^1_{i,j} + \sum_{i,j=1}^{m-1,k-1}\delta^2_{i,j}\Delta^2_{i,j}E^2.$$

Dla $j = 1, \dots, k-1$ oraz $i = 1, \dots, k$ definiujemy pomocnicze funkcjonały

$$\delta^2_{m,j} = -\sum_{i=1}^{m-1} \delta^2_{i,j} \quad \text{oraz} \quad \Delta^2_{i,k} = -\sum_{j=1}^{k-1} \Delta^2_{i,j},$$

$$\delta^2_{i,k} = 0 \quad \text{oraz} \quad \Delta^2_{m,j} = 0,$$

a także definiujemy

$$\delta^2_{m,k} = 0 \quad \text{oraz} \quad \Delta^2_{m,k} = 0.$$

Przyjmijmy, że przekształcenie

$$\mathcal{F}' : (k[\varepsilon]^{m\times 1} \times k[\varepsilon]^{1\times(m-1)(k-1)}) \times (k[\varepsilon]^{1\times k} \times k[\varepsilon]^{(m-1)(k-1)\times 1}) \to k[\varepsilon]^{m\times k} \times k[\varepsilon]^{1\times 1}$$

jest zdefiniowane wzorem

$$\mathcal{F}' = \sum_{i,j=1}^{m,k} (\delta^1_i + \varepsilon\delta^2_{i,j})(\Delta^1_j + \varepsilon\Delta^2_{i,j})(\varepsilon^2 E^1_{i,j} + E^2) - (\sum_{i=1}^{m} \delta^1_i)(\sum_{j=1}^{k} \Delta^1_i)E^2.$$

Oczywiście, rząd $\mathcal{F}'$ jest $\leq mk+1$. Sprowadzimy teraz $\mathcal{F}'$ do postaci wielomianu. Wyraz wolny $\mathcal{F}'$ jest równy

$$\sum_{i,j=1}^{m,k} \delta^1_i \Delta^1_j E^2 - (\sum_{i=1}^{m} \delta^1_i)(\sum_{j=1}^{k} \Delta^1_i)E^2 = 0.$$

Współczynnik przy ε ma postać

$$\begin{aligned}
&\sum_{i,j=1}^{m,k} (\delta^2_{i,j}\Delta^1_j + \delta^1_i \Delta^2_{i,j})E^2 = \\
&= \sum_{i,j=1}^{m-1,k-1} (\delta^2_{i,j}\Delta^1_j + \delta^1_i \Delta^2_{i,j})E^2 + \sum_{j=1}^{k-1} (\delta^2_{m,j}\Delta^1_j + \delta^1_m \Delta^2_{m,j})E^2 + \\
&\quad + \sum_{i=1}^{m-1} (\delta^2_{i,k}\Delta^1_k + \delta^1_i \Delta^2_{i,k})E^2 + (\delta^2_{m,k}\Delta^1_k + \delta^1_m \Delta^2_{m,k})E^2 = \\
&= \sum_{i,j=1}^{m-1,k-1} (\delta^2_{i,j}\Delta^1_j + \delta^1_i \Delta^2_{i,j})E^2 + \sum_{j=1}^{k-1} \delta^2_{m,j}\Delta^1_j E^2 + \sum_{i=1}^{m-1} \delta^1_i \Delta^2_{i,k} E^2 = \\
&= \sum_{i,j=1}^{m-1,k-1} (\delta^2_{i,j}\Delta^1_j + \delta^1_i \Delta^2_{i,j})E^2 - \sum_{j=1}^{k-1} (\sum_{i=1}^{m-1} \delta^2_{i,j})\Delta^1_j E^2 - \sum_{i=1}^{m-1} \delta^1_i (\sum_{j=1}^{k-1} \Delta^2_{i,j})E^2 = 0
\end{aligned}$$

W końcu, współczynnik przy ε^2 jest równy

$$\begin{aligned}
&\sum_{i,j=1}^{m,k} (\delta^1_i \Delta^1_j E^1_{i,j} + \delta^2_{i,j}\Delta^2_{i,j}E^2) = \\
&= \mathcal{F} + \sum_{j=1}^{k-1} \delta^2_{m,j}\Delta^2_{m,j}E^2 + \sum_{i=1}^{m-1} \delta^2_{i,k}\Delta^2_{i,k}E^2 + \delta^2_{m,k}\Delta^2_{m,k}E^2 = \mathcal{F}
\end{aligned}$$

Dowiedliśmy więc, że

$$\mathcal{F}' = \varepsilon^2 \mathcal{F} + \varepsilon^3 \sum_{i,j=1}^{m,k} (\delta^1_i \Delta^2_{i,j} + \delta^2_{i,j}\Delta^1_j)E^1_{i,j}.$$

Wniosek 8.8 *Rząd graniczny* $\tilde{R}_3(\mathcal{F})$ *przekształcenia* $\mathcal{F}$ *jest* $\leq mk + 1$. $\square$

Wniosek 8.9 *Dla dowolnych dodatnich liczb naturalnych* m, k *zachodzi nierówność*

$$(mk)^{\omega/3} + ((m-1)(k-1))^{\omega/3} \leq mk + 1.$$

Dowód. Jest to oczywista konsekwencja twierdzenia 8.1 i wniosku 8.8. $\square$

Jeżeli do podanego wzoru podstawimy $m = k = 4$, to otrzymamy

$$16^{\omega/3} + 9^{\omega/3} \leq 17.$$

Dla $\omega = 2,5479$ lewa strona wzoru przyjmuje wartość < 17, a dla $\omega = 2,548$ jej wartość jest > 17. Wobec tego mamy

Wniosek 8.10 (Schönhage (1981)) *Współczynnik* ω *jest mniejszy niż 2,548.* $\square$

Wiadomo jeszcze, że

Twierdzenie 8.11 (Coppersmith, Winograd (1987)) *Współczynnik* ω *jest mniejszy niż 2,39.*