PROSTE A POZYTECZNE NIERÓWNOśCI MACIERZOWE

TOMASZ TKOCZ

STRESZCZENIE. Tekst zawiera notatki do referatu z Seminarium Fizyki Teoretycznej. Jest to w zasadzie tłumaczenie (z uzupełnionymi pewnymi skrótami myślowymi i innym podejściem do triku Weyla) odpowiedniego rozdziału książki [3].

1. WSTĘP

W fizyce statystycznej kluczowe znaczenie ma *suma statystyczna*. W mechanice kwantowej ma ona postać $Z = \operatorname{tr} e^{-\beta(\hat{H}-\mu\hat{N})}$ ($\beta = \frac{1}{kT}$, $\hat{H}$ — hamiltonian układu, μ — potencjał chemiczny, $\hat{N}$ — operator liczby cząstek — jeśli w układzie jest ona zachowana). Zatem bardzo ważne znaczenie jest umieć szacować wyrażenia $\operatorname{tr} e^A$, gdzie A jest macierzą i niech to będzie dla nas motywacją, choć nie wiem jaka ona była naprawdę, żeby sobie dzisiaj obejrzeć kilka klasycznych nierówności. Przyjrzymy się nierównościom:

- Peierlsa – Bogolubowa
- Höldera
- Goldena – Thompsona

O ich zastosowaniach nie będzie tu wspomniane. Aby jednak wszystkich (pragmatyków?) uspokoić: nierówności te stosują się! Hasłowo, niektóre zastosowania, to

- dowody wypukłości różnych ważnych funkcji termodynamicznych, np. energii swobodnej
- dowody istnienia granicy termodynamicznej
- oszacowania i porównywanie sumy statystycznej różnych modeli (np. Isinga i Heisenberga — przykład z [3])
- dowód braku uporządkowania dalekiego zasięgu w modelu Hubbarda (w wymiarach 1 i 2 i temperaturach dodatnich) — twierdzenie Komy–Tasakiego (por. [2])
- ostro też te nierówności stosuje Pan Wojtkiewicz podając oszacowania (z góry i z dołu) sumy statystycznej w modelu Hubbarda (por. [4])

Chciałbym podziękować Panu J. Wojtkiewiczowi — opiekunowi referatu — za pokazanie mi tytułowych nierówności.

2. Nierówność Peierlsa–Bogolubowa

Rozważać będziemy operatory liniowe $A\colon \mathbb{C}^n \longrightarrow \mathbb{C}^n$ (o macierzy A) oraz na $\mathbb{C}^n$ mamy zadany iloczyn skalarny $<,>$. Przypomnijmy sobie kilka faktów z algebry liniowej.

Twierdzenie 1 (spektralne)**.** *Jeśli A jest samosprzężony ($A^\dagger = A$), to istnieje baza ortonormalna $u_1, \ldots, u_n$ oraz liczby rzeczywiste $\lambda_1, \ldots \lambda_n$ takie, że*

$$Ax = \sum_{k=1}^{n} \lambda_k < x, u_k > u_k.$$

Definicja 1. Niech A — operator samosprzężony. Dla $f\colon R \to R$ definiujemy operator $f(A)\colon \mathbb{C}^n \to \mathbb{C}^n$

$$f(A)x := \sum_{k=1}^{n} f(\lambda_k) < x, u_k > u_k.$$

Uwaga 1. *Jeśli f rozwija się w otoczeniu 0 w szereg ($f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k$, $|x| \leqslant R$), to $f(A) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} A^k$, dla A takiego, że $\|A\| \leqslant R$. Zatem podana* ad hoc *definicja $f(A)$ dla pożądnych f pokrywa się z tą znaną skądinąd (np. dla funkcji $e, \ln, \sqrt{\ }$).*

Dowód. Rozwijamy $f(\lambda_k)$ w szereg mając na uwadze, że $\|A\| = max_{1 \leqslant i \leqslant n} |\lambda_i| \leqslant R$

$$\begin{aligned} f(A)x &= \sum_{k=1}^{n} f(\lambda_k) < x, u_k > u_k = \sum_{k=1}^{n} \sum_{l=0}^{\infty} \frac{f^{(l)}(0)}{l!} \lambda_k^l < x, u_k > u_k \\ &= \sum_{l=0}^{\infty} \frac{f^{(l)}(0)}{l!} \sum_{k=1}^{n} \lambda_k^l < x, u_k > u_k = \sum_{l=0}^{\infty} \frac{f^{(l)}(0)}{l!} A^l. \end{aligned}$$

□

Koniec przypomnienia. Przechodzimy do sedna tego paragrafu.

Twierdzenie 2. *Niech $f\colon \mathbb{R} \to \mathbb{R}$ będzie funkcją wypukłą, A operatorem samosprzężonym. Wówczas dla dowolnego x, $\|x\| = 1$ zachodzi*

$$f(< Ax, x >) \leqslant < f(A)x, x > .$$

Dowód. Mamy wobec rozkładu spektralnego dla A, że

$$\begin{aligned} f(< Ax, x > &= f\left(< \sum_{k=1}^{n} \lambda_k < x, u_k > \right) = f\left(\sum_{k=1}^{n} | < x, u_k > |^2 \lambda_k \right) \\ &\leqslant \sum_{k=1}^{n} | < x, u_k > |^2 f(\lambda_k) = < f(A)x, x >, \end{aligned}$$

gdzie skorzystaliśmy z nierówności Jensena (dla wag $| < x, u_k > |^2$ sumujących się do $\|x\|^2 = 1$). □

Z powyższego mamy natychmiast

Wniosek 1. *Jeśli* (v_k) *jest bazą ortonormalną, to*

$$\operatorname{tr} f(A) \geqslant \sum_{k=1}^{n} f(<Av_k, v_k>).$$

Dowód. Wystarczy zauważyć, że ogólnie $<Av_k, v_k>$ jest współrzędną (k,k) macierzy operatora A w bazie (v_k), więc

$$\operatorname{tr} f(A) = \sum_{k=1}^{n} <f(A)v_k, v_k> \geqslant \sum_{k=1}^{n} f(<Av_k, v_k>).$$

□

Twierdzenie 3 (Nierówność Peierlsa–Bogolubowa)**.** *Dla operatorów samosprzężonych* A, B *i funkcji wypukłej* $f\colon \mathbb{R} \to \mathbb{R}$ *zachodzi*

$$\frac{\operatorname{tr} f(A)e^B}{\operatorname{tr} e^B} \geqslant f\left(\frac{\operatorname{tr} Ae^B}{\operatorname{tr} e^B}\right).$$

Dowód. Niech (v_k) będzie bazą orotonormalną z twierdzenia spektralnego dla B, zaś $\mu_1, \ldots, \mu_k$ odpowiadającymi im wartościami własnymi. Mamy wobec nierówności z twierdzenia 2 i nierówności Jensena, że

$$\begin{aligned}
\frac{\operatorname{tr} f(A)e^B}{\operatorname{tr} e^B} &= \frac{1}{\sum e^{\mu_k}} \sum_{k=1}^{n} <f(A)e^B v_k, v_k> \\
&= \sum_{k=1}^{n} \frac{e^{\mu_k}}{\sum e^{\mu_k}} <f(A)v_k, v_k> \geqslant \sum_{k=1}^{n} \frac{e^{\mu_k}}{\sum e^{\mu_k}} f(<Av_k, v_k>) \\
&\geqslant f\left(\sum_{k=1}^{n} \frac{e^{\mu_k}}{\sum e^{\mu_k}} <Av_k, v_k>\right) = f\left(\frac{\operatorname{tr} Ae^B}{\operatorname{tr} e^B}\right).
\end{aligned}$$

□

Identycznie dowodzimy (ale nie wynika ono z powyższego twierdzenia, bo niekoniecznie $e^{A+B} = e^A e^B$)

Twierdzenie 4 (Nierówność Peierlsa–Bogolubowa')**.** *Dla operatorów samosprzężonych* A, B *zachodzi*

$$\frac{\operatorname{tr} e^{A+B}}{\operatorname{tr} e^B} \geqslant \exp\left(\frac{\operatorname{tr}(Ae^B)}{\operatorname{tr} e^B}\right).$$

Dowód. Z poprzedniego dowodu widać, że $\operatorname{tr} e^B$ jest tylko czynnikiem normującym, więc można bez utraty ogólności założyć, że $\operatorname{tr} e^B = 1$ (w razie czego rozważając zamiast B operator $B + cI$, dla odpowiedniej stałej c).

Wówczas ((v_k), (μ_k) ma to samo znaczenie co powyżej)

$$\begin{aligned}
\operatorname{tr} e^{A+B} &= \sum_{k=1}^{n} < e^{A+B} v_k, v_k > \geqslant \sum_{k=1}^{n} e^{<(A+B)v_k, v_k>} \\
&= \sum_{k=1}^{n} e^{\mu_k} e^{<Av_k, v_k>} \geqslant \exp\left(\sum_{k=1}^{n} e^{\mu_k} < Av_k, v_k >\right) \\
&= \exp\Big(\operatorname{tr}(Ae^B)\Big).
\end{aligned}$$

□

Jako ciekawostkę zauważmy na zakończenie tego paragrafu takie przyjemne zastosowanie głównego narzędzia — twierdzenia 2

Wniosek 2. *Niech* $C : \mathbb{C}^n \to \mathbb{C}^n$ *będzie operatorem dodatnio określonym i samosprzężonym. Wówczas*

$$\det C \leqslant \prod_{k=1}^{n} c_{kk},$$

gdzie $[c_{ij}]$ *jest macierzą* C *w bazie standardowej* (e_k)*.*

Dowód. Dodatnia określoność oznacza, że dla każdego $x \neq 0$ $< Cx, x >> 0$, więc w szczególności wartości własne C są liczbami rzeczywistymi dodatnimi oraz $c_{kk} > 0$. Biorąc $f(x) = -\ln x$ w twierdzeniu 2 mamy

$$- < \ln Ce_k, e_k > = < f(C)e_k, e_k > \geqslant f(< Ce_k, e_k >) = -\ln c_k k,$$

więc $\ln c_{kk} \geqslant < \ln Ce_k, e_k >$ i powysumowaniu

$$\ln \prod_{k=1}^{n} c_{kk} \geqslant \operatorname{tr}\, \ln C = \sum_{k=1}^{n} \ln \lambda_k = \ln \prod_{k=1}^{n} \lambda_k = \ln \det C,$$

czyli teza. □

3. Nierówność Höldera

Będziemy chcieli odejść na chwilę od operatorów samosprzężonych. Ale w ogólności jest też coś na kształt twierdzenia spektralnego (gdy $\dim V = \infty$ i operator jest zwarty nazywa się to rozkładem Schaudera). Niech więc A będzie teraz dowolnym endomorfizmem przestrzeni $\mathbb{C}^n$. Bierzemy $T := A^\dagger A$. Wtedy T jest oczywiście samosprzężony, ale też nieujemny, bo

$$< Tx, x > = < Ax, Ax > = \|Ax\|^2 \geqslant 0$$

(widać po co $A^\dagger A$ zamiast $AA^\dagger$). Zatem z twierdzenia spektralnego istnieje baza ortonormalna (u_k) i wartości własne nieujemne (λ_k) dla T. Możemy więc zdefiniwać

$$|A| := \sqrt{T} = \sqrt{A^\dagger A}.$$

Mamy

$$
\begin{aligned}
(1) \qquad \| \, |A| \, x\|^2 &= <|A|x, |A|x> = <|A|^\dagger |A|x, x> = <|A|^2 x, x> \\
&= <A^\dagger Ax, x> = <Ax, Ax> = \|Ax\|^2,
\end{aligned}
$$

czyli definiując $U : \mathrm{im}|A| \to \mathrm{im}A$ wzorem $U(|A|x) := Ax$ mamy poprawnie określoną (bo $\ker|A| = \ker A$ — wobec (1)) izometrię na swój obraz, którą łatwo rozszerzamy do izometrii $\mathbb{C}^n$. Zatem

$$
\begin{aligned}
Ax &= U(|A|x) = U\left(\sum_{k=1}^{n} \lambda_k <x, u_k> u_k\right) \\
&= \sum_{k=1}^{n} \lambda_k <x, u_k> Uu_k = \sum_{k=1}^{n} \lambda_k <x, u_k> v_k,
\end{aligned}
$$

gdzie (v_k) pewna baza ortonormalna (obraz bazy (u_k) przy U). Zatem zachodzi

Wniosek 3. *Dla $A\colon \mathbb{C}^n \to \mathbb{C}^n$ istnieją bazy ortonormalne $(u_k), (v_k)$ i liczby $\mu_1(A) \geqslant \ldots \geqslant \mu_n(A) \geqslant 0$ — wartości własne $|A|$ zwane* ***wartościami osobliwymi*** *A takie, że*

$$
Ax = \sum_{k=1}^{n} \mu_k(A) <x, u_k> v_k
$$

Uwaga 2. $\mu_k(A^\dagger) = \mu_k(A)$

Dowód. Wobec $U|A| = A$ jest $A^\dagger = |A|U^\dagger$ i mamy

$$
AA^\dagger v_k = A|A|U^\dagger v_k = A|A|u_k = A\mu_k(A)u_k = \mu_k(A)^2 v_k,
$$

więc $|A^\dagger|v_k = \mu_k(A)v_k$, co oznacza, że $|A|$, $|A^\dagger$ mają te same wartości własne. □

Definicja 2. $\|A\|_p := (\mathrm{tr}\,|A|^p)^{1/p}, \quad p \in [1, \infty]$.

Uwaga 3.

$$
\begin{aligned}
(\mathrm{tr}\,|A|^p)^{1/p} &= \left(\sum_{k=1}^{n} \mu_k(A)^p\right)^{1/p} = \|(\mu_1(A), \ldots, \mu_n(A))\|_p \xrightarrow[p\to\infty]{} \max_k \mu_k(A) \\
&= \mu_1(A) = \| \, |A| \, \| = \|A\|,
\end{aligned}
$$

czyli $\|A\|_\infty = \|A\|$.

Możemy już teraz sformułować główne twierdzenie tego rozdziału

Twierdzenie 5 (Nierówność Höldera)**.** *Dla $r, p, q \in [1, \infty]$ takich, że $\frac{1}{r} = \frac{1}{p} + \frac{1}{q}$ i dowolnych operatorów $A, B\colon \mathbb{C}^n \to \mathbb{C}^n$ zachodzi*

$$
\|AB\|_r \leqslant \|A\|_p \|B\|_q
$$

Dowód można łatwo przeprowadzić w oparciu o zwykłą (dla p-tych norm ciągów) nierówność Höldera jeśli tylko udowodnimy, że

$$\sum_{k=1}^{n} \mu_k(AB)^r \leqslant \sum_{k=1}^{n} \mu_k(A)^r \mu_k(B)^r, \tag{2}$$

bowiem wówczas stąd i uwagi 3 mamy

$$\begin{aligned} \|AB\|_r^r &= \sum_{k=1}^{n} \mu_k(AB)^r \leqslant \sum_{k=1}^{n} \mu_k(A)^r \mu_k(B)^r \\ &\leqslant \left(\sum_{k=1}^{n} \mu_k(A)^r)^{p/r}\right)^{r/p} \left(\sum_{k=1}^{n} \mu_k(B)^r)^{p/r}\right)^{r/p} = \|A\|_p^r \|B\|_q^r. \end{aligned}$$

Dowód (2) zaś wynika z dwóch rzeczy

Lemat 1. *Dla liczb rzeczywistych $0 \leqslant a_1 \leqslant \ldots \leqslant a_n, 0 \leqslant b_1 \leqslant \ldots \leqslant b_n$, takich, że dla każdego $k = 1, \ldots, n$ $\prod_{j=1}^{k} b_j \leqslant \prod_{j=1}^{k} a_j$ zachodzi*

$$\sum_{j=1}^{k} b_j \leqslant \sum_{j=1}^{n} a_j,$$

dla każdego $k = 1, \ldots, n$.

Dowód. Jest to w zasadzie zasada majoryzacyjna Hardy'ego – Littlewooda – Póly zastosowana do funkcji exp. Dowód można znaleźć w [1] lub (bardziej geometryczny) w [3]. $\square$

Lemat 2. *Dla dowolnych operatorów $A, B \colon \mathbb{C}^n \to \mathbb{C}^n$ i każdego $k = 1, \ldots, n$ zachodzi*

$$\prod_{j=1}^{k} \mu_j(AB) \leqslant \prod_{j=1}^{k} \mu_j(A)\mu_j(B).$$

Dowód. Użyjemy sztuczki Weyla z algebrą przekształceń antysymetrycznych, po to żeby zgrabnie wydobyć iloczyn $\prod_{j=1}^{k} \mu_j(A)$. W przestrzeni odwzorowań k-liniowych antysymetrycznych $\bigwedge^k(\mathbb{C}^n)^*$ definiujemy dla operatora $A \colon \mathbb{C}^n \to \mathbb{C}^n$ jego **k-tą potęgę dziubkową** $A^{\wedge k} \colon \bigwedge^k(\mathbb{C}^n)^* \to \bigwedge^k(\mathbb{C}^n)^*$ wzorem

$$(A^{\wedge k}\omega)(v_1, \ldots, v_k) := \omega(Av_1, \ldots, Av_k),$$

dla $\varphi \in \bigwedge^k(\mathbb{C}^n)^*$, $v_1, \ldots, v_k \in \mathbb{C}^n$. (Można o tym myśleć np. jak o *pool–backu* stałej formy ω przy przekształceniu A). Zauważmy też, że dla $k = 1$ ta potęga dziubkowa operatora A staje się jego zwykłym sprzężeniem Banachowskim.

Po co to wszystko? Mamy

$$\begin{aligned} \mu_1(A^{\wedge k}) &= \text{największa wartość własna}|A^{\wedge k}| \\ &\overset{(5)}{=} \text{największa wartość własna}|A^\dagger|^{\wedge k} \\ &\overset{(4)}{=} \mu_1(A^\dagger) \cdot \ldots \cdot \mu_k(A^\dagger) = \mu_1(A) \cdot \ldots \cdot \mu_k(A), \end{aligned} \tag{3}$$

skąd

$$\begin{aligned}\prod_{j=1}^{k}\mu_j(AB) &\overset{(3)}{=} \mu_1\left((AB)^{\wedge k}\right) = \mu_1(B^{\wedge k}A^{\wedge k})\\ &= \|B^{\wedge k}A^{\wedge k}\| \leqslant \|B^{\wedge k}\|\|^{\wedge k}\| = \mu_1(B^{\wedge k})\mu_1(A^{\wedge k})\\ &\overset{(3)}{=} \prod_{j=1}^{k}\mu_j(A)\prod_{j=1}^{k}\mu_j(B),\end{aligned}$$

czyli teza lematu. Pozostaje jeszcze udowdnić

$$\text{największa wartość własna}|A|^{\wedge k} = \mu_1(A)\cdot\ldots\cdot\mu_k(A). \tag{4}$$

$$|A^{\wedge k}| = |A^{\dagger}|^{\wedge k} \tag{5}$$

Dla dowodu (4) weźmy operator S samosprzężony nieujemny (takim jest zawsze $|A|$) o wartościach własnych $\lambda_1 \geqslant \ldots \geqslant \lambda_n \geqslant 0$ dla bazy ortonormalnej (u_k) (tzn. $Su_k = \lambda_k u_k$). Zauważmy, że (pamiętamy, że jeśli $(\mathrm{d}u_k)$ oznacza bazę dualną do (u_k), to $\mathrm{d}u_{i_1}\wedge\ldots\wedge\mathrm{d}u_{i_k}$, $1\leqslant i_1 < \ldots i_k \leqslant n$ jest bazą ortonormalną $\bigwedge^k(\mathbb{C}^n)^*$)

$$S^{\wedge k}\mathrm{d}u_{i_1}\wedge\ldots\wedge\mathrm{d}u_{i_k} = (S^*\mathrm{d}u_{i_1})\wedge\ldots\wedge(S^*\mathrm{d}u_{i_k}),$$

ale $S^*\mathrm{d}u_l = \lambda_l\mathrm{d}u_l$, więc

$$S^{\wedge k}\mathrm{d}u_{i_1}\wedge\ldots\wedge\mathrm{d}u_{i_k} = \lambda_{i_1}\cdot\ldots\cdot\lambda_{i_k}\mathrm{d}u_{i_1}\wedge\ldots\wedge\mathrm{d}u_{i_k},$$

czyli $S^{\wedge k}$ ma wektory własne $\mathrm{d}u_{i_1}\wedge\ldots\wedge\mathrm{d}u_{i_k}$ o wartościach własnych $\lambda_{i_1}\cdot\ldots\cdot\lambda_{i_k}$. Ponieważ jest ich tyle co $\dim\bigwedge^k(\mathbb{C}^n)^*$, więc są to wszystkie wektory własne, więc i wszystkie wartości własne, a największa, to $\lambda_1\cdot\ldots\cdot\lambda_k$ i to trzeba było tu pokazać.

(5) udowodnimy w dwóch małych krokach. Po pierwsze $(A^{\wedge k})^{\dagger} = (A^{\dagger})^{\wedge k}$, bo wystarczy sprawdzić, że

$$< A^{\wedge k}\omega, \eta > = < \omega, A^{\wedge k}\eta >, \quad \omega,\eta \in \bigwedge^k(\mathbb{C}^n)^*,$$

ale tę równość wystarczy sprawdzić na bazie ortonormalnej $\omega = \mathrm{d}x_I$, $\eta = \mathrm{d}x_J$ dualnej do bazy standardowej (e_k) (używamy dalej multiindeksów $I = (i_1,\ldots,i_k)$); wówczas, skoro $(\det_I(v_1,\ldots,v_k)$ oznacza wyznacznik macierzy powstałej przez wybranie wierszy $i_1,\ldots,i_k$ z kolum $v_1,\ldots v_k)$

$$A^{\wedge k}\mathrm{d}x_I = \sum_L \det{}_L(Ae_{i_1},\ldots,Ae_{i_k})\mathrm{d}x_L,$$

to

$$\begin{aligned}< \mathrm{d}x_I, (A^{\dagger})^{\wedge k}\mathrm{d}x_J > &= \overline{\det{}_I(A^{\dagger}e_{j_1},\ldots,A^{\dagger}e_{j_k})} = \det{}_I(A^T e_{j_1},\ldots,A^T e_{j_k})\\ &= \det{}_J(Ae_{i_1},\ldots,Ae_{i_k}) = < A^{\wedge k}\mathrm{d}x_I, \mathrm{d}x_J >.\end{aligned}$$

Po drugie

$$\begin{aligned}|A^{\wedge k}| &= \sqrt{(A^{\wedge k})^\dagger (A^{\wedge k})} = \sqrt{(A^\dagger)^{\wedge k} A^{\wedge k}} \\ &= \sqrt{(AA^\dagger)^{\wedge k}} \overset{4}{=} (\sqrt{AA^\dagger})^{\wedge k} = |A^\dagger|^{\wedge k}.\end{aligned}$$

□

4. Nierówność Goldena–Thompsona

Celem tutaj jest dowód następującego

Twierdzenie 6 (Nierówność Goldena–Thompsona)**.** *Dla operatorów samosprzężonych $A, B\colon C^n \to \mathbb{C}^n$ zachodzi*

$$\operatorname{tr} e^{A+B} \leqslant \operatorname{tr}(e^A e^B).$$

Potrzebne będzie kilka lematów.

Lemat 3. *Dla operatorów samosprzężonych $A, B\colon C^n \to \mathbb{C}^n$ dodatnich (tzn. $< Ax, x >> 0$, dla każdego $x \neq 0$) zachodzi*

$$\|AB\|_{sp}^p \leqslant \|A^p B^p\|_s, \quad p \geqslant 1, s \geqslant 2.$$

Dowód. Ustalmy p, s jak w założeniach i oznaczmy $C := A^p$, $D := B^p$, $f(\alpha) := \|C^\alpha D^\alpha\|_{s/p}$ — funkcja ciągła na $(0,1]$. Sprawdźmy czy przedłuża się ona na $[0,1]$, tzn. chcemy obliczyć $\lim_{\alpha\to 0+} f(\alpha)$. Szacujemy

$$\begin{aligned}\|C^\alpha D^\alpha\|_{s/\alpha} &\overset{\text{Hölder}}{\leqslant} \|C^\alpha\|_{2s/\alpha} \|D^\alpha\|_{2s/\alpha} \\ &= (\operatorname{tr} C \operatorname{tr} D)^{\alpha/2s} \xrightarrow[\alpha\to 0+]{} 1,\end{aligned}$$

dalej, biorąc za v dowolny jednostkowy wektor własny dla D o wartości własnej μ mamy z drugiej strony

$$\begin{aligned}\|C^\alpha D^\alpha\|_{s/p} &\overset{\text{porównanie p-tych norm ciągów}}{\geqslant} \|C^\alpha D \alpha\|_\infty = \|C^\alpha D^\alpha\| \\ &\geqslant \|C^\alpha D^\alpha v\| = \left\| \sum_{k=1}^n \mu^\alpha \lambda_k^\alpha < v, u_k > u_k \right\| \\ &\geqslant (|\mu| min_{1\leqslant k\leqslant n} |\lambda_k|)^\alpha \xrightarrow[\alpha\to 0+]{} 1,\end{aligned}$$

gdzie korzystaliśmy jeszcze z twierdzenia spektralnego dla C (oczywiście (u_k) oznacza bazę ortonormalna wektorów własnych odpowiadających wartościom własnym (λ_k)). Zatem $f(0) = 1$. Zauważmy, że teza, to nierówność

$$f\left(\frac{1}{p}\right) \leqslant f(1) = f(1) f(0),$$

co równoważnie można zapisać jako

$$f\left(\frac{1}{p} \cdot 1 + \left(1 - \frac{1}{p}\right) \cdot 0\right) \leqslant f(1)^{1/p} f(0)^{1-1/p},$$

czyli wystarczy udowodnić wypukłość funkcji $\ln f$. Wobec jej ciągłości wyniknie, to z nierówności

$$f\left(\frac{\alpha+\beta}{2}\right) \leqslant f(\alpha)^{1/2} f(\beta)^{1/2}. \tag{6}$$

Do jej dowodu przydadzą się trzy rzeczy.

$$\|X\|_{pq}^p = (\operatorname{tr}|X|^{pq})^{1/q} = \||X|^p\|_q, \quad p, q \geqslant 1 \quad X\text{— dowolny} \tag{7}$$

$$\|XY\|_p \leqslant \|YX\|_p, \quad X, Y\text{— dowolne takie, że } XY \text{ samosprzężony,} \tag{8}$$

bo wobec lematu 1 (zasady majoryzacyjnej) i definicji p-tej normy macierzowej wystarczy pokazać, że dla każdego $k = 1, \ldots, n$

$$\prod_{j=1}^{k} \mu_j XY \leqslant \prod_{j=1}^{k} \mu_j YX.$$

Jednak mamy

$$\begin{aligned}
\prod_{j=1}^{k} \mu_j(XY) &= \mu_1(XY)^{\wedge k}) = \max \text{ modułów wartości własnych} Y^{\wedge k} X^{\wedge k} \\
&= r(Y^{\wedge k} X^{\wedge k}) = r(X^{\wedge k} Y^{\wedge k}) = r((YX)^{\wedge k}) \\
&\leqslant \|(YX)^{\wedge k}\| = \mu_1\left((YX)^{\wedge k}\right) = \prod_{j=1}^{k} \mu_j YX,
\end{aligned}$$

gdzie przez $r(X)$ oznaczamy promień spektralny operatora X i korzystamy ze znanych faktów, że ogólnie $r(XY) = r(YX)$, bo XY i YX mają te same wartości własne oraz $r(X) \leqslant \|X\|$, bo jeśli u jest jednostkowym wektorem własnym o wartości własnej λ, to $|\lambda| = \|Xu\| \leqslant \|X\|$.

$$\|X^\dagger\|_p = \|X\|_p, \tag{9}$$

oczywiste, wobec uwagi 2.

Wracamy do dowodu (6). Bez utraty ogólności załóżmy, że $\beta \geqslant \alpha$. Mamy

$$\begin{aligned}
f\left(\frac{\alpha+\beta}{2}\right) &= \|C^{\frac{\alpha+\beta}{2}} D^{\frac{\alpha+\beta}{2}}\|_{2s/(\alpha+\beta)} \overset{(7)}{=} \||C^{\frac{\alpha+\beta}{2}} D^{\frac{\alpha+\beta}{2}}|^2\|_{s/(\alpha+\beta)} \\
&= \|D^{\frac{\alpha+\beta}{2}} C^{\alpha+\beta} C^{\frac{\alpha+\beta}{2}}\|_{s/(\alpha+\beta)}^{1/2} = \|D^\alpha D \frac{\beta-\alpha}{2} C^{\alpha+\beta} D^{\frac{\alpha+\beta}{2}}\| \\
&\overset{(8)}{=} \|D^\alpha C^{\alpha+\beta} D^\beta\|_{s/(\alpha+\beta)}^{1/2} \overset{\text{Hölder}}{\leqslant} \|D^\alpha C^\alpha\|_{s/\alpha}^{1/2} \|C^\beta D^\beta\|_{s/\beta}^{1/2} \\
&\overset{(9)}{=} f(\alpha)^{1/2} f(\beta)^{1/2}.
\end{aligned}$$

□

Przy okazji można stąd udowodnić taki przyjemny

Wniosek 4. *Przy założeniach lematu 3 zachodzi także nierówność*

$$\operatorname{tr}(AB)^n \leqslant \operatorname{tr}(A^n B^n).$$

Dowód. Korzystając z cykliczności śladu otrzymujemy

$$\begin{aligned}
\operatorname{tr}(AB)^n \quad &= \quad \operatorname{tr}\left(A^{1/2}\underbrace{(A^{1/2}BA^{1/2})\cdot\ldots\cdot(A^{1/2}BA^{1/2})}_{n-1}A^{1/2}B\right) \\
&= \quad \operatorname{tr}(A^{1/2}BA^{1/2}) = \operatorname{tr}|B^{1/2}A^{1/2}|^{2n} = \|B^{1/2}A^{1/2}\|_{2n}^{2n} \\
&\overset{\text{lemat 3}}{\leqslant} \quad \|B^{n/2}A^{n/2}\|_2^2 = \operatorname{tr}(A^{n/2}B^nA^{n/2}) = \operatorname{tr}(A^nB^n).
\end{aligned}$$

□

Lemat 4. *Dla dowolnych operatorów $X, Y\colon \mathbb{C}^n \to \mathbb{C}^n$ zachodzi*

$$e^{X+Y} = \lim_{n\to\infty}(e^{\frac{1}{n}X}e^{\frac{1}{n}Y})^n = \lim_{n\to\infty}(e^{\frac{1}{2n}X}e^{\frac{1}{n}Y}e^{\frac{1}{2n}X})^n,$$

gdzie zbieżność jest w normie operatorowej.

Dowód. Niech $A_n := e^{\frac{1}{n}X}e^{\frac{1}{n}Y}, B_n := e^{\frac{1}{n}(X+Y)}$. Chemy udowodnić, że $\|A_n^n - B_n^n\| \xrightarrow[n\to\infty]{} 0$, ale

$$\begin{aligned}
\|A_n\|, \|B_n\| \quad &\leqslant \quad e^{\frac{1}{n}(\|X\|+\|Y\|)} \\
\|A_n - B_n\| \quad &\leqslant \quad \frac{C}{n^2},
\end{aligned}$$

bo pierwsze wyrazy odpowiednich szeregów się redukują i stała C zależy tylko od operatorów X, Y (nie zależy od n). Mamy stąd

$$\begin{aligned}
\|A_n^n - B_n^n\| \quad &= \quad \|A_n^{n-1}(A_n - B_n) + A_n^{n-2}(A_n - B_n)B_n + A_n^{n-3}(A_n - B_n)B_n^2 \\
&\quad + \ldots + A_n(A_n - B_n)B_n^{n-2} + (A_n - B_n)B_n^{n-1}\| \\
&\leqslant \quad ne^{\|X\|+\|Y\|}\|A_n - B_n\| \leqslant \frac{C}{n}e^{\|X\|+\|Y\|} \xrightarrow[n\to\infty]{} 0.
\end{aligned}$$

Dowód drugiej równości jest identyczny. □

Lemat 5. *Dla operatorów samosprzężonych $A, B\colon \mathbb{C}^n \to \mathbb{C}^n$ zachodzi*

$$\|e^{A+B}\|_p \leqslant \|e^Ae^B\|_p, \quad p \geqslant 1.$$

Dowód. Niech najpierw $p \geqslant 2$; wtedy

$$\|e^{\frac{1}{2n}A}e^{\frac{1}{2n}B}\|_{2np}^{2n} \overset{\text{lemat 3}}{\leqslant} \|e^Ae^B\|_p,$$

ale lewa strona wynosi

$$\||e^{\frac{1}{2n}A}e^{\frac{1}{2n}B}|^{2n}\|_p = \|(e^{\frac{1}{2n}B}e^{\frac{1}{n}A}e^{\frac{1}{2n}B})^n\|_p \xrightarrow[n\to\infty]{\text{lemat 4}} \|e^(A+B)\|,$$

skąd teza w tym przypadku.

Jeśli $1 \leqslant p < 2$, to mamy z poprzedniego przypadku

$$\begin{aligned} \|e^{A+B}\|_p &= \|e^{\frac{A+B}{2}}\|_{2p}^2 \leqslant \|e^{\frac{1}{2}A}e^{\frac{1}{2}B}\|_{2p}^2 \\ &= \||e^{\frac{1}{2}A}e^{\frac{1}{2}B}|^2\|_p = \|e^{\frac{1}{2}B}e^{A}e^{\frac{1}{2}B}\|_p \\ &\overset{(8)}{\leqslant} \|e^A e^B\|_p. \end{aligned}$$

□

Teraz łatwo zakończyć dowód wyjściowego twierdzenia.

Dowód twierdzenia 6. Pamiętając, że założenia $A+B$ jest też samosprzężony, mamy

$$\begin{aligned} \operatorname{tr} e^{A+B} &= \|e^{\frac{1}{2}(A+B)}\|_2^2 \overset{\text{lemat 5}}{\leqslant} \|e^{\frac{1}{2}A}e^{\frac{1}{2}B}\|_2^2 \\ &= \operatorname{tr}\left(e^{\frac{1}{2}B}e^{A}e^{\frac{1}{2}B}\right) = \operatorname{tr}(e^A e^B). \end{aligned}$$

□

Jako zastosowanie wszystkiego podamy jeszcze bardzo ważny

Wniosek 5. *Funkcja* $f\colon \{\text{operatory samosprzężone } \mathbb{C}^n \to \mathbb{C}^n\} \longrightarrow \mathbb{R}$ *zadana wzorem* $f(A) := \ln\left(\operatorname{tr} e^A\right)$ *jest wypukła.*

Dowód. Po prostu sprawdzamy definicję wypukłości

$$\begin{aligned} f(\alpha A + (1-\alpha)B) &= \ln\left(\operatorname{tr} e^{\alpha A + (1-\alpha)B}\right) \\ &\overset{\text{Golden–Thompson}}{\leqslant} \ln\left(\operatorname{tr}(e^{\alpha A}e^{(1-\alpha)B})\right) \\ &\overset{\text{Hölder}}{\leqslant} \ln\left((\operatorname{tr} e^A)^{\alpha}(\operatorname{tr} e^B)^{1-\alpha}\right) \\ &= \alpha f(A) + (1-\alpha) f(B). \end{aligned}$$

□

Literatura

[1] M. Kuczma, *An introduction to the theory of functional equations and inequalities. Cauchy's equation and Jensen's inequality*, Prace naukowe Uniwersytetu Śląskiego w Katowicach, nr 489, Państwowe Wydawnictwo Naukowe i Uniwersytet Śląski, Warszawa Kraków Katowice 1985

[2] T. Koma, H. Tasaki, *Decay of Superconducting and Magnetic Correlations in One and Two–Dimensional Hubbard Models*, Phys. Rev. Lett. **68**, 2348(1992)

[3] B. Simon, *The tatistical Mechanics of Lattice Gases*, Princeton, New Jersey, Princeton University Press

[4] J. Wojtkiewicz, *Estimations of free energy for the Hubbard model*