Uniwersytet Warszawski
Wydział Matematyki, Informatyki i Mechaniki

Paweł Wolff

Hiperkontrakcja i nierówności funkcyjne dla miar produktowych

rozprawa doktorska

Promotor rozprawy
dr hab. Krzysztof Oleszkiewicz, prof. UW
Instytut Matematyki UW

Czerwiec 2008

Oświadczenie autora rozprawy:
oświadczam, że niniejsza rozprawa została napisana przeze mnie samodzielnie.

. . . .

data *Paweł Wolff*

Oświadczenie promotora rozprawy:
niniejsza rozprawa jest gotowa do oceny przez recenzentów.

. . . .

data *dr hab. Krzysztof Oleszkiewicz, prof. UW*

Streszczenie

Niniejsza rozprawa poświęcona jest kilku aspektom hiperkontraktywnych półgrup Markowa, hiperkontraktywnych rzeczywistych zmiennych losowych oraz nierówności funkcyjnych typu entropii-energii. Rezultaty związane z hiperkontraktywnością obejmują: oszacowania optymalnych stałych hiperkontrakcji dla półgrupy Markowa działającej na skończonej, dyskretnej przestrzeni probabilistycznej; pełną charakteryzację hiperkontraktywnych zmiennych losowych w terminach ich rozkładu oraz geometrii współistniejącej przestrzeni liniowej unormowanej. W kontekście nierówności funkcyjnych typu entropii-energii, badana jest własność tensoryzacji funkcjonałów φ-entropii oraz ich iteracji. Udowodniona jest także pewna nowa nierówność koncentracyjna dla miary gaussowskiej.

Słowa kluczowe: hiperkontrakcja, koncentracja miary, miara produktowa, nierówność funkcyjna, własność tensoryzacji

2000 Mathematics Subject Classification: Primary: 60E15; Secondary: 28A35, 28D10, 46B20, 60A10.

Abstract

The thesis is devoted to several aspects of hypercontractive semigroups, hypercontractive real-valued random variables and functional inequalities of entropy-energy type. The results concerned with the hypercontractivity include: estimates of the optimal hypercontractivity constants for a natural Markov semigroup acting on a discrete finite probability space; full characterization of hypercontractive random variables in terms of their distribution and geometry of underlying linear normed space. In the context of the functional inequalities of entropy-energy type, the tensorization property of φ-entropy functionals and their iterations is investigated. Inspired by the latter, some new example of a concentration inequality for Gaussian measure is established.

Keywords: functional inequality, hypercontractivity, measure concentration, product measure, tensorization property

2000 Mathematics Subject Classification: Primary: 60E15; Secondary: 28A35, 28D10, 46B20, 60A10.

Spis treści

Wstęp

Wiele klasycznych zagadnień Teorii Prawdopodobieństwa zajmuje się badaniem własności produktów miar probabilistycznych. Wystarczy wspomnieć twierdzenia graniczne dla sum niezależnych zmiennych losowych: Prawo Wielkich Liczb, Centralne Twierdzenie Graniczne czy Prawo Iterowanego Logarytmu. Cechą wspólną dla tych zagadnień jest to, że rozważane produkty miar nie mają z góry ustalonego wymiaru, a często wręcz wymiar dąży do nieskończoności. Kluczowym narzędziem do badania własności produktów miar są nierówności probabilistyczne. Mówiąc bardzo ogólnie, niniejsza rozprawa traktuje o różnych aspektach nierówności hiperkontrakcyjnych, nierówności funkcyjnych typu entropii-energii (do których zalicza się nierówność Poincaré i logarytmiczną nierówność Sobolewa) i koncentracji miary.

Rozprawa podzielona jest na dwie części. Część pierwsza, „Hiperkontrakcja", składa się z trzech rozdziałów. Rozdział 1, wprowadzający do tej części, zawiera definicje rozważanych dalej pojęć, ich podstawowe własności oraz omawia znane wyniki, jak również ich zastosowania. Kolejne dwa rozdziały składają się już w znacznej mierze z wyników własnych autora (chyba, że w danym miejscu zaznaczone jest inaczej). Rozdział 2 traktuje o hiperkontrakcji w kontekście półgrup Markowa. Zamieszczone tam wyniki mają charakter wyraźnie ilościowy i dotyczą oszacowań tzw. stałych hiperkontrakcji. Rozdział 3 dotyczy hiperkontraktywnych zmiennych losowych. Wyniki w tym rozdziale są natury jakościowej i przedstawiają pewne warunki konieczne i dostateczne dla hiperkontraktywności zmiennych losowych.

Część druga rozprawy, „Nierówności funkcyjne dla miar produktowych", także składa się z trzech rozdziałów. Tu również Rozdział 4 stanowi wprowadzenie, zaś dwa kolejne rozdziały przedstawiają wyniki autora. I tak, w Rozdziale 5 badana jest własność tensoryzacji pewnej rodziny nierówności funkcyjnych, zwanych nierównościami entropii-energii. W Rozdziale 6 omówiona jest pewna nowa nierówność koncentracyjna dla miary gaussowskiej.

Podziękowania. Chciałbym w tym miejscu wyrazić swoją wdzięczność Promotorowi, prof. Krzysztofowi Oleszkiewiczowi, za lata sprawowanej nade mną opieki naukowej. Niniejsza rozprawa nie powstałaby bez jego pomocy: począwszy od propozycji problemów badawczych, poprzez dyskusje i wskazówki na ich temat, aż po szczegółowe uwagi nt. tekstu pracy doktorskiej. Chciałbym także podziękować prof. Stanisławowi Kwapieniowi za wskazówki do publikowanych przeze mnie prac, m.in. na bazie których powstała ta rozprawa. Podziękowania kieruję także do prof. Rafała Latały. Dużą część mojej wiedzy z Teorii Prawdopodobieństwa i Analizy zawdzięczam właśnie jemu. Wiele zawdzięczam także rozmowom z moimi kolegami, byłymi doktorantami: dr. Radosławem Adamczakiem i dr. Witoldem Bednorzem.

Osobno chciałbym podziękować Mariuszowi Baryło, doktorantowi Wydziału Matematyki, Informatyki i Mechaniki UW, za nieocenioną pomoc przy ostatnich etapach przygotowywania rozprawy.

Praca doktorska powstała przy wsparciu Ministerstwa Nauki i Szkolnictwa Wyższego z grantu promotorskiego N N201 0226 33 oraz z grantu PO3A 012 29.

Oznaczenia

W rozprawie będziemy używali następujących oznaczeń:
— $\mathbb{R}, \mathbb{N}$ — zbiory liczb rzeczywistych, naturalnych;
— $\#A$ oznacza moc zbioru A;
— $|\cdot|$ oznacza normę euklidesową;
— $\mathbf{1}_A(\omega) = \begin{cases} 1, & \text{gdy } \omega \in A, \\ 0, & \text{gdy } \omega \notin A; \end{cases}$
— dla przestrzeni topologicznej X, $\mathcal{B}(X)$ oznacza rodzinę zbiorów borelowskich w X; jeśli X jest zwarta, przez $C(X)$ będziemy oznaczać przestrzeń funkcji ciągłych z metryką supremum;
— $L^p(\Omega, \mathcal{F}, \mu)$ $(1 \leqslant p \leqslant \infty)$ — przestrzeń L^p (ze standardową normą) funkcji określonych na przestrzeni miarowej $(\Omega, \mathcal{F}, \mu)$ i o wartościach rzeczywistych; często będziemy stosować krótsze oznaczenie $L^p(\mu)$, jeśli z kontekstu będzie jasne, z jaką przestrzenią miarową mamy do czynienia;
— ℓ_p, ℓ_p^n $(1 \leqslant p \leqslant \infty, n \in \mathbb{N})$ — przestrzenie ciągowe (rzeczywiste), odpowiednio nieskończenie- i n-wymiarowe;
— przez δ_a rozumiemy miarę probabilistyczną skupioną w punkcie a (deltę Diraca w a);
— jeśli μ jest miarą probabilistyczną na pewnej przestrzeni mierzalnej $(\Omega, \mathcal{F})$, to przez $\mathbb{E}_\mu$ będziemy oznaczać funkcjonał wartości średniej $L^1(\Omega, \mathcal{F}, \mu) \to \mathbb{R}$, tj. $\mathbb{E}_\mu f = \int f\, d\mu$;
— w przestrzeni probabilistycznej $(\Omega, \mathcal{F}, \mathbb{P})$, wartość oczekiwaną oznaczamy symbolem $\mathbb{E}$;
— dla zmiennej losowej X (o wartościach w ośrodkowej przestrzeni liniowo-topologicznej) przez $\mathcal{L}(X)$ rozumiemy jej rozkład (czyli odpowiednią miarę probabilistyczną borelowską);
— przez F będziemy oznaczać przestrzeń liniową unormowaną, z normą $\|\cdot\|$;
— dla zmiennej losowej X przyjmującej wartości w ośrodkowej przestrzeni liniowej unormowanej F, przez $\|X\|_p$ (dla $0 < p < \infty$) rozumiemy $\mathbb{E}\Big(\|X\|^p\Big)^{1/p}$, zaś $\|X\|_\infty = \operatorname{ess\,sup} \|X\|$; dla $p \geqslant 1$ zachodzi nierówność trójkąta: $\|X + Y\|_p \leqslant \|X\|_p + \|Y\|_p$;
— Lin — powłoka liniowa zadanego zbioru wektorów;
— przez $\mathcal{C}^k$ będziemy oznaczać klasę funkcji k-krotnie różniczkowalnych w sposób ciągły; funkcje, o których będziemy mówić, że należą do klasy $\mathcal{C}^k$, będą miały jako dziedzinę podzbiór otwarty przestrzeni $\mathbb{R}^n$ lub, nieco ogólniej, przestrzeni liniowej (unormowanej) V wymiaru skończonego;
— dla funkcji $F\colon V \to \mathbb{R}$ klasy $\mathcal{C}^1$ (V — jak wyżej) i $x_0 \in V$, przez $\mathrm{D}F(x_0)$ będziemy

oznaczać różniczkę F w punkcie x_0, rozumianą jako element przestrzeni V^* (wtedy $\big(\mathrm{D}F(x_0)\big)(v)$ dla $v \in V$ jest pochodną kierunkową F w punkcie x_0, w kierunku v);
— dla funkcji $F\colon V \to \mathbb{R}$ klasy $\mathcal{C}^k$ (V — jak wyżej) i $x_0 \in V$, przez $\mathrm{D}^k F(x_0)$ będziemy oznaczać różniczkę rzędu k z funkcji F w punkcie x_0, rozumianą jako przekształcenie k-liniowe $V \times \cdots \times V \to \mathbb{R}$ (wtedy $\big(\mathrm{D}^k F(x_0)\big)(v_1, \ldots, v_k)$ dla $v_i \in V$ jest pochodną kierunkową rzędu k z funkcji F w punkcie x_0, w kierunkach $v_1, \ldots, v_n$);
— dla funkcji $F \in \mathbb{R}^n \to \mathbb{R}$ i $x_0 \in \mathbb{R}^n$, przez $\mathrm{Hess} F(x_0)$ rozumiemy macierz pochodnych cząstkowych rzędu 2 funkcji F w x_0;
— jeśli A jest macierzą (w szczególności wektorem), to przez A^t rozumiemy jej transpozycję; współrzędne wektora w $\mathbb{R}^n$ będą reprezentowane w macierzy wymiaru $n \times 1$ (czyli jako wektor pionowy);
— dla $a, b \in \mathbb{R}$, $a \vee b = \max(a, b)$, $a \wedge b = \min(a, b)$;
— jeśli dla $x, y \geqslant 0$ oraz $c \geqslant 1$ spełniona jest nierówność

$$c^{-1} x \leqslant y \leqslant cx, \tag{1}$$

to będziemy pisać $x \overset{c}{\simeq} y$; z reguły w takim kontekście x i y będą funkcjami pewnej liczby zmiennych, natomiast c będzie stałą numeryczną lub stałą zależącą od ściśle określonego parametru; tak więc np. jeśli piszemy $x \overset{c}{\simeq} y$, przy czym c jest stałą zależną tylko od pewnego parametru ε, to oznacza to tyle, że dla każdej możliwej wartości ε istnieje stała c taka, że nierówność (1) zachodzi dla wszystkich (z wyjątkiem ε, który jest w tym momencie ustalony) wartości zmiennych, od których zależą x i y; w przypadku, gdy stała c będzie stałą numeryczną (nie zależącą od żadnych parametrów), często nie będziemy jej jawnie podawać pisząc $x \simeq y$;
— podobnie, dla $x, y \geqslant 0$, pisząc $x \lesssim y$ będziemy rozumieć, że istnieje stała numeryczna $c > 0$, dla której zachodzi nierówność $x \leqslant cy$, dla wszystkich możliwych wartości zmiennych, od których zależą x i y.

Część I

Hiperkontrakcja

Rozdział 1

Wprowadzenie

Hiperkontrakcja, w klasycznym rozumieniu, jest własnością półgrupy operatorów. Jeden ze standardowych kontekstów dla tego pojęcia jest następujący (patrz np. [2, Ch. 2], [17, Ch. 1]). Niech Ω będzie przestrzenią polską[1], a $(P_t)_{t\geqslant 0}$ —- półgrupą operatorów Markowa z przestrzenią stanów Ω, działającą na przestrzeni funkcji borelowskich ograniczonych o wartościach rzeczywistych, zadaną przez jedno-parametrową rodzinę stochastycznych jąder przejścia $(Q_t(x,dy))_{t\geqslant 0}$:

$$P_t f(x) = \int_\Omega Q_t(x,dy)\, f(y). \tag{1.1}$$

Przypomnijmy, że od $Q_t(x,dy)$ żądamy następujących własności:
(i) dla każdego $x\in\Omega$, $Q_t(x,\cdot)$ jest borelowską miarą probabilistyczną na Ω, przy czym $Q_0(x,\cdot)=\delta_x$,
(ii) dla każdego $A\in\mathcal{B}(\Omega)$, $Q_t(\cdot,A)$ jest funkcją borelowską,
(iii) dla każdych $x\in\Omega$, $A\in\mathcal{B}(\Omega)$ i $s,t\geqslant 0$ zachodzi równanie Chapmana-Kołmogorowa:

$$Q_{s+t}(x,A) = \int_\Omega Q_s(x,dy)\, Q_t(y,A).$$

Z powyższych własności wynika, że $(P_t)_{t\geqslant 0}$ jest istotnie półgrupą operatorów (tj. $P_0 = Id$, $P_{s+t} = P_s\circ P_t$) i dla każdego $t\geqslant 0$,
(iv) $P_t 1 = 1$,
(v) jeśli $f\geqslant 0$, to $P_t f\geqslant 0$ (tj. operator P_t *zachowuje nieujemność*).
Przypuśćmy ponadto, że istnieje dokładnie jedna probabilistyczna miara borelowska μ na Ω, która jest miarą niezmienniczą dla półgrupy $(P_t)_{t\geqslant 0}$, tzn. dla dowolnej funkcji borelowskiej ograniczonej $f\colon\Omega\to\mathbb{R}$ oraz dowolnego $t\geqslant 0$,

$$\int P_t f\, d\mu = \int f\, d\mu. \tag{1.2}$$

W takiej sytuacji, półgrupę $(P_t)_{t\geqslant 0}$ możemy rozszerzyć do półgrupy działającej na przestrzeni $L^p(\mu)$ dla dowolnego $1\leqslant p\leqslant\infty$. Co więcej, otrzymane półgrupy są półgrupami operatorów kontrakcji. Istotnie, dla dowolnej funkcji borelowskiej ograniczonej $f\colon\Omega\to\mathbb{R}$, używając (iv) i (v) pokazujemy, że dla każdej funkcji wypukłej $\varphi\colon\mathbb{R}\to\mathbb{R}$, $P_t\varphi(f)\geqslant\varphi(P_t f)$ (punktowo). Biorąc $\varphi(u)=|u|^p$ $(1\leqslant p<\infty)$ i korzystając z (1.2), otrzymujemy $\|P_t f\|_{L^p(\mu)}\leqslant\|f\|_{L^p(\mu)}$. Dla $p=\infty$ wystarczy zauważyć, że dla dowolnej funkcji borelowskiej ograniczonej $f\colon\Omega\to\mathbb{R}$, $f\geqslant 0$ μ-p.w. implikuje $P_t f\geqslant 0$ μ-p.w.

[1] Rozważane przykłady przestrzeni Ω będą obejmowały przypadki: $\mathbb{R},\mathbb{R}^n$ oraz skończony zbiór dyskretny.

Dla uproszczenia notacji, rozszerzenia półgrupy $(P_t)_{t\geqslant 0}$ do półgrup działających na przestrzeniach $L^p(\mu)$ będziemy nadal oznaczali jako $(P_t)_{t\geqslant 0}$.

W dalszych rozważaniach, na półgrupę $(P_t)_{t\geqslant 0}$ będziemy nakładali mocniejszy warunek aniżeli (1.2): dla dowolnych funkcji borelowskich ograniczonych $f, g\colon \Omega \to \mathbb{R}$,

$$\int f P_t g \, d\mu = \int g P_t f \, d\mu. \tag{1.3}$$

W takiej sytuacji, półgrupa $P_t\colon L^2(\mu) \to L^2(\mu)$ jest półgrupą operatorów samosprzężonych.

Często półgrupy $(P_t)_{t\geqslant 0}$ będziemy definiowali w oparciu o generator infinitezymalny, czyli wzorem $P_t = e^{tL}$. Z reguły będziemy jednak też podawać bardziej jawną reprezentację $(P_t)_{t\geqslant 0}$, taką jak w (1.1), z której będzie wynikać, że $(P_t)_{t\geqslant 0}$ ma wszystkie wymienione wyżej własności.

W opisanej powyżej sytuacji, tj. gdy półgrupa operatorów $(P_t)_{t\geqslant 0}$ ma własności (iv), (v), a wraz z pewną miarą probabilistyczną μ spełnia (1.3), możemy wprowadzić następującą definicję:

Definicja 1.1. Półgrupa $(P_t)_{t\geqslant 0}$ jest (p, q)-*hiperkontraktywna* $(1 \leqslant q < p \leqslant \infty)$, jeśli istnieje stała $t_0 = t_0(p, q) \geqslant 0$, dla której operator P_{t_0} rozszerza się do operatora ciągłego $P_{t_0}\colon L^q(\mu) \to L^p(\mu)$, o normie 1.

W dalszym ciągu mówiąc, że dana półgrupa jest *hiperkontraktywna*, będziemy rozumieć, że jest (p, q)-hiperkontraktywna dla pewnych $1 \leqslant q < p \leqslant \infty$.

1.1. Krótki rys historyczny. Przegląd klasycznych wyników i ich zastosowań

Pierwsze przykłady hiperkontraktywnych półgrup Markowa pochodzą z fizyki matematycznej — półgrupa Ornsteina-Uhlenbecka, oraz z analizy harmonicznej — półgrupa operatorów splotowych na kostce dyskretnej $\{-1, 1\}^n$.

Przez (standardową) półgrupę Ornsteina-Uhlenbecka w przestrzeni euklidesowej $\mathbb{R}^n$ będziemy rozumieć półgrupę operatorów $S_t = e^{tL}$, gdzie generator infinitezymalny L dany jest wzorem

$$Lf = \Delta f - \langle x, \nabla f \rangle.$$

Przypomnijmy, że kanoniczna miara gaussowska $\gamma_n(dx) = (2\pi)^{-n/2} e^{-|x|^2/2} dx$ jest odwracalną miarą niezmienniczą dla półgrupy $(S_t)_{t\geqslant 0}$, oraz że znana jest jawna reprezentacja S_t w takiej postaci jak (1.1), która w pewnej wersji, zwanej często *wzorem Mehlera*, wygląda następująco:

$$S_t f(x) = \int_{\mathbb{R}^n} f(e^{-t}x + \sqrt{1 - e^{-2t}}y) \, \gamma_n(dy).$$

Powyższy wzór dobrze definiuje S_t jako operator $L^p(\gamma_n) \to L^p(\gamma_n)$ dla wszystkich $p \in [1, \infty]$.

Hiperkontraktywność półgrupy Ornsteina-Uhlenbecka pochodzi od E. Nelsona i I. Segala. Była ona jednym z podstawowych faktów w rozwijanej wówczas kwantowej teorii pola. W pracy [34], Nelson pokazał, że dla dostatecznie dużych $t_0 = t_0(p)$, operator S_{t_0} rozszerza się do operatora ciągłego $L^2(\gamma_n) \to L^p(\gamma_n)$, z normą niezależącą od wymiaru przestrzeni n. Hiperkontraktywność (a więc i normę 1 dla operatora $S_{t_0}\colon L^q \to L^p$) uzyskał Segal [41, Theorem 1], lecz ograniczenie na $t_0 = t_0(p,q)$ nie było optymalne. Optymalny wynik kilka lat później otrzymał Nelson [35] — ograniczenie na t_0 to

$$e^{-t_0} \leqslant \sqrt{(q-1)/(p-1)}. \tag{1.4}$$

To samo zjawisko hiperkontraktywności półgrupy operatorów pojawiło się, niezależnie, w pracy A. Bonami [8] dotyczącej analizy harmonicznej na grupach dyskretnych. Badała ona półgrupę operatorów splotowych $(T_t)_{t\geqslant 0}$ na kostce dyskretnej $\mathbb{Z}_2^n \cong \{-1,1\}^n$,

$$T_t f = f * \mu_{e^{-t}}^{\otimes n}, \tag{1.5}$$

gdzie $\mu_a^{\otimes n}$ oznacza n-krotny produkt miar dwupunktowych $\frac{1+a}{2}\delta_1 + \frac{1-a}{2}\delta_{-1}$. Oczywiście odwracalną miarą niezmienniczą μ jest w tym przypadku produkt miar dwupunktowych $\frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1}$. Tak zdefiniowaną półgrupę $(T_t)_{t\geqslant 0}$, na potrzeby niniejszej części rozprawy, będziemy nazywali *(symetryczną) półgrupą Bernoulliego*. Wynik [8, Ch. 3, Théorème 3] mówi, że dla t_0 spełniającego (1.4), $T_{t_0}\colon L^q(\mu) \to L^p(\mu)$ jest operatorem ciągłym, o normie 1.

Ten sam wynik (jak również pewną wersję dla operatorów T_s, zdefiniowanych formalnie także wzorem (1.5), lecz dla $s = t - i\pi/2$, $t \geqslant 0$) uzyskał później także Beckner [4] w pracy poświęconej optymalnym stałym w nierówności Hausdorffa-Younga. Przedstawił on też argument oparty o Centralne Twierdzenie Graniczne, który pozwala wydedukować np. hiperkontrakcję półgrupy Ornsteina-Uhlenbecka z hiperkontrakcji półgrupy Bernoulliego.

L. Gross w pracy [16] badał infinitezymalną wersję nierówności hiperkontrakcyjnej

$$\|P_t f\|_{L^p(\mu)} \leqslant \|f\|_{L^q(\mu)},$$

tzw. logarytmiczną nierówność Sobolewa:

$$\mathbb{E}_\mu |f|^r \ln |f|^r - \mathbb{E}_\mu |f|^r \ln \mathbb{E}_\mu |f|^r \leqslant C(r)\, \mathbb{E}_\mu |f|^{r-1} L f, \tag{1.6}$$

gdzie $-L$ jest generatorem półgrupy Markowa $(P_t)_{t\geqslant 0}$, tzn. $P_t = e^{-tL}$. Mówiąc w skrócie, logarytmiczna nierówność Sobolewa (z wykładnikiem r), związana z danym generatorem półgrupy Markowa, jest konsekwencją hiperkontraktywności półgrupy pomiędzy przestrzeniami $L^r(\mu)$ i $L^{r+\varepsilon}(\mu)$, przy $\varepsilon \to 0$ ([16, Theorem 2]). Z kolei wiedząc, że logarytmiczna nierówność Sobolewa zachodzi dla wykładników $r \in (a,b)$, gdzie $1 \leqslant a < b \leqslant \infty$, [16, Theorem 1] pokazuje, że półgrupa jest (p,q)-hiperkontraktywna dla dowolnych $p, q \in (a,b)$, $p > q$. Wspomniane twierdzenia podają ponadto precyzyjny, ilościowy związek optymalnej stałej $C(r)$ w (1.6) z zachowaniem najlepszej (czyli najmniejszej) stałej $t_0 = t_0(p,q)$, dla której $P_{t_0}\colon L_q \to L_p$ jest kontrakcją.

Użyteczność pojęcia hiperkontrakcji wynika m.in. ze stabilności na „tensorowanie". Zachodzi bowiem następujący, abstrakcyjny fakt, zwany często lematem Segala [41, Lemma 1.4], choć akurat przytoczona tutaj wersja pochodzi z pracy Bonami [8, Ch. 3, Lemme 1]).

Stwierdzenie 1.1 ([41, 8]). *Dla $i = 1, 2$, niech $(\Omega_i, \mathcal{F}_i, \mu_i)$ i $(\Sigma_i, \mathcal{M}_i, \nu_i)$ będą przestrzeniami miarowymi (z miarami dodatnimi, σ-skończonymi), zaś $T^{(i)} \colon L^q(\mu_i) \to L^p(\nu_i)$ operatorem ciągłym, o normie A_i. Przypuśćmy, że $p \geqslant q \geqslant 1$. Dalej będziemy rozważać przestrzenie: $L^q(\mu_1 \otimes \mu_2) = L^q(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2, \mu_1 \otimes \mu_2)$ i $L^p(\nu_1 \otimes \nu_2) = L^p(\Sigma_1 \times \Sigma_2, \mathcal{M}_1 \otimes \mathcal{M}_2, \nu_1 \otimes \nu_2)$. Niech $Y \subset L^q(\mu_1 \otimes \mu_2)$ będzie gęstą podprzestrzenią złożoną z funkcji postaci*

$$f(\omega_1, \omega_2) = \sum_{j \leqslant n} g_j(\omega_1) h_j(\omega_2), \text{ gdzie } n \in \mathbb{N},\ \omega_1 \in \Omega_1,\ \omega_2 \in \Omega_2,\ g_j \in L^q(\mu_1),\ h_j \in L^q(\mu_2).$$

Zdefiniujmy operator $T \colon Y \to L^p(\nu_1 \otimes \nu_2)$ jako produkt operatorów: $T = T^{(1)} \otimes T^{(2)}$, co wyraża się wzorem

$$(Tf)(\sigma_1, \sigma_2) = \sum_{j \leqslant n} \left(T^{(1)} g_j\right)(\sigma_1) \cdot \left(T^{(2)} h_j\right)(\sigma_2),$$

gdzie f jest takiej postaci jak wyżej.

Wówczas T rozszerza się (jednoznacznie) do operatora ciągłego $T \colon L^q(\mu_1 \otimes \mu_2) \to L^p(\nu_1 \otimes \nu_2)$ o normie nie większej niż $A_1 A_2$.

Dowód podany w pracy [8] opiera się na zastosowaniu nierówności Minkowskiego w odpowiedniej przestrzeni $L^{p/q}$.

Z powyższego stwierdzenia wynika natychmiast, że hiperkontraktywność zarówno półgrupy Ornsteina-Uhlenbecka, jak i półgrupy Bernoulliego wystarczy badać w wymiarze $n = 1$, gdyż obydwie mają strukturę produktową. Dokładniej, jeśli $(S_t)_{t \geqslant 0}$ jest półgrupą Ornsteina-Uhlenbecka na prostej, to n-krotny produkt, $S_t^{\otimes n}$, jest operatorem z półgrupy Ornsteina-Uhlenbecka na $\mathbb{R}^n$.

D. Bakry i M. Émery [3] podali warunek dostateczny (tzw. *kryterium* Γ_2) na to, aby zadana półgrupa Markowa, będąca w klasie tzw. półgrup dyfuzji, była hiperkontraktywna. Rozważmy najprostszą sytuację, kiedy przestrzeń stanów to $\mathbb{R}^n$, a półgrupa dyfuzji $(P_t)_{t \geqslant 0}$ zadana jest przez operator infinitezymalny

$$L = \Delta - \langle \nabla U, \nabla \rangle,$$

gdzie $U \colon \mathbb{R}^n \to \mathbb{R}$ jest funkcją gładką. Jeśli $e^{-U(x)} dx$ jest miarą skończoną, to

$$\mu(dx) = \frac{1}{Z} e^{-U(x)} dx,$$

dla odpowiedniego $Z > 0$ jest miarą niezmienniczą dla $(P_t)_{t \geqslant 0}$, spełniającą (1.3). Ponadto, jeśli istnieje stała $\lambda > 0$, dla której

$$\forall_{x \in \mathbb{R}^n} \quad \operatorname{Hess} U(x) \geqslant \lambda Id, \tag{1.7}$$

to półgrupa $(P_t)_{t \geqslant 0}$ jest hiperkontraktywna.

Nieco wcześniej, C. Borell [10] zauważył, że wynik Bonami o hiperkontraktywności półgrupy Bernoulliego może być wykorzystany od badania chaosów rademacherowych, również tych o współczynnikach będących wektorami przestrzeni Banacha. Praca W. Krakowiaka i J. Szulgi [21] stanowi rozwinięcie idei zawartych w pracach Borella [10, 11] na przypadek form wieloliniowych od innych zmiennych niż zmienne Bernoulliego (zob. także [23, Ch. 3]). Właśnie w pracy [21] zostało wprowadzone pojęcie hiperkontraktywnej zmiennej losowej:

Definicja 1.2. Niech $0 < q < p < \infty$, F będzie przestrzenią liniową unormowaną, zaś θ — niezdegenerowaną zmienną losową (tzn. taką, której rozkład nie jest skupiony w jednym punkcie) o wartościach rzeczywistych. Powiemy, że θ jest (p, q)-*hiperkontraktywna (ze stałą* $\sigma > 0$*) w przestrzeni* F, jeśli $\mathbb{E}|\theta|^p < \infty$ oraz

$$\forall_{x,y \in F} \quad \|x + \sigma\theta y\|_p \leqslant \|x + \theta y\|_q. \tag{1.8}$$

Hiperkontraktywne zmienne losowe posiadają analogiczną własność do tej, która dla hiperkontraktywnych półgrup operatorów wynika ze Stwierdzenia 1.1. Własność ta mówi w skrócie, że nierówność (1.8) przenosi się np. na sumy niezależnych kopii zmiennej θ lub na formy wieloliniowe, których argumentami są niezależne kopie zmiennej θ. Na przykład dla sum, odpowiedni fakt można sformułować następująco:

Stwierdzenie 1.2 ([22], [23, Ch. 3] — w obydwu sformułowanie nieco ogólniejsze)**.** *Niech* $0 < q < p < \infty$*. Jeśli* $(\theta_k)_{k=1}^n$ *jest ciągiem niezależnych zmiennych losowych, z których każda jest* (p, q)*-hiperkontraktywna w przestrzeni unormowanej* F *ze stałą* $\sigma > 0$*, to wówczas dla dowolnych wektorów* $x, y_1, \ldots, y_n \in F$*,*

$$\Big\|x + \sigma \sum_{k=1}^n \theta_k y_k\Big\|_p \leqslant \Big\|x + \sum_{k=1}^n \theta_k y_k\Big\|_q.$$

Kładąc w powyższej nierówności $x = 0$ otrzymamy nierówność typu Chinczyna-Kahane'a. Dla form wieloliniowych, odpowiednia nierówność, wynikająca z hiperkontraktywności ciągu zmiennych θ_k, wygląda następująco: dla dowolnych wektorów $x \in F$ i $y_I \in F$, gdzie $I \subset \{1, \ldots, n\}$, $\#I = m$ $(m \leqslant n)$,

$$\Big\|x + \sigma^m \sum_{\substack{I \subset \{1,\ldots,n\} \\ \#I = m}} y_I \prod_{i \in I} \theta_i\Big\|_p \leqslant \Big\|x + \sum_{\substack{I \subset \{1,\ldots,n\} \\ \#I = m}} y_I \prod_{i \in I} \theta_i\Big\|_q.$$

S. Kwapień i J. Szulga [22] uzyskali pewną charakteryzację hiperkontraktywnych zmiennych losowych wśród zmiennych o rozkładzie symetrycznym. Charakteryzacja ta obejmuje dwa skrajne przypadki: hiperkontraktywność w przestrzeni $F = \mathbb{R}$ oraz hiperkontraktywność w dowolnej przestrzeni liniowej unormowanej F (równoważnie, w przestrzeni $F = \ell_\infty$). R. Latała w swojej pracy magisterskiej [24] rozszerzył te wyniki, między innymi pozbywając się założenia symetryczności.

Nierówności hiperkontrakcyjne, które jak już wspomnieliśmy, wywodzą się z fizyki matematycznej i analizy harmonicznej, mają zastosowanie również w wielu innych działach matematyki. Wspomnieliśmy już o zastosowaniu idei hiperkontrakcji do badania własności zmiennych losowych, sum niezależnych kopii czy też form wieloliniowych od takich zmiennych (patrz np. [10], [21], [22]). Hiperkontraktywne oszacowania dla półgrup Markowa (czy też logarytmiczne nierówności Sobolewa) są silnym narzędziem do badania ich tempa zbieżności do miary stacjonarnej (patrz np. [14]). Ciekawe zastosowanie hiperkontraktywności półgrupy $(T_t)_{t\geqslant 0}$ zawiera praca [20]. Podano tam optymalne co do rzędu rozwiązanie pewnego problemu z kombinatoryki ekstremalnej. Wynik ten uzyskano przy użyciu dyskretnej analizy harmonicznej, gdzie centralnym wynikiem była właśnie nierówność hiperkontrakcyjna pochodząca od Bonami. Dzięki tym narzędziom uzyskano wiele innych rezultatów natury kombinatorycznej. W tym momencie wymienię tylko wynik M. Talagranda [45] oraz niedawny wynik I. Dinur, E. Friedguta i O. Regeva [15], w którym znalazł zastosowanie odpowiednik nierówności Bonami dla *niesymetrycznej półgrupy Bernoulliego* (czyli takiej, której miara stacjonarna jest produktem niekoniecznie symetrycznych miar dwupunktowych). Tę ostatnią nierówność, wraz z optymalnymi stałymi $t_0(p,q)$, gdy p lub q jest równe 2 (patrz Definicja 1.1), uzyskał K. Oleszkiewicz [38].

1.2. Wyniki zawarte w tej części rozprawy

Zawartość kolejnych rozdziałów z tej części rozprawy wpisuje się w kontekst nakreślony w poprzednim punkcie.

Rozdział 2 zawiera wyniki dotyczące oszacowań stałych hiperkontrakcji $t_0(p,q)$ dla pewnej prostej półgrupy Markowa. Rozpatrywane przypadki obejmują niesymetryczną półgrupę Bernoulliego oraz półgrupy związane z probabilistycznymi miarami dyskretnymi o skończonej liczbie atomów. Podane oszacowania stałych $\sigma_0 = e^{-t_0(p,q)}$, dla parametrów p, q leżących po jednej stronie 2, są optymalne z dokładnością do czynnika uniwersalnego. Wyniki z tego rozdziału zostały opublikowane w pracy [46].

Natomiast rezultaty przedstawione w Rozdziale 3 stanowią uogólnienie wspomnianych wyników Kwapienia i Szulgi [22] oraz Latały [24]. Głównym tematem tego rozdziału jest bowiem charakteryzacja zmiennych losowych, które są hiperkontraktywne w ustalonej, lecz dowolnej, przestrzeni liniowej unormowanej. Uzyskana charakteryzacja, zdaniem autora, lepiej pokazuje związki hiperkontrakcji zmiennych losowych z geometrią przestrzeni (głównie z jej izometrycznymi niezmiennikami, takimi jak *moduł wypukłości*).

Rozdział 2

Oszacowania stałych hiperkontrakcji dla rozkładów dyskretnych

2.1. Definicje i podstawowe fakty

W niniejszym rozdziale przyjmiemy następujące oznaczenia. Ω będzie dowolnym zbiorem skończonym, zaś μ będzie miarą probabilistyczną określoną na wszystkich podzbiorach Ω. Aby nie rozważać przypadków trywialnych, umówmy się, że μ ma zawsze przynajmniej dwa atomy oraz miara każdego atomu jest niezerowa. W takim kontekście, przestrzenie $L^p(\Omega, \mu) = L^p(\mu)$ mają oczywiście skończony wymiar (równy $\#\Omega$).

Niech L będzie rzutem ortogonalnym w $L^2(\mu)$ na podprzestrzeń funkcji o zerowej średniej, tzn. $L = Id - \mathbb{E}_\mu$. Będziemy rozważać półgrupy operatorów $(T_t)_{t\geqslant 0}$, zadane wzorem $T_t = e^{-tL}$. Zauważmy, że ze względu na to, że L jest rzutem ortogonalnym, operatory półgrupy $(T_t)_{t\geqslant 0}$ wyrażają się prostą formułą

$$T_t = \mathbb{E}_\mu + e^{-t}L = e^{-t}Id + (1 - e^{-t})\mathbb{E}_\mu, \tag{2.1}$$

co po przepisaniu do postaci (1.1) daje

$$T_t f(x) = \int_\Omega Q_t(x, dy)\, f(y), \text{ gdzie } Q_t(x, \cdot) = e^{-t}\delta_x + (1 - e^{-t})\mu. \tag{2.2}$$

Zatem półgrupa $(T_t)_{t\geqslant 0}$ wraz z miarą μ spełniają warunki (iv), (v) przedstawione na początku Rozdziału 1 oraz równanie (1.3), co w szczególności oznacza, że $(T_t)_{t\geqslant 0}$, jako półgrupa działająca na $L^2(\mu)$, jest samosprzężona.

Wykładniki p i q zawsze będą spełniały $1 < q < p < \infty$, natomiast α i β, występujące jako wagi atomów miary dwupunktowej, będą w związku $\alpha + \beta = 1$. Ponadto będziemy zakładać, że $\alpha \in (0, 1/2]$.

Przypomnijmy, że zgodnie z Definicją 1.1, półgrupa $(T_t)_{t\geqslant 0}$ jest (p, q)-*hiperkontraktywna*, jeśli istnieje takie $t_0 \geqslant 0$, że dla dowolnego $f \in L^q(\mu)$,

$$\|T_{t_0} f\|_{L^p(\mu)} \leqslant \|f\|_{L^q(\mu)}\,. \tag{2.3}$$

Jeśli t_0 o powyższej własności istnieje i jest najmniejsze możliwe (minimum istnieje, choćby dlatego, że $(T_t)_{t\geqslant 0}$ jest półgrupą mocno ciągłą w L^p), to przez *stałą* (p, q)-*hiperkontrakcji dla miary* μ będziemy rozumieć stałą

$$\sigma_{p,q}(\mu) := e^{-t_0}.$$

Zauważmy, że z nietrywialności miary μ wynika, iż $t_0 > 0$, a więc $\sigma_{p,q}(\mu) \in (0, 1)$.

Oczywiście hiperkontraktywność półgrupy $(T_t)_{t \geqslant 0}$ i związane z tym stałe $\sigma_{p,q}(\mu)$ nie zależą od samego zbioru Ω, ale jedynie od wag atomów miary μ. W szczególności, jeśli $\#\Omega = 2$, to możemy przyjąć, że $\mu = \beta\delta_{-\alpha} + \alpha\delta_\beta$ dla pewnego α i w związku z tym będziemy używali oznaczenia

$$\sigma_{p,q}(\alpha) := \sigma_{p,q}\left(\beta\delta_{-\alpha} + \alpha\delta_\beta\right).$$

Ponieważ operator T_t jest samosprzężony (jako operator działający na $L^2(\mu)$), to mamy następującą równość norm

$$\|T_t\|_{L^q(\mu) \to L^p(\mu)} = \|T_t\|_{L^{p'}(\mu) \to L^{q'}(\mu)},$$

gdzie $p' = \frac{p}{p-1}$ i $q' = \frac{q}{q-1}$. Stąd już wynika, że

$$\sigma_{p,q}(\mu) = \sigma_{q',p'}(\mu). \tag{2.4}$$

Kolejna obserwacja wynika z faktu, że operatory T_t zachowują nieujemność. Jeśli $f\colon \Omega \to \mathbb{R}$ zapiszemy w postaci $f = a + g$, gdzie $a \in \mathbb{R}$, zaś $\mathbb{E}_\mu g = 0$, to z formuły (2.1) wynika, że $T_t f = a + e^{-t} g$. Operator T_t zachowuje nieujemność, więc $T_t|f| \geqslant T_t f$ oraz $T_t|-f| \geqslant -T_t f$, czyli $T_t|f| \geqslant |T_t f|$. W związku z tym wystarczy sprawdzać nierówność (2.3) tylko dla $f = a + g \geqslant 0$. Wówczas, pomijając trywialny przypadek $g \equiv 0$, a musi być ściśle dodatnie, a więc ze względu na jednorodność nierówności (2.3), możemy zakładać, że $a = 1$. Tym samym

$$\begin{aligned}\sigma_{p,q}(\mu) = \max\{\sigma \in (0,1) \,\big|\, & \|1 + \sigma g\|_{L^p(\mu)} \leqslant \|1 + g\|_{L^q(\mu)} \\ & \text{dla wszystkich takich } g \text{, że } \mathbb{E}_\mu g = 0 \text{ i } 1 + g \geqslant 0\},\end{aligned} \tag{2.5}$$

co w przypadku miar dwupunktowych $\mu = \beta\delta_{-\alpha} + \alpha\delta_\beta$ oznacza, że

$$\begin{aligned}\sigma_{p,q}(\alpha) = \max\{\sigma \in (0,1) \,\big|\, & (\beta|1 - \sigma\alpha u|^p + \alpha|1 + \sigma\beta u|^p)^{1/p} \\ & \leqslant (\beta|1 - \alpha u|^q + \alpha|1 + \beta u|^q)^{1/q} \\ & \text{dla wszystkich } u \in [-1/\beta, 1/\alpha]\}.\end{aligned} \tag{2.6}$$

Na koniec podamy stwierdzenie wiążące hiperkontraktywność półgrupy $(T_t)_{t \geqslant 0}$ z hiperkontraktywnością odpowiadającej jej zmiennej losowej.

Stwierdzenie 2.1. *Przypuśćmy, że μ jest dyskretną miarą probabilistyczną na prostej rzeczywistej, o średniej zero, zaś θ jest zmienną losową o rozkładzie μ. Wówczas θ jest (p, q)-hiperkontraktywna ze stałą $\sigma_{p,q}(\mu)$ w dowolnej przestrzeni liniowej unormowanej F.*

Dowód. Sprawdzamy, że (1.8) zachodzi ze stałą $\sigma = \sigma_{p,q}(\mu)$. Niech t_0 będzie takie, że $e^{-t_0} = \sigma$. Rozważamy półgrupę $(T_t)_{t \geqslant 0}$, odpowiadającą mierze μ, tj. $T_t = e^{-t(Id - \mathbb{E}_\mu)}$, oraz funkcję wypukłą $f(s) = \|x + sy\|$, dla dowolnych, ustalonych wektorów $x, y \in F$. Wówczas, stosując nierówność Jensena, otrzymamy

$$\left(T_{t_0} f\right)(s) = (1 - \sigma)\mathbb{E}_\mu f + \sigma f(s) \geqslant (1 - \sigma) f(0) + \sigma f(s) \geqslant f(\sigma s),$$

więc $\|x + \sigma\theta y\|_p = \|f(\sigma\theta)\|_p \leqslant \|(T_{t_0} f)(\theta)\|_p \leqslant \|f(\theta)\|_q = \|x + \theta y\|_q$. $\square$

2.2. Rozkłady dwupunktowe — sformułowania wyników

Klasyczny rezultat pochodzący od Bonami [8, Ch. 3, Théorème 3] i Becknera [4] orzeka, że dla dowolnych $p > q > 1$,

$$\sigma_{p,q}(1/2) = \sqrt{\frac{q-1}{p-1}}. \tag{2.7}$$

Optymalne stałe hiperkontrakcji dla niesymetrycznych zmiennych dwupunktowych zostały znalezione przez Oleszkiewicza [38] w przypadku, gdy p lub q jest równe 2:

$$\sigma_{2,q}(\alpha) = \left(\frac{\beta^{2-2/q} - \alpha^{2-2/q}}{\alpha^{1-2/q}\beta - \beta^{1-2/q}\alpha}\right)^{1/2}, \quad \sigma_{p,2}(\alpha) = \left(\frac{\beta^{2/p} - \alpha^{2/p}}{\alpha^{2/p-1}\beta - \beta^{2/p-1}\alpha}\right)^{1/2},$$

gdzie $\beta = 1 - \alpha$. Łatwo sprawdzić, że mają miejsce naturalne zależności:

$$\lim_{\alpha\to 1/2^-} \sigma_{2,q}(\alpha) = \sigma_{2,q}(1/2) \quad \text{oraz} \quad \lim_{\alpha\to 1/2^-} \sigma_{p,2}(\alpha) = \sigma_{p,2}(1/2).$$

Powyższe rezultaty pokazują, że wielkość $\sigma_{p,q}(\alpha)$ jest dobrze zdefiniowana dla dowolnych $p > q > 1$ i $\alpha \in (0,1)$ (tj. odpowiednia półgrupa operatorów jest (p,q)-hiperkontraktywna). W oczywisty sposób zachodzą bowiem nierówności $\sigma_{p,q}(\alpha) \geqslant \sigma_{2,q}(\alpha)$ dla $p \leqslant 2$, $\sigma_{p,q}(\alpha) \geqslant \sigma_{p,2}(\alpha)$ dla $q \geqslant 2$, oraz $\sigma_{p,q}(\alpha) \geqslant \sigma_{2,q}(\alpha)\sigma_{p,2}(\alpha)$ dla $q < 2 < p$.

Elementarny rachunek pokazuje asymptotyczne zachowanie $\sigma_{2,q}(\alpha)$ oraz $\sigma_{p,2}(\alpha)$ przy $\alpha \to 0$ lub $q \to 1$ ($p \to \infty$) (patrz [38, Theorem 2.1]):

$$\sigma_{2,q}(\alpha) \simeq \begin{cases} \alpha^{\frac{1}{q}-\frac{1}{2}}, & \text{gdy } \frac{1}{q-1} \leqslant \ln\frac{1}{\alpha}, \\ \sqrt{(q-1)\alpha\ln(1/\alpha)}, & \text{gdy } \frac{1}{q-1} > \ln\frac{1}{\alpha}, \end{cases}$$

$$\sigma_{p,2}(\alpha) \simeq \begin{cases} \alpha^{\frac{1}{2}-\frac{1}{p}}, & \text{gdy } p \leqslant \ln\frac{1}{\alpha}, \\ \sqrt{\frac{1}{p}\alpha\ln(1/\alpha)}, & \text{gdy } p > \ln\frac{1}{\alpha}. \end{cases}$$

W Punkcie 2.3 potrzebna nam będzie monotoniczność funkcji $\alpha \mapsto \sigma_{2,q}(\alpha)$:

Lemat 2.2. *Dla* $1 < q < 2 < p < \infty$*, funkcje* $(0,1/2] \ni \alpha \mapsto \sigma_{2,q}(\alpha)$ *oraz* $(0,1/2] \ni \alpha \mapsto \sigma_{p,2}(\alpha)$ *są ściśle rosnące.*

Dowód powyższego lematu, jak i pozostałych rezultatów sformułowanych w tym punkcie zamieszczamy w Punkcie 2.4

Główny wynik bieżącego punktu dotyczy oszacowań stałych $\sigma_{p,q}(\alpha)$ dla dowolnych wartości parametrów α, p, q z wyjątkiem przypadku $q < 2 < p$.

Twierdzenie 2.3. $\sigma_{p,q}(\alpha) \simeq \tilde{\sigma}_{p,q}(\alpha)$ *dla dowolnych* $1 < q < p \leqslant 2$, $\alpha \in I := (0,1/2]$, *gdzie*

$$\tilde{\sigma}_{p,q}(\alpha) = \begin{cases} \alpha^{\frac{1}{q}-\frac{1}{p}} & \textit{dla } \alpha \in I_1 = \left\{\alpha \in I \,\middle|\, \frac{e^{\frac{1}{q-1}}}{q-1} \leqslant \frac{1}{\alpha}\ln\frac{1}{\alpha}\right\}, \\ \left((q-1)\ln\frac{1}{\alpha}\right)^{\frac{1}{p}} \alpha^{1-\frac{1}{p}} & \textit{dla } \alpha \in I_2 = \left\{\alpha \in I \,\middle|\, \frac{e^{\frac{1}{p-1}}}{q-1} \leqslant \frac{1}{\alpha}\ln\frac{1}{\alpha} < \frac{e^{\frac{1}{q-1}}}{q-1}\right\}, \\ \frac{q-1}{p-1}\frac{\ln(1/\alpha)}{1+\ln\left(\frac{q-1}{p-1}\frac{1}{\alpha}\ln\frac{1}{\alpha}\right)} & \textit{dla } \alpha \in I_3 = \left\{\alpha \in I \,\middle|\, \frac{p-1}{q-1} \leqslant \frac{1}{\alpha}\ln\frac{1}{\alpha} < \frac{e^{\frac{1}{p-1}}}{q-1}\right\}, \\ \sqrt{\frac{q-1}{p-1}\alpha\ln(1/\alpha)} & \textit{dla } \alpha \in I_4 = \left\{\alpha \in I \,\middle|\, \frac{1}{\alpha}\ln\frac{1}{\alpha} < \frac{p-1}{q-1}\right\}. \end{cases}$$

Podkreślmy, że zbiory I_j ($j = 1, \dots, 4$) z powyższego twierdzenia stanowią rozbicie $(0, 1/2]$ na rozłączne przedziały (I_4 może być pusty), oraz jeśli $\alpha_j \in I_j$ (dla $j = 1, \dots, 4$), to $\alpha_1 < \alpha_2 < \alpha_3 < \alpha_4$.

Na mocy (2.4), jako natychmiastowy wniosek z poprzedniego twierdzenia dostajemy

Twierdzenie 2.4. *$\sigma_{p,q}(\alpha) \simeq \tilde{\sigma}_{p,q}(\alpha)$ dla dowolnych $2 \leqslant q < p < \infty$, $\alpha \in I := (0, 1/2]$, gdzie*

$$\tilde{\sigma}_{p,q}(\alpha) = \begin{cases} \alpha^{\frac{1}{q}-\frac{1}{p}} & \text{dla } \alpha \in I_1' = \{\alpha \in I \mid pe^p \leqslant \frac{1}{\alpha}\ln\frac{1}{\alpha}\}, \\ \left(\frac{\ln(1/\alpha)}{p}\right)^{1-\frac{1}{q}} \alpha^{\frac{1}{q}} & \text{dla } \alpha \in I_2' = \{\alpha \in I \mid pe^q \leqslant \frac{1}{\alpha}\ln\frac{1}{\alpha} < pe^p\}, \\ \frac{q}{p}\frac{\ln(1/\alpha)}{1+\ln\left(\frac{q}{p}\frac{1}{\alpha}\ln\frac{1}{\alpha}\right)} & \text{dla } \alpha \in I_3' = \{\alpha \in I \mid \frac{p}{q} \leqslant \frac{1}{\alpha}\ln\frac{1}{\alpha} < pe^q\}, \\ \sqrt{\frac{q}{p}\alpha\ln(1/\alpha)} & \text{dla } \alpha \in I_4' = \{\alpha \in I \mid \frac{1}{\alpha}\ln\frac{1}{\alpha} < \frac{p}{q}\}. \end{cases}$$

Tak jak w przypadku stałych $\sigma_{2,q}(\alpha)$ i $\sigma_{p,2}(\alpha)$, w Punkcie 2.3 potrzebny będzie nam

Lemat 2.5. *Dla dowolnych p, q takich, że $1 < q < p \leqslant 2$ lub $2 \leqslant q < p < \infty$, $\sigma_{p,q}(\alpha_1) \lesssim \sigma_{p,q}(\alpha_2)$, gdzie $0 < \alpha_1 < \alpha_2 \leqslant 1/2$.*

Oczywiście, Twierdzenie 2.4 wraz ze Stwierdzeniami 2.1 i 1.2 implikują, że jeśli (θ_k) jest ciągiem niezależnych zmiennych losowych o wspólnych rozkładzie $\beta\delta_{-\alpha} + \alpha\delta_\beta$, to dla dowolnej przestrzeni unormowanej F i wektorów $v_1, \dots, v_n \in F$, oznaczając $S = \sum_{k=1}^n \theta_k v_k$, mamy

$$\|S\|_q \gtrsim \tilde{\sigma}_{p,q}(\alpha)\, \|S\|_p\,. \tag{2.8}$$

A priori nie jest jasne, czy stała w powyższej nierówności dla momentów sumy S nie może być lepsza (tzn. większa) co do rzędu (cały czas zaniedbujemy czynnik $\simeq 1$). Innymi słowy nie jest jasne, czy metoda hiperkontrakcji musi dawać optymalne (co do rzędu) stałe w takich nierównościach jak (2.8). Jednak wynik Latały [25], dotyczący oszacowań momentów sum niezależnych zmiennych losowych pozwala wykazać, że faktycznie rzędu stałej w nierówności (2.8) nie można zwiększyć:

Twierdzenie 2.6. *Niech $(\theta_k)_{k=1}^\infty$ będzie ciągiem niezależnych zmiennych losowych o wspólnym rozkładzie $\beta\delta_{-\alpha} + \alpha\delta_\beta$. Połóżmy $n(\alpha, p) = \lceil p/\ln(1/\alpha)\rceil$ oraz $S = \sum_{k=1}^{n(\alpha,p)} \theta_k$. Wówczas dla dowolnych $2 \leqslant q < p < \infty$,*

$$\|S\|_q \simeq \tilde{\sigma}_{p,q}(\alpha)\, \|S\|_p\,.$$

W powyższym twierdzeniu warto zwrócić uwagę, że opisywany przez nie przypadek równości (z dokładnością do czynnika $\simeq 1$) w nierówności (2.8) występuje choćby na prostej (tj. $F = \mathbb{R}$), gdy $v_1 = \dots = v_n = 1$ oraz $n = \lceil p/\ln(1/\alpha)\rceil$. Wydaje się interesującym fakt, że n nie musi zależeć od q (czyli niższego momentu). Ponadto, jak zobaczymy w dowodzie Twierdzenia 2.6, n można także zdefiniować jako największą liczbę całkowitą, dla której $\|\sum_{k=1}^n \theta_k\|_p \overset{c}{\simeq} \|\sum_{k=1}^n \theta_k\|_\infty$, gdzie $c \geqslant 1$ jest pewną, dostatecznie dużą, lecz konkretną stałą.

2.3. Rozkłady dyskretne

Twierdzenie 2.7. *Niech μ będzie dyskretną miarą probabilistyczną, której masa najmniejszego atomu jest równa $\alpha_* > 0$. Wówczas dla dowolnych $1 < q < p < \infty$,*

$$\inf_{\alpha\in[\alpha_*,1/2]} \sigma_{p,q}(\alpha) \leqslant \sigma_{p,q}(\mu) \leqslant \sigma_{p,q}(\alpha_*).$$

Dowód. Druga nierówność jest dość oczywista (patrz np. (2.5)). Dalej będziemy rozważać jedynie przypadek $p \geqslant 2$, gdyż przypadek $1 < q < p < 2$ wyniknie już z (2.4). Aby udowodnić pierwszą nierówność z tezy, wystarczy pokazać, że dla każdego $t \geqslant 0$,

$$\|T_t\|_{L^q(\mu)\to L^p(\mu)} \leqslant \sup_{\alpha\in[\alpha_*,1/2]} \left\|S_t^{(\alpha)}\right\|_{L^q(\nu^{(\alpha)})\to L^p(\nu^{(\alpha)})}, \tag{2.9}$$

gdzie $(T_t)_{t\geqslant 0}$ jest półgrupą związaną z miarą μ, zaś $(S_t^{(\alpha)})_{t\geqslant 0}$ jest półgrupą związaną z miarą dwupunktową $\nu^{(\alpha)} = \beta\delta_{-\alpha} + \alpha\delta_\beta$. Argument wariacyjny, który poniżej zastosujemy, jest podobny do tego, który został użyty w pracy Diaconis i Saloff-Coste [14], a pochodzi od Rothausa [40]. Główna różnica polega na tym, że argument ten stosujemy do nierówności hiperkontrakcyjnej, podczas gdy w cytowanych pracach argument ten pojawiał się w kontekście logarytmicznych nierówności Sobolewa.

Ustalmy $t \geqslant 0$ i połóżmy

$$A = \|T_t\|_{L^q(\mu)\to L^p(\mu)} = \sup\left\{\|T_t f\|_{L^p(\mu)} \;\middle|\; f \in L^q(\mu), \|f\|_{L^q(\mu)} = 1\right\}.$$

Oczywiście, jeśli $A = 1$, to (2.9) w sposób trywialny zachodzi. Załóżmy więc, że $A > 1$. Wymiar przestrzeni $L^q(\mu)$ jest skończony, więc $A < \infty$ oraz istnieje f_0, dla którego supremum definiujące wielkość A jest osiągane. Zdefiniujmy dwa funkcjonały działające na $L^q(\mu)$: $I(f) = \|T_t f\|^p_{L^p(\mu)}$ oraz $J(f) = \|f\|^q_{L^q(\mu)}$. Wówczas w f_0 funkcjonał I przyjmuje maksimum warunkowe, pod warunkiem $J(f) = 1$. Obydwa funkcjonały I oraz J są klasy $\mathcal{C}^1$, a także $\mathrm{D}J(f_0) \neq 0$ (bo $f_0 \neq 0$, gdyż $J(f_0) = 1$). Zatem, na mocy twierdzenia o mnożnikach Lagrange'a, dla pewnej stałej λ zachodzi równość

$$\mathrm{D}I(f_0) = \lambda \mathrm{D}J(f_0). \tag{2.10}$$

Zbadamy teraz, jakie własności musi mieć funkcja f_0. Po pierwsze, z faktu, że operator T_t zachowuje nieujemność, wynika, że f_0 mogliśmy wybrać tak, aby $f_0 \geqslant 0$. Ponadto założenie $A > 1$ wyklucza możliwość, że f_0 jest funkcją stałą. Teraz pokażemy, że f_0 może przyjmować co najwyżej dwie wartości. W tym celu policzmy najpierw pochodne występujące w równaniu (2.10) (poniżej $\langle f, g\rangle_\mu$ oznacza $\mathbb{E}_\mu fg$):

$$\mathrm{D}J(f_0) = q\langle {f_0}^{q-1}, \cdot\rangle_\mu,$$

$$\begin{aligned}\mathrm{D}I(f_0) &= \mathrm{D}\,\|\cdot\|^p_{L^p(\mu)}(T_t f_0) \circ \mathrm{D}T_t(f_0) = p\langle (T_t f_0)^{p-1}, \cdot\rangle_\mu \circ T_t \\ &= p\langle (T_t f_0)^{p-1}, T_t(\cdot)\rangle_\mu = p\langle T_t\left((T_t f_0)^{p-1}\right), \cdot\rangle_\mu\end{aligned}$$

(ostatnia równość wynika z faktu, że T_t jest operatorem samosprzężonym w $L^2(\mu)$, czyli właśnie względem iloczynu skalarnego $\langle\cdot,\cdot\rangle_\mu$). Zatem równość (2.10) oznacza, że

$$T_t\left((T_t f_0)^{p-1}\right) = \tilde{\lambda}^{q-1} {f_0}^{q-1} \tag{2.11}$$

dla pewnej stałej $\tilde{\lambda} \geqslant 0$. Ponieważ $f_0 \geqslant 0$, także $g_0 := T_t f_0 \geqslant 0$ oraz $T_t g_0^{p-1} \geqslant 0$. Zauważmy, że ze wzoru (2.1) wynika, że $f_0 = (T_t)^{-1} g_0 = e^t g_0 + (1 - e^t)\mathbb{E}_\mu g_0$, co po wstawieniu do (2.11) daje

$$\left(e^{-t} g_0^{p-1} + (1 - e^{-t}) \|g_0\|_{L^{p-1}(\mu)}^{p-1}\right)^{\frac{1}{q-1}} = \tilde{\lambda}\left(e^t g_0 + (1 - e^t)\mathbb{E}_\mu g_0\right).$$

Ustalając argument funkcji g_0 i oznaczając wartość g_0 na tym argumencie przez u ($u \geqslant 0$), otrzymujemy

$$\left(\left(a u^{p-1} + b v^{p-1}\right)^{\frac{1}{p-1}}\right)^{\frac{p-1}{q-1}} = cu + d, \tag{2.12}$$

gdzie $a > 0$, $b, v \geqslant 0$ i $c, d \in \mathbb{R}$ są pewnymi stałymi, które nie zależą od dokonanego wyboru argumentu funkcji g_0 (a więc tym samym nie zależą od u). Prawa strona (2.12) jest funkcją liniową zmiennej u, podczas gdy lewa strona (2.12) jest ściśle wypukłą funkcją u. Istotnie, jest to bowiem złożenie wypukłej i ściśle rosnącej funkcji $[0, \infty) \ni u \mapsto (a u^{p-1} + b v^{p-1})^{\frac{1}{p-1}}$ (ze względu na założenie $p \geqslant 2$, ostatnie wyrażenie jest po prostu normą L^{p-1} na pewnej przestrzeni dwupunktowej) oraz funkcji $[0, \infty) \ni z \mapsto z^{\frac{p-1}{q-1}}$, która jest ściśle wypukła i ściśle rosnąca. W związku z powyższym równanie (2.12) ma co najwyżej dwa rozwiązania $u \in [0, \infty)$, a stąd g_0 (oraz f_0) przyjmuje dokładnie dwie wartości (funkcję stałą już wykluczyliśmy). Oznaczmy te wartości przez u_1 i u_2 w taki sposób, że $\mu(f_0^{-1}(u_1)) = \alpha$ i $\mu(f_0^{-1}(u_2)) = 1 - \alpha$ dla pewnego $\alpha \in (0, 1/2]$. Oczywiście $\alpha \in [\alpha_*, 1/2]$, i jak łatwo się przekonać, $\|T_t\|_{L^q(\mu) \to L^p(\mu)} = \|S_t^{(\alpha)}\|_{L^q(\nu^{(\alpha)}) \to L^p(\nu^{(\alpha)})}$, co implikuje żądaną równość (2.9).

□

Wniosek 2.8. *Niech μ i α_* będą jak w Twierdzeniu 2.7. Wówczas dla dowolnych $1 < q < 2 < p < \infty$,*

$$\sigma_{2,q}(\mu) = \sigma_{2,q}(\alpha_*) \quad i \quad \sigma_{p,2}(\mu) = \sigma_{p,2}(\alpha_*),$$

oraz dla dowolnych $1 < q < p < \infty$, spełniających $q, p \leqslant 2$ lub $q, p \geqslant 2$,

$$\sigma_{p,q}(\mu) \simeq \sigma_{p,q}(\alpha_*).$$

Dowód. Pierwsza część wynika z Twierdzenia 2.7 i Lematu 2.2, druga zaś z Twierdzenia 2.7 i Lematu 2.5. □

Warto w tym miejscu może wspomnieć, że Twierdzenie 2.7 pozwala łatwo uogólnić pewną nierówność typu Sobolewa, rozważaną przez Latałę i Oleszkiewicza [27, Remark 2], z przypadku przestrzeni dwupunktowych na przypadek skończonych przestrzeni dyskretnych.

Wniosek 2.9. *Niech μ będzie takie jak w Twierdzeniu 2.7, zaś α będzie masą najmniejszego atomu miary μ. Zdefiniujmy formę kwadratową $\mathcal{E}(f) = \mathbb{E}_\mu f L f$, gdzie $L = Id - \mathbb{E}_\mu$ (wtedy $\mathcal{E}(f)$ jest po prostu wariancją f względem μ). Wówczas dla dowolnego $1 < q < 2$ i $f \in L^2(\mu)$,*

$$\mathbb{E}_\mu f^2 - (\mathbb{E}_\mu |f|^q)^{2/q} \leqslant C_q(\alpha)\mathcal{E}(f), \tag{2.13}$$

gdzie $C_q(\alpha) = \frac{\alpha^{1-2/q} - \beta^{1-2/q}}{\alpha^{1-2/q}\beta - \beta^{1-2/q}\alpha}$, a $\beta = 1 - \alpha$.

Dowód. Rozważamy półgrupę $T_t = e^{-tL}$. Jest to półgrupa operatorów samosprzężonych w $L^2(\mu)$. Wniosek 2.8 implikuje, że dla t_0, dla którego $e^{-t_0} = \sigma_{2,q}(\alpha)$, zachodzi $\|T_{t_0}\|_{L^q(\mu)\to L^2(\mu)} \leqslant 1$, czyli dla dowolnego $f \in L^2(\mu)$,

$$\begin{aligned}\|f\|^2_{L^q(\mu)} \geqslant \|T_{t_0}f\|^2_{L^2(\mu)} &= \langle T_{t_0}f, T_{t_0}f\rangle_\mu = \langle f, T_{2t_0}f\rangle_\mu \\ &= e^{-2t_0}\mathbb{E}_\mu f^2 + (1-e^{-2t_0})(\mathbb{E}_\mu f)^2 = \mathbb{E}_\mu f^2 - (1-e^{-2t_0})\mathcal{E}(f) \\ &= \mathbb{E}_\mu f^2 - C_q(\alpha)\mathcal{E}(f),\end{aligned}$$

gdyż jak łatwo sprawdzić, $C_q(\alpha) = 1 - \sigma^2_{2,q}(\alpha)$. □

Dzieląc obie strony nierówności (2.13) przez $2-q$ i przechodząc do granicy $q \to 2^-$, odzyskujemy (z optymalną stałą) logarytmiczną nierówność Sobolewa dla rozkładów dyskretnych [14, Theorem A.1]:

$$\mathbb{E}_\mu f^2 \ln f^2 - \mathbb{E}_\mu f^2 \ln \mathbb{E}_\mu f^2 \leqslant \frac{\ln\beta - \ln\alpha}{\beta - \alpha}\mathcal{E}(f).$$

Na koniec tego punktu pokażemy, jak z rezultatu sformułowanego we Wniosku 2.8 wyprowadzić nierówność dla momentów sum niezależnych zmiennych losowych dyskretnych o wartościach w przestrzeni liniowej unormowanej.

Wniosek 2.10. *Niech $(X_k)_{k=1}^n$ będzie ciągiem niezależnych dyskretnych zmiennych losowych o wartościach w przestrzeni unormowanej F takich, że $\mathbb{E}X_k = 0$. Połóżmy*

$$\alpha_* := \min\Big\{\mathbb{P}(X_k = v) \;\Big|\; \mathbb{P}(X_k = v) \neq 0, v \in F, k = 1, \dots n\Big\}$$

oraz $S = \sum_{k=1}^n X_k$. Wówczas dla dowolnych $2 \leqslant q < p < \infty$,

$$\|S\|_p \leqslant C\sigma^{-1}_{p,q}(\alpha_*)\,\|S\|_q\,,$$

gdzie $C > 0$ jest stałą uniwersalną, która może być równa 1, gdy $q = 2$.

Dowód. Gdyby zmienne losowe X_k były postaci $\theta_k y_k$, gdzie θ_k jest rzeczywistą zmienną losową, zaś $y_k \in F$, wystarczyłoby, obok Wniosku 2.8, użyć Stwierdzeń 2.1 i 1.2.

W ogólnym przypadku można się odwołać m.in. do wersji Stwierdzenia 1.2, jaka znajduje się np. w [22]. Uogólnić trzeba by także sformułowanie Stwierdzenia 2.1.

Możemy też bezpośrednio użyć hiperkontraktywności odpowiednich półgrup, w podobny sposób jak w [38, Theorem 2.1]. Otóż niech μ_k będzie rozkładem zmiennej X_k, Ω_k będzie nośnikiem μ_k (który jest zbiorem skończonym), natomiast $(T_t^{(k)})_{t\geqslant 0}$ będzie półgrupą związaną z miarą μ_k (patrz wzór (2.1)). Niech t_k spełnia $e^{-t_k} = \sigma_{p,q}(\mu_k)$ i połóżmy $t_0 = \max_{1\leqslant k\leqslant n} t_k$. Wówczas dla każdego $k = 1, \dots, n$, $\|T_{t_0}^{(k)}\|_{L^q(\mu_k)\to L^p(\mu_k)} \leqslant 1$. Wobec tego, na mocy Stwierdzenia 1.1 wiemy, że operator

$$\boldsymbol{T} = T_{t_0}^{(1)} \otimes \cdots \otimes T_{t_0}^{(n)},$$

jako operator z $L^q(\mu_1 \otimes \cdots \otimes \mu_n)$ w $L^p(\mu_1 \otimes \cdots \otimes \mu_n)$ ma normę $\leqslant 1$. Ponadto $\boldsymbol{T}$ zachowuje nieujemność, o czym można się przekonać np. reprezentując $T_t^{(k)}$ tak jak we wzorze (2.2):

$$T_t^{(k)} f = \int_{\Omega_k} Q_t^{(k)}(x, dy)\, f(y),$$

gdzie $Q_t^{(k)}(x, \cdot)$ jest miarą probabilistyczną $e^{-t}\delta_x + (1 - e^{-t})\mu_k$. Wówczas

$$\boldsymbol{Tf}(x_1, \cdots, x_n) = \int_{\Omega_1} \cdots \int_{\Omega_n} Q_{t_0}^{(n)}(x_n, dy_n) \cdots Q_{t_0}^{(1)}(x_1, dy_1)\, \boldsymbol{f}(x_1, \ldots, x_n)$$

a dla dowolnych $x_1 \in \Omega_1, \ldots, x_n \in \Omega_n$ miara produktowa $Q_{t_0}^{(n)}(x_n, dy_n) \cdots Q_{t_0}^{(1)}(x_1, dy_1)$ jest oczywiście nieujemna.

Powyższe własności operatora $\boldsymbol{T}$ wykorzystamy następująco. Dla dowolnego funkcjonału $v^* \in F^*$ o normie 1, rozważamy funkcję $\boldsymbol{f}_{v^*} \colon \Omega_1 \times \cdots \times \Omega_n \to \mathbb{R}$,

$$\boldsymbol{f}_{v^*}(x_1, \ldots, x_n) = v^*(x_1 + \ldots + x_n).$$

Wprowadźmy także $\boldsymbol{g} \colon \Omega_1 \times \cdots \times \Omega_n \to \mathbb{R}$,

$$\boldsymbol{g}(x_1, \ldots, x_n) = \|x_1 + \ldots + x_n\|.$$

Oczywiście $\boldsymbol{f}_{v^*} \leqslant \boldsymbol{g}$ (punktowo), zatem z faktu, że operator $\boldsymbol{T}$ zachowuje nieujemność, wynika, iż

$$\boldsymbol{Tf}_{v^*} \leqslant \boldsymbol{Tg}.$$

Biorąc supremum po v^*, $\|v^*\| = 1$, oraz zauważając, że

$$\begin{aligned}\boldsymbol{Tf}_{v^*}(x_1, \ldots, x_n) &= \sum_{k=1}^{n} T_{t_0}^{(k)}\Big(v^*(x_k)\Big) = \sum_{k=1}^{n} \Big(e^{-t_0} v^*(x_k) + (1 - e^{-t_0})\mathbb{E}_{\mu_k} v^*(x_k)\Big) \\ &= e^{-t_0} \boldsymbol{f}_{v^*}(x_1, \ldots, x_n),\end{aligned}$$

gdzie w ostatniej równości skorzystaliśmy z tego, że miary μ_k mają średnią zero, otrzymujemy

$$e^{-t_0} \sup_{\|v^*\|=1} \boldsymbol{f}_{v^*} = e^{-t_0} \boldsymbol{g} \leqslant \boldsymbol{Tg}.$$

Zatem

$$e^{-t_0} \|\boldsymbol{g}\|_{L^p(\mu_1 \otimes \cdots \otimes \mu_n)} \leqslant \|\boldsymbol{Tg}\|_{L^p(\mu_1 \otimes \cdots \otimes \mu_n)} \leqslant \|\boldsymbol{g}\|_{L^q(\mu_1 \otimes \cdots \otimes \mu_n)}.$$

Wystarczy teraz skorzystać z Wniosku 2.8, aby dostać, że $c\sigma_{p,q}(\alpha_*) \leqslant e^{-t_0} \leqslant \sigma_{p,q}(\alpha_*)$ dla pewnej stałej uniwersalnej $c \in (0, 1]$, która może być równa 1, gdy $q = 2$. □

Oczywiście, Twierdzenie 2.4 pozwala zamienić w nierówności z Wniosku 2.10 stałą $\sigma_{p,q}^{-1}(\alpha_*)$ na $C'\tilde{\sigma}_{p,q}^{-1}(\alpha_*)$, gdzie $C' > 0$ jest pewną stałą numeryczną. Wszystkie te stałe co do rzędu są optymalne, z uwagi na tezę Twierdzenia 2.6.

2.4. Rozkłady dwupunktowe — dowody

Dowód Lematu 2.2. Ze względu na (2.4), wystarczy pokazać, że funkcja $f(\alpha) = \sigma_{2,q}^2(\alpha)$ jest rosnąca na przedziale $\alpha \in (0, 1/2)$. Dla parametru $t \in (0, 1)$ oraz dwóch zmiennych $x, y > 0$, $x \neq y$ określmy funkcję

$$\varphi(x, y) = \frac{x^{1-t} - y^{1-t}}{xy^{-t} - yx^{-t}}.$$

Oczywiście $\varphi(x, y) = \varphi(y, x)$ oraz $\varphi(x, y) = \varphi(\lambda x, \lambda y)$ dla $\lambda > 0$, więc $\partial_x \varphi(x_0, y_0) = \partial_y \varphi(y_0, x_0)$ oraz $\partial_x \varphi(x_0, y_0) = \lambda \partial_x \varphi(\lambda x_0, \lambda y_0)$. Przyjmując $t = (2/q) - 1$, dostajemy $f(\alpha) = \varphi(1 - \alpha, \alpha)$, a zatem

$$\begin{aligned} f'(\alpha) &= \partial_y \varphi(1-\alpha, \alpha) - \partial_x \varphi(1-\alpha, \alpha) \\ &= \frac{1}{1-\alpha} \partial_x \varphi\left(\frac{\alpha}{1-\alpha}, 1\right) - \frac{1}{\alpha} \partial_x \varphi\left(\frac{1-\alpha}{\alpha}, 1\right). \end{aligned}$$

Dowód będzie więc zakończony, jeśli pokażemy, że

$$\partial_x \varphi(x, 1) > 0 \text{ dla } x \in (0, 1) \quad \text{oraz} \quad \partial_x \varphi(x, 1) < 0 \text{ dla } x \in (1, \infty).$$

Oznaczając

$$\psi(x, s) = \frac{1}{s}\left(x^s - x^{-s}\right) = \frac{\ln x}{s} \int_0^s \left(x^u + x^{-u}\right) du \tag{2.14}$$

dla $x > 0$ i $s \in (0, 1]$, otrzymujemy

$$\begin{aligned} \partial_x \varphi(x, 1) &= \frac{1}{(x - x^{-t})^2} \left((1-t) x^{-t} \left(x - x^{-t}\right) - \left(x^{-t+1} - 1\right)\left(1 + t x^{-t-1}\right)\right) \\ &= \frac{t x^{-t}}{(x - x^{-t})^2} \left(\psi(x, t) - \psi(x, 1)\right). \end{aligned}$$

Teraz wystarczy już tylko pokazać, że funkcja $(0, 1] \ni s \mapsto \psi(x, s)$ jest rosnąca dla $x > 1$ oraz malejąca dla $x \in (0, 1)$. Z uwagi na tożsamość $\psi(x, s) = -\psi(1/x, s)$, możemy rozważać jedynie przypadek $x > 1$. Ustalmy $0 < s_1 < s_2 \leqslant 1$ i połóżmy $\lambda = s_1/s_2 < 1$. Korzystając z (2.14) i używając nierówności $x^{\lambda u} + x^{-\lambda u} < x^u + x^{-u}$, która wynika wprost z wypukłości funkcji $u \mapsto x^u$, otrzymujemy

$$\begin{aligned} \psi(x, s_1) &= \frac{\ln x}{s_1} \int_0^{s_2} \left(x^{\lambda u} + x^{-\lambda u}\right) \lambda du \\ &< \frac{\ln x}{s_2} \int_0^{s_2} \left(x^u + x^{-u}\right) du = \psi(x, s_2). \end{aligned}$$

□

Teraz sformułujemy i udowodnimy kilka lematów, które wykorzystamy w dowodzie Twierdzenia 2.3.

Lemat 2.11. *Dla dowolnych* $1 < q < p < \infty$, $0 < \alpha_1 < \alpha_2 \leqslant 1/2$ *i* $\beta_i = 1 - \alpha_i$ $(i = 1, 2)$ *prawdziwe jest oszacowanie* $\sigma_{p,q}(\alpha_1) \overset{c}{\simeq} \sigma_{p,q}(\alpha_2)$, *gdzie* $c = \frac{\alpha_2}{\alpha_1}\frac{\beta_1}{\beta_2}$.

Dowód. Po pierwsze, zauważmy, że dla dowolnej funkcji wypukłej $\varphi\colon \mathbb{R} \to \mathbb{R}$ spełnione są nierówności

$$\beta_1 \varphi(-\alpha_1 y) + \alpha_1 \varphi(\beta_1 y) \leqslant \beta_2 \varphi(-\alpha_2 \frac{\beta_1}{\beta_2} y) + \alpha_2 \varphi(\beta_2 \frac{\beta_1}{\beta_2} y), \tag{2.15}$$

$$\beta_2 \varphi(-\alpha_2 y) + \alpha_2 \varphi(\beta_2 y) \leqslant \beta_1 \varphi(-\alpha_1 \frac{\alpha_2}{\alpha_1} y) + \alpha_1 \varphi(\beta_1 \frac{\alpha_2}{\alpha_1} y), \tag{2.16}$$

dla wszystkich $y \in \mathbb{R}$.

Ustalmy dowolne $u \in \mathbb{R}$ oraz $r \geqslant 1$ i rozważmy funkcję wypukłą $\varphi_r(x) = |1+xu|^r$. Kładąc $\sigma = \sigma_{p,q}(\alpha_2)$ i korzystając kolejno z (2.15), (2.6) i (2.16), otrzymujemy

$$(\beta_1\varphi_p(-\alpha_1\sigma) + \alpha_1\varphi_p(\beta_1\sigma))^{1/p} \leqslant \left(\beta_2\varphi_p(-\alpha_2\frac{\beta_1}{\beta_2}\sigma) + \alpha_2\varphi_p(\beta_2\frac{\beta_1}{\beta_2}\sigma)\right)^{1/p}$$

$$\leqslant \left(\beta_2\varphi_q(-\alpha_2\frac{\beta_1}{\beta_2}) + \alpha_2\varphi_q(\beta_2\frac{\beta_1}{\beta_2})\right)^{1/q} \leqslant (\beta_1\varphi_q(-c\alpha_1) + \alpha_1\varphi_q(c\beta_1))^{1/q},$$

co wobec dowolności wyboru $u \in \mathbb{R}$ dowodzi, że $\sigma_{p,q}(\alpha_2) \leqslant c\sigma_{p,q}(\alpha_1)$. Podobnie, kładąc $\sigma = \sigma_{p,q}(\alpha_1)$ i korzystając kolejno z (2.16), (2.6) i (2.15), dostaniemy

$$(\beta_2\varphi_p(-\alpha_2\sigma) + \alpha_2\varphi_p(\beta_2\sigma))^{1/p} \leqslant \left(\beta_1\varphi_p(-\alpha_1\frac{\alpha_2}{\alpha_1}\sigma) + \alpha_1\varphi_p(\beta_1\frac{\alpha_2}{\alpha_1}\sigma)\right)^{1/p}$$

$$\leqslant \left(\beta_1\varphi_q(-\alpha_1\frac{\alpha_2}{\alpha_1}) + \alpha_1\varphi_q(\beta_1\frac{\alpha_2}{\alpha_1})\right)^{1/q} \leqslant (\beta_2\varphi_q(-c\alpha_2) + \alpha_2\varphi_q(c\beta_2))^{1/q},$$

co dowodzi, że $\sigma_{p,q}(\alpha_1) \leqslant c\sigma_{p,q}(\alpha_2)$. □

Lemat 2.12. *Dla każdego $C > 1$ istnieje $D = D(C) > 1$ takie, że dla wszystkich $1 < q < p \leqslant 2$ i $\alpha_1, \alpha_2 \in (0, 1/2]$, $\alpha_1 \overset{C}{\simeq} \alpha_2$ implikuje $\tilde{\sigma}_{p,q}(\alpha_1) \overset{D}{\simeq} \tilde{\sigma}_{p,q}(\alpha_2)$.*

Dowód. W przypadku, gdy $\alpha_1, \alpha_2 \in I_j$ $(j = 1, \ldots, 4)$, teza jest dość oczywista — wystarczy skorzystać z tego, że $\ln(1/\alpha_1) \overset{D'(C)}{\simeq} \ln(1/\alpha_2)$, a przypadku $\alpha_1, \alpha_2 \in I_3$ dodatkowo zauważamy, że $1+\ln\left(\frac{q-1}{p-1}\frac{1}{\alpha_1}\ln\frac{1}{\alpha_1}\right) \overset{D''(C)}{\simeq} 1+\ln\left(\frac{q-1}{p-1}\frac{1}{\alpha_2}\ln\frac{1}{\alpha_2}\right)$. Aby otrzymać tezę dla dowolnej pary α_1, α_2 spełniającej założenia lematu, wystarczy pokazać, że $\tilde{\sigma}_{p,q}(\alpha) \simeq \lim_{\alpha'\to\alpha^+}\tilde{\sigma}_{p,q}(\alpha')$ dla $\alpha = \sup I_j$ $(j = 1, 2, 3)$. Dla $j = 1$,

$$\lim_{\alpha'\to\alpha^+}\tilde{\sigma}_{p,q}(\alpha') = \alpha^{1-\frac{1}{p}} \simeq \alpha^{\frac{1}{q}-\frac{1}{p}} = \tilde{\sigma}_{p,q}(\alpha),$$

gdyż $\alpha^{(1/q)-1} = e^{1/q} \simeq 1$. Dla $j = 2$,

$$\lim_{\alpha'\to\alpha^+}\tilde{\sigma}_{p,q}(\alpha') = \frac{q-1}{p-1}\frac{\ln(1/\alpha)}{1+\ln(1/(p-1))+1/(p-1)} \simeq (q-1)\ln(1/\alpha), \tag{2.17}$$

$$\tilde{\sigma}_{p,q}(\alpha) = (q-1)\ln(1/\alpha)\left((q-1)\frac{1}{\alpha}\ln\frac{1}{\alpha}\right)^{\frac{1}{p}-1} = (q-1)\ln(1/\alpha)e^{-1/p},$$

i wreszcie, dla $j = 3$, $\tilde{\sigma}_{p,q}(\alpha) = \alpha = \lim_{\alpha'\to\alpha^+}\tilde{\sigma}_{p,q}(\alpha')$. □

W kolejnych dwóch lematach i w dowodzie Twierdzenia 2.3 będziemy pracowali z następującymi funkcjami. Niech $1 < q < p \leqslant 2$ i $\alpha \in (0, 1/2]$ oraz $\beta = 1 - \alpha$. Dla $x \in [-1/\beta, 1/\alpha]$ definiujemy

$$L_{\alpha,p}(x) = \beta(1-\alpha x)^p + \alpha(1+\beta x)^p - 1,$$

$$R_{\alpha,q,p}(x) = (\beta(1-\alpha x)^q + \alpha(1+\beta x)^q)^{p/q} - 1.$$

Jeśli $\alpha \leqslant e^{-2}$, określimy funkcję

$$\tilde{L}_{\alpha,p}(x) = \begin{cases} (p-1)\alpha x^2 & \text{dla } x \in [0,e], \\ e(p-1)\alpha x \ln x & \text{dla } x \in (e, e^{1/(p-1)} \wedge (1/\alpha)], \\ \alpha x^p & \text{dla } x \in (e^{1/(p-1)} \wedge (1/\alpha), 1/\alpha], \end{cases}$$

a gdy dodatkowo $1/\alpha \notin (e^{1/(q-1)}, e^{q/(q-1)}]$, zdefiniujemy także funkcję

$$\tilde{R}_{\alpha,q,p}(x) = \begin{cases} (q-1)\alpha x^2 & \text{dla } x \in [0,e], \\ e(q-1)\alpha x \ln x & \text{dla } x \in (e, 1/\alpha], \end{cases} \qquad \text{gdy } 1/\alpha \leqslant e^{1/(q-1)},$$

i

$$\tilde{R}_{\alpha,q,p}(x) = \begin{cases} (q-1)\alpha x^2 & \text{dla } x \in [0,e], \\ e(q-1)\alpha x \ln x & \text{dla } x \in (e, e^{1/(q-1)}], \\ \alpha x^q & \text{dla } x \in (e^{1/(q-1)}, (1/\alpha)^{1/q}], \\ \alpha^{p/q} x^p & \text{dla } x \in ((1/\alpha)^{1/q}, 1/\alpha], \end{cases} \qquad \text{gdy } 1/\alpha > e^{q/(q-1)}.$$

Zauważmy, że funkcje $\tilde{L}_{\alpha,p}$ i $\tilde{R}_{\alpha,q,p}$ są ciągłe, różniczkowalne (z wyjątkiem przypadku $1/\alpha > e^{q/(q-1)}$, kiedy $\tilde{R}_{\alpha,q,p}$ nie jest różniczkowalna w punkcie $(1/\alpha)^{1/q}$) oraz rosnące.

Ponadto będziemy wykorzystywać następujące przybliżenia: dla dowolnego $\varepsilon > 0$ istnieją stałe $c_1, c_2, c_3 > 0$ (każda zależna jedynie od ε) takie, że

$$e^t - 1 \overset{c_1}{\simeq} t \quad \text{dla } 0 \leqslant t \leqslant \varepsilon, \tag{2.18}$$

$$e^t - 1 \overset{c_2}{\simeq} e^t \quad \text{dla } t \geqslant \varepsilon, \tag{2.19}$$

$$\ln(1+t) \overset{c_3}{\simeq} t \quad \text{dla } 0 \leqslant t \leqslant \varepsilon. \tag{2.20}$$

Lemat 2.13. *Dla $\alpha \leqslant e^{-2}$ i $x \in [-1/\beta, 1/\alpha]$, $\tilde{L}_{\alpha,p}(|x|) \simeq L_{\alpha,p}(x)$.*

Dowód. Po pierwsze, łatwo widać, że jeśli $C \geqslant 1$, to dla wszystkich $x_1, x_2 \in [0, 1/\alpha]$ spełniających $x_1 \overset{C}{\simeq} x_2$ zachodzi także relacja $\tilde{L}_{\alpha,p}(x_1) \overset{D}{\simeq} \tilde{L}_{\alpha,p}(x_2)$, gdzie $D = D(C) \geqslant 1$ jest pewną stałą, zależną jedynie od C.

Po drugie, $L_{\alpha,p}$ jest funkcją wypukłą, klasy C^∞ (na przedziale $(-1/\beta, 1/\alpha)$), zatem $L_{\alpha,p}$ jest malejąca na $[-1/\beta, 0]$ oraz rosnąca na $[0, 1/\alpha]$, gdyż $L'_{\alpha,p}(0) = 0$.

Te dwie powyższe obserwacje pokazują, że wystarczy udowodnić tezę jedynie dla $x \in \{-1/\beta\} \cup [-1/(2\beta), 1/(2\beta)] \cup [e^2, 1/\alpha]$.

Przypadek $x \in [-1/(2\beta), 1/(2\beta)]$. Wystarczy pokazać, że

$$L''_{\alpha,p}(x) \simeq (p-1)\alpha. \tag{2.21}$$

Istotnie, całkując stronami otrzymamy $L'_{\alpha,p}(x) \simeq (p-1)\alpha x$, gdyż $L'_{\alpha,p}(0) = 0$, a całkując raz jeszcze dostaniemy $L_{\alpha,p}(x) \simeq (p-1)\alpha x^2$, ponieważ $L_{\alpha,p}(0) = 0$. Teraz,

$$L''_{\alpha,p}(x) = p(p-1)\alpha\beta \left(\alpha(1-\alpha x)^{p-2} + \beta(1+\beta x)^{p-2}\right).$$

Oczywiście $1-\alpha x \simeq 1$ i $1+\beta x \simeq 1$. Ponieważ $p-2 \in (-1, 0]$, to również $(1-\alpha x)^{p-2} \simeq 1$ i $(1+\beta x)^{p-2} \simeq 1$, co pociąga (2.21).

Przypadek $x = -1/\beta$. Korzystając ze wzorów (2.18) i (2.20) (np. z $\varepsilon = 1$) otrzymujemy, że

$$\begin{aligned} L_{\alpha,p}(-1/\beta) &= \beta(1+\alpha/\beta)^p - 1 = (1+\alpha/\beta)^{p-1} - 1 = e^{(p-1)\ln(1+\alpha/\beta)} - 1 \\ &\simeq (p-1)\ln(1+\alpha/\beta) \simeq (p-1)\frac{\alpha}{\beta} \simeq \tilde{L}_{\alpha,p}(1/\beta), \end{aligned}$$

gdyż $(p-1)\ln(1+\alpha/\beta) \leqslant (p-1)\frac{\alpha}{\beta} < 1$ oraz $\alpha/\beta < 1$.

Przypadek $x \in [e^2, 1/\alpha]$. Będziemy przybliżać wyrażenie

$$\begin{aligned} L_{\alpha,p}(x) &= \beta(1-\alpha x)\left((1-\alpha x)^{p-1} - 1\right) + \alpha(1+\beta x)\left((1+\beta x)^{p-1} - 1\right) \\ &= M_1 + M_2. \end{aligned}$$

Pierwszy składnik jest równy zero, gdy $x = 1/\alpha$. Poza tym jest ujemny i możemy wówczas zastosować następujące oszacowanie

$$\begin{aligned} M_1 &= \beta(1-\alpha x)\left(e^{(p-1)\ln(1-\alpha x)} - 1\right) \geqslant \beta(1-\alpha x)(p-1)\ln(1-\alpha x) \\ &\geqslant \beta(p-1)(-\alpha x) > -(p-1)\alpha x, \end{aligned}$$

w którym druga z nierówności wynika stąd, że $y\ln y \geqslant y - 1$ dla $y > 0$. Drugi składnik, M_2, możemy oszacować następująco:

$$\begin{aligned} M_2 &= \alpha(1+\beta x)\left(e^{(p-1)\ln(1+\beta x)} - 1\right) \geqslant \alpha(1+\beta x)(p-1)\ln(1+\beta x) \\ &\geqslant \alpha x(p-1)\ln x \geqslant 2(p-1)\alpha x, \end{aligned}$$

gdyż $1 + \beta x \geqslant x$ i $\ln x \geqslant 2$. Z powyższych oszacowań wynika, że $M_2 \geqslant M_1 + M_2 = (M_1 + M_2/2) + M_2/2 \geqslant M_2/2$, zatem $L_{\alpha,p}(x) \simeq M_2$ i dalej będziemy zajmować się tylko przybliżaniem wyrażenia M_2. Ponieważ $1 + \beta x \geqslant x \geqslant e^2$, to $1 + \beta x \simeq x$, $\ln(1+\beta x) \simeq \ln x$ i $(1+\beta x)^{p-1} \simeq x^{p-1}$. Zatem, korzystając ze wzorów (2.18) i (2.19), otrzymujemy

$$M_2 \simeq \alpha x\left(e^{(p-1)\ln(1+\beta x)} - 1\right) \simeq \alpha x \begin{cases} (p-1)\ln x & \text{dla } x \leqslant e^{\frac{1}{p-1}}, \\ x^{p-1} & \text{dla } x > e^{\frac{1}{p-1}}. \end{cases}$$

□

Lemat 2.14. *Dla* $x \in [-1/\beta, 1/\alpha]$, $\tilde{R}_{\alpha,q,p}(|x|) \simeq R_{\alpha,q,p}(x)$, *o ile* $\tilde{R}_{\alpha,q,p}$ *jest określone.*

Dowód. Ponieważ $p/q \in (1,2)$, to $(1+t)^{p/q} - 1 \simeq (1+t)^{p/q} \simeq t^{p/q}$ dla $t \geqslant 1$, oraz $(1+t)^{p/q} - 1 \simeq t$ dla $t \in [0,1]$. Stąd

$$R_{\alpha,q,p}(x) = (1 + L_{\alpha,q}(x))^{p/q} - 1 \simeq L_{\alpha,q}(x) \vee L_{\alpha,q}(x)^{p/q},$$

podczas gdy bezpośrednio z określenia funkcji $\tilde{R}_{\alpha,q,p}$ i $\tilde{L}_{\alpha,q}$ wynika, że $\tilde{R}_{\alpha,q,p}(x) \simeq \tilde{L}_{\alpha,q}(x) \vee \tilde{L}_{\alpha,q}(x)^{p/q}$. Pozostaje więc zastosować Lemat 2.13. □

Dowód Twierdzenia 2.3. Optymalność stałej $\tilde\sigma_{p,q}(\alpha)$, tzn. nierówność $\sigma_{p,q}(\alpha) \lesssim \tilde\sigma_{p,q}(\alpha)$, łatwo wynika z Twierdzenia 2.6. Kładąc $p' = p/(p-1)$, $q' = q/(q-1)$ i biorąc S takie jak w Twierdzeniu 2.6, dostajemy $\|S\|_{q'} \,/\, \|S\|_{p'} \simeq \tilde\sigma^{-1}_{q',p'}(\alpha)$. Tłumacząc (q', p') na (p, q) we wzorze na $\tilde\sigma_{q',p'}(\alpha)$ w Twierdzeniu 2.4, z pomocą Lematu 2.12 pokażemy, że $\tilde\sigma_{q',p'}(\alpha) \simeq \tilde\sigma_{p,q}(\alpha)$. Z drugiej strony, Stwierdzenia 2.1 i 1.2 implikują, że $\|S\|_{q'} \,/\, \|S\|_{p'} \leqslant \sigma^{-1}_{q',p'}(\alpha) = \sigma^{-1}_{p,q}(\alpha)$.

W dalszym ciągu dowodzić będziemy nierówności przeciwnej: $\tilde\sigma_{p,q}(\alpha) \lesssim \sigma_{p,q}(\alpha)$. Dzięki Lematom 2.11 i 2.12 możemy założyć, że

$$\alpha \notin (e^{-2}, 1/2] \quad \text{oraz} \quad 1/\alpha \notin (e^{1/(q-1)}, e^{q/(q-1)}],$$

gdyż $e^{-2} \simeq 1/2$ oraz $e^{1/(q-1)} \simeq e^{q/(q-1)}$. Z uwagi na (2.6) wystarczy pokazać, że

$$L_{\alpha,p}(\sigma x) \leqslant R_{\alpha,q,p}(x) \quad \text{dla wszystkich } x \in [-1/\beta, 1/\alpha], \tag{2.22}$$

gdzie $\sigma \simeq \tilde\sigma := \tilde\sigma_{p,q}(\alpha)$. Do tego z kolei wystarczy, że pokażemy, iż istnieje stała numeryczna $C_1 > 0$, dla której

$$\tilde L_{\alpha,p}(\tilde\sigma x) \leqslant C_1 \tilde R_{\alpha,q,p}(x) \quad \text{dla wszystkich } x \in [0, 1/\alpha]. \tag{2.23}$$

Istotnie, stosując (2.23) wraz z Lematami 2.13 i 2.14 otrzymamy, że

$$L_{\alpha,p}(\tilde\sigma x) \leqslant C_2 R_{\alpha,q,p}(x) \quad \text{dla wszystkich } x \in [-1/\beta, 1/\alpha], \tag{2.24}$$

gdzie C_2 jest pewną stałą numeryczną. Oczywiście, możemy założyć, że $C_2 \geqslant 1$. Ponieważ lewa strona (2.24) jest wypukłą funkcją zmiennej x, przyjmującą wartość zero dla $x = 0$, to $L_{\alpha,p}(C_2^{-1}\tilde\sigma x) \leqslant C_2^{-1} L_{\alpha,p}(\tilde\sigma x)$, co wraz z (2.24) dowodzi (2.22) ze stałą $\sigma = C_2^{-1}\tilde\sigma$.

Teraz sprawdzimy, że funkcja ciągła $(0, 1/\alpha) \ni x \mapsto \frac{\tilde L_{\alpha,p}(\tilde\sigma x)}{\tilde R_{\alpha,q,p}(x)}$ jest niemalejąca. Ponieważ funkcja ta jest także różniczkowalna (z wyjątkiem punktu $(1/\alpha)^{1/q}$, o ile $1/\alpha > e^{q/(q-1)}$), to wystarczy, że pokażemy, iż jej pochodna jest nieujemna lub, równoważnie,

$$\hat L_{\alpha,p}(\tilde\sigma x) \geqslant \hat R_{\alpha,q,p}(x) \quad \text{dla wszystkich } x \in (0, 1/\alpha), \tag{2.25}$$

(z wyłączeniem $x = (1/\alpha)^{1/q}$, gdy $1/\alpha > e^{q/(q-1)}$), gdzie

$$\hat L_{\alpha,p}(x) = x\frac{\tilde L'_{\alpha,p}(x)}{\tilde L_{\alpha,p}(x)} \qquad \text{i} \qquad \hat R_{\alpha,q,p}(x) = x\frac{\tilde R'_{\alpha,q,p}(x)}{\tilde R_{\alpha,q,p}(x)}.$$

Prosty rachunek pokazuje, że

$$\hat L_{\alpha,p}(x) = \begin{cases} 2 & \text{dla } x \in (0, e], \\ 1 + \frac{1}{\ln x} & \text{dla } x \in (e, e^{1/(p-1)} \wedge 1/\alpha), \\ p & \text{dla } x \in [e^{1/(p-1)} \wedge 1/\alpha, 1/\alpha), \end{cases}$$

$$\hat R_{\alpha,q,p}(x) = \hat L_{\alpha,q}(x), \qquad \text{gdy } 1/\alpha \leqslant e^{\frac{1}{q-1}},$$

$$\hat R_{\alpha,q,p}(x) = \begin{cases} 2 & \text{dla } x \in (0, e], \\ 1 + \frac{1}{\ln x} & \text{dla } x \in (e, e^{1/(q-1)}], \\ q & \text{dla } x \in (e^{1/(q-1)}, (1/\alpha)^{1/q}), \\ p & \text{dla } x \in ((1/\alpha)^{1/q}, 1/\alpha), \end{cases} \qquad \text{gdy } 1/\alpha > e^{\frac{q}{q-1}}.$$

Funkcja $\hat{L}_{\alpha,p}$ jest nierosnąca, więc $\hat{L}_{\alpha,p}(\tilde{\sigma}x) \geqslant \hat{L}_{\alpha,p}(x)$, gdyż $\tilde{\sigma} \leqslant 1$ (ostatnią nierówność można łatwo sprawdzić, chociażby posługując się fragmentami dowodu Lematu 2.5, poniżej). Ponadto, $\hat{L}_{\alpha,p}(x) = \hat{R}_{\alpha,q,p}(x)$ dla $x \in (0, e^{1/(p-1)} \wedge 1/\alpha)$ oraz $\hat{L}_{\alpha,p}(x) = p \geqslant \hat{R}_{\alpha,q,p}(x)$ dla $x \in [e^{1/(p-1)} \wedge 1/\alpha, 1/\alpha)$ (znów wyłączając $x = (1/\alpha)^{1/q}$, gdy $1/\alpha > e^{q/(q-1)}$), co dowodzi (2.25).

Dowód będzie zakończony, gdy sprawdzimy (2.23) dla $x = 1/\alpha$. Rozważmy cztery przypadki (występujące we wzorze na $\tilde{\sigma}_{p,q}(\alpha)$ ze sformułowania Twierdzenia 2.3).

Przypadek $\alpha \in I_1$. Oczywiście $\tilde{\sigma}/\alpha = (1/\alpha)^{1-1/q+1/p} > (1/\alpha)^{1/p}$. Ponieważ założyliśmy, że $1/\alpha \notin (e^{1/(q-1)}, e^{q/(q-1)}]$, to $1/\alpha > e^{q/(q-1)} > e^{p/(p-1)}$, a przez to $\tilde{\sigma}/\alpha > e^{1/(p-1)}$, więc $\tilde{L}_{\alpha,p}(\tilde{\sigma}/\alpha) = \alpha(\tilde{\sigma}/\alpha)^p = (1/\alpha)^{p-p/q} = \tilde{R}_{\alpha,q,p}(1/\alpha)$.

Przypadek $\alpha \in I_2$. Oczywiście $\tilde{R}_{\alpha,q,p}(1/\alpha) = e(q-1)\ln(1/\alpha)$. Jeśli $\tilde{\sigma}/\alpha > e^{1/(p-1)}$, to $\tilde{L}_{\alpha,p}(\tilde{\sigma}/\alpha) = \alpha(\tilde{\sigma}/\alpha)^p = (q-1)\ln(1/\alpha)$, a jeśli $\tilde{\sigma}/\alpha \leqslant e^{1/(p-1)}$, to z monotoniczności $\tilde{L}_{\alpha,p}(x)$ i $x \mapsto x\ln x$ (na $[e^{-1}, \infty)$) będziemy mieli

$$\begin{aligned}\tilde{L}_{\alpha,p}(\tilde{\sigma}/\alpha) &\leqslant \tilde{L}_{\alpha,p}\left(e^{1/(p-1)} \wedge (1/\alpha)\right)\\ &\leqslant e(p-1)\alpha e^{1/(p-1)} \ln\left(e^{1/(p-1)}\right) = e\alpha e^{1/(p-1)} \leqslant e(q-1)\ln(1/\alpha),\end{aligned}$$

gdzie ostatnia nierówność wynika z faktu, że $\alpha \in I_2$.

Przypadek $\alpha \in I_3$. $\tilde{R}_{\alpha,q,p}(1/\alpha)$ jest tym samym, co w poprzednim przypadku. Pokażemy, że $\tilde{\sigma}/\alpha \leqslant e^{1/(p-1)}$. Kładąc

$$u_0 := \ln\left(\frac{q-1}{p-1}\frac{1}{\alpha}\ln\frac{1}{\alpha}\right),$$

dostaniemy $\tilde{\sigma}/\alpha = \frac{e^{u_0}}{1+u_0}$. Ponieważ $\alpha \in I_3$, to

$$0 \leqslant u_0 \leqslant \frac{1}{p-1} + \ln\frac{1}{p-1}.$$

Funkcja $h(u) = \frac{e^u}{1+u}$ jest rosnąca dla $u \geqslant 0$ ($h'(u) = ue^u/(1+u)^2$), więc

$$\tilde{\sigma}/\alpha \leqslant h\left(\frac{1}{p-1} + \ln\frac{1}{p-1}\right) < e^{\frac{1}{p-1}}.$$

Teraz oszacujemy iloraz $\frac{\tilde{L}_{\alpha,p}(\tilde{\sigma}/\alpha)}{\tilde{R}_{\alpha,q,p}(1/\alpha)}$. Jeśli $\tilde{\sigma}/\alpha > e$, to $\tilde{L}_{\alpha,p}(\tilde{\sigma}/\alpha) = e(p-1)\tilde{\sigma}\ln(\tilde{\sigma}/\alpha)$, w przeciwnym razie, z monotoniczności $\tilde{L}_{\alpha,p}(x)$ wnioskujemy, że $\tilde{L}_{\alpha,p}(\tilde{\sigma}/\alpha) \leqslant (p-1)\alpha e^2$. W pierwszym wariancie szacowany iloraz jest równy

$$\frac{e(p-1)\tilde{\sigma}\ln(\tilde{\sigma}/\alpha)}{e(q-1)\ln(1/\alpha)} = \frac{u_0 - \ln(1+u_0)}{1+u_0} < 1.$$

W drugim wariancie ten sam iloraz jest nie większy niż

$$\frac{(p-1)\alpha e^2}{e(q-1)\ln(1/\alpha)} \leqslant e, \quad \text{gdyż} \quad \frac{p-1}{q-1} \leqslant \frac{1}{\alpha}\ln\frac{1}{\alpha} \quad (\text{bo } \alpha \in I_3).$$

Przypadek $\alpha \in I_4$. $\tilde{R}_{\alpha,q,p}(1/\alpha)$ znów jest tym samym, co w poprzednim przypadku. Oczywiście $\tilde{\sigma}/\alpha \in (0,1)$, więc

$$\frac{\tilde{L}_{\alpha,p}(\tilde{\sigma}/\alpha)}{\tilde{R}_{\alpha,q,p}(1/\alpha)} = \frac{(p-1)\tilde{\sigma}^2/\alpha}{e(q-1)\ln(1/\alpha)} = e^{-1}.$$

Powyższe rozważania dowodzą (2.23) ze stałą $C_1 = e$. □

Dowód Twierdzenia 2.4. Wynika natychmiast z Twierdzenia 2.3, z (2.4) oraz z faktu, że $\tilde{\sigma}_{p,q}(\alpha) \simeq \tilde{\sigma}_{q',p'}(\alpha)$ (gdzie $p' = p/(p-1)$, $q' = q/(q-1)$), który został skomentowany na początku dowodu Twierdzenia 2.3. □

Dowód Lematu 2.5. Na mocy (2.4) możemy rozpatrywać jedynie przypadek $1 < q < p \leqslant 2$. Twierdzenie 2.3 implikuje, że wystarczy sprawdzić, iż

$$\tilde{\sigma}_{p,q}(\alpha_1) \lesssim \tilde{\sigma}_{p,q}(\alpha_2).$$

Z kolei na mocy Lematu 2.12 wystarczy powyższą nierówność pokazać jedynie dla $\alpha_1, \alpha_2 \in I_j$, $j = 1, \ldots, 4$.

Przypadek $\alpha_1, \alpha_2 \in I_1$. Funkcja $\alpha \mapsto \alpha^{1/q-1/p}$ jest rosnąca, gdyż $1/q - 1/p > 0$.

Przypadek $\alpha_1, \alpha_2 \in I_2$. Sprawdzamy, że funkcja $h(\alpha) = (\ln(1/\alpha))^{1/p}\alpha^{1-1/p}$ jest rosnąca dla $\alpha \leqslant e^{-1/(p-1)}$, czyli w szczególności dla $\alpha \in I_2$. Istotnie,

$$\begin{aligned} h'(\alpha) &= -\frac{1}{p}\left(\alpha\ln(1/\alpha)\right)^{-1} h(\alpha) + \left(1 - \frac{1}{p}\right)\alpha^{-1}h(\alpha) \\ &= \alpha^{-1}h(\alpha)\left(1 - \frac{1}{p}\left(1 + (\ln(1/\alpha))^{-1}\right)\right), \end{aligned}$$

co jest nieujemne, gdyż $1 + (\ln(1/\alpha))^{-1} \leqslant p$.

Przypadek $\alpha_1, \alpha_2 \in I_3$. Sprawdzamy, że funkcja

$$h(\alpha) = \frac{\ln(1/\alpha)}{1 + \ln\left(\frac{q-1}{p-1}\frac{1}{\alpha}\ln\frac{1}{\alpha}\right)} = \frac{h_1(\alpha)}{1 + \ln h_2(\alpha)}$$

jest rosnąca, gdy $\ln(1/\alpha) \leqslant \frac{p-1}{q-1}$, a w przeciwnym wypadku — malejąca. Istotnie, znak pochodnej funkcji h jest taki, jak znak

$$\begin{aligned} h_1'(\alpha)(1 + \ln h_2(\alpha)) - h_1(\alpha)\frac{h_2'(\alpha)}{h_2(\alpha)} &= -\alpha^{-1}(1 + \ln h_2(\alpha)) + h_1(\alpha)\frac{1 + \ln(1/\alpha)}{\alpha\ln(1/\alpha)} \\ &= -\alpha^{-1}(\ln h_2(\alpha) - \ln(1/\alpha)) = -\alpha^{-1}\ln\left(\frac{q-1}{p-1}\ln\frac{1}{\alpha}\right). \end{aligned}$$

Zatem jeśli $\{\alpha \mid \ln(1/\alpha) > \frac{p-1}{q-1}\} \cap I_3 = \emptyset$, to rozpatrywanie tego przypadku będzie zakończone. Załóżmy przeciwnie. Wówczas na niepustym przedziale $(\underline{\alpha}, \overline{\alpha}]$ (gdzie $\underline{\alpha} =$

$\inf I_3$ i $\ln(1/\overline{\alpha}) = (p-1)/(q-1))$ funkcja $h(\alpha)$ jest malejąca. Ponieważ $e \leqslant 1/\overline{\alpha} < 1/\underline{\alpha}$, to $\frac{1}{\overline{\alpha}} \ln \frac{1}{\overline{\alpha}} < \frac{1}{\underline{\alpha}} \ln \frac{1}{\underline{\alpha}}$, więc

$$\frac{p-1}{q-1} e^{(p-1)/(q-1)} < \frac{1}{q-1} e^{1/(p-1)},$$

czyli

$$\frac{1}{q-1} < \frac{1}{(p-1)^2} + \frac{1}{p-1} \ln \frac{1}{p-1} < \frac{2}{(p-1)^2},$$

co daje

$$\ln \frac{1}{q-1} < \ln 2 + 2 \ln \frac{1}{p-1} < \frac{3}{p-1}. \tag{2.26}$$

Pozostaje pokazać, że dla pewnej stałej numerycznej $C \geqslant 1$ zachodzi $\lim_{\alpha \to \underline{\alpha}^+} \tilde{\sigma}_{p,q}(\alpha) \leqslant C\tilde{\sigma}_{p,q}(\overline{\alpha})$. Oczywiście $\frac{1}{\overline{\alpha}} \ln \frac{1}{\overline{\alpha}} \geqslant \frac{p-1}{q-1}$, więc $\overline{\alpha} \notin I_4$. Skoro $\overline{\alpha} > \underline{\alpha}$, to $\overline{\alpha} \in I_3$ i mamy

$$\tilde{\sigma}_{p,q}(\overline{\alpha}) = \frac{1}{1 + \frac{p-1}{q-1}} \simeq \frac{q-1}{p-1},$$

podczas gdy z (2.17) wynika, że

$$\lim_{\alpha \to \underline{\alpha}^+} \tilde{\sigma}_{p,q}(\alpha) \simeq (q-1) \ln \frac{1}{\underline{\alpha}} \simeq (q-1) \left(\frac{1}{p-1} + \ln \frac{1}{q-1} \right),$$

jako że

$$\ln \frac{1}{\underline{\alpha}} \simeq \ln \left(\frac{1}{\underline{\alpha}} \ln \frac{1}{\underline{\alpha}} \right) \simeq \ln \frac{e^{\frac{1}{p-1}}}{q-1}.$$

Zatem z (2.26) otrzymujemy

$$\lim_{\alpha \to \underline{\alpha}^+} \tilde{\sigma}_{p,q}(\alpha) \simeq \frac{q-1}{p-1}.$$

Przypadek $\alpha_1, \alpha_2 \in I_4$. Funkcja $\alpha \mapsto \alpha \ln(1/\alpha)$ jest rosnąca dla $\alpha \leqslant e^{-1}$, gdyż jej pochodna jest równa $\ln(1/\alpha) - 1$. Natomiast z Lematu 2.12 wiemy, że $\tilde{\sigma}_{p,q}(\alpha) \simeq \tilde{\sigma}_{p,q}(e^{-1})$, gdy $\alpha \in (e^{-1}, 1/2]$.

□

Dowód Twierdzenia 2.6. Niech $(\theta'_k)_{k=1}^{\infty}$ będzie niezależną kopią ciągu $(\theta_k)_{k=1}^{\infty}$. Połóżmy $\eta_k = \theta_k - \theta'_k$, $S' = \sum_{k=1}^{n(\alpha,p)} \theta'_k$, i $S'' = S - S'$, a także $\theta = \theta_1$, $\eta = \eta_1$. Zmienne losowe η_k mają rozkład symetryczny $\alpha\beta\delta_{-1} + (1-2\alpha\beta)\delta_0 + \alpha\beta\delta_1$. Ponadto, dla każdego $r \geqslant 1$ mamy $\|\theta\|_r \leqslant \|\eta\|_r \leqslant 2\|\theta\|_r$ (pierwsza nierówność wynika z nierówności Jensena i z faktu, że $E\theta = 0$, podczas gdy druga — z nierówności trójkąta dla normy w przestrzeni L^r). Oczywiście, takie same nierówności mają miejsce w przypadku zmiennych S i S''. W związku z tym, dalej będziemy pokazywać, że $\|S''\|_q / \|S''\|_p \simeq \tilde{\sigma}_{p,q}(\alpha)$.

Po pierwsze zauważmy, że jeśli $p \leqslant \ln(1/\alpha)$ (równoważnie, $n(\alpha, p) = 1$), to $\alpha \in I'_1$, podczas gdy $\|S''\|_q = \|\eta\|_q = (2\alpha\beta)^{1/q} \simeq \alpha^{1/q}$ i $\|S''\|_p \simeq \alpha^{1/p}$.

Dalej będziemy zakładać, że $p > \ln(1/\alpha)$ (równoważnie, $n := n(\alpha,p) \geqslant 2$). Następujący prosty argument pokazuje, że $\|S''\|_p \simeq n \simeq p/\ln(1/\alpha)$. Ponieważ $n/2 < p/\ln(1/\alpha)$, to $(\alpha\beta)^n \geqslant \alpha^{2n} > e^{-4p}$, więc $\mathbb{P}(S'' = n) > e^{-4p}$. Stąd

$$ne^{-4} < n\mathbb{P}(S'' = n)^{1/p} \leqslant \|S''\|_p \leqslant \|S''\|_\infty = n.$$

Teraz zajmiemy się przybliżeniem $\|S''\|_q$. Odpowiednią formułę przybliżającą podał R. Latała [nieopublikowane]. Dla kompletności dowodu wzór ten wyprowadzimy z ogólniejszego wyniku [25, Corollary 2], który stwierdza, że dla $q \geqslant 2$ i ciągu symetrycznych zmiennych losowych o jednakowym rozkładzie $\xi, \xi_1, \ldots, \xi_n$,

$$\|\xi_1 + \ldots + \xi_n\|_q \simeq \sup\left\{ \frac{q}{s}\left(\frac{n}{q}\right)^{1/s} \|\xi\|_s \;\middle|\; 2 \vee \frac{q}{n} \leqslant s \leqslant q \right\}.$$

W naszym przypadku powyższy wynik implikuje, że

$$\frac{\|S''\|_q}{\|S''\|_p} \simeq \bar{f} := \sup\left\{ f(s) = \frac{1}{s}\left(\frac{q}{p}\ln\frac{1}{\alpha}\right)^{1-1/s} \alpha^{1/s} \;\middle|\; 2 \vee \frac{q}{\lceil p/\ln(1/\alpha)\rceil} \leqslant s \leqslant q \right\},$$

gdyż $n^{(1/s)-1} \simeq (p/\ln(1/\alpha))^{(1/s)-1}$. Funkcja f jest różniczkowalna dla $s \in (0,\infty)$ oraz

$$f'(s) = \left(\frac{q}{p}\ln\frac{1}{\alpha}\right)\left(\frac{1}{s}\left(\frac{q}{p}\frac{1}{\alpha}\ln\frac{1}{\alpha}\right)^{-1/s}\right)' = \frac{1}{s^2}f(s)\left(\ln\left(\frac{q}{p}\frac{1}{\alpha}\ln\frac{1}{\alpha}\right) - s\right),$$

skąd widać, że f jest rosnąca na przedziale $(0, s_0]$ i malejąca na $[s_0,\infty)$, gdzie $s_0 = \ln\left(\frac{q}{p}\frac{1}{\alpha}\ln\frac{1}{\alpha}\right)$. Rozpatrzmy trzy przypadki:

Przypadek $(p/q)e^q \leqslant (1/\alpha)\ln(1/\alpha) < pe^p$. Wówczas $s_0 \geqslant q$, zatem

$$\bar{f} = f(q) = q^{-1/q}\left(\frac{\ln(1/\alpha)}{p}\right)^{1-1/q}\alpha^{1/q} \simeq \left(\frac{\ln(1/\alpha)}{p}\right)^{1-1/q}\alpha^{1/q}.$$

Jeśli $pe^q \leqslant \frac{1}{\alpha}\ln\frac{1}{\alpha} < pe^p$ (tzn. $\alpha \in I_2'$), to $\bar{f} \simeq \tilde{\sigma}_{p,q}(\alpha)$. Natomiast jeśli $\frac{p}{q}e^q \leqslant \frac{1}{\alpha}\ln\frac{1}{\alpha} < pe^q$ (czyli $\alpha \in I_3'$), to

$$\frac{\tilde{\sigma}_{p,q}(\alpha)}{\left(\frac{\ln(1/\alpha)}{p}\right)^{1-1/q}\alpha^{1/q}} = \frac{q}{1 + \ln\left(\frac{q}{p}\frac{1}{\alpha}\ln\frac{1}{\alpha}\right)}\left(\frac{1}{p}\frac{1}{\alpha}\ln\frac{1}{\alpha}\right)^{1/q} \simeq 1,$$

gdyż zarówno pierwszy, jak i drugi czynnik, z uwagi na nierówności $q \leqslant \ln\left(\frac{q}{p}\frac{1}{\alpha}\ln\frac{1}{\alpha}\right) < q + \ln q$ i $q^{-1}e^q \leqslant \frac{1}{p}\frac{1}{\alpha}\ln\frac{1}{\alpha} < e^q$ (odpowiednio), są $\simeq 1$.

Przypadek $(p/q)e^2 \leqslant (1/\alpha)\ln(1/\alpha) < (p/q)e^q$. Najpierw sprawdźmy, że $2 \vee \frac{q}{\lceil p/\ln(1/\alpha)\rceil} \leqslant s_0 < q$. Oczywiście $2 \leqslant s_0 < q$. Gdyby $s_0 < q/\lceil p/\ln(1/\alpha)\rceil$, mielibyśmy $2 \leqslant s_0 < (q/p)\ln(1/\alpha)$, więc z monotoniczności $\ln$ wynikałoby, że

$$\ln\frac{1}{\alpha} < \ln\frac{2}{\alpha} < \ln\left(\frac{q}{p}\frac{1}{\alpha}\ln\frac{1}{\alpha}\right) = s_0 < \frac{q}{p}\ln\frac{1}{\alpha},$$

co przeczy nierówności $q < p$. Zatem $\bar{f} = f(s_0)$. Teraz już pokażemy, że $f(s_0) \simeq \tilde{\sigma}_{p,q}(\alpha)$. Otóż $s_0 \geqslant 2$, więc $s_0 \simeq 1 + s_0$, a stąd

$$f(s_0) \simeq \frac{q}{p}\frac{\ln(1/\alpha)}{1+s_0}\left(\frac{q}{p}\frac{1}{\alpha}\ln\frac{1}{\alpha}\right)^{-1/s_0} = \tilde{\sigma}_{p,q}(\alpha)\,(e^{s_0})^{-1/s_0} \simeq \tilde{\sigma}_{p,q}(\alpha).$$

Przypadek $(1/\alpha)\ln(1/\alpha) < (p/q)e^2$. Wówczas $s_0 < 2$ i

$$\frac{q}{\lceil p/\ln(1/\alpha)\rceil} \leqslant \frac{q}{p}\ln\frac{1}{\alpha} < \alpha e^2.$$

Jeśli $\alpha e^2 \leqslant 2$, to $q/\lceil p/\ln(1/\alpha)\rceil < 2$. Zauważmy, że gdy $q/\lceil p/\ln(1/\alpha)\rceil \leqslant 2$, to $\bar{f} = f(2) = \frac{1}{2}\sqrt{\frac{q}{p}\alpha\ln(1/\alpha)}$, co jest $\simeq \tilde{\sigma}_{p,q}(\alpha)$, gdy $\alpha \in I_4'$. Gdy zaś $\alpha \in I_3'$, tj. $p/q \leqslant (1/\alpha)\ln(1/\alpha) < (p/q)e^2$, to

$$\bar{f} \simeq \alpha \simeq \frac{q}{p}\frac{\ln(1/\alpha)}{1+\ln\left(\frac{q}{p}\frac{1}{\alpha}\ln\frac{1}{\alpha}\right)} = \tilde{\sigma}_{p,q}(\alpha).$$

Pozostaje rozważyć sytuację $\alpha \in (2e^{-2}, 1/2]$ i $s_1 := q/\lceil p/\ln(1/\alpha)\rceil > 2$. Wówczas $\bar{f} = f(s_1)$ i $2 < s_1 \leqslant (q/p)\ln(1/\alpha) < \alpha e^2 \leqslant \frac{1}{2}e^2$. Stąd $\bar{f} \simeq 1$, a także $q/p \simeq 1$ (jako że $\ln\frac{1}{\alpha} \simeq 1$), co implikuje, że $\tilde{\sigma}_{p,q}(1/2) \simeq 1$. Teraz wystarczy skorzystać z Lematu 2.12, by wykazać, że $\tilde{\sigma}_{p,q}(\alpha) \simeq \tilde{\sigma}_{p,q}(1/2) \simeq 1 \simeq \bar{f}$.

$\square$

Rozdział 3

Hiperkontrakcja zmiennych losowych a geometria przestrzeni unormowanych

3.1. Definicje i lematy pomocnicze

W niniejszym rozdziale przez p i q będziemy oznaczali wykładniki spełniające warunek $p > q > 0$, F będzie przestrzenią liniową unormowaną o wymiarze $\geqslant 1$, a θ — rzeczywistą zmienną losową. Ponadto przyjmiemy umowę, że rozważając zmienną losową θ będziemy mieli na myśli zmienną losową niezdegenerowaną, tzn. taką, której rozkład nie jest skupiony w jednym punkcie.

Przypomnijmy definicję hiperkontraktywnej zmiennej losowej, podaną w Rozdziale 1:

Definicja 3.1. Powiemy, że θ jest (p, q)-*hiperkontraktywna ze stałą* $\sigma > 0$ *w przestrzeni* F, jeśli $\mathbb{E}|\theta|^p < \infty$ oraz

$$\forall_{x,y \in F} \quad \|x + \sigma\theta y\|_p \leqslant \|x + \theta y\|_q. \tag{3.1}$$

Zbiór wszystkich rozkładów zmiennych losowych θ, które są (p, q)-hiperkontraktywne ze stałą σ w przestrzeni F, będziemy oznaczać przez $\mathcal{HC}(\sigma, p, q, F)$.

Przytoczymy teraz dwa proste fakty dotyczące pojęcia hiperkontraktywności zmiennej losowej, istotne z punktu widzenia głównych wyników niniejszego rozdziału.

Fakt 3.1. *Jeśli* $\mathcal{HC}(\sigma, p, q, F) \neq \emptyset$, *to* $\sigma < 1$.

Dowód. Kładąc $x = 0$ oraz dowolne $y \neq 0$ w nierówności (3.1), dostajemy $\sigma\|\theta\|_p \leqslant \|\theta\|_q$. Ponieważ $p > q$, nierówność Höldera (monotoniczność norm L^p) implikuje, że $\sigma \leqslant 1$, przy czym σ może być 1 jedynie, gdy zmienna losowa $|\theta|$ jest zdegenerowana. Jednak gdyby $\sigma = 1$, to biorąc $x = y \neq 0$ w nierówności (3.1), dostalibyśmy $\|1+\theta\|_p = \|1 + \theta\|_q$, czyli zmienna $|1 + \theta|$ także byłaby zdegenerowana, co wraz z poprzednim przeczy umowie, że θ jest niezdegenerowana. □

Fakt 3.2. *Jeśli* $\mathcal{L}(\theta) \in \mathcal{HC}(\sigma, p, q, F)$ *dla pewnego* $\sigma > 0$ *oraz* $\mathbb{E}|\theta| < \infty$, *to* $\mathbb{E}\theta = 0$.

Dowód. Rozważmy dowolny wektor jednostkowy x oraz wektor $y = tx$ dla $t \in \mathbb{R}$. Jeśli przyjmiemy oznaczenie $g_s(t) = (\mathbb{E}|1 + t\theta|^s)^{1/s}$ ($s \in \{p, q\}$), to nierówność (3.1) stwierdza, że $g_p(\sigma t) \leqslant g_q(t)$ dla wszystkich $t \in \mathbb{R}$. Ponieważ $g_s(0) = 1$, po zróżniczkowaniu otrzymujemy, że $\sigma g_p'(0) = g_q'(0)$. W celu sprawdzenia, że funkcja g_s istotnie jest różniczkowalna w 0 i że pochodna ta wynosi $\mathbb{E}\theta$, stosujemy twierdzenie Lebesgue'a o zbieżności zmajoryzowanej (szczegóły pozostawiamy Czytelnikowi). Do zakończenia dowodu wystarczy użyć Faktu 3.1. □

Ponieważ rozważania niniejszego rozdziału będą miały charakter wyraźnie jakościowy (w przeciwieństwie do wyników z Rozdziału 2), celowe będzie przyjęcie następującego oznaczenia:

$$\mathcal{HC}(p,q,F) = \bigcup_{\sigma>0} \mathcal{HC}(\sigma,p,q,F).$$

Sformułujmy jeszcze jeden fakt, oczywisty na mocy samej Definicji 3.1:

Fakt 3.3. *Niech E, F będą liniowymi przestrzeniami unormowanymi. Jeśli dla każdej dwuwymiarowej podprzestrzeni $E_2 \subset E$ istnieje izometryczne włożenie $\imath\colon E_2 \hookrightarrow F$, to $\mathcal{HC}(p,q,E) \supseteq \mathcal{HC}(p,q,F)$.*

W dalszych dociekaniach ograniczymy się do przypadku $p > q > 1$. Motywowane jest to następującym faktem [22, Proposition 2.3]: jeśli $0 < q \leqslant 1$, $q < p$, to $\mathcal{HC}(p,q,F,\sigma) = \mathcal{HC}(p,q,\mathbb{R},\sigma)$ dla dowolnej przestrzeni liniowej unormowanej F. Tym samym, w przypadku $q \leqslant 1$ geometria przestrzeni F nie odgrywa żadnej roli. Co więcej, już przy założeniu $\mathbb{E}|\theta|^2 < \infty$, jeśli $\mathcal{L}(\theta) \in \mathcal{HC}(p,q,\mathbb{R})$ dla pewnych $p > q > 0$ to koniecznie $q > 1$ (patrz np. [43, Corollary 2.4]).

Poniższy lemat, kluczowy dla dalszych rozważań z tego rozdziału, pochodzi z pracy magisterskiej R. Latały [24]. Dla kompletności przytaczamy go wraz z dowodem.

Lemat 3.4 ([24])**.** *Ustalmy $p > q > 1$ oraz F. Załóżmy, że $\|\theta\|_p < \infty$ oraz $\mathbb{E}\theta = 0$. Wówczas następujące warunki są równoważne:*

(i) $\mathcal{L}(\theta) \in \mathcal{HC}(p,q,F)$, *czyli* $\exists_{\sigma>0} \forall_{x,y\in F} \quad \|x+\sigma\theta y\|_p \leqslant \|x+\theta y\|_q$,

(ii) $\exists_{\sigma,\varepsilon,K>0} \forall_{x,y\in F, \|x\|=1, \|y\|\leqslant\varepsilon} \quad \mathbb{E}\|x+\sigma\theta y\|^p - 1 \leqslant K(\mathbb{E}\|x+\theta y\|^q - 1)$,

(iii) $\exists_{\varepsilon,K>0} \forall_{x,y\in F, \|x\|=1, \|y\|\leqslant\varepsilon} \quad \mathbb{E}\|x+\theta y\|^p - 1 \leqslant K(\mathbb{E}\|x+\theta y\|^q - 1)$,

(iv) $\exists_{\varepsilon,K,L>0} \forall_{x,y\in F, \|x\|=1, \|y\|\leqslant\varepsilon} \quad \mathbb{E}\|\theta y\|^p \mathbf{1}_{\{\|\theta y\|\geqslant L\}} \leqslant K(\mathbb{E}\|x+\theta y\|^q - 1)$.

Dowód. (i) $\Longrightarrow$ (ii) Rozważmy σ, o którym mówi (i). Nierówność z warunku (ii) udowodnimy dla $\varepsilon = 1$. Ustalmy dowolne wektory x, y spełniające $\|x\| = 1, \|y\| \leqslant 1$. Oznaczmy

$$a = \mathbb{E}\|x+\sigma\theta y\|^p - 1, \qquad b = \mathbb{E}\|x+\theta y\|^q - 1.$$

Nierówność z warunku (i) można przepisać równoważnie:

$$a \leqslant (1+b)^{p/q} - 1. \tag{3.2}$$

Z uwagi na wypukłość funkcji $t \mapsto (1+t)^{p/q} - 1$ i oszacowanie $b \leqslant \mathbb{E}(1+|\theta|)^q - 1 \leqslant 2^q(1+\mathbb{E}|\theta|^q) = C = C(q, \|\theta\|_q)$ otrzymujemy, że

$$(1+b)^{p/q} - 1 \leqslant Kb$$

dla pewnej stałej $K = K(C, p/q)$, co w połączeniu z (3.2) daje (ii).

(ii) $\Longrightarrow$ (iv) Weźmy $\sigma, \varepsilon, K > 0$, których istnienie przewiduje (ii). Wystarczy pokazać, że dla pewnych stałych $K', L > 0$,

$$\forall_{x,y\in F, \|x\|=1, \|y\|\leqslant\varepsilon} \quad \mathbb{E}\|\theta y\|^p \mathbf{1}_{\{\|\theta y\|\geqslant L\}} \leqslant K'(\mathbb{E}\|x+\sigma\theta y\|^p - 1).$$

Zauważmy, że dla dostatecznie małego $c = c(p) > 0$ zachodzi nierówność

$$(1+t)^p \geqslant 1 + pt + c|t|^p \mathbf{1}_{\{t \geqslant 1\}}$$

dla wszystkich $t \geqslant -1$. Ustalmy dowolne $\|x\| = 1$ i $\|y\| \leqslant \varepsilon$. Wstawiając do powyższej nierówności $t := \|x + \sigma\theta y\| - 1$, i pamiętając, że $\mathbb{E}\theta = 0$, a więc na mocy nierówności Jensena $\mathbb{E}\|x + \sigma\theta y\| \geqslant 1$, otrzymujemy

$$\begin{aligned} \mathbb{E}\|x + \sigma\theta y\|^p - 1 &\geqslant p(\mathbb{E}\|x + \sigma\theta y\| - 1) + c\mathbb{E}(\|x + \sigma\theta y\| - 1)^p \mathbf{1}_{\{\|x+\sigma\theta y\| \geqslant 2\}} \\ &\geqslant c\mathbb{E}(\|\sigma\theta y\| - 2)^p \mathbf{1}_{\{\|\sigma\theta y\| \geqslant 3\}} \geqslant c(\sigma/3)^p \mathbb{E}\|\theta y\|^p \mathbf{1}_{\{\|\theta y\| \geqslant 3/\sigma\}}. \end{aligned}$$

(iv) $\Longrightarrow$ (iii) Niech $\varepsilon, K, L > 0$ będą takie jak w (iv). Oczywiście możemy założyć, że $L \geqslant 2$ i ewentualnie jeszcze zwiększyć L, aby

$$\forall_{t \geqslant L-1} \quad qt \leqslant t^q - 1. \tag{3.3}$$

Ustalmy dowolne $\|x\| = 1$, $\|y\| \leqslant \varepsilon$. Oznaczmy $A = \{\|\theta y\| < L\}$, $B = \{\|\theta y\| \geqslant L\}$. Lewą stronę nierówności z warunku (iii) możemy zapisać jako sumę dwóch składników:

$$\mathbb{E}\|x + \theta y\|^p - 1 = \mathbb{E}(\|x + \theta y\|^p - 1)\mathbf{1}_A + \mathbb{E}(\|x + \theta y\|^p - 1)\mathbf{1}_B. \tag{3.4}$$

Odpowiednie oszacowanie drugiego składnika wynika łatwo z nierówności z warunku (iv):

$$\begin{aligned} \mathbb{E}(\|x + \theta y\|^p - 1)\mathbf{1}_B &\leqslant \mathbb{E}(1 + \|\theta y\|)^p \mathbf{1}_B \leqslant 2^p \mathbb{E}\|\theta y\|^p \mathbf{1}_B \\ &\leqslant 2^p K(\mathbb{E}\|x + \theta y\|^q - 1). \end{aligned}$$

Podobne oszacowanie pierwszego składnika prawej strony (3.4), choć nie korzysta z (iv), jest nieco bardziej żmudne. Po pierwsze, wykorzystamy nierówność

$$\forall_{0 \leqslant t \leqslant L+1} \quad t^p - 1 - p(t-1) \leqslant C(t^q - 1 - q(t-1)),$$

w której stała $C > 0$ zależy tylko od p, q i L. Ewentualnie powiększając C, możemy założyć, że $Cq > p$. Stosując powyższą nierówność dla $t := \|x + \theta y\|$ i całkując ją stronami po zbiorze A otrzymamy

$$\mathbb{E}(\|x+\theta y\|^p - 1)\mathbf{1}_A - p\mathbb{E}(\|x+\theta y\| - 1)\mathbf{1}_A \leqslant C\Big(\mathbb{E}(\|x+\theta y\|^q - 1)\mathbf{1}_A - q\mathbb{E}(\|x+\theta y\| - 1)\mathbf{1}_A\Big). \tag{3.5}$$

Rozważmy teraz dwa przypadki:

Przypadek 1: $\mathbb{E}(\|x + \theta y\| - 1)\mathbf{1}_A \geqslant 0$. Z uwagi na fakt, że $p - Cq < 0$, nierówność (3.5) implikuje w tym przypadku, iż

$$\mathbb{E}(\|x + \theta y\|^p - 1)\mathbf{1}_A \leqslant C\mathbb{E}(\|x + \theta y\|^q - 1)\mathbf{1}_A \leqslant C\mathbb{E}(\|x + \theta y\|^q - 1),$$

gdzie ostatnia nierówność wynika stąd, że $(\|x + \theta y\|^q - 1)\mathbf{1}_B \geqslant \Big((L-1)^q - 1\Big)\mathbf{1}_B \geqslant 0$.

Przypadek 2: $\mathbb{E}(\|x+\theta y\|-1)\mathbf{1}_A < 0$. Wówczas

$$\mathbb{E}(\|x+\theta y\|^p - 1)\mathbf{1}_A \leqslant \mathbb{E}(\|x+\theta y\|^p - 1)\mathbf{1}_A - p\mathbb{E}(\|x+\theta y\| - 1)\mathbf{1}_A$$

i, na mocy nierówności (3.5),

$$\begin{aligned} &\leqslant C\Big(\mathbb{E}(\|x+\theta y\|^q - 1)\mathbf{1}_A + q\Big(\mathbb{P}(A) - \mathbb{E}\|x+\theta y\|\mathbf{1}_A\Big)\Big) \\ &\leqslant C\Big(\mathbb{E}(\|x+\theta y\|^q - 1)\mathbf{1}_A + q\mathbb{E}\|x+\theta y\|\mathbf{1}_B\Big), \end{aligned}$$

przy czym ostatnia nierówność wynika z tego, że $\mathbb{P}(A) \leqslant 1 \leqslant \mathbb{E}\|x+\theta y\|$. Ponieważ $\|x+\theta y\|\mathbf{1}_B \geqslant (L-1)\mathbf{1}_B$, możemy zastosować (3.3) do $t := \|x+\theta y\|$ i po scałkowaniu po zbiorze B otrzymamy

$$\mathbb{E}(\|x+\theta y\|^p - 1)\mathbf{1}_A \leqslant C\Big(\mathbb{E}(\|x+\theta y\|^q - 1)\mathbf{1}_A + \mathbb{E}(\|x+\theta y\|^q - 1)\mathbf{1}_B\Big) = C\Big(\mathbb{E}\|x+\theta y\|^q - 1\Big).$$

(iii) $\Longrightarrow$ (i) Niech $\varepsilon, K > 0$ będą takie jak w (iii). Nierówność z punktu (i) jest w oczywisty sposób spełniona dla $x = 0$ (dla dowolnego $0 < \sigma < \|\theta\|_q / \|\theta\|_p$). Z uwagi na jednorodność wystarczy dalej rozważyć $\|x\| = 1$. Na mocy nierówności trójkąta $\|x + \sigma\theta y\|_p \leqslant \sigma\|y\|\|\theta\|_p + 1$ oraz $\|x+\theta y\|_q \geqslant \|y\|\|\theta\|_q - 1$, więc dla y o dużej normie nierówność z punktu (i) będzie spełniona dla $\sigma < \|\theta\|_q/(2\|\theta\|_p)$. W związku z tym dalej możemy zakładać, że $\|y\| \leqslant C = C(\|\theta\|_p, \|\theta\|_q)$. Niech $\sigma > 0$ nie przekracza $\|\theta\|_q/(2\|\theta\|_p)$, a ponadto $\sigma K \leqslant 1$ i $\sigma C \leqslant \varepsilon$. Stosując nierówność z punktu (iii) dla σy zamiast y, otrzymujemy

$$\mathbb{E}\|x+\sigma\theta y\|^p - 1 \leqslant K\Big(\mathbb{E}\|x+\sigma\theta y\|^q - 1\Big) \leqslant \sigma K\Big(\mathbb{E}\|x+\theta y\|^q - 1\Big),$$

gdzie ostatnia nierówność wynika z wypukłości funkcji $t \mapsto \mathbb{E}\|x+t\theta y\|^q - 1$ znikającej w 0. Ponieważ $\sigma K \leqslant 1$, w rezultacie

$$\mathbb{E}\|x+\sigma\theta y\|^p \leqslant \mathbb{E}\|x+\theta y\|^q \leqslant (\mathbb{E}\|x+\theta y\|^q)^{p/q},$$

gdyż $\mathbb{E}\|x+\theta y\|^q \geqslant \|x + \mathbb{E}\theta y\|^q = 1$.

□

Na zakończenie niniejszego punktu przytoczymy definicję ortogonalności Jamesa (zwaną też ortogonalnością Birkhoffa) w kontekście przestrzeni liniowych unormowanych. Pojęcie to posłuży do sformułowania głównych wyników bieżącego rozdziału.

Definicja 3.2 (Ortogonalność Jamesa [19, Definition 1.2])**.** Niech $x, y \in F$. Powiemy, że wektor x jest *ortogonalny (w sensie Jamesa)* do y, jeśli

$$\forall_{t\in\mathbb{R}} \; \|x+ty\| \geqslant \|x\|.$$

Fakt ten będziemy zapisywać: $x \perp y$. Ponadto zbiór wektorów, do których wektor x jest ortogonalny, czyli $\Big\{y \in F \;\Big|\; x \perp y\Big\}$, będziemy oznaczać jako $x^\perp$.

Uwaga 3.5. Można podać także nieco inną, lecz równoważną definicję zbioru $x^{\perp}$:

$$x^{\perp} = \bigcup \left\{ \ker x^* \;\middle|\; x^* \in F^*, x^*(x) = \|x\|, \|x^*\| = 1 \right\}.$$

Jeśli $y \in \ker x^*$, to $\|x + ty\| \geqslant x^*(x + ty) = x^*(x) = \|x\|$. Z drugiej strony, gdy $y \neq 0$ oraz $\|x + ty\| \geqslant \|x\|$ dla wszystkich $t \in \mathbb{R}$, to x i y nie mogą być współliniowe. Ponadto możemy zdefiniować funkcjonał x^* na 2-wymiarowej przestrzeni $\mathrm{Lin}\,\{x, y\} \subset F$ spełniający $x^*(x) = \|x\|$ oraz $x^*(y) = 0$. Ponieważ $\|x + ty\| \geqslant \|x\| = x^*(x + ty)$, norma x^* na $\mathrm{Lin}\,\{x, y\}$ wynosi 1. Na mocy twierdzenia Hahna-Banacha możemy rozszerzyć x^* do funkcjonału określonego na F z zachowaniem normy.

Na mocy powyższej uwagi oczywiste jest, że $x^{\perp} \neq \{0\}$, o ile $\dim F \geqslant 2$. Warto także wspomnieć, że relacja $\perp$ nie jest symetryczna (wystarczy wziąć $F = \ell^2_\infty$, a wtedy $(1, 1) \perp (1, 0)$, lecz nie na odwrót). M.in. z tego powodu będziemy raczej stosować zapis $y \in x^{\perp}$, aniżeli $x \perp y$.

Wprowadźmy oznaczenie:

$$\Gamma = \left\{ (x, y) \in F \times F \;\middle|\; \|x\| = \|y\| = 1, y \in x^{\perp} \right\}.$$

Jedynie w przypadku $\dim F = 1$, $\Gamma = \emptyset$. Dla wszystkich par wektorów $(x, y) \in \Gamma$ zdefiniujmy funkcję

$$\tau_{x,y}(t) = \begin{cases} \|x + ty\| - 1, & \text{gdy } |t| < 1, \\ 1, & \text{gdy } |t| \geqslant 1. \end{cases}$$

Oczywiście $0 \leqslant \tau_{x,y} \leqslant 1$, $\tau_{x,y}$ znika w 0 i jest funkcją wypukłą na przedziale $[-1, 1]$, a stąd jest ona niemalejąca na $\mathbb{R}_+$ i nierosnąca na $\mathbb{R}_-$.

Można także zdefiniować funkcję $\upsilon_{x,y}(t) = (\|x + ty\| - 1) \wedge 1$. Tak samo jak w przypadku funkcji $\tau_{x,y}$, $0 \leqslant \upsilon_{x,y} \leqslant 1$ oraz $\upsilon_{x,y}$ jest niemalejąca na R_+ i nierosnąca na R_-. Ponadto

$$\upsilon_{x,y}(t) \leqslant \tau_{x,y}(t) \leqslant \upsilon_{x,y}(3t). \tag{3.6}$$

Pierwsza nierówność jest oczywista. Druga zaś, dla $|t| < 1$, wynika z tego, że $\upsilon_{x,y}(3t) \geqslant \upsilon_{x,y}(t) = \tau_{x,y}(t)$ ($\tau_{x,y} < 1$ na $(-1, 1)$). Natomiast dla $|t| \geqslant 1$, $\|x + 3ty\| - 1 \geqslant 3|t|\|y\| - \|x\| - 1 \geqslant 1$, czyli $\upsilon_{x,y}(3t) = 1$.

Nierówności (3.6) pozwolą na zastąpienie w przedstawianych dalej wynikach funkcji $\tau_{x,y}$ przez funkcję $\upsilon_{x,y}$.

3.2. Charakteryzacja zmiennych losowych hiperkontraktywnych w ustalonej przestrzeni unormowanej

Główny wynik tego rozdziału jest następujący:

Twierdzenie 3.6. *Niech $p > q > 1$ oraz θ będzie zmienną losową o średniej zero i skończonym p-tym momencie. Wówczas następujące warunki są równoważne:*

(i) $\mathcal{L}(\theta) \in \mathcal{HC}(p, q, F)$,

(ii) *istnieją* $K, t_0 > 0$ *takie, że dla każdego* $t \in [-t_0, t_0]$

$$\mathbb{E}|t\theta|^p \mathbf{1}_{\{|t\theta| \geqslant 1\}} \leqslant K\mathbb{E}\tau_{x,y}(t\theta) \quad \textit{dla dowolnych } (x,y) \in \Gamma\textit{, a także}$$

$$\mathbb{E}|t\theta|^p \mathbf{1}_{\{|t\theta| \geqslant 1\}} \leqslant K\mathbb{E}\left((t\theta)^2 \wedge 1\right).$$

Warunek (ii) charakteryzujący hiperkontraktywność zmiennej losowej jest podobny do warunków uzyskanych przez Kwapienia i Szulgę [22] oraz Latałę [24]. Jednak tamte wyniki obejmowały jedynie przypadki niektórych przestrzeni unormowanych, tj. ℓ_∞, $\mathbb{R}$, przestrzeń Hilberta $\mathbf{H}$ czy przestrzenie L^r. Konkretnie, uzyskane zostały tam następujące równoważności: dla dowolnych $p > q > 1$, zmienna losowa θ, dla której $\mathbb{E}\theta = 0$ i $\mathbb{E}|\theta|^p < \infty$, jest (p,q)-hiperkontraktywna w przestrzeni F wtedy i tylko wtedy, gdy[1]

$$\exists_{K,t_0>0}\ \forall_{|t|<t_0}\ \mathbb{E}|t\theta|^p \mathbf{1}_{\{|t\theta| \geqslant 1\}} \leqslant K\mathbb{P}(|t\theta| \geqslant 1) \quad (\text{przypadek } F = \ell_\infty),$$

$$\exists_{K,t_0>0}\ \forall_{|t|<t_0}\ \mathbb{E}|t\theta|^p \mathbf{1}_{\{|t\theta| \geqslant 1\}} \leqslant K\mathbb{E}\Big((t\theta)^2 \wedge 1\Big) \quad (\text{przypadek } F = \mathbb{R}, \mathbf{H}, L^r \text{ dla } r \in (1,2]),$$

$$\exists_{K,t_0>0}\ \forall_{|t|<t_0}\ \mathbb{E}|t\theta|^p \mathbf{1}_{\{|t\theta| \geqslant 1\}} \leqslant K\mathbb{E}\Big(|t\theta|^r \wedge 1\Big) \quad (\text{przypadek } F = L^r, \text{ dla } r \in (2,\infty)).$$

Twierdzenie 3.6 uogólnia wyżej przytoczone wyniki na przypadek dowolnej, ustalonej przestrzeni unormowanej F, pokazując jednocześnie, że geometryczny aspekt hiperkontrakcji sprowadza się do zachowania funkcji $\tau_{x,y}$ dla $(x,y) \in \Gamma$. Jako jeden z wniosków wypływających z Twierdzenia 3.6 podamy także prostszy i bardziej zwarty warunek wystarczający na to, aby zmienna losowa była hiperkontraktywna w przestrzeni F. Warunek ten będzie odwoływał się do modułu wypukłości przestrzeni F.

Dowód Twierdzenia 3.6. (i) $\Longrightarrow$ (ii) (Idee tej części dowodu są dość podobne do analogicznego, mniej ogólnego wyniku z pracy [24]). Niech $\varepsilon, K, L > 0$ będą takie jak w warunku (iv) z Lematu 3.4. Niech $\|x\| = 1$, $\|y\| \leqslant \varepsilon$ i połóżmy $t = \|y\|$. Korzystając z oczywistej nierówności

$$\mathbb{E}|t\theta|^p \mathbf{1}_{\{|t\theta| \geqslant 1\}} \leqslant L^p \mathbb{P}(|t\theta| \geqslant 1) + \mathbb{E}|t\theta|^p \mathbf{1}_{\{|t\theta| \geqslant L\}}, \tag{3.7}$$

z nierówności z warunku (iv) z Lematu 3.4 wnioskujemy, że

$$\mathbb{E}|t\theta|^p \mathbf{1}_{\{|t\theta| \geqslant 1\}} \leqslant L^p \mathbb{P}(|t\theta| \geqslant 1) + K\mathbb{E}(\|x + \theta y\|^q - 1). \tag{3.8}$$

Najpierw wykażemy drugą nierówność z warunku (ii). Wybierając $y = tx$, (3.8) przekształcamy do postaci

$$\mathbb{E}|t\theta|^p \mathbf{1}_{\{|t\theta| \geqslant 1\}} \leqslant L^p \mathbb{P}(|t\theta| \geqslant 1) + K\mathbb{E}(|1 + t\theta|^q - 1).$$

Stosując nierówność $\forall_{u\in\mathbb{R}}\ |1+u|^q \leqslant 1 + qu + C\left(u^2 \mathbf{1}_{\{|u|<1\}} + |u|^q \mathbf{1}_{\{|u| \geqslant 1\}}\right)$, która jest prawdziwa dla pewnej stałej $C = C(q) > 0$, i pamiętając, że $\mathbb{E}\theta = 0$, otrzymujemy

$$\mathbb{E}|t\theta|^p \mathbf{1}_{\{|t\theta| \geqslant 1\}} \leqslant L^p \mathbb{P}(|t\theta| \geqslant 1) + CK\mathbb{E}|t\theta|^2 \mathbf{1}_{\{|t\theta|<1\}} + CK\mathbb{E}|t\theta|^q \mathbf{1}_{\{|t\theta| \geqslant 1\}}.$$

[1] W odniesieniu do ostatniego wymienionego przypadku, w pracy [24, Wniosek 3] zostało pokazane nawet nieco więcej — nierówność z tego przypadku jest równoważna (p,q)-hiperkontraktywności zmiennej θ w każdej przestrzeni F, której moduł wypukłości $\delta_F(\varepsilon) \simeq \varepsilon^r$ (patrz Definicja 3.3).

Po zastosowaniu do ostatniego składnika po prawej stronie nierówności Höldera dostajemy

$$\mathbb{E}|t\theta|^p \mathbf{1}_{\{|t\theta|\geqslant 1\}} \leqslant K'\mathbb{E}(|t\theta|^2 \wedge 1) + K'' \left(\mathbb{E}|t\theta|^p \mathbf{1}_{\{|t\theta|\geqslant 1\}}\right)^{q/p} \mathbb{P}(|t\theta| \geqslant 1)^{1-\frac{q}{p}},$$

co już łatwo implikuje drugą nierówność z warunku (ii). Jeżeli bowiem $\mathbb{E}|t\theta|^p \mathbf{1}_{\{|t\theta|\geqslant 1\}} > 2K'\mathbb{E}(|t\theta|^2 \wedge 1)$, to

$$\frac{1}{2}\mathbb{E}|t\theta|^p \mathbf{1}_{\{|t\theta|\geqslant 1\}} \leqslant K'' \left(\mathbb{E}|t\theta|^p \mathbf{1}_{\{|t\theta|\geqslant 1\}}\right)^{q/p} \left(\mathbb{E}(|t\theta|^2 \wedge 1)\right)^{1-\frac{q}{p}}.$$

Wykazując pierwszą nierówność z warunku (ii), oczywiście zakładamy, że $\dim F > 1$. Ustalmy dowolną parę wektorów $(x, y') \in \Gamma$. Nierówność (3.8) dla wektorów x i $y = ty'$ mówi, iż

$$\begin{aligned} \mathbb{E}|t\theta|^p \mathbf{1}_{\{|t\theta|\geqslant 1\}} \leqslant L^p \mathbb{P}(|t\theta| \geqslant 1) &+ K\mathbb{E}(\|x + t\theta y'\|^q - 1)\mathbf{1}_{\{|t\theta|\geqslant 1\}} \\ &+ K\mathbb{E}\Big((1 + \tau_{x,y'}(t\theta))^q - 1\Big)\mathbf{1}_{\{|t\theta|<1\}}. \end{aligned} \tag{3.9}$$

Do drugiego składnika po prawej stronie stosujemy nierówność trójkąta dla normy oraz nierówność Höldera w podobny sposób, jak przed chwilą:

$$\begin{aligned} \mathbb{E}(\|x + t\theta y'\|^q - 1)\mathbf{1}_{\{|t\theta|\geqslant 1\}} &\leqslant \mathbb{E}\Big((1 + |t\theta|)^q - 1\Big)\mathbf{1}_{\{|t\theta|\geqslant 1\}} \leqslant 2^q \mathbb{E}|t\theta|^q \mathbf{1}_{\{|t\theta|\geqslant 1\}} \\ &\leqslant 2^q \left(\mathbb{E}|t\theta|^p \mathbf{1}_{\{|t\theta|\geqslant 1\}}\right)^{q/p} \mathbb{P}(|t\theta| \geqslant 1)^{1-\frac{q}{p}}. \end{aligned} \tag{3.10}$$

Trzeci składnik prawej strony nierówności (3.9) szacujemy używając nierówności $\forall_{u\in[0,1]}$ $(1 + u)^q - 1 \leqslant (2^q - 1)u$:

$$\mathbb{E}\Big((1 + \tau_{x,y'}(t\theta))^q - 1\Big)\mathbf{1}_{\{|t\theta|<1\}} \leqslant (2^q - 1)\mathbb{E}\tau_{x,y'}(t\theta)\mathbf{1}_{\{|t\theta|<1\}}. \tag{3.11}$$

Wstawiając (3.10) i (3.11) do (3.9) i korzystając z oczywistej równości $\mathbb{P}(|t\theta| \geqslant 1) + \mathbb{E}\tau_{x,y'}(t\theta)\mathbf{1}_{\{|t\theta|<1\}} = \mathbb{E}\tau_{x,y'}(t\theta)$, otrzymamy

$$\mathbb{E}|t\theta|^p \mathbf{1}_{\{|t\theta|\geqslant 1\}} \leqslant K'\mathbb{E}\tau_{x,y'}(t\theta) + K'' \left(\mathbb{E}\,|t\theta|^p\, \mathbf{1}_{\{|t\theta|\geqslant 1\}}\right)^{q/p} (\mathbb{E}\tau_{x,y'}(t\theta))^{1-\frac{q}{p}},$$

co, w podobny sposób jak poprzednio, implikuje pierwszą nierówność z warunku (ii).

(ii) $\Longrightarrow$ (i) Pokażemy, że istnieje stała $c > 0$ taka, że dla dowolnych $\|x\| = \|y\| = 1$ zachodzi alternatywa:

$$\forall_{t\in\mathbb{R}} \quad \mathbb{E}\|x + 4t\theta y\|^q - 1 \geqslant c\mathbb{E}\left((t\theta)^2 \wedge 1\right) \tag{3.12}$$

lub

$$\forall_{t\in\mathbb{R}}\ \exists_{y_\perp\in x^\perp, \|y_\perp\|=1} \quad \mathbb{E}\|x + 4t\theta y\|^q - 1 \geqslant c\mathbb{E}\tau_{x,y_\perp}(t\theta). \tag{3.13}$$

To już zakończy dowód, albowiem spełniony będzie warunek (iv) z Lematu 3.4 (dla $\varepsilon = 4t_0$, $L = 4$).

Ustalmy dowolne wektory jednostkowe x i y. Jeśli $y \notin \text{Lin}\{x\}$, wówczas możemy znaleźć wektor jednostkowy $y_\perp \in x^\perp \cap \text{Lin}\{x, y\}$ (patrz Uwaga 3.5). Biorąc ewentualnie $-y_\perp$ zamiast $y_\perp$, wektor y możemy przedstawić jako

$$y = \mu x + \lambda y_\perp, \quad \text{gdzie } \mu \in \mathbb{R}, \lambda \geqslant 0. \tag{3.14}$$

W przypadku gdy $y \in \text{Lin}\{x\}$, (3.14) także zachodzi, dla $|\mu| = 1$ i $\lambda = 0$, więc nawet nie ma potrzeby określania wektora $y_\perp$ w tym przypadku.
Rozważmy teraz dwa przypadki:

Przypadek 1: $|\mu| \geqslant \frac{1}{8}$. (Tutaj dopuszczamy sytuację, że $y \in \text{Lin}\{x\}$). Pokażemy, że wówczas zachodzi (3.12).
Ponieważ $1 = \|y\| = \|\mu x + \lambda y_\perp\| \geqslant \|\mu x\|$, to $|\mu| \leqslant 1$. Następnie zauważmy, że dla dostatecznie małych stałych $c_1 = c_1(q) > 0$ i $c_2 = c_2(q) > 0$ zachodzą nierówności

$$\forall_{u\in\mathbb{R}} \quad |1+u|^q \geqslant 1 + qu + c_1 u^2 \mathbf{1}_{\{|u|<1\}},$$
$$\forall_{u\in\mathbb{R}} \quad |1+u|^q \geqslant 1 + qu + c_2 \mathbf{1}_{\{|u|\geqslant 1/8\}}.$$

Z uwagi na to, że $|\mu| \in [1/8, 1]$, zachodzą także dwie podobne nierówności, z tymi samymi stałymi c_1, c_2 co powyżej:

$$\forall_{u\in\mathbb{R}} \quad |1+u|^q \geqslant 1 + qu + c_1 u^2 \mathbf{1}_{\{|u|<|\mu|\}},$$
$$\forall_{u\in\mathbb{R}} \quad |1+u|^q \geqslant 1 + qu + c_2 \mathbf{1}_{\{|u|\geqslant|\mu|\}}.$$

Z nich otrzymujemy

$$\forall_{u\in\mathbb{R}} \quad |1+u|^q \geqslant 1 + qu + c_1 u^2 \mathbf{1}_{\{|u|<|\mu|\}} + c_2 \mathbf{1}_{\{|u|\geqslant|\mu|\}}. \tag{3.15}$$

Z uwagi na (3.14), dla dowolnego $s \in \mathbb{R}$, $\|x + sy\| = \|(1 + s\mu)x + s\lambda y_\perp\| \geqslant |1 + s\mu|$. Pamiętając, że funkcja $s \mapsto \|x + sy\|^q - 1$ jest wypukła, dostajemy

$$\frac{1}{4}\left(\mathbb{E}\|x + 4t\theta y\|^q - 1\right) \geqslant \mathbb{E}\|x + t\theta y\|^q - 1 \geqslant \mathbb{E}|1 + t\mu\theta|^q - 1$$

i stosując (3.15)

$$\geqslant q\mathbb{E}(t\mu\theta) + c_1 \mathbb{E}(t\mu\theta)^2 \mathbf{1}_{\{|t\theta|<1\}} + c_2 \mathbb{P}(|t\theta| \geqslant 1),$$

a ponieważ $\mathbb{E}\theta = 0$ i $|\mu| \geqslant 1/8$, to

$$\frac{1}{4}\left(\mathbb{E}\|x + 4t\theta y\|^q - 1\right) \geqslant \Big((c_1/64) \wedge c_2\Big)\mathbb{E}\left((t\theta)^2 \wedge 1\right).$$

Przypadek 2: $|\mu| < \frac{1}{8}$. Tym razem pokażemy, że prawdziwa jest nierówność (3.13). Ponieważ $1 = \|y\| = \|\mu x + \lambda y_\perp\| \leqslant |\mu| + \lambda$, to $\lambda > \frac{7}{8}$. Oszacujemy z dołu wielkość $\|x + sy\| - 1$, oddzielnie dla $|s| < 4$ i $|s| \geqslant 4$.

Dla $|s| < 4$ zachodzi $\frac{1}{2} < 1 + \mu s < \frac{3}{2}$. Z (3.14) dostajemy, że

$$x + sy = (1 + \mu s)\left(x + \frac{\lambda}{1 + \mu s} sy_\perp\right),$$

przy czym $\frac{\lambda}{1+\mu s} > \frac{7}{8} \cdot \frac{2}{3} > \frac{1}{4}$. Wobec tego $\|x + sy\| = (1 + \mu s)\|x + \frac{\lambda}{1+\mu s} sy_\perp\|$ i

$$\begin{aligned} \|x + sy\| - 1 &= (1 + \mu s)\left(\left\|x + \frac{\lambda}{1 + \mu s} sy_\perp\right\| - 1\right) + \mu s \\ &\geqslant \frac{1}{2}\left(\left\|x + \frac{1}{4} sy_\perp\right\| - 1\right) + \mu s = \frac{1}{2}\tau_{x,y_\perp}(s/4) + \mu s, \end{aligned} \tag{3.16}$$

gdyż funkcja $t \mapsto \|x + ty_\perp\| - 1$ jest nieujemna i niemalejąca na $[0, \infty)$.
Dla $|s| \geqslant 4$, nierówność trójkąta dla normy prowadzi do oszacowania

$$\begin{aligned} \|x + sy\| - 1 &= \|(1 + \mu s)x + \lambda sy_\perp\| - 1 \geqslant \|\lambda sy_\perp\| - \|(1 + \mu s)x\| - 1 \\ &\geqslant |\lambda s| - |\mu s| - 2 = (\lambda - 2|\mu|)|s| - 2 + |\mu s| \\ &> (7/8 - 1/4) \cdot 4 - 2 + \mu s = \frac{1}{2} + \mu s. \end{aligned} \tag{3.17}$$

Nierówność $|1 + u|^q \geqslant 1 + qu$ zastosowana z $u := \|x + 4t\theta y\| - 1$ implikuje, że

$$\frac{1}{q}\left(\mathbb{E}\|x + 4t\theta y\|^q - 1\right) \geqslant \mathbb{E}\left(\|x + 4t\theta y\| - 1\right)\mathbf{1}_{\{|t\theta| < 1\}} + \mathbb{E}\left(\|x + 4t\theta y\| - 1\right)\mathbf{1}_{\{|t\theta| \geqslant 1\}},$$

a z (3.16) i (3.17) zastosowanymi z $s := 4t\theta$,

$$\geqslant \mathbb{E}\left(\frac{1}{2}\tau_{x,y_\perp}(t\theta) + 4t\mu\theta\right)\mathbf{1}_{\{|t\theta| < 1\}} + \mathbb{E}\left(\frac{1}{2} + 4t\mu\theta\right)\mathbf{1}_{\{|t\theta| \geqslant 1\}}$$

i wobec tego, że $\mathbb{E}\theta = 0$,

$$\geqslant \frac{1}{2}\mathbb{E}\tau_{x,y_\perp}(t\theta).$$

$\square$

Uwaga 3.7. Obie nierówności z warunku (ii) powyższego twierdzenia mogą zostać nieznacznie zmodyfikowane, zachowując przy tym równoważność z warunkiem (i). Po pierwsze, (3.7) pokazuje, że po lewej stronie tych nierówności może stać wyrażenie $\mathbb{E}|t\theta|^p\mathbf{1}_{\{|t\theta| \geqslant L\}}$, w którym $L \geqslant 1$. Po drugie, w pierwszej z nierówności z warunku (ii) można użyć funkcji $\upsilon_{x,y}$ zamiast $\tau_{x,y}$, zmieniając jednocześnie wyrażenie po lewej stronie na $\mathbb{E}|t\theta|^p\mathbf{1}_{\{|t\theta| \geqslant L\}}$, z dowolnym $L \geqslant 3$. Równoważność łatwo wynika z (3.6) i (3.7).

Poniżej przedstawimy kilka wniosków wynikających z udowodnionego przed chwilą twierdzenia charakteryzacyjnego.

Wniosek 3.8. *Dla ustalonego* $p > 1$ *oraz przestrzeni* F*, zbiory* $\mathcal{HC}(p, q, F)$ *dla* $q \in (1, p)$ *są sobie równe.*

Dowód. Oczywisty, ponieważ Fakt 3.2 zapewnia, że rozważane zmienne losowe mają średnią 0, a w Twierdzeniu 3.6, w warunku (ii), parametr q nie występuje. □

W tym momencie uzasadnione jest przyjęcie oznaczenia $\mathcal{HC}(p, F)$ na dowolny zbiór $\mathcal{HC}(p, q, F)$, gdzie $q \in (1, p)$. Jednocześnie zmienną losową θ, dla której $\mathcal{L}(\theta) \in \mathcal{HC}(p, F)$, będziemy nazywać *$p$-hiperkontraktywną w przestrzeni F.*

Tak jak zapowiadaliśmy, Twierdzenie 3.6 jest uogólnieniem twierdzeń charakteryzacyjnych podanych w pracach [22, 24]. Na przykład natychmiast uzyskujemy następujący

Wniosek 3.9 ([22, 24])**.** *Niech* $p > 1$ *i* θ *będzie zmienną losową o średniej zero i skończonym* p*-tym momencie. Jeśli*

$$\exists_{K,t_0>0}\ \forall_{|t|<t_0} \quad \mathbb{E}|t\theta|^p \mathbf{1}_{\{|t\theta|\geqslant 1\}} \leqslant K\mathbb{P}(|t\theta| \geqslant 1), \tag{3.18}$$

to $\mathcal{L}(\theta) \in \mathcal{HC}(p, F)$ *dla dowolnej przestrzeni* F*.*

Dowód. Wynika natychmiast z nierówności: $\mathbf{1}_{[1,\infty)}(t) \leqslant \tau_{x,y}(t)$, $\mathbf{1}_{[1,\infty)}(t) \leqslant t^2 \wedge 1$. □

Z Twierdzenia 3.6 wynika także implikacja odwrotna [22, 24]: jeśli θ jest p-hiperkontraktywna w każdej przestrzeni F (lub chociażby w ℓ_∞, czy nawet ℓ_∞^2), wówczas zachodzi (3.18). Poniżej udowodnimy nawet nieco mocniejszy fakt. Wcześniej jednak przytoczymy definicję modułu wypukłości przestrzeni liniowej unormowanej.

Definicja 3.3 (zob. np. [31, Rozdział 1.e])**.** *Modułem wypukłości* przestrzeni F nazywamy funkcję $\delta_F\colon [0, 2] \to \mathbb{R}_+$, zadaną wzorem

$$\delta_F(\varepsilon) = \inf\left\{1 - \|u + v\|/2 \;\middle|\; u, v \in F, \|u\| = \|v\| = 1, \|u - v\| = \varepsilon\right\}.$$

Ponadto F jest *jednostajnie wypukła*, jeśli $\delta_F(\varepsilon) > 0$ dla $\varepsilon \in (0, 2)$.

Uwaga 3.10. Równoważną definicję otrzymalibyśmy biorąc kres dolny po wektorach jednostkowych u, v spełniających $\|u - v\| \geqslant \varepsilon$ (patrz np. [31, uwaga za definicją 1.e.1]). Z takiej definicji natychmiast wynika, że funkcja δ_F jest niemalejąca.

Wniosek 3.11. *Niech* $p > 1$*. Niech* F *będzie dowolną przestrzenią unormowaną, która nie jest jednostajnie wypukła. Wówczas* $\mathcal{HC}(p, F) = \mathcal{HC}(p, \ell_\infty)$*. Innymi słowy, jeśli pewna zmienna losowa jest* p*-hiperkontraktywna w* F*, to jest* p*-hiperkontraktywna w dowolnej przestrzeni unormowanej.*

Dowód. Ustalmy $0 < \varepsilon_0 < 2$, dla którego $\delta_F(\varepsilon_0) = 0$. Weźmy zmienną losową θ, dla której $\mathcal{L}(\theta) \in \mathcal{HC}(p, F)$. Wystarczy, że wykażemy, że dla pewnych $K, t_0, s_0 > 0$,

$$\forall_{|t|\leqslant t_0}\ \mathbb{E}|t\theta|^p \mathbf{1}_{\{|t\theta|\geqslant 1\}} \leqslant K\mathbb{P}(|t\theta| \geqslant s_0), \tag{3.19}$$

gdyż wówczas

$$\mathbb{E}|t\theta|^p \mathbf{1}_{\{|t\theta|\geqslant s_0\}} \leqslant \mathbb{P}(|t\theta| \geqslant s_0) + \mathbb{E}|t\theta|^p \mathbf{1}_{\{|t\theta|\geqslant 1\}} \leqslant (K + 1)\mathbb{P}(|t\theta| \geqslant s_0),$$

co oznacza, że spełniona jest nierówność (3.18) z Wniosku 3.9. Z kolei nierówność (3.19) wyniknie natychmiast z Twierdzenia 3.6, jeśli tylko pokażemy, że dla dowolnego $\delta_0 > 0$ istnieje para wektorów $(x, y) \in \Gamma$, dla której

$$\tau_{x,y}(\varepsilon) \leqslant \delta_0 \text{ dla wszystkich } \varepsilon \in [0, s_0], \tag{3.20}$$

gdyż pierwsza nierówność z warunku (ii) Twierdzenia 3.6 implikuje wówczas

$$\mathbb{E}|t\theta|^p \mathbf{1}_{\{|t\theta|\geqslant 1\}} \leqslant K\left(\delta_0 + \mathbb{P}(|t\theta| \geqslant s_0)\right).$$

W tym momencie wystarczy przejść do granicy $\delta_0 \to 0$.

Ponieważ $\delta_F(\varepsilon_0) = 0$, możemy dobrać wektory jednostkowe u, v, dla których $\|u - v\| = \varepsilon_0$, zaś $\delta := 1 - \|u+v\|/2$ jest mniejsze niż dowolna, z góry zadana liczba dodatnia. Bez kłopotu można więc założyć, że $\delta < \varepsilon_0/16$. Następnie, niech $x = (u+v)/\|u+v\|$, x^* będzie funkcjonałem o normie 1 i takim, że $x^*(x) = 1$, a wtedy $x^*(u), x^*(v) \leqslant 1$. Rozważmy wektor jednostkowy $y \in \ker x^* \cap \mathrm{Lin}\,\{u, v\}$. Oczywiście, na mocy Uwagi 3.5, $y \in x^\perp$. Wprowadźmy także funkcjonały u^* i v^* takie, że $u^*(u) = u^*(x) = 1$ i $v^*(v) = v^*(x) = 1$. Ponieważ $(u+v)/2 = (1-\delta)x$, to

$$2 - 2\delta = v^*(v+u) = 1 + v^*(u) \quad \text{oraz} \quad 2 - 2\delta = x^*(v+u) \leqslant 1 + x^*(u),$$

zatem

$$x^*(u) \geqslant 1 - 2\delta = v^*(u).$$

Niech $\lambda_u \in \mathbb{R}$ będzie takie, że $x + \lambda_u y \in \mathrm{Lin}\,\{u\}$. Ponieważ wiemy, iż $1 \geqslant x^*(u) \geqslant 1 - 2\delta > 0$ i $x^*(x + \lambda_u y) = 1$, to $x + \lambda_u y = c_u u$ dla pewnego $c_u \geqslant 1$. Ponadto

$$\|x + \lambda_u y\| = \frac{x^*(x + \lambda_u y)}{x^*(u)} \leqslant \frac{1}{1 - 2\delta} \leqslant 1 + 3\delta,$$

przy czym ostatnia nierówność zachodzi dzięki temu, że $0 \leqslant \delta < \varepsilon_0/16 \leqslant 1/8$. Zatem

$$\|x + \lambda_u y\| - 1 \leqslant 3\delta. \tag{3.21}$$

Analogicznie określamy λ_v jako liczbę, dla której $x + \lambda_v y \in \mathrm{Lin}\,\{v\}$, i pokazujemy, że $x + \lambda_v y = c_v v$ dla pewnego $c_v \geqslant 1$, a także $\|x + \lambda_v y\| - 1 \leqslant 3\delta$.

Pokażemy teraz, że λ_u i λ_v mają różne znaki. Rozważymy w tym celu funkcjonał z^* taki, że $z^*(x) = 0$ i $z^*(u) = 1$, a wtedy $z^*(v) = -1$. Wówczas

$$1 = z^*(u) \leqslant z^*(x + \lambda_u y) = \lambda_u z^*(y) \quad \text{oraz} \quad -1 = z^*(v) \geqslant z^*(x + \lambda_v y) = \lambda_v z^*(y).$$

Teraz wykażemy dolne oszacowanie na $|\lambda_u|$ i $|\lambda_v|$. Nierówność trójkąta i (3.21) implikują, że

$$\begin{aligned}\frac{\varepsilon_0}{2} &= \|u - (u+v)/2\| \leqslant \|u - (x + \lambda_u y)\| + \|(x + \lambda_u y) - x\| + \|x - (u+v)/2\| \\ &= \|x + \lambda_u y\| - 1 + |\lambda_u| + \delta \leqslant |\lambda_u| + 4\delta,\end{aligned}$$

wobec czego $|\lambda_u| \geqslant \varepsilon_0/2 - 4\delta \geqslant \varepsilon_0/4$, gdyż $\delta < \varepsilon_0/16$. Tak samo pokazujemy, że $|\lambda_v| \geqslant \varepsilon_0/4$.

Dzięki monotoniczności funkcji $t \mapsto \|x + ty\| - 1$ na $[0, \infty)$ i $(-\infty, 0]$, powyższe rozważania dowodzą (3.20) dla $\delta_0 = 3\delta$ i $s_0 = \varepsilon_0/4$. $\square$

Poniższe twierdzenie podaje warunek dostateczny dla p-hiperkontraktywności w przestrzeni F. W tym warunku, geometria przestrzeni F zredukowana jest do modułu wypukłości.

Twierdzenie 3.12. *Niech* $p > 1$, $\mathbb{E}\theta = 0$, $\mathbb{E}|\theta|^p < \infty$ *i* $\dim F \geqslant 2$. *Jeśli istnieją stałe* $K, t_0 > 0$ *takie, że*

$$\forall_{|t|\leqslant t_0} \quad \mathbb{E}|t\theta|^p \mathbf{1}_{\{|t\theta|\geqslant 1\}} \leqslant K\mathbb{E}\bar{\delta}_F(|t\theta|), \tag{3.22}$$

gdzie

$$\bar{\delta}_F(\varepsilon) = \begin{cases} \delta_F(\varepsilon), & \textit{gdy } \varepsilon \in [0,1), \\ 1, & \textit{gdy } \varepsilon \geqslant 1, \end{cases}$$

to $\mathcal{L}(\theta) \in \mathcal{HC}(p, F)$.

Dowód. Najpierw zauważmy, że na mocy Twierdzenia 3.6 wystarczy pokazać, że

$$\forall_{(x,y)\in\Gamma}\ \forall_{s\in\mathbb{R}} \quad \bar{\delta}_F\left(|s|/4\right) \leqslant \tau_{x,y}(s) \tag{3.23}$$

oraz

$$\forall_{s\in\mathbb{R}} \quad \bar{\delta}_F\left(|s|/4\right) \leqslant s^2 \wedge 1. \tag{3.24}$$

Istotnie, kładąc w nierówności (3.22) $t/4$ zamiast t, otrzymamy, że

$$\forall_{|t|\leqslant 4t_0}\ 4^{-p}\mathbb{E}|t\theta|^p \mathbf{1}_{\{|t\theta|\geqslant 4\}} \leqslant K\mathbb{E}\bar{\delta}_F(|t\theta|/4). \tag{3.25}$$

Nierówności z warunku (ii) z Twierdzenia 3.6 dostajemy składając nierówność

$$\mathbb{E}|t\theta|^p \mathbf{1}_{\{|t\theta|\geqslant 1\}} \leqslant \mathbb{E}|t\theta|^p \mathbf{1}_{\{|t\theta|\geqslant 4\}} + 4^p\mathbb{P}(|t\theta| \geqslant 1)$$

wraz z (3.25) i z (3.23)/(3.24).

Nierówności (3.23) i (3.24) są oczywiste dla $|s| \geqslant 1$. Dalej będziemy zakładać $|s| < 1$. Nierówność (3.24) wynika ze znanego rezultatu [36], który mówi, że moduł wypukłości dowolnej przestrzeni unormowanej jest nie lepszy niż moduł wypukłości przestrzeni Hilberta. Konkretniej, $\delta_F(\varepsilon) \leqslant 1 - \sqrt{1 - \frac{\varepsilon^2}{4}}$, przy czym ostatnia wielkość, na mocy wypukłości funkcji $[0,1] \ni u \mapsto 1 - \sqrt{1-u}$, szacuje się z góry przez $\varepsilon^2/4$, co w zupełności wystarcza.

Dla dowodu nierówności (3.23) ustalmy dowolne $(x, y) \in \Gamma$ oraz $|s| < 1$. Wprowadźmy wektory $u = x, v = (x + sy)/\|x + sy\|, w = (u + v)/\|u + v\|$ oraz funkcjonał x^* o normie 1, dla którego $x^*(x) = 1$, $x^*(y) = 0$ (taki funkcjonał istnieje, gdyż $y \in x^\perp$ — patrz Uwaga 3.5). Wówczas

$$x^*(v) = \frac{x^*(x+sy)}{\|x+sy\|} = \frac{1}{1+\tau_{x,y}(s)} \geqslant 1 - \tau_{x,y}(s), \tag{3.26}$$

więc $x^*((u+v)/2) \geqslant 1 - \tau_{x,y}(s)/2$, gdyż $x^*(u) = 1$. To oznacza, że

$$1 \geqslant x^*(w) = \frac{x^*((u+v)/2)}{\|(u+v)/2\|} \geqslant \frac{1-\tau_{x,y}(s)/2}{\|(u+v)/2\|},$$

czyli
$$\delta_F(\|u-v\|) \leqslant 1 - \|(u+v)/2\| \leqslant \tau_{x,y}(s)/2.$$
Teraz wystarczy już tylko pokazać, że $\|u-v\| \geqslant |s|/4$, jako że funkcja δ_F jest niemalejąca (patrz Uwaga 3.10). Rozpatrzmy w tym celu dwa przypadki. Jeśli $\tau_{x,y}(s) \leqslant |s|/2$, to z uwagi na równość $x + sy = (1 + \tau_{x,y}(s))v$,
$$\|u-v\| = \|sy - \tau_{x,y}(s)v\| \geqslant |s| - \tau_{x,y}(s) \geqslant |s|/2.$$
Jeśli zaś $\tau_{x,y}(s) > |s|/2$, to korzystając z równości $x^*(v) = 1/(1+\tau_{x,y}(s))$ (patrz (3.26)) oraz nierówności $1/(1+a) \leqslant 1 - a/2$, prawdziwej dla $a \in [0,1]$, otrzymujemy
$$\|u-v\| \geqslant x^*(u-v) = 1 - x^*(v) = 1 - \frac{1}{1+\tau_{x,y}(s)} \geqslant \tau_{x,y}(s)/2 > |s|/4.$$

□

3.3. Przykład

W niniejszym punkcie skonstruujemy zmienną losową θ oraz przestrzeń F o następującej własności: $\mathcal{L}(\theta) \notin \mathcal{HC}(2,F)$ a równocześnie $\mathcal{L}(\theta) \in \mathcal{HC}(2,E)$ dla dowolnej skończenie wymiarowej podprzestrzeni $E \subset F$.

Niech θ ma rozkład taki, jak w przykładzie z pracy [22], w którym podana zmienna losowa była 2-hiperkontraktywna w $\mathbb{R}$, ale nie 2-hiperkontraktywna w ℓ_∞. Konkretnie, niech rozkład θ będzie symetryczny, zadany wzorem $\mathbb{P}(|\theta| \geqslant x) = c(x\ln x)^{-2}$ dla $x \geqslant 2$, gdzie $c = 4(\ln 2)^2$. Aby przekonać się, że θ ma żądane własności, dla $0 < t \leqslant 1/4$ obliczymy następujące wielkości:
$$\begin{aligned}
\mathbb{E}|t\theta|^2\mathbf{1}_{\{|t\theta|\geqslant 1\}} &= t^2\mathbb{E}|\theta|^2\mathbf{1}_{\{|\theta|\geqslant 1/t\}} = t^2\int_{1/t}^{\infty} 2x\mathbb{P}(|\theta| \geqslant x)\,dx \\
&= 2ct^2\int_{1/t}^{\infty}\frac{1}{x(\ln x)^2}\,dx + \mathbb{P}|\theta| \geqslant 1/t, \quad \text{i podstawiając } y = \ln x, \\
&= 2ct^2\int_{\ln(1/t)}^{\infty} y^{-2}\,dy + ct^2(\ln(1/t))^{-2} \simeq t^2(\ln(1/t))^{-1};
\end{aligned}$$
$$\mathbb{P}(|t\theta| \geqslant 1) = ct^2(\ln(1/t))^{-2};$$
$$\begin{aligned}
\mathbb{E}\Big(|t\theta|^2 \wedge 1\Big) &= (2t)^2 + 2ct^2\int_2^{1/t}\frac{1}{x(\ln x)^2}\,dx \\
&= 4t^2 + 2ct^2\Big((\ln 2)^{-1} - (\ln(1/t))^{-1}\Big) \simeq t^2.
\end{aligned}$$
Wystarczy teraz zauważyć, że nierówność (3.18) nie zachodzi, co na mocy komentarza pod Wnioskiem 3.9 pokazuje, że $\mathcal{L}(\theta) \notin \mathcal{HC}(2,\ell_\infty)$, a także $\mathcal{L}(\theta) \notin \mathcal{HC}(2,\ell_\infty^2)$. Natomiast spełniona jest druga nierówność z warunku (ii) z Twierdzenia 3.6, zatem istotnie θ jest 2-hiperkontraktywna na prostej (a także w dowolnej przestrzeni Hilberta, np. ℓ_2 czy też ℓ_2^2).

Właśnie na powyższym przykładzie dwuwymiarowych przestrzeni ℓ^2_∞ i ℓ^2_2, oczywiście izomorficznych, widać, że hiperkontraktywność zmiennej losowej nie jest własnością stabilną ze względu na przenormowanie przestrzeni. Istotną rolę grają izometryczne własności przestrzeni, wyrażone poprzez funkcję $\tau_{x,y}$ dla $(x, y) \in \Gamma$.

Teraz wskażemy nieskończenie wymiarową przestrzeń F, dla której $\mathcal{L}(\theta) \notin \mathcal{HC}(2, F)$, lecz $\mathcal{L}(\theta) \in \mathcal{HC}(2, E)$, dla dowolnej skończenie wymiarowej podprzestrzeni $E \subset F$. To pokaże, że choć definicja hiperkontraktywności „testuje” nierówność (3.1) na podprzestrzeniach wymiaru 2, może się zdarzyć, że „przeszkody” wcale nie stanowią izometryczne własności konkretnej dwuwymiarowej podprzestrzeni.

Niech F będzie przestrzenią funkcji ciągłych $C([0, 1])$ z normą

$$\|f\| = \left(\|f\|_\infty^2 + \|f\|_2^2\right)^{1/2},$$

gdzie $\|f\|_\infty = \sup |f|$ i $\|f\|_2 = (\int_0^1 f(x)^2 dx)^{1/2}$. Przestrzeń F jest wprawdzie ściśle wypukła (tj. jej kula jednostkowa jest zbiorem ściśle wypukłym), jednak nie jest jednostajnie wypukła — wskażemy dwa ciągi $(f_n), (g_n) \subset F$ funkcji o normie 1, dla których $\|f_n - g_n\| \geqslant 1$, lecz $1 - \|(f_n + g_n)/2\| \to 0$ dla $n \to \infty$ (por. Uwaga 3.10). Niech f_n będzie funkcją o nośniku zawartym w zbiorze $[0, 1/n] \cup [1 - 1/n, 1]$, przyjmującą największą wartość, $1 - \varepsilon_n$, w punktach 0 i 1 ($\varepsilon_n > 0$ ma być tak dobrane, aby $\|f_n\| = 1$; tym samym $\varepsilon_n \to 0$ przy $n \to \infty$). Funkcję g_n określmy wzorem

$$g_n(x) = \begin{cases} f_n(x), & \text{gdy } x \in [0, 1/2), \\ -f_n(x), & \text{gdy } x \in [1/2, 1]. \end{cases}$$

Pozostaje zauważyć, że $\|f_n - g_n\| \geqslant \|f_n - g_n\|_\infty = 2(1 - \varepsilon_n)$ oraz $\|(f_n + g_n)/2\| \geqslant 1 - \varepsilon_n$.

Zatem na mocy Wniosku 3.11 dostajemy, że $\mathcal{L}(\theta) \notin \mathcal{HC}(2, F)$. W celu pokazania, że dla dowolnej skończenie wymiarowej podprzestrzeni $E \subset F$ zachodzi $\mathcal{L}(\theta) \in \mathcal{HC}(2, E)$, skorzystamy z Twierdzenia 3.12. Wpierw jednak udowodnimy oszacowanie $\delta_E(\varepsilon) \geqslant c_E \varepsilon^2$, ze stałą $c_E > 0$ zależącą jedynie od wyboru podprzestrzeni E. To już wystarczy, gdyż wtedy sprawdzenie nierówności (3.22) sprowadza sie do rachunków, które przeprowadziliśmy, by pokazać, że $\mathcal{L}(\theta) \in \mathcal{HC}(2, \mathbb{R})$.

Ponieważ E jest skończonego wymiaru, normy $\|\cdot\|$ i $\|\cdot\|_2$ są równoważne na E. W szczególności istnieje stała c_E taka, że

$$\forall_{f \in E} \; \|f\|_2 \geqslant c_E \|f\|.$$

Ustalmy teraz dowolne $f, g \in E$, spełniające $\|f\| = \|g\| = 1$. Na mocy tożsamości równoległoboku,

$$\left\|\frac{f+g}{2}\right\|_2^2 + \left\|\frac{f-g}{2}\right\|_2^2 = \frac{\|f\|_2^2 + \|g\|_2^2}{2} = 1 - \frac{\|f\|_\infty^2 + \|g\|_\infty^2}{2} \leqslant 1 - \left\|\frac{f+g}{2}\right\|_\infty^2,$$

gdzie na końcu skorzystaliśmy z nierówności $(a^2 + b^2)/2 \geqslant \left((a + b)/2\right)^2$. Zatem

$$\left\|\frac{f+g}{2}\right\|^2 \leqslant 1 - \left\|\frac{f-g}{2}\right\|_2^2 \leqslant 1 - c_E^2 \left\|\frac{f-g}{2}\right\|^2,$$

czyli

$$\left\|\frac{f+g}{2}\right\| \leqslant \left(1 - c_E^2\left\|\frac{f-g}{2}\right\|^2\right)^{1/2} \leqslant 1 - \frac{c_E^2}{2}\left\|\frac{f-g}{2}\right\|^2.$$

Tym samym otrzymaliśmy $1 - \|\frac{f+g}{2}\| \geqslant \frac{c_E^2}{8}\|f-g\|^2$ dla dowolnych $\|f\| = \|g\| = 1$, a zatem $\delta_E(\varepsilon) \geqslant \frac{c_E^2}{8}\varepsilon^2$.

Część II

Nierówności funkcyjne dla miar produktowych

Rozdział 4

Wprowadzenie

Jednym z eleganckich i efektywnych podejść do zagadnień koncentracji miary i izoperymetrii jest podejście poprzez nierówności funkcyjne. Dla danego zagadnienia, np. dla nierówności izoperymetrycznej, odpowiadająca jej nierówność funkcyjna powinna mieć następującą własność: jeśli miara μ spełnia ową nierówność funkcyjną, to μ spełnia też odpowiednią nierówność izoperymetryczną. Pożądana jest też jeszcze inna własność takiej nierówności funkcyjnej, którą będziemy nazywać *„własnością tensoryzacji”*: jeśli miary μ_i $(i = 1, 2)$ ją spełniają, to ich produkt, $\mu_1 \otimes \mu_2$, także tę nierówność spełnia. Jest to szczególnie pożyteczne, gdy badana miara jest miarą produktową, a jednocześnie rozważany problem „własności tensoryzacji” nie posiada (jak to jest w przypadku koncentracji miary czy izoperymetrii). Wystarczy wówczas udowodnić nierówność funkcyjną dla każdego czynnika miary μ z osobna, aby otrzymać rozwiązanie wyjściowego problemu dla miary μ. Z takim samym zjawiskiem mieliśmy do czynienia w przypadku hiperkontrakcji — patrz np. Stwierdzenia 1.1 i 1.2; z tego drugiego natychmiast wynika nierówność typu Chinczyna-Kahane'a, która (w ogólności) nie jest konsekwencją trywialnego przypadku jednowymiarowego (porównywalności momentów pojedynczej zmiennej losowej).

Wśród spektakularnych przykładów realizacji wyżej nakreślonego podejścia należy wymienić *nierówność Bobkowa* zaproponowaną w pracy [6]. Bobkov pokazał, że kanoniczna miara gaussowska na $\mathbb{R}^n$ spełnia tę nierówność (dowodząc to dla $n = 1$, a dalej korzystając z własności tensoryzacji tej nierówności), wobec czego jako wniosek otrzymał nierówność izoperymetryczną dla miary gaussowskiej. Ta ostatnia pochodzi (niezależnie) od V.N. Sudakova i B.S. Tsirel'sona [42] oraz od C. Borella [9]. Innym przykładem jest dowód B. Maureya [32] nierówności koncentracyjnej dla produktu dwustronnych miar wykładniczych, pochodzącej od M. Talagranda [44]. Jeszcze inny dowód wyniku Talagranda podali S.G. Bobkov i M. Ledoux [7].

Wśród najbardziej znanych nierówności funkcyjnych, mających zastosowanie do zagadnienia koncentracji miary, należy wymienić *nierówność Poincaré* oraz *logarytmiczną nierówność Sobolewa*. Mówimy, że probabilistyczna miara borelowska μ na $\mathbb{R}^n$ spełnia *nierówność Poincaré* ze stałą $C > 0$, jeśli dla dowolnej funkcji $f\colon \mathbb{R}^n \to \mathbb{R}$ klasy $\mathcal{C}^1$

$$\int f^2\,d\mu - \left(\int f\,d\mu\right)^2 \leqslant C\int |\nabla f|^2\,d\mu. \tag{4.1}$$

Podobnie, miara μ spełnia *logarytmiczną nierówność Sobolewa* ze stałą $C > 0$, jeśli dla dowolnej funkcji $f\colon \mathbb{R}^n \to \mathbb{R}$ klasy $\mathcal{C}^1$

$$\int f^2 \ln f^2\,d\mu - \int f^2\,d\mu\,\ln\int f^2\,d\mu \leqslant 2C\int |\nabla f|^2\,d\mu.$$

Wśród przykładów miar należy wymienić kanoniczną miarę gaussowską — w dowolnym wymiarze spełnia obie nierówności ze stałą 1, patrz [16], miarę wykładniczą z parametrem 1 (a także jej produkt) — spełnia nierówność Poincaré ze stałą 4 [7], miary na $\mathbb{R}^n$ o logarytmicznie jednostajnie wklęsłej gęstości (patrz (1.7)) — spełniają obie nierówności [3]. Należy jeszcze zaznaczyć, że zachodzi ogólny fakt: jeśli μ spełnia logarytmiczną nierówność Sobolewa ze stałą C, to z tą samą stałą spełnia też nierówność Poincaré (patrz np. [29, Sec. 5.1]).

Z punktu widzenia tematyki poruszanej w tej części rozprawy, ważną konsekwencją faktu, że miara μ spełnia jedną z powyższych nierówności jest *własność koncentracji* miary μ. Kwestię tę omówimy w Punkcie 4.2.

Wspomnijmy jeszcze nierówność funkcyjną pochodzącą od Maureya i Pisiera [39]. Ponieważ tego rodzaju nierówność jest znana głównie w przypadku miary gaussowskiej, ograniczymy się do następującego sformułowania: niech $X, \bar{X}$ będą niezależnymi kopiami wektora gaussowskiego w $\mathbb{R}^n$ o średniej zero, zaś $\Psi\colon \mathbb{R} \to \mathbb{R}$ będzie funkcją wypukłą. Wówczas dla dowolnej funkcji $f\colon \mathbb{R}^n \to \mathbb{R}$ klasy $\mathcal{C}^1$ zachodzi nierówność:

$$\mathbb{E}\,\Psi\Big(f(X) - \mathbb{E}f(X)\Big) \leqslant \mathbb{E}\,\Psi\Big((\pi/2)\langle \nabla f(X), \bar{X}\rangle\Big). \tag{4.2}$$

4.1. Tensoryzacja

Tak jak już wspomnieliśmy na początku tego rozdziału, zaletą podejścia poprzez nierówności funkcyjne do takich problemów jak koncentracja miary czy nierówności izoperymetryczne jest fakt, że odpowiednie nierówności funkcyjne dla miar produktowych na $\mathbb{R}^n$ często można dowodzić przez indukcję po wymiarze.

W Rozdziale 5 przyjrzymy się przyczynom, dla których pewna klasa nierówności funkcyjnych, zwanych *nierównościami entropii-energii*, zawierająca m.in. nierówność Poincaré czy logarytmiczną nierówność Sobolewa, w przypadku miary produktowej dopuszcza dowód indukcyjny, po wymiarze przestrzeni produktowej. Przykłady takich nierówności, inne niż dotychczas dwa wymienione, omawia praca K. Oleszkiewicza i R. Latały [27], gdzie przedyskutowane zostało także ich zastosowanie do koncentracji miary[1]. Cechą wspólną tych wszystkich nierówności funkcyjnych jest m.in. postać *funkcjonału entropii*:

$$\Psi^\varphi(Z) = \mathbb{E}\varphi(Z) - \varphi(\mathbb{E}Z).$$

Dla przykładu, jeśli przez X oznaczymy zmienną losową o rozkładzie μ, to w nierówności Poincaré (dla miary μ) mamy do czynienia z $\Psi^\varphi(Z)$ dla $Z = f(X)$ i $\varphi(x) = x^2$; w logarytmicznej nierówności Sobolewa $Z = f(X)^2$ i $\varphi(x) = x \ln x$, natomiast w nierównościach rozważanych w [27], $Z = |f(X)|^p$ i $\varphi(x) = x^{2/p}$ ($x \geqslant 0$), przy czym parametr $p \in [1, 2]$. W dalszym ciągu, funkcjonał $Z \mapsto \Psi^\varphi(Z)$ będziemy nazywali funkcjonałem *φ-entropii*.

Własność tensoryzacji wspomnianych nierówności funkcyjnych opiera się w istocie na odpowiedniej własności dla funkcjonału $\Psi^\varphi(Z)$, która z kolei opiera się na pewnej

[1] Podobne nierówności, lecz w szczególnym kontekście miary gaussowskiej i miary Haara na sferze, rozważał już wcześniej W. Beckner [5].

własności funkcji φ: dla dowolnej przestrzeni probabilistycznej $(\Omega, \mathcal{F}, \mathbb{P})$, gdzie $\mathbb{P} = \mathbb{P}_1 \otimes \mathbb{P}_2$, i dowolnej całkowalnej zmiennej losowej Z,

$$\mathbb{E}\varphi(Z) - \varphi(\mathbb{E}Z) \leqslant \mathbb{E}\Big(\big(\mathbb{E}_1\varphi(Z) - \varphi(\mathbb{E}_1 Z)\big) + \big(\mathbb{E}_2\varphi(Z) - \varphi(\mathbb{E}_2 Z)\big)\Big) \tag{4.3}$$

lub równoważnie

$$\mathbb{E}\varphi(Z) - \mathbb{E}_1\varphi(\mathbb{E}_2 Z) - \mathbb{E}_2\varphi(\mathbb{E}_1 Z) + \varphi(\mathbb{E}Z) \geqslant 0,$$

gdzie $\mathbb{E}_i$ oznacza całkowanie względem $\mathbb{P}_i$. Powyższą własność funkcjonału φ-entropii będziemy nazywać *własnością tensoryzacji* (czasami też mówi się o *podaddytywności*).

Wiedząc, że dla ustalonego φ nierówność (4.3) jest spełniona, możemy pokazać stabilność nierówności entropii-energii ze względu na tensorowanie. Załóżmy, że miary μ_1, μ_2 (określone na $\mathbb{R}^{n_1}$ i $\mathbb{R}^{n_2}$, odpowiednio) spełniają takową nierówność z pewną stałą $C > 0$, tzn. dla dowolnej funkcji $f\colon \mathbb{R}^{n_i} \to \mathbb{R}$ klasy $\mathcal{C}^1$,

$$\Psi^\varphi(Z_i) \leqslant C\,\mathbb{E}|\nabla f(X_i)|^2,$$

gdzie X_i, dla $i = 1, 2$, jest zmienną losową o rozkładzie μ_i, natomiast Z_i jest pewną (ustaloną, właściwą dla danej nierówności) funkcją zmiennej $f(X_i)$ (np. $Z_i = f(X_i)$ w przypadku nierówności Poincaré). Jeśli zdefiniujemy zmienne losowe X_1 i X_2 tak, aby były niezależne, to (X_1, X_2) ma rozkład $\mu_1 \otimes \mu_2$. Dla dowolnej funkcji $f\colon \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \to \mathbb{R}$ klasy $\mathcal{C}^1$, niech Z będzie zmienną losową odpowiadającą zmiennej $f(X_1, X_2)$ (w takim sensie jak w przypadku Z_i). Wówczas, oznaczając gradient funkcji $x_i \mapsto f(x_1, x_2)$ przez $\nabla_i f$,

$$\begin{aligned}
\Psi^\varphi(Z) &\leqslant \mathbb{E}\Big(\big(\mathbb{E}_1\varphi(Z) - \varphi(\mathbb{E}_1 Z)\big) + \big(\mathbb{E}_2\varphi(Z) - \varphi(\mathbb{E}_2 Z)\big)\Big) \\
&\leqslant C\,\mathbb{E}\big(\mathbb{E}_1|\nabla_1 f(X_1, X_2)|^2 + \mathbb{E}_2|\nabla_2 f(X_1, X_2)|^2\big) \\
&= C\,\mathbb{E}\big(|\nabla_1 f(X_1, X_2)|^2 + |\nabla_2 f(X_1, X_2)|^2\big) = C\,\mathbb{E}|\nabla f(X_1, X_2)|^2.
\end{aligned}$$

Własność tensoryzacji dla klasycznego funkcjonału entropii (tj. $\Psi^\varphi(Z)$ dla $\varphi(x) = x \ln x$), w przeciwieństwie do tej samej własności dla funkcjonału wariancji, nie jest oczywista i wywodzi się z Teorii Informacji. Nieco szersze ujęcie tematu tensoryzacji funkcjonałów φ-entropii można znaleźć w [28, Sec. 4]. Zaprezentowano tam pochodzące od Bobkova podejście wariacyjne do nierówności (4.3), jednolite w pewnej klasie funkcji φ (zawierającej x^2, $x \ln x$). Abstrakcyjne podejście do problemu tensoryzacji funkcjonałów φ-entropii zawiera praca [27]. Zdefiniowana tam została klasa $\Phi \subset \mathcal{C}^2\big((0,\infty); \mathbb{R}\big)$, zawierająca dokładnie funkcje afiniczne oraz funkcje φ, dla których $\varphi'' > 0$ oraz $1/\varphi''$ jest funkcją wklęsła. Pokazane zostało [27, Corollary 3], że jeśli $\varphi \in \Phi$, to spełniona jest nierówność tensoryzacyjna (4.3)[2]. (Przedstawione rozumowanie daje się natychmiast odwrócić, przynajmniej przy dodatkowym założeniu, że funkcja φ, dla której zakładamy

[2] Podobny problem, w kontekście nierówności hiperkontrakcyjnych, został rozwiązany przez Oleszkiewicza [37]. Stwierdzenie (1.2) zostało tam uogólnione z norm L^p/L^q na pary ogólniejszych funkcjonałów, zawierających choćby momenty wykładnicze.

(4.3), jest klasy $\mathcal{C}^2$). Jednak ściśle rzecz biorąc, zamieszczony w [27] dowód pokazuje jedynie, że jeśli $\varphi \in \Phi$, to funkcjonał φ-entropii jest wypukły, tzn. dla dowolnego $\lambda \in (0,1)$ i dowolnych zmiennych losowych Z_1, Z_2,

$$\Psi^\varphi(\lambda Z_1 + (1-\lambda)Z_2) \leqslant \lambda\Psi^\varphi(Z_1) + (1-\lambda)\Psi^\varphi(Z_2).$$

Powyższa nierówność jest w sposób oczywisty równoważna nierówności (4.3), o ile w tamtej $\mathbb{P}_1$ lub $\mathbb{P}_2$ jest miarą dwupunktową (z atomami o wagach λ i $1-\lambda$), a trochę ogólniej — miarą dyskretną. Jednak ogólnie wydaje się, że do przejścia od wypukłości φ-entropii do nierówności (4.3) nierówność Jensena nie wystarcza — wszak funkcjonał φ-entropii, o którym zakładamy wypukłość, ma zazwyczaj dziedzinę będącą nieskończenie wymiarową przestrzenią liniową. Na ten problem zwrócono uwagę w pracy [12, Sec. 3]. Zasugerowano tam, że taka implikacja może nie zachodzić, jeżeli abstrahuje się od tego, że funkcja φ pochodzi ze szczególnej klasy funkcji (klasy Φ).

W Punkcie 5.2 pokażemy jednak, że raczej standardowy argument aproksymacyjny pozwala przezwyciężyć wspomnianą trudność. Stosowne twierdzenie, które charakteryzuje wszystkie funkcje φ (w klasie funkcji ciągłych), dla których funkcjonał φ-entropii ma własność tensoryzacji, znajduje się w Punkcie 5.3. Główna idea jest taka sama jak w pracy [27], jednak do ścisłej charakteryzacji potrzebujemy wspomnianych wyników z Punktu 5.2, jak również pewnego argumentu regularyzacyjnego z Punktu 5.3, który pozwala z nierówności (4.3) wywnioskować gładkość φ.

Ponadto, w Punkcie 5.3, dokonana została analogiczna charakteryzacja iterowanych funkcjonałów φ-entropii, posiadających własność tensoryzacji (już w nieco innym sensie niż (4.3)). Ten problem został postawiony w pracy [27, str. 167], gdzie sformułowano pewne hipotezy, jak również zaanonsowano częściowe rezultaty. Przedstawione w Punkcie 5.3 wyniki pokazują, że wbrew przypuszczeniom, odpowiednia klasa funkcjonałów okazuje się być bardzo uboga — zawiera jedynie tzw. *funkcjonały iterowanych wariancji*.

4.2. Koncentracja miary

Zjawisko koncentracji miary, wspomniane na początku tego rozdziału, w sposób abstrakcyjny formułuje się zazwyczaj następująco. Niech (X, d, μ) będzie przestrzenią metryczną z borelowską miarą probabilistyczną. Dla podzbioru borelowskiego $A \subset X$, rozważać będziemy zbiór

$$A_r = \left\{ x \in X \;\middle|\; d(x, A) < r \right\}.$$

Z przestrzenią (X, d, μ) będziemy wiązać funkcję

$$\alpha_{(X,d,\mu)}(r) := \sup \left\{ 1 - \mu(A_r) \;\middle|\; \mu(A) \geqslant 1/2 \right\}, \quad \text{dla } r > 0,$$

zwaną *funkcją koncentracji*. Poza nielicznymi, choć istotnymi przypadkami, funkcja $\alpha_{(X,d,\mu)}$ nie jest znana explicite[3]. Jednak dla zastosowań wystarczają często „dobre” oszacowania

$$\alpha_{(X,d,\mu)}(r) \leqslant \alpha(r),$$

[3] Wyrażalność explicite funkcji koncentracji jest bardzo blisko związana z rozwiązaniem problemu izoperymetrycznego w danej przestrzeni.

gdzie $\alpha(r)$ jest pewną konkretną funkcją. Przepisując powyższą nierówność i tłumacząc po drodze czym jest funkcja koncentracji otrzymamy, że gdy $\mu(A) \geqslant 1/2$, to

$$\mu(A_r^c) \leqslant \alpha(r), \quad \text{dla dowolnego } r > 0. \tag{4.4}$$

Nierówność (4.4), zamiast w języku zbiorów i ich otoczeń, można wyrazić w języku funkcji lipschitzowskich i odchyleń ich wartości od mediany:

$$\mu\left(\left\{x \in X \;\middle|\; f(x) \geqslant M + r\right\}\right) \leqslant \alpha(r), \qquad \text{dla dowolnej funkcji } f\colon X \to \mathbb{R}, \text{ 1-lipschitzowskiej}, \tag{4.5}$$

gdzie M jest medianą funkcji f, tzn. $\mu\left(\{f \geqslant M\}\right), \mu\left(\{f \leqslant M\}\right) \geqslant 1/2$. (Równoważność (4.4) i (4.5) jest dość oczywista; dowód znajduje się np. w [29, Proposition 2.6]). Podobnie można rozważać oszacowania odchyleń funkcji f od jej wartości średniej. Te ostatnie są wnioskami ze wspomnianych nierówności: Poincaré, logarytmicznej Sobolewa, a w przypadku miary gaussowskiej także Maureya-Pisiera. Jeśli miara μ na $\mathbb{R}^n$ spełnia nierówność Poincaré ze stałą $C > 0$, to dla dowolnej funkcji 1-lipschitzowskiej $f\colon \mathbb{R}^n \to \mathbb{R}$ oraz $r > 0$,

$$\mu\left(\left\{x \in \mathbb{R}^n \;\middle|\; f(x) \geqslant \int f\, d\mu + r\right\}\right) \leqslant K e^{-kr/\sqrt{C}}. \tag{4.6}$$

dla pewnych stałych numerycznych $K, k > 0$ (patrz np. [1]). W przypadku logarytmicznej nierówności Sobolewa otrzymujemy oszacowanie typu gaussowskiego (zamiast eksponencjalnego):

$$\mu\left(\left\{x \in \mathbb{R}^n \;\middle|\; f(x) \geqslant \int f\, d\mu + r\right\}\right) \leqslant e^{-r^2/2C}. \tag{4.7}$$

Wyprowadzenie powyższego oszacowania z logarytmicznej nierówności Sobolewa (patrz np. [13], także [29, Theorem 5.3]) pochodzi od I. Herbsta. Nierówność Maureya-Pisiera (4.2) implikuje podobne oszacowanie dla miary gaussowskiej, jednak stałe nie są optymalne.

Dziś istnieje bardzo bogata teoria koncentracji miary, którą przez wiele lat rozwijali m.in.: Talagrand, Ledoux, Bobkov, Gromov, Milman, Maurey i wielu innych. Doskonałą pozycją monograficzną na ten temat jest książka M. Ledoux [29]. Teoria koncentracji miary znalazła bardzo liczne zastosowania nie tylko w asymptotycznej geometrii przestrzeni Banacha (z której się wywodzi), czy też, bardziej współcześnie, w asymptotycznej analizie geometrycznej (ang.: *asymptotic geometric analysis*). Obecnie teoria koncentracji miary zajmuje istotne miejsce w Teorii Prawdopodobieństwa, będąc ważnym jej narzędziem, dzięki któremu rozwiązano szereg problemów leżących w głównym nurcie tej teorii: uzyskano szereg użytecznych nierówności probabilistycznych, uogólniających np. znane nierówności dla odchyleń sum niezależnych zmiennych losowych na dowolne funkcje lipschitzowskie od niezależnych zmiennych; otrzymano szereg twierdzeń granicznych dla zmiennych losowych o wartościach w przestrzeniach Banacha (patrz np. [30]), etc.

W Rozdziale 6 zajmiemy się problemem inspirowanym rozważaniami z Rozdziału 5, a dokładniej — z Punktu 5.4. Udowodnimy pewną nierówność koncentracyjną w przestrzeni Gaussa, wyrażoną w języku funkcji, podobnie jak to ma miejsce w przypadku nierówności (4.5). Jednak ten rodzaj koncentracji będzie miał miejsce dla innej klasy funkcji aniżeli funkcje lipschitzowskie. Okazuje się, że do tego celu doskonale nadaje się pewien wariant nierówności (4.2), który, naśladując oryginalny dowód Maureya-Pisiera, udowodnimy w Punkcie 6.3. Dalej, przy użyciu oszacowań momentów chaosów gaussowskich, wyprowadzimy wyżej wspomnianą nierówność koncentracyjną dla pewnej klasy funkcji nielipschitzowskich.

Rozdział 5

Funkcjonały z własnością tensoryzacji

5.1. Notacja i definicje

W bieżącym rozdziale, d i n będą liczbami całkowitymi dodatnimi, U będzie otwartym, wypukłym podzbiorem $\mathbb{R}^d$, zaś $\varphi\colon U \to \mathbb{R}$ będzie funkcją ciągłą. Przestrzenie probabilistyczne będziemy oznaczać przez $(\Omega, \mathcal{F}, \mathbb{P})$, $(\Omega_k, \mathcal{F}_k, \mathbb{P}_k)$, etc. W kontekście przestrzeni produktowej $(\Omega, \mathcal{F}, \mathbb{P}) = \bigotimes_{k=1}^n (\Omega_k, \mathcal{F}_k, \mathbb{P}_k)$, przez $\mathbb{E}_K$ dla dowolnego $K \subset \{1, \ldots, n\}$ będziemy rozumieć wartość oczekiwaną względem miary produktowej $\bigotimes_{k\in K} \mathbb{P}_k$; w zależności od potrzeb, $\mathbb{E}_K Z$ będziemy traktowali albo jako zmienną losową na przestrzeni $\bigotimes_{k\notin K} (\Omega_k, \mathcal{F}_k, \mathbb{P}_k)$, albo jako odpowiednią warunkową wartość oczekiwaną, czyli zmienną losową określoną na wyjściowym produkcie. Dla uproszczenia notacji, dla $k \in \{1, \ldots, n\}$ będziemy pisać $\mathbb{E}_k$ zamiast $\mathbb{E}_{\{k\}}$.

Dla $V \subseteq \mathbb{R}^d$, pisząc $Z\colon (\Omega, \mathcal{F}, \mathbb{P}) \to V$, będziemy mieli na myśli, że Z jest zmienną losową o wartościach w $\mathbb{R}^d$ oraz że $\mathbb{P}(Z \in V) = 1$.

Dla ustalonych $U \subseteq \mathbb{R}^d$, $\varphi\colon U \to \mathbb{R}$ i $(\Omega, \mathcal{F}, \mathbb{P}) = \bigotimes_{k=1}^n (\Omega_k, \mathcal{F}_k, \mathbb{P}_k)$ będziemy rozważać funkcjonał Ψ_n^φ o dziedzinie

$$\mathrm{Dom}(\Psi_n^\varphi) = \Big\{ Z\colon (\Omega, \mathcal{F}, \mathbb{P}) \to V \;\Big|\; V \subset U \text{ jest zwarty, wypukły} \Big\}, \tag{5.1}$$

określony wzorem

$$\Psi_n^\varphi(Z) = \sum_{K\subseteq\{1,\ldots,n\}} (-1)^{\#K} \mathbb{E}_{K^c} \varphi(\mathbb{E}_K Z). \tag{5.2}$$

Oczywiście $\mathrm{Dom}(\Psi_n^\varphi)$ jest wypukłym podzbiorem przestrzeni $L^0(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$ — zmiennych losowych określonych na $(\Omega, \mathcal{F}, \mathbb{P})$ o wartościach w $\mathbb{R}^d$. Zwróćmy uwagę na fakt, że określenie funkcjonału Ψ_n^φ zależy nie tylko od d, n, U i φ, ale także od wybranej przestrzeni produktowej $(\Omega, \mathcal{F}, \mathbb{P})$, której z powodów notacyjnych nie będziemy jednak umieszczać przy oznaczeniu funkcjonału. Pośrednio informację o przestrzeni $(\Omega, \mathcal{F}, \mathbb{P})$ niesie ze sobą dziedzina Ψ_n^φ. Należy jednak zaznaczyć, że czasami będziemy rozważać funkcjonał Ψ_n^φ na dziedzinie, którą będziemy definiować wprost, zamiast odwoływać się do (5.1). Jednak i w takich przypadkach będzie jasne, z jaką przestrzenią produktową $(\Omega, \mathcal{F}, \mathbb{P})$ mamy do czynienia, i w kontekście tejże przestrzeni rozważana dziedzina będzie tożsama z dziedziną zdefiniowaną poprzez (5.1).

Poniżej definiujemy główny obiekt, którym będziemy się zajmować w niniejszym rozdziale (patrz także [27]).

Definicja 5.1. Mówimy, że φ *należy do klasy* $C_n(U)$ (w skrócie $\varphi \in C_n(U)$) wtedy i tylko wtedy, gdy dla dowolnej przestrzeni produktowej $(\Omega, \mathcal{F}, \mathbb{P}) = \bigotimes_{k=1}^n (\Omega_k, \mathcal{F}_k, \mathbb{P}_k)$ odpowiedni funkcjonał Ψ_n^φ jest nieujemny, tzn. dla każdej zmiennej losowej $Z \in \mathrm{Dom}(\Psi_n^\varphi)$,

$$\Psi_n^\varphi(Z) \geqslant 0.$$

Uwaga 5.1. Oczywiście $C_n(U)$ jest stożkiem wypukłym, tzn. jeśli $\varphi_1, \varphi_2 \in C_n(U)$, $\lambda_1, \lambda_2 \geqslant 0$, to $\lambda_1\varphi_1 + \lambda_2\varphi_2 \in C_n(U)$.

Uwaga 5.2. Nadużywając nieco wprowadzonej notacji, możemy podać prostą formułę indukcyjną definiującą rodzinę funkcjonałów Ψ_n^φ $(n = 1, 2, \ldots)$ dla ustalonego ciągu przestrzeni probabilistycznych $(\Omega_n, \mathcal{F}_n, \mathbb{P}_n)$ $(n = 1, 2, \ldots)$:

$$\Psi_n^\varphi(Z) = \mathbb{E}_n \Psi_{n-1}^\varphi(Z) - \Psi_{n-1}^\varphi(\mathbb{E}_n Z), \tag{5.3}$$

dla $Z\colon \bigotimes_{k=1}^n (\Omega_k, \mathcal{F}_k, \mathbb{P}_k) \to V$, gdzie $V \subset U$ jest zwarty i wypukły. Nadużycie notacji w powyższym wzorze polega na tym, że aplikujemy zmienną Z do funkcjonału Ψ_{n-1}^φ. Oczywiście rozumiemy to w ten sposób, że n-tą współrzędną produktu ustalamy i stosujemy Ψ_{n-1}^φ do Z warunkowo. W wyniku otrzymujemy zmienną określoną na $(\Omega_n, \mathcal{F}_n, \mathbb{P}_n)$, do której przykładamy wartość oczekiwaną $\mathbb{E}_n$. Ze składnikiem $\Psi_{n-1}^\varphi(\mathbb{E}_n Z)$ nie ma problemu, jeśli $\mathbb{E}_n Z$ traktujemy jako zmienną określoną na produkcie wszystkich przestrzeni z wyjątkiem n-tej.

Na funkcjonały Ψ_n^φ można więc patrzeć jako na iteracje funkcjonału *φ-entropii* $\mathbb{E}\varphi(Z) - \varphi(\mathbb{E}Z)$. Widać ponadto, że jest pewien związek między nieujemnością funkcjonału Ψ_n^φ a wypukłością Ψ_{n-1}^φ. Precyzyjne sformułowanie tej obserwacji znajduje się w Stwierdzeniu 5.6 ((i) $\iff$ (ii')).

Uwaga 5.3. Funkcjonał Ψ_n^φ można rozszerzyć do funkcjonału $\tilde{\Psi}_n^\varphi$ działającego na nieco większej klasie zmiennych losowych, których wartości nie muszą być ograniczone prawie na pewno do zwartego, wypukłego zbioru $V \subset U$. Jednak pewne założenia dotyczące całkowalności trzeba przyjąć, aby zapewnić, że prawa strona wzoru (5.2) ma sens. Po pierwsze, założymy, że funkcja φ, za pomocą której definiujemy funkcjonały $\tilde{\Psi}_n^\varphi$, jest wypukła. Po drugie, $\tilde{\Psi}_n^\varphi$ będzie zdefiniowany na dziedzinie

$$\mathrm{Dom}(\tilde{\Psi}_n^\varphi) = \left\{ Z\colon (\Omega, \mathcal{F}, \mathbb{P}) \to U \;\middle|\; \mathbb{E}|Z| < \infty, \mathbb{E}|\varphi(Z)| < \infty \right\}.$$

Wówczas z nierówności Jensena wynika, że dla każdego $K \subseteq \{1, \ldots, n\}$,

$$a\mathbb{E}_K Z + b \leqslant \varphi(\mathbb{E}_K Z) \leqslant \mathbb{E}_K \varphi(Z) \text{ p.n.}$$

dla pewnych $a, b \in \mathbb{R}$. Ponieważ zarówno dolne jak i górne oszacowanie $\varphi(\mathbb{E}_K Z)$ jest całkowalne względem $\mathbb{E}_{K^c}$, to każdy składnik sumy ze wzoru (5.2) jest dobrze zdefiniowany i skończony. Ponadto zaznaczmy, że dzięki założeniu wypukłości φ, dziedzina funkcjonału $\tilde{\Psi}_n^\varphi$ jest wypukłym podzbiorem $L^0(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$ (wynika to z wypukłości warunku $\mathbb{E}|\varphi(Z)| < \infty$).

W kontekście klas $C_n(U)$ założenie, że φ jest wypukła, nie jest ograniczające, gdyż, jak się przekonamy, klasy $C_n(U)$ mogą zawierać jedynie funkcje wypukłe. Ponadto, prosty argument aproksymacyjny pokaże, że nieujemność funkcjonału $\tilde{\Psi}_n^\varphi$ jest konsekwencją nieujemności Ψ_n^φ (implikacja w przeciwną stronę jest oczywista; patrz Stwierdzenie 5.6, (i) $\iff$ (iii)).

Przykład 5.4. Z nierówności Jensena natychmiast wynika, że $C_1(U)$ zawiera dokładnie funkcje wypukłe (a więc jednocześnie i ciągłe) na U.

Przykład 5.5. Klasa $C_2((0,\infty))$ zawiera dokładnie te funkcje φ, dla których funkcjonały φ-entropii mają własność podaddytywności. Najważniejsze przykłady to $\varphi(x) = x^p$ dla $p \in (1,2]$ oraz $\varphi(x) = x \ln x$. W Punkcie 4.1 wspomnieliśmy, że w [27] pokazano, iż $\Phi \subseteq C_2((0,\infty))$. Natomiast w Twierdzeniu 5.12 udowodnimy, że klasy te są tożsame.

W kolejnych punktach będziemy używali jeszcze następującego oznaczenia, upraszczającego notację. Otóż przez $\{-1,1\}^n_\lambda$ będziemy rozumieć dyskretną przestrzeń probabilistyczną, będącą n-krotnym produktem miar dwupunktowych $\lambda\delta_1 + (1-\lambda)\delta_{-1}$; jeśli λ w powyższym oznaczeniu będzie pominięte, oznacza to, że $\lambda = 1/2$.

5.2. Podstawowe własności klas C_n

Zaczniemy od stwierdzenia mówiącego, iż pewne warianty definicji klasy C_n są faktycznie równoważne.

Stwierdzenie 5.6. *Niech $\varphi\colon U \to \mathbb{R}$ będzie funkcją ciągłą. Wówczas następujące warunki są równoważne:*

(i) $\varphi \in C_n(U)$,

(ii) *dla dowolnej zmiennej losowej $Z\colon \{-1,1\}^n \to U$ zachodzi $\Psi^\varphi_n(Z) \geqslant 0$,*

(ii') *dla dowolnej pary zmiennych losowych $Z_1, Z_2\colon \{-1,1\}^{n-1} \to U$,*

$$\frac{1}{2}\Psi^\varphi_{n-1}(Z_1) + \frac{1}{2}\Psi^\varphi_{n-1}(Z_2) \geqslant \Psi^\varphi_{n-1}\left(\frac{Z_1+Z_2}{2}\right),$$

(iii) *φ jest funkcją wypukłą oraz dla dowolnej przestrzeni $(\Omega,\mathcal{F},\mathbb{P}) = \bigotimes_{k=1}^n(\Omega_k,\mathcal{F}_k,\mathbb{P}_k)$, odpowiedni funkcjonał $\tilde{\Psi}^\varphi_n$ jest nieujemny, tzn. dla każdej zmiennej losowej $Z \in \mathrm{Dom}(\tilde{\Psi}^\varphi_n)$ zachodzi $\tilde{\Psi}^\varphi_n(Z) \geqslant 0$.*

W dowodzie powyższego stwierdzenia będziemy używać następujących lematów:

Lemat 5.7. *Niech $(\Omega,\mathcal{F},\mathbb{P}) = (\Omega_1,\mathcal{F}_1,\mathbb{P}_1) \otimes (\Omega_2,\mathcal{F}_2,\mathbb{P}_2)$ będzie produktową przestrzenią probabilistyczną, zaś V będzie zwartym, wypukłym podzbiorem $\mathbb{R}^d$. Wówczas dla dowolnej zmiennej losowej $Z\colon (\Omega,\mathcal{F},\mathbb{P}) \to V$ i $\varepsilon > 0$ istnieje zmienna losowa $\tilde{Z}\colon (\Omega,\mathcal{F},\mathbb{P}) \to V$ będąca postaci*

$$\tilde{Z} = \sum_{i=1}^{M}\sum_{j=1}^{N} a_{ij}\mathbf{1}_{A_i \times B_j},$$

gdzie $a_{ij} \in V$ oraz $(A_i)_{i=1}^M, (B_j)_{j=1}^N$ stanowią mierzalne rozbicie $(\Omega_1,\mathcal{F}_1,\mathbb{P}_1)$ i $(\Omega_2,\mathcal{F}_2,\mathbb{P}_2)$ (odpowiednio), a także spełniająca $\mathbb{P}(|\tilde{Z} - Z| \geqslant \varepsilon) < \varepsilon$.

Dowód. Ustalmy dowolne $\varepsilon > 0$ i wybierzmy skończone pokrycie V kulami otwartymi $U_i = B(a_i,\varepsilon)$ $(i = 1,\ldots,L)$, których środki należą do V (tj. $a_i \in V$). Definiujemy mierzalne (względem $\mathcal{F}_1 \otimes \mathcal{F}_2$) i rozłączne zbiory $C_i = Z^{-1}\left(U_i \setminus \bigcup_{j<i} U_j\right)$. Teraz każdy zbiór C_i chcemy przybliżyć sumą skończenie wielu rozłącznych prostokątów

mierzalnych (tj. zbiorów postaci $A \times B$, gdzie $A \in \mathcal{F}_1$, $B \in \mathcal{F}_2$) w taki sposób, aby miara różnicy symetrycznej tej sumy i zbioru C_i była mała. Ponieważ $\mathbb{P}_1 \otimes \mathbb{P}_2$ jest miarą produktową, to istnieje przeliczalnie wiele zbiorów $A_{i,j} \in \mathcal{F}_1$, $B_{i,j} \in \mathcal{F}_2$ $(j = 1, 2, \ldots)$, dla których zbiory $A_{i,j} \times B_{i,j}$ $(j = 1, 2, \ldots)$ są parami rozłączne (przy ustalonym i), $C_i \subseteq \bigcup_{j=1}^{\infty} (A_{i,j} \times B_{i,j})$ i

$$\mathbb{P}(C_i) + \varepsilon/L^2 > \sum_{j=1}^{\infty} \mathbb{P}_1(A_{i,j})\mathbb{P}_2(B_{i,j}).$$

Biorąc m_i takie, że ogon powyższego szeregu dla $j > m_i$ jest mniejszy niż ε/L^2 i kładąc $\tilde{C}_i = \bigcup_{j=1}^{m_i} (A_{i,j} \times B_{i,j})$, otrzymamy

$$\mathbb{P}(C_i \setminus \tilde{C}_i) \leqslant \mathbb{P}\left(\bigcup_{j > m_i} (A_{i,j} \times B_{i,j}) \right) < \varepsilon/L^2, \tag{5.4}$$

$$\mathbb{P}(\tilde{C}_i \setminus C_i) \leqslant \mathbb{P}\left(\bigcup_{j=1}^{\infty} (A_{i,j} \times B_{i,j}) \right) - \mathbb{P}(C_i) < \varepsilon/L^2. \tag{5.5}$$

Wprowadźmy

$$D_i = \tilde{C}_i \setminus \bigcup_{i' \neq i} \tilde{C}_{i'} \qquad \text{dla } i = 1, \ldots, L.$$

Oczywiście zbiory D_i są parami rozłączne i każdy z nich jest skończoną sumą rozłącznych prostokątów mierzalnych. Kładąc $D_0 = \Omega \setminus \sum_{i=1}^{L} D_i$ (co jest też skończoną sumą rozłącznych prostokątów mierzalnych) i ustalając dowolny punkt $a_0 \in V$ wnioskujemy, że $\tilde{Z} = \sum_{i=0}^{L} a_i \mathbf{1}_{D_i}$ jest oczekiwanej postaci (formalnie, należy wziąć jeszcze wspólny podpodział Ω_1 i Ω_2 generowany przez wszystkie, lecz w skończonej liczbie, prostokąty mierzalne z $D_0, D_1, \ldots, D_L$).

Dowód lematu zakończymy pokazując, że $\mathbb{P}(|\tilde{Z} - Z| \geqslant \varepsilon) < \varepsilon$. Dla każdego i,

$$\left\{ |\tilde{Z} - Z| \geqslant \varepsilon \right\} \cap C_i \subseteq C_i \setminus D_i = (C_i \setminus \tilde{C}_i) \cup \bigcup_{i' \neq i} (C_i \cap \tilde{C}_{i'}) \subseteq (C_i \setminus \tilde{C}_i) \cup \bigcup_{i' \neq i} (\tilde{C}_{i'} \setminus C_{i'}),$$

gdyż $C_i \cap \tilde{C}_{i'} = (\tilde{C}_{i'} \setminus C_{i'}) \cap C_i \subseteq \tilde{C}_{i'} \setminus C_{i'}$. Stąd dla każdego $i = 1, \ldots, L$, (5.4) i (5.5) implikują, że

$$\mathbb{P}(\left\{ |\tilde{Z} - Z| \geqslant \varepsilon \right\} \cap C_i) \leqslant \mathbb{P}(C_i \setminus \tilde{C}_i) + \sum_{i' \neq i} \mathbb{P}(\tilde{C}_{i'} \setminus C_{i'}) < L \cdot \varepsilon/L = \varepsilon/L.$$

□

Lemat 5.8. *Niech V będzie zwartym, wypukłym podzbiorem $\mathbb{R}^d$, a $\varphi\colon V \to \mathbb{R}$ funkcją ciągłą. Jeśli ciąg zmiennych losowych $Z_k\colon (\Omega_1, \mathcal{F}_1, \mathbb{P}_1) \otimes (\Omega_2, \mathcal{F}_2, \mathbb{P}_2) \to V$ zbiega według prawdopodobieństwa do Z, to $\mathbb{E}_1\varphi(\mathbb{E}_2 Z_k) \to \mathbb{E}_1\varphi(\mathbb{E}_2 Z)$.*

Dowód. Weźmy $R > 0$ takie, że $V \subseteq B(0, R)$. Dalej ustalmy dowolne $\varepsilon > 0$ oraz k takie, że $\mathbb{P}(|Z_k - Z| \geqslant \varepsilon) < \varepsilon$. Rozważmy następujące zbiory mierzalne:

$$A = \{|Z_k - Z| \geqslant \varepsilon\} \subseteq \Omega_1 \times \Omega_2,$$

$$A_{\omega_1} = \left\{ \omega_2 \;\middle|\; (\omega_1, \omega_2) \in A \right\} \subseteq \Omega_2 \quad \text{dla } \omega_1 \in \Omega_1,$$

$$B = \left\{ \omega_1 \;\middle|\; \mathbb{P}_2(A_{\omega_1}) \geqslant \sqrt{\varepsilon} \right\} \subseteq \Omega_1.$$

Na mocy twierdzenia Fubiniego

$$\varepsilon > \mathbb{E}\mathbf{1}_A = \int_{\Omega_1} \mathbb{P}_2(A_{\omega_1})\,\mathbb{P}_1(d\omega_1) \geqslant \sqrt{\varepsilon}\mathbb{P}_1(B),$$

więc $\mathbb{P}_1(B) < \sqrt{\varepsilon}$. Wobec tego, pisząc

$$\begin{aligned}|\mathbb{E}_1\varphi(\mathbb{E}_2 Z_k) - \mathbb{E}_1\varphi(\mathbb{E}_2 Z)| &\leqslant \int_B |\varphi(\mathbb{E}_2 Z_k(\omega_1,\cdot) - \varphi(\mathbb{E}_2 Z(\omega_1,\cdot)|\,\mathbb{P}_1(d\omega_1)\\ &\quad + \int_{\Omega_1\setminus B} |\varphi(\mathbb{E}_2 Z_k(\omega_1,\cdot) - \varphi(\mathbb{E}_2 Z(\omega_1,\cdot)|\,\mathbb{P}_1(d\omega_1),\end{aligned}$$

możemy oszacować pierwszy składnik po prawej stronie przez

$$2\mathbb{P}_1(B)\sup_V|\varphi| < 2\sqrt{\varepsilon}\sup_V|\varphi|.$$

Natomiast dla $\omega_1 \notin B$,

$$\begin{aligned}|\mathbb{E}_2 Z_k(\omega_1,\cdot) - \mathbb{E}_2 Z(\omega_1,\cdot)| &\leqslant \mathbb{E}_2(\|Z_k\|_\infty + \|Z\|_\infty)\mathbf{1}_{A_{\omega_1}} + \varepsilon\mathbb{E}_2\mathbf{1}_{\Omega_2\setminus A_{\omega_1}}\\ &< 2R\sqrt{\varepsilon} + \varepsilon\end{aligned}$$

i jednostajna ciągłość φ implikuje, że

$$|\mathbb{E}_1\varphi(\mathbb{E}_2 Z_k) - \mathbb{E}_1\varphi(\mathbb{E}_2 Z)| < 2\sqrt{\varepsilon}\sup_V|\varphi| + \delta(2R\sqrt{\varepsilon}+\varepsilon) \to 0, \text{ gdy } \varepsilon \to 0,$$

gdzie $\delta(\varepsilon)$ jest modułem ciągłości funkcji φ. □

Dowód Stwierdzenia 5.6. Implikacje (i) $\Longrightarrow$ (ii), (iii) $\Longrightarrow$ (i) oraz (ii) $\Longleftrightarrow$ (ii') są oczywiste. Implikację (i) $\Longrightarrow$ (iii) pokażemy po udowodnieniu Stwierdzenia 5.9, które wprawdzie korzysta ze Stwierdzenia 5.6, lecz jedynie z tej części, która mówi o równoważności warunków (i) oraz (ii).

Pokażemy więc, że (ii) $\Longrightarrow$ (i). Wystarczy w tym celu pokazać, że dla dowolnego zwartego, wypukłego zbioru $V \subset U$ i ciągu przestrzeni probabilistycznych $(\Omega_k,\mathcal{F}_k,\mathbb{P}_k)$, $k = 1,\ldots,n-1$, zachodzi

$$\begin{aligned}&\Psi_n^\varphi(Z) \geqslant 0 \text{ dla każdego } Z\colon \bigotimes_{k=1}^{n-1}(\Omega_k,\mathcal{F}_k,\mathbb{P}_k)\otimes\{-1,1\} \to V\\ &\Longrightarrow\ \Psi_n^\varphi(Z) \geqslant 0 \text{ dla każdego } (\Omega_n,\mathcal{F}_n,\mathbb{P}_n) \text{ i } Z\colon \bigotimes_{k=1}^{n}(\Omega_k,\mathcal{F}_k,\mathbb{P}_k) \to V,\end{aligned} \tag{5.6}$$

co inaczej można wysłowić, że wypukłość funkcjonału Ψ_{n-1}^φ (nawet $1/2$-wypukłość) pociąga nieujemność funkcjonału Ψ_n^φ. Stosując potem (5.6) n razy, otrzymamy (i).

Najpierw pokażmy, że implikacja (5.6) zachodzi przynajmniej dla $(\Omega_n,\mathcal{F}_n,\mathbb{P}_n) = \{-1,1\}_\lambda$ $(\lambda \in (0,1))$. Poprzednik implikacji (5.6) orzeka, że dla dowolnej pary zmiennych losowych $Z_1, Z_2\colon \bigotimes_{k=1}^{n-1}(\Omega_k,\mathcal{F}_k,\mathbb{P}_k) \to V$ zachodzi

$$\lambda\Psi_{n-1}^\varphi(Z_1) + (1-\lambda)\Psi_{n-1}^\varphi(Z_2) \geqslant \Psi_{n-1}^\varphi(\lambda Z_1 + (1-\lambda)Z_2) \tag{5.7}$$

z $\lambda = 1/2$. To łatwo implikuje (5.7) także dla $\lambda = j_i 2^{-i}$ ($0 < j_i < 2^i$), a biorąc ciąg $\lambda_i \to \lambda$, gdzie wszystkie λ_i są dwójkowo-wymierne, dostajemy (5.7) dla dowolnego $\lambda \in [0,1]$. Trzeba tylko zobaczyć, że skoro $X_i := \lambda_i Z_1 + (1-\lambda_i)Z_2 \to \lambda Z_1 + (1-\lambda)Z_2 =: X$ p.n., to $\mathbb{E}_K X_i \to \mathbb{E}_K X$ p.n. (ciąg (X_i) jest ograniczony p.n.), a także $\mathbb{E}_{K^c}\varphi(\mathbb{E}_K X_i) \to \mathbb{E}_{K^c}\varphi(\mathbb{E}_K X)$ (φ jest ciągła i ograniczona na V).

Teraz pokażemy (5.6) w pełnej ogólności, tzn. $(\Omega_n, \mathcal{F}_n, \mathbb{P}_n)$ może być dowolną przestrzenią probabilistyczną. Ustalmy dowolne $Z\colon \bigotimes_{k=1}^n (\Omega_k, \mathcal{F}_k, \mathbb{P}_k) \to V$. Lemat 5.7 stwierdza, że dla dowolnego $\varepsilon > 0$ możemy znaleźć $\tilde{Z}\colon \bigotimes_{k=1}^n (\Omega_k, \mathcal{F}_k, \mathbb{P}_k) \to V$ takie, że $\mathbb{P}(|\tilde{Z} - Z| \geqslant \varepsilon) < \varepsilon$ i

$$\tilde{Z}(\omega', \omega_n) = \sum_{j=1}^N \tilde{Z}_j(\omega')\mathbf{1}_{B_j}(\omega_n),$$

gdzie $\tilde{Z}_j\colon \bigotimes_{k=1}^{n-1} (\Omega_k, \mathcal{F}_k, \mathbb{P}_k) \to V$, $\omega' \in \prod_{k=1}^{n-1} \Omega_k$, a $(B_j)_{j=1}^N$ jest skończonym rozbiciem mierzalnym przestrzeni $(\Omega_n, \mathcal{F}_n, \mathbb{P}_n)$, $\omega_n \in \Omega_n$. Stosując nierówność (5.7) $N-1$ razy, otrzymamy

$$\begin{aligned}\mathbb{E}_n \Psi_{n-1}^{\varphi}(\tilde{Z}) &= \sum_{j=1}^N \mathbb{P}_n(B_j)\Psi_{n-1}^{\varphi}(\tilde{Z}_j) \geqslant \Psi_{n-1}^{\varphi}\left(\sum_{j=1}^N \mathbb{P}_n(B_j)\tilde{Z}_j\right) \\ &= \Psi_{n-1}^{\varphi}(\mathbb{E}_n \tilde{Z}),\end{aligned}$$

skąd, za pomocą wzoru (5.3), dostajemy $\Psi_n^{\varphi}(\tilde{Z}) \geqslant 0$. Lemat 5.8 implikuje z kolei, że dla każdego $K \subseteq \{1, \ldots, n\}$,

$$|\mathbb{E}_{K^c}\varphi(\mathbb{E}_K \tilde{Z}) - \mathbb{E}_{K^c}\varphi(\mathbb{E}_K Z)| \to 0, \text{ gdy } \varepsilon \to 0,$$

czyli także $\Psi_n^{\varphi}(Z) \geqslant 0$. □

Stwierdzenie 5.9. $C_{n+1}(U) \subseteq C_n(U)$.

Dowód. Niech $\varphi \in C_{n+1}(U)$. Na mocy Stwierdzenia 5.6 wystarczy pokazać, że $\Psi_n^{\varphi}(Z) \geqslant 0$ dla Z określonych na $\Omega = \{-1,1\}^n$ i o wartościach w U. Zdefiniujmy nową zmienną losową $\overline{Z}$ na produkcie $(n+1)$ przestrzeni $\{-1,1\}^n \times \Omega$ (ostatnie Ω traktujemy jako jedną przestrzeń, będącą ostatnim, $(n+1)$-szym czynnikiem w produkcie):

$$\overline{Z}(\varepsilon_1, \ldots, \varepsilon_n, \overline{\varepsilon}) = Z(\varepsilon_1\overline{\varepsilon}_1, \ldots, \varepsilon_n\overline{\varepsilon}_n),$$

gdzie $\varepsilon_k \in \{-1,1\}$ i $\overline{\varepsilon} = (\overline{\varepsilon}_1, \ldots, \overline{\varepsilon}_n) \in \Omega$. Ponieważ $\varphi \in C_{n+1}(U)$, to $\Psi_{n+1}^{\varphi}(\overline{Z}) = \mathbb{E}_{n+1}\Psi_n^{\varphi}(\overline{Z}) - \Psi_n^{\varphi}(\mathbb{E}_{n+1}\overline{Z}) \geqslant 0$. Jednak $\Psi_n^{\varphi}(\overline{Z}(\cdot, \overline{\varepsilon}))$ nie zależy od $\overline{\varepsilon}$ i jest zawsze równa $\Psi_n^{\varphi}(Z)$. Podobnie $\mathbb{E}_{n+1}\overline{Z}(\varepsilon_1, \ldots, \varepsilon_n, \cdot)$ nie zależy od ε_k i jest zawsze równa $\mathbb{E}Z$, więc $\Psi_{n+1}^{\varphi}(\overline{Z}) = \Psi_n^{\varphi}(Z)$. □

Teraz dokończymy

Dowód Stwierdzenia 5.6, (i) $\Longrightarrow$ (iii). Weźmy dowolne $(\Omega, \mathcal{F}, \mathbb{P}) = \bigotimes_{k=1}^n (\Omega_k, \mathcal{F}_k, \mathbb{P}_k)$, $\varphi \in C_n(U)$, dla których rozważymy odpowiadające im funkcjonały Ψ^{φ} oraz $\tilde{\Psi}^{\varphi}$ (ten drugi jest dobrze określony, bowiem na mocy Stwierdzenia 5.9 $\varphi \in C_1(U)$, tzn. φ

jest funkcją wypukłą). Ustalmy wstępujący ciąg (V_i) wypukłych, niepustych, zwartych podzbiorów U, dla których $\bigcup_i V_i = U$. Wybierzmy dowolne $v_0 \in V_1$ i zdefiniujmy ciąg zmiennych losowych

$$Z_i = Z\mathbf{1}_{Z\in V_i} + v_0\mathbf{1}_{Z\notin V_i},$$

zbieżny do Z p.n.. Pokażemy, że

$$\mathbb{E}_{K^c}\varphi(\mathbb{E}_K Z_i) \to \mathbb{E}_{K^c}\varphi(\mathbb{E}_K Z), \tag{5.8}$$

co natychmiast implikuje, że $\Psi_n^\varphi(Z_i) \to \tilde{\Psi}_n^\varphi(Z)$, a tym samym kończy dowód.

Ponieważ $|Z_i| \leqslant |Z|+|v_0|$ oraz $\mathbb{E}_K|Z| < \infty$ p.n., to na mocy twierdzenia Lebesgue'a o zbieżności zmajoryzowanej $\mathbb{E}_K Z_i \to \mathbb{E}_K Z$ p.n., a także $\varphi(\mathbb{E}_K Z_i) \to \varphi(\mathbb{E}_K Z)$ p.n., gdyż φ jest ciągła. Z kolei dzięki wypukłości φ,

$$a\mathbb{E}_K Z_i + b \leqslant \varphi(\mathbb{E}_K Z_i) \leqslant \mathbb{E}_K\varphi(Z_i)$$

dla pewnych $a, b \in \mathbb{R}$. Ponieważ

$$\mathbb{E}_K\varphi(Z_i) \leqslant \mathbb{E}_K|\varphi(Z)| + \varphi(v_0),$$

$$|a\mathbb{E}_K Z_i + b| \leqslant |a|(\mathbb{E}_K|Z| + |v_0|) + |b|$$

i obydwa oszacowania górne są całkowalne względem $\mathbb{E}_{K^c}$, to znów stosując twierdzenie Lebesgue'a o zbieżności zmajoryzowanej otrzymamy (5.8). □

W związku z powyższym, dalej będziemy pisać Ψ_n^φ nawet wtedy, gdy faktycznie będziemy mieli na myśli rozszerzenie do $\tilde{\Psi}_n^\varphi$.

Powinniśmy jeszcze zaznaczyć, że np. w przypadku funkcji φ z klasy $C_2((0,\infty))$, możemy oczekiwać, że $\Psi_2^\varphi(Z) \geqslant 0$ nie tylko dla $Z > 0$ p.n., ale także dla Z, które mają atom w 0, o ile φ przedłuża się do funkcji ciągłej na $[0,\infty)$ (patrz Przykład 5.5). Ogólniejsze, a zarazem ścisłe, sformułowanie tego faktu zawiera

Uwaga 5.10. Jeśli $\varphi\colon U \to \mathbb{R}$ należy do klasy $C_n(U)$ i rozszerza się do funkcji ciągłej $\overline{\varphi}\colon \overline{U} \to \mathbb{R}$, to dla ustalonej przestrzeni produktowej $(\Omega, \mathcal{F}, \mathbb{P}) = \bigotimes_{k=1}^n(\Omega_k, \mathcal{F}_k, \mathbb{P}_k)$ możemy określić funkcjonał $\Psi_n^{\overline{\varphi}}$ na dziedzinie

$$\mathrm{Dom}(\Psi_n^{\overline{\varphi}}) = \Big\{Z\colon (\Omega, \mathcal{F}, \mathbb{P}) \to \overline{U} \;\Big|\; \mathbb{E}|Z| < \infty, \mathbb{E}|\varphi(Z)| < \infty\Big\}$$

wzorem (5.2), w którym zamiast φ kładziemy $\overline{\varphi}$. (Powyższa definicja $\Psi_n^{\overline{\varphi}}$ jest poprawna, albowiem φ, na mocy Stwierdzenia 5.9 jest wypukła, więc $\overline{\varphi}$ też jest wypukła; dalej argumentujemy jak w Uwadze 5.3). Wówczas $\Psi_n^{\overline{\varphi}} \geqslant 0$. Istotnie, wybierając dowolny punkt $v_0 \in U$ i kładąc $Z_\varepsilon = Z\mathbf{1}_{\{Z\notin\partial U\}} + ((1-\varepsilon)Z + \varepsilon v_0)\mathbf{1}_{\{Z\in\partial U\}}$ dla $\varepsilon \in (0,1)$, otrzymamy rodzinę zmiennych losowych Z_ε o wartościach w U, które przy $\varepsilon \to 0$ zbiegają p.n. do Z. Uzasadnienie, że $\Psi_n^{\overline{\varphi}}(Z_\varepsilon) = \Psi_n^\varphi(Z_\varepsilon) \to \Psi_n^{\overline{\varphi}}(Z)$, gdy $\varepsilon \to 0$, jest analogiczne do uzasadnienia zbieżności (5.8).

Następne stwierdzenie wyjaśnia, co oznacza „własność tensoryzacji” w przypadku klas $C_n(U)$.

Stwierdzenie 5.11. *Niech $\varphi \in C_{n+1}(U)$ ($n \geqslant 1$). Niech $\Omega_{k,0}$, $\Omega_{k,1}$ dla $k = 1, 2, \ldots, n$ będą przestrzeniami probabilistycznymi*[1]*. Wówczas dla dowolnego $Z\colon \bigotimes_{k=1}^{n}(\Omega_{k,0} \otimes \Omega_{k,1}) \to U$ takiego, że $\mathbb{E}|Z| < \infty$ i $\mathbb{E}|\varphi(Z)| < \infty$, zachodzi*

$$\Psi_n^{\varphi}(Z) \leqslant \mathbb{E} \sum_{A \subseteq \{1,\ldots,n\}} \Psi_{n,A}^{\varphi}(Z),$$

gdzie $\Psi_{n,A}^{\varphi}(Z)$ oznacza zastosowanie funkcjonału Ψ_n^{φ} do zmiennej Z, traktowanej jako zmienna losowa określona na produkcie $\bigotimes_{k=1}^{n} \Omega_{k,\mathbf{1}_A(k)}$, poprzez ustalenie $\omega_{k,1-\mathbf{1}_A(k)}$.

Dowód. Pokażemy, że dla $Z\colon (\Omega_{1,0} \otimes \Omega_{1,1}) \otimes \Omega_2 \otimes \ldots \otimes \Omega_n \to U$, odpowiednio całkowalnych, zachodzi

$$\Psi_n^{\varphi}(Z) \leqslant \mathbb{E}\left(\Psi_{n,0}^{\varphi}(Z) + \Psi_{n,1}^{\varphi}(Z)\right),$$

gdzie $\Psi_{n,0}^{\varphi}(Z)$ oznacza zastosowanie Ψ_n^{φ} do zmiennej Z, traktowanej jako zmienna losowa określona na produkcie $\Omega_{1,0} \otimes \Omega_2 \otimes \ldots \otimes \Omega_n$, poprzez ustalenie ω_1^1. Podobnie interpretujemy $\Psi_{n,1}^{\varphi}(Z)$. Numerując współrzędne $\omega_{1,0}, \omega_{1,1}, \omega_2, \ldots, \omega_n$ kolejno $1_0, 1_1, 2, \ldots, n$, będziemy mieli

$$\begin{aligned}
\Psi_n^{\varphi}(Z) &= \sum_{\substack{K \subseteq \{1_0,1_1,2,\ldots,n\} \\ \#(K \cap \{1_0,1_1\}) \neq 1}} (-1)^{\#(K \setminus \{1_1\})} \mathbb{E}_{K^c} \varphi(\mathbb{E}_K Z), \\
\mathbb{E}\Psi_{n,0}^{\varphi}(Z) &= \sum_{K \subseteq \{1_0,2,\ldots,n\}} (-1)^{\#K} \mathbb{E}_{\{1_1\} \cup K^c} \varphi(\mathbb{E}_K Z), \\
\mathbb{E}\Psi_{n,1}^{\varphi}(Z) &= \sum_{K \subseteq \{1_1,2,\ldots,n\}} (-1)^{\#K} \mathbb{E}_{\{1_0\} \cup K^c} \varphi(\mathbb{E}_K Z).
\end{aligned}$$

Teraz łatwo sprawdzimy, że $\mathbb{E}\Psi_{n,0}^{\varphi}(Z) + \mathbb{E}\Psi_{n,1}^{\varphi}(Z) - \Psi_n^{\varphi}(Z) = \Psi_{n+1}^{\varphi}(Z) \geqslant 0$, gdzie w wyrażeniu $\Psi_{n+1}^{\varphi}(Z)$ traktujemy Z jako zmienną losową określoną na produkcie $n+1$ przestrzeni: $\Omega_{1,0}, \Omega_{1,1}, \Omega_2, \ldots, \Omega_n$.

Pozostaje powyższy argument zastosować indukcyjnie. □

5.3. Opis klas C_n

Twierdzenie 5.12. *Niech $U = (a, b) \subseteq \mathbb{R}$ (dopuszczamy $a = -\infty$ lub $b = \infty$), zaś $\varphi\colon U \to \mathbb{R}$ będzie funkcją ciągłą. $\varphi \in C_2(U)$ wtedy i tylko wtedy, gdy*

φ jest funkcją afiniczną lub (5.9)
φ jest dwukrotnie różniczkowalna, $\varphi'' > 0$ i $1/\varphi''$ jest funkcją wklęsłą.

Dowód. Dostateczność warunku (5.9) udowodniono w [27][2]. Ściślej rzecz biorąc, pokazana została tam wypukłość funkcjonału Ψ_1^{φ} [27, Lemma 4], czyli warunek (ii') ze Stwierdzenia 5.6.

[1] W celu skrócenia notacji rezygnujemy tutaj z pisania $(\Omega, \mathcal{F}, \mathbb{P})$.

[2] Omawiany tam faktycznie przypadek to $a = 0$, $b = \infty$, lecz przypadek ogólny nie wymaga żadnych istotnych zmian.

W dalszym ciągu będziemy dowodzili implikacji przeciwnej. Najpierw załóżmy, że $\varphi \in C_2(U) \cap \mathcal{C}^2$. Wówczas możemy naśladować dowód [27, Lemma 3]. Niech $F\colon U \times U \to \mathbb{R}$ będzie zdefiniowana wzorem

$$F(x,y) = \frac{\varphi(x)+\varphi(y)}{2} - \varphi\left(\frac{x+y}{2}\right).$$

Zauważmy, że jeśli $Z\colon \{-1,1\} \to U$ przyjmuje dwie wartości x i y, to $\Psi_1^{\varphi}(Z) = F(x,y)$. Zatem Stwierdzenie 5.6 ((i) $\Longrightarrow$ (ii')) implikuje, że F jest wypukła. Ponieważ F jest klasy $\mathcal{C}^2$, to $\mathrm{Hess}F$ jest nieujemnie określony. W szczególności

$$\frac{\partial^2 F}{\partial x^2}(x,y) = \frac{1}{2}\varphi''(x) - \frac{1}{4}\varphi''\left(\frac{x+y}{2}\right) \geqslant 0,$$

a ponieważ $\varphi'' \geqslant 0$ ($\varphi \in C_2(U) \subseteq C_1(U)$, czyli φ jest wypukła), to wnosimy stąd, że jeśli $\varphi''(x_0) = 0$ dla pewnego $x_0 \in U$, to $\varphi''(x) = 0$ dla $x \in (\frac{a+x_0}{2}, \frac{b+x_0}{2})$. Stosując ten argument indukcyjnie otrzymamy $\varphi'' \equiv 0$, tzn. φ jest afiniczna. W związku z tym dalej będziemy zakładać, że $\varphi'' > 0$. Nieujemność $\mathrm{Hess}F$ oznacza też, że

$$\frac{\partial^2 F}{\partial x^2}\frac{\partial^2 F}{\partial y^2} \geqslant \left(\frac{\partial^2 F}{\partial x \partial y}\right)^2,$$

co, jak łatwo sprawdzić, jest równoważne wklęsłości funkcji $1/\varphi''$ (rozważanej w punktach x, y i $\frac{x+y}{2}$).

Pozostaje pokazać, że założenie $\varphi \in C_2(U)$ pociąga za sobą $\varphi \in \mathcal{C}^2$. Dla $\varepsilon > 0$ oznaczmy $U^\varepsilon = (a+\varepsilon, b-\varepsilon)$ i zdefiniujmy funkcję gładką $\varphi_\varepsilon\colon U^\varepsilon \to \mathbb{R}$ jako splot $\varphi_\varepsilon = \varphi * \eta_\varepsilon$, gdzie $\eta_\varepsilon \geqslant 0$ jest gładką aproksymacją δ_0 (przy $\varepsilon \to 0$), przy czym $\mathrm{supp}(\eta_\varepsilon) \subseteq (-\varepsilon, \varepsilon)$. Ponieważ $C_2(U)$ jest stożkiem wypukłym, to $\varphi_\varepsilon \in C_2(U^\varepsilon)$.

Stosując do φ_ε poprzednią część dowodu wnioskujemy, że φ_ε jest afiniczna albo φ_ε ma ściśle dodatnią drugą pochodną i $1/\varphi_\varepsilon''$ jest funkcją wklęsłą. Zatem φ_ε'' jest wypukła. Istotnie, przypadek afiniczny jest trywialny, natomiast jeśli $\varphi_\varepsilon'' > 0$, to z wklęsłości $1/\varphi_\varepsilon''$ dla punktów x, y i $(x+y)/2$ otrzymujemy

$$\varphi_\varepsilon''\left(\frac{x+y}{2}\right) \leqslant \frac{2\varphi_\varepsilon''(x)\varphi_\varepsilon''(y)}{\varphi_\varepsilon''(x)+\varphi_\varepsilon''(y)} \leqslant \frac{\varphi_\varepsilon''(x)+\varphi_\varepsilon''(y)}{2}.$$

Stąd wiemy, że $\varphi_\varepsilon'' \geqslant 0$ oraz dla pewnego $x_0 \in \mathbb{R} \cup \{-\infty, +\infty\}$, φ_ε'' jest nierosnąca na $(-\infty, x_0] \cap U$ oraz niemalejąca na $[x_0, \infty) \cap U$, czyli φ_ε' jest niemalejąca i wklęsło-wypukła, tj. wklęsła na $(-\infty, x_0] \cap U$ i wypukła na $[x_0, \infty) \cap U$.

Najpierw pokażemy, że $\varphi \in \mathcal{C}^1$. Niech

$$\mathcal{D}_\varphi = \left\{x_0 \in U \,\middle|\, \varphi \text{ jest różniczkowalna w } x_0\right\}, \text{ oraz } \mathcal{ND}_\varphi = U \setminus \mathcal{D}_\varphi.$$

Ponieważ φ jest wypukła, to z teorii funkcji wypukłych (rzeczywistych) wiemy, że $\mathcal{ND}_\varphi$ jest co najwyżej przeliczalny, a więc $\mathcal{ND}_\varphi$ ma miarę Lebesgue'a zero oraz $\mathcal{D}_\varphi$ jest gęsty w U. Ponadto funkcja $\varphi'\colon \mathcal{D}_\varphi \to \mathbb{R}$ jest ciągła w punktach zbioru $\mathcal{D}_\varphi$, a φ — lokalnie lipschitzowska. Z powyższych faktów oraz z twierdzenia Lebesgue'a o zbieżności zmajoryzowanej wynika, że

$$\begin{aligned}\varphi_\varepsilon'(x) &= \lim_{h\to 0}\int \frac{\varphi(x-y+h)-\varphi(x-y)}{h}\eta_\varepsilon(y)dy \\ &= (\varphi' * \eta_\varepsilon)(x) \quad \text{dla } x \in U^\varepsilon\end{aligned} \tag{5.10}$$

(φ' jest zdefiniowana na $\mathcal{D}_\varphi$, czyli p.w., dlatego nie ma kłopotu z określeniem splotu $\varphi' * \eta_\varepsilon$). Przechodząc z $\varepsilon \to 0$, z ciągłości φ' w $\mathcal{D}_\varphi$ dostaniemy

$$\lim_{\varepsilon \to 0} \varphi'_\varepsilon(x) = \varphi'(x) \quad \text{dla } x \in \mathcal{D}_\varphi. \tag{5.11}$$

Ustalmy teraz ciąg $(\varepsilon_k)_{k=0}^\infty$ malejący do zera, którego początkowy wyraz ε_0 jest „mały". Poniżej, wszystkie funkcje φ_{ε_k} będziemy obcinać do wspólnej dziedziny U^{ε_0}. Jak już powiedzieliśmy wcześniej, funkcje φ'_{ε_k} są niemalejące i wklęsło-wypukłe. To powoduje, że zbieżność punktowa (5.11) na zbiorze $U^{\varepsilon_0} \cap \mathcal{D}_\varphi$, gęstym w U^{ε_0}, implikuje zbieżność niemal jednostajną ciągu $(\varphi'_{\varepsilon_k})$ na U^{ε_0}. Istotnie, rozważmy dowolny przedział domknięty $[a_0, b_0] \subset U^{\varepsilon_0}$. Biorąc dowolne $a_i, b_i \in U^{\varepsilon_0} \cap \mathcal{D}_\varphi$ $(i = 1, 2)$ takie, że $a_1 < a_2 \leqslant a_0$ i $b_0 \leqslant b_1 < b_2$, widzimy, że dla dostatecznie dużego k_0, stała Lipschitza funkcji φ'_{ε_k} (dla $k \geqslant k_0$) na przedziale $[a_0, b_0]$ jest mniejsza niż

$$\max\left(\frac{\varphi'(a_2) - \varphi'(a_1) + 1}{a_2 - a_1}, \frac{\varphi'(b_2) - \varphi'(b_1) + 1}{b_2 - b_1}\right).$$

Zatem na mocy twierdzenia Arzeli-Ascoliego istnieje podciąg (ε_{k_l}) taki, że $\varphi'_{\varepsilon_{k_l}}$ zbiega jednostajnie na $[a_0, b_0]$ do pewnej funkcji ciągłej, która jednocześnie musi być pochodną funkcji φ (na mocy twierdzenia o różniczkowalności granicy ciągu funkcji różniczkowalnych, których pochodne zbiegają jednostajnie), a więc pochodna φ we wszystkich punktach przedziału (a_0, b_0) istnieje i jest ciągła. Przechodząc do granicy $\varepsilon_0 \to 0$ oraz $a_0 \to a$, $b_0 \to b$ wnioskujemy, że φ na całej swojej dziedzinie (a, b) jest klasy $\mathcal{C}^1$. Wiemy także, że φ' również jest niemalejąca i wklęsło-wypukła.

Dowód faktu, że $\varphi \in \mathcal{C}^2$, jest podobny, dlatego poniżej przedstawiamy tylko szkic. Różniczkując równość (5.10) stronami widzimy, że $\varphi''_\varepsilon = (\varphi' * \eta_\varepsilon)'$. Następnie, stosując (5.11) do φ' zamiast φ (na co pozwala nam fakt, że φ' jest wklęsło-wypukła, a przytoczone wcześniej własności pochodnej funkcji wypukłej pozostają w mocy również w przypadku pochodnej funkcji wklęsło-wypukłej), dostaniemy

$$\varphi''_\varepsilon(x) = (\varphi' * \eta_\varepsilon)'(x) \to \varphi''(x) \quad \text{dla } x \in \mathcal{D}_{\varphi'}.$$

Używając teraz wypukłości funkcji φ''_ε, podobnie jak przed chwilą pokażemy, że rodzina $(\varphi_\varepsilon)_\varepsilon$ jest równociągła na przedziałach zwartych. W konsekwencji, pewien ciąg $\varphi''_{\varepsilon_k}$ jest niemal jednostajnie zbieżny do pewnej funkcji ciągłej, która jednocześnie musi być pochodną φ'. □

Twierdzenie 5.13. *Niech $U \subseteq \mathbb{R}^d$ będzie zbiorem otwartym, wypukłym. Dla dowolnego $n \geqslant 3$,*

$$C_n(U) = \left\{\varphi \colon U \to \mathbb{R} \;\middle|\; \varphi(x) = Q(x) + v^*(x) + c\right\},$$

gdzie Q oznacza nieujemnie określoną formę kwadratową na $\mathbb{R}^d$, v^* — funkcjonał liniowy na $\mathbb{R}^d$, a $c \in \mathbb{R}$.

Dowód. Zawieranie $\supseteq$ jest łatwe. Ponieważ wartości oczekiwane występujące we wzorze (5.2) są przemienne z v^*, to bez straty ogólności możemy założyć, że $\varphi(x) = Q(x)$. Ponadto możemy przyjąć, że $U = \mathbb{R}^d$, gdyż jeśli $\varphi \in C_n(U)$, a $U' \subseteq U$, to $\varphi_{|U'} \in C_n(U')$.

Najpierw pokażemy, że jeśli $\varphi(x) = Q(x)$ jest formą kwadratową, to

$$\Psi_n^\varphi(Z) = \Psi_n^\varphi(Z - \mathbb{E}_n Z). \tag{5.12}$$

Istotnie, oznaczając przez $Q(x, y)$ formę dwuliniową pochodzącą od $Q(x)$ i stosując wzór (5.3), dostajemy

$$\begin{aligned}
&\Psi_n^\varphi(Z - \mathbb{E}_n Z) = \mathbb{E}_n \Psi_{n-1}^\varphi(Z - \mathbb{E}_n Z) - \Psi_{n-1}^\varphi(0) \\
&= \mathbb{E}_n \sum_{K \subseteq \{1,\dots,n-1\}} (-1)^{\#K} \mathbb{E}_{K^c} Q(\mathbb{E}_K(Z - \mathbb{E}_n Z)) \\
&= \sum_{K \subseteq \{1,\dots,n-1\}} (-1)^{\#K} \mathbb{E}_{K^c} \mathbb{E}_n \Big(Q(\mathbb{E}_K Z) - 2Q(\mathbb{E}_K Z, \mathbb{E}_{K \cup \{n\}} Z) + Q(\mathbb{E}_{K \cup \{n\}} Z)\Big) \\
&= \sum_{K \subseteq \{1,\dots,n-1\}} (-1)^{\#K} \mathbb{E}_{K^c} \Big(\mathbb{E}_n Q(\mathbb{E}_K Z) - 2Q(\mathbb{E}_n \mathbb{E}_K Z, \mathbb{E}_{K \cup \{n\}} Z) + Q(\mathbb{E}_{K \cup \{n\}} Z)\Big) \\
&= \sum_{K \subseteq \{1,\dots,n-1\}} (-1)^{\#K} \Big(\mathbb{E}_{K^c \cup \{n\}} Q(\mathbb{E}_K Z) - \mathbb{E}_{K^c} Q(\mathbb{E}_{K \cup \{n\}} Z)\Big) = \Psi_n^\varphi(Z).
\end{aligned}$$

Teraz indukcyjnie pokazujemy, że $\Psi_n^\varphi \geqslant 0$, czyli $Q \in C_n(\mathbb{R}^d)$. Oczywiście $\Psi_1^\varphi \geqslant 0$. Natomiast na mocy (5.12) i (5.3) otrzymujemy

$$\Psi_n^\varphi(Z) = \Psi_n^\varphi(Z - \mathbb{E}_n Z) = \mathbb{E}_n \Psi_{n-1}^\varphi(Z - \mathbb{E}_n Z) - \Psi_{n-1}^\varphi(0) \geqslant 0,$$

gdyż na mocy założenia indukcyjnego wiemy, iż $\Psi_{n-1}^\varphi(Z - \mathbb{E}_n Z) \geqslant 0$ p.n.

Zawieranie $\subseteq$ jest mniej oczywiste. Stwierdzenie 5.9 pozwala nam rozpatrywać jedynie przypadek $n = 3$. Elegancki argument przytoczony poniżej pochodzi od K. Oleszkiewicza. (Oryginalny argument autora był nieco bardziej skomplikowany i działał w mniejszej ogólności — np. dla $U = (0, \infty) \subseteq \mathbb{R}$, ale już nie dla skończonych przedziałów).

Najpierw załóżmy dodatkowo, że funkcja $\varphi \in C_3(U)$ jest gładka. Zdefiniujmy zmienną losową $X\colon \{-1, 1\}^3 \to \mathbb{R}$ wzorem

$$X(\varepsilon_1, \varepsilon_2, \varepsilon_3) = \begin{cases} 3, & \text{gdy } |\varepsilon_1 + \varepsilon_2 + \varepsilon_3| = 3, \\ -1 & \text{w przeciwnym przypadku.} \end{cases}$$

Ustalmy $a \in U$ oraz $v \in \mathbb{R}^d$. Dla $t \in \mathbb{R}$ wprowadźmy zmienną losową $Z_t = a + vtX$. Jeśli $|t|$ jest wystarczająco mały, to Z_t ma wartości w U. Z założenia wiemy, że $\Psi_3^\varphi(Z_t) \geqslant 0$. Z drugiej strony, kładąc $f(x) = \varphi(a + vx)$ dla x z pewnego otwartego przedziału zawierającego 0, otrzymujemy

$$\begin{aligned}
\Psi_3^\varphi(Z_t) &= \sum_{K \subseteq \{1,2,3\}} (-1)^{\#K} \mathbb{E}_{K^c} f(t \mathbb{E}_K X) \\
&= \frac{1}{4} f(3t) - \frac{3}{2} f(t) + 2f(0) - \frac{3}{4} f(-t).
\end{aligned} \tag{5.13}$$

Zauważmy, że ostatnie wyrażenie znika, gdy za $f(x)$ wstawimy 1, x lub x^2, natomiast daje $6t^3$, gdy wstawimy x^3. Ponieważ f jest gładka, to wstawiając do wzoru (5.13) rozwinięcie Taylora $f(x) = f(0) + f'(0)x + \frac{1}{2} f''(0)x^2 + \frac{1}{6} f'''(0)x^3 + o(x^3)$ otrzymamy

$$\lim_{t \to 0} \frac{\Psi_3^\varphi(Z_t)}{t^3} = f'''(0).$$

Ponieważ $\frac{\Psi_3^\varphi(Z_t)}{t^3} \geqslant 0$ dla $t > 0$ oraz $\frac{\Psi_3^\varphi(Z_t)}{t^3} \leqslant 0$ dla $t < 0$, to $f'''(0) = 0$, zatem $\mathrm{D}^3\varphi(a)(v,v,v) = 0$ dla dowolnego $v \in \mathbb{R}^d$ i $a \in U$, czyli $D^3\varphi \equiv 0$. Elementarne rozumowanie (np. korzystające z twierdzenia Lagrange'a o wartości średniej) pokazuje, że φ jest żądanej postaci. Podobny rezultat, tyle że dla funkcji określonych na nieskończenie wymiarowej przestrzeni wektorowej[3], jest również prawdziwy. Np. w [33] udowodnione zostało, że jeśli funkcja wzdłuż każdej prostej jest wielomianem (jednej zmiennej) stopnia co najwyżej k, to cała funkcja jest wielomianem stopnia co najwyżej k.

Przypadek ogólny (bez założenia gładkości φ) prosto wynika z poprzedniego. Dla $\varepsilon > 0$ zdefiniujmy

$$U^\varepsilon = \left\{ x \in U \;\middle|\; \overline{B}(x,\varepsilon) \subseteq U \right\}.$$

Oczywiście U^ε jest otwartym, wypukłym podzbiorem U. Niech $\varphi_\varepsilon \colon U^\varepsilon \to \mathbb{R}$ będzie splotem $\varphi_\varepsilon = \varphi * \eta_\varepsilon$, gdzie $\eta_\varepsilon \geqslant 0$ jest gładką aproksymacją δ_0, przy czym $\operatorname{supp}(\eta_\varepsilon) \subseteq B(0,\varepsilon)$. Ponieważ $C_3(U)$ jest stożkiem wypukłym, to $\varphi_\varepsilon \in C_3(U^\varepsilon)$, więc z poprzedniej części dowodu wynika, iż φ_ε jest „funkcją kwadratową". Rozważając $\varepsilon \to 0$ wnioskujemy, że także φ jest „funkcją kwadratową". □

5.4. Iterowana nierówność Poincaré

Z Twierdzenia 5.13 wynika, że w kontekście iterowanych funkcjonałów φ-entropii, tj. Ψ_n^φ dla $n \geqslant 2$, jedynie pewien odpowiednik nierówności Poincaré (patrz (4.1)) może posiadać własność tensoryzacji. Poniżej wyjaśnimy tę kwestię w przypadku $n = 2$, z uwagi na prostotę notacji.

Niech μ i ν będą borelowskimi miarami probabilistycznymi na $\mathbb{R}^m$ i $\mathbb{R}^n$, odpowiednio, zaś X i Y — niezależnymi wektorami losowymi o rozkładzie μ i ν (odp.). Będziemy mówić, że dla *pary miar* μ i ν spełniona jest *iterowana nierówność Poincaré* ze stałą $C > 0$, jeśli dla dowolnej funkcji $f \colon \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$ klasy $\mathcal{C}^2$,

$$\Psi_2^{x \mapsto x^2}\Big(f(X,Y)\Big) \leqslant C\,\mathbb{E} \sum_{1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n} \left| \frac{\partial^2 f}{\partial x_i \partial y_j}(X,Y) \right|^2. \tag{5.14}$$

Ze Stwierdzenia 5.11 i z uwagi na postać prawej strony nierówności (5.14) natychmiast wynika, że powyższa nierówność funkcyjna ma własność tensoryzacji w następującym sensie: niech μ_k (dla $k = 1, \ldots, M$) będą borelowskimi miarami probabilistycznymi na $\mathbb{R}^{m_k}$, zaś ν_l (dla $l = 1, \ldots, N$) będą borelowskimi miarami probabilistycznymi na $\mathbb{R}^{n_l}$. Jeśli dla każdych $1 \leqslant k \leqslant M$ i $1 \leqslant l \leqslant N$ iterowana nierówność Poincaré spełniona jest dla pary miar μ_k i ν_l (ze stałą C), to z tą samą stałą C nierówność ta spełniona jest także dla pary miar $\mu_1 \otimes \cdots \otimes \mu_M$ i $\nu_1 \otimes \cdots \otimes \nu_N$.

Niestety, autorowi rozprawy nie udało się z nierówności takiej jak (5.14) wyprowadzić koncentracji dla pary miar μ i ν, jak to ma miejsce w przypadku zwykłej nierówności Poincaré (patrz (4.6)). Jednakże, w przypadku gdy μ i ν są miarami gaussowskimi, pewna (optymalna) nierówność koncentracyjna została udowodniona inną metodą, co przedstawimy w Rozdziale 6.

[3] W przestrzeni wektorowej pojęcie wielomianu rzędu k można zdefiniować np. poprzez zerowanie się wszystkich operatorów różnicowych rzędu k.

Rozdział 6

Koncentracja dla funkcji nielipschitzowskich

6.1. Motywacja

Zacznijmy od prostej obserwacji. Niech $f\colon \mathbb{R}\times\mathbb{R}\to\mathbb{R}$ będzie funkcją klasy $\mathcal{C}^2$. Połóżmy $g(x,y)=f(x,0)+f(0,y)-f(0,0)$. Jeśli $|\frac{\partial^2 f}{\partial x\partial y}|\leqslant L$ dla pewnej stałej $L>0$, to dla dowolnego $t\geqslant 0$ spełniona jest nierówność

$$(\gamma\otimes\gamma)\Big(\Big\{(x,y)\in\mathbb{R}\times\mathbb{R}\,\Big|\,|f(x,y)-g(x,y)|\geqslant t\Big\}\Big)\leqslant Ce^{-ct/L}, \qquad (6.1)$$

gdzie γ jest standardową miarą gaussowską na $\mathbb{R}$, zaś $C,c>0$ są pewnymi stałymi numerycznymi. Istotnie, z równości

$$f(x,y)-g(x,y)=f(x,y)-f(x,0)-f(0,y)+f(0,0)=\int_0^x\int_0^y\frac{\partial^2 f}{\partial x\partial y}(u,v)\,du\,dv$$

oraz z ograniczenia na pochodną mieszaną f dostajemy

$$|f(x,y)-g(x,y)|\leqslant L|xy|.$$

Nierówność (6.1) wynika teraz z faktu, że zmienna losowa $Z=|g_1g_2|$, gdzie g_1,g_2 są niezależnymi zmiennymi losowymi o rozkładzie $\mathcal{N}(0,1)$, ma ogony malejące wykładniczo. Ściślej, na mocy nierówności Czebyszewa, dla każdego $p\geqslant 1$,

$$\mathbb{P}(Z\geqslant e\|Z\|_p)\leqslant e^{-p}, \qquad (6.2)$$

a ponieważ $\|Z\|_p=\|g_1\|_p^2\leqslant Cp$, to $\mathbb{P}(Z\geqslant t)\leqslant C'e^{-ct}$ dla dowolnego $t\geqslant 0$.

Badanie miary odchylenia funkcji f od funkcji takiej postaci jak g można umotywować następująco. Rozważmy $f(x,y)=f_1(x)+f_2(y)$, gdzie f_1,f_2 są klasy $\mathcal{C}^2$. Ponieważ $\frac{\partial^2 f}{\partial x\partial y}\equiv 0$, to w takim przypadku (6.1) może zachodzić jedynie dla $g\equiv f$. Przykład ten pokazuje tym samym, że nie można spodziewać się np. stałej w miejscu funkcji g w nierówności (6.1) (np. średniej z f względem miary $\gamma\otimes\gamma$, etc.).

W dalszym ciągu niniejszego rozdziału pokażemy, że nierówność (6.1) można uogólnić m.in. na przypadek funkcji $f\colon\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}$ i miary $\gamma_n\otimes\gamma_n$, gdzie γ_n jest kanoniczną miarą gaussowską na $\mathbb{R}^n$.

6.2. Notacja

Niech $X, Y, \bar{X}, \bar{Y}$ będą niezależnymi wektorami losowymi na $\mathbb{R}^n$, o rozkładzie γ_n. Rozważmy przestrzeń probabilistyczną $(\Omega, \mathcal{F}, \mathbb{P})$ z wyróżnionymi σ-ciałami $\mathcal{F}_1$, $\mathcal{F}_2$, które są niezależne oraz spełniają

$$\mathcal{F} = \sigma(\mathcal{F}_1, \mathcal{F}_2). \tag{6.3}$$

Ponadto niech $\mathcal{F}_1$ i $\mathcal{F}_2$ będą dostatecznie bogate, aby niezależne wektory losowe $X, Y, \bar{X}, \bar{Y}$ zdefiniować na $(\Omega, \mathcal{F}, \mathbb{P})$ w taki sposób, żeby $X, \bar{X}$ były $\mathcal{F}_1$-mierzalne, zaś $Y, \bar{Y}$ były $\mathcal{F}_2$-mierzalne.

Dla dowolnej całkowalnej zmiennej losowej $V\colon (\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$ zdefiniujmy

$$\Pi_1 V = \mathbb{E}\Big[V\Big|\mathcal{F}_1\Big] + \mathbb{E}\Big[V\Big|\mathcal{F}_2\Big] - \mathbb{E}V.$$

Operator Π_1 ograniczony do przestrzeni $L^2(\Omega, \mathcal{F}, \mathbb{P})$ jest rzutem ortogonalnym na podprzestrzeń rozpiętą przez zmienne losowe będące albo $\mathcal{F}_1$-, albo $\mathcal{F}_2$-mierzalne. Na przykład dla funkcji $f\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, dla której $\mathbb{E}|f(X,Y)| < \infty$,

$$\Pi_1 f(X,Y) = \mathbb{E}\Big[f(X,Y)\Big|X\Big] + \mathbb{E}\Big[f(X,Y)\Big|Y\Big] - \mathbb{E}f(X,Y) \quad \text{p.n.}$$

Ustalmy funkcję $f\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ klasy $\mathcal{C}^2$ oraz $(x,y) \in \mathbb{R}^n \times \mathbb{R}^n$. Przez macierz pochodnych mieszanych drugiego rzędu funkcji f w punkcie (x,y) będziemy rozumieć

$$\partial^2_M f(x,y) = \left(\frac{\partial^2 f}{\partial x_i \partial y_j}(x,y)\right)_{1 \leqslant i,j \leqslant n}.$$

Podkreślmy, że powyższa macierz zazwyczaj nie jest symetryczna. Macierz $\partial^2_M f(x,y)$ będziemy też utożsamiać z formą dwuliniową w tym sensie, że dla pary wektorów $\bar{x}, \bar{y} \in \mathbb{R}^n$ będziemy pisać

$$\Big(\partial^2_M f(x,y)\Big)(\bar{x},\bar{y}) = \bar{x}^t\Big(\partial^2_M f(x,y)\Big)\bar{y} = \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial y_j}(x,y)\bar{x}_i\bar{y}_j.$$

Dla macierzy A rozmiaru $n \times n$ będziemy rozważać jej normę operatorową oraz normę Hilberta-Schmidta:

$$\|A\|_{\mathrm{op}} = \sup\Big\{x^t A y \;\Big|\; x,y \in \mathbb{R}^n, |x| = |y| = 1\Big\},$$
$$\|A\|_{\mathrm{HS}} = \sqrt{\sum_{i,j=1}^n a_{ij}^2} = \sqrt{\mathrm{tr}(A^t A)}.$$

6.3. Nierówność koncentracyjna

Poniższe twierdzenie stanowi zapowiedziane uogólnienie nierówności (6.1):

Twierdzenie 6.1. *Niech $f\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ będzie klasy $\mathcal{C}^2$ oraz $\mathbb{E}|f(X,Y)| < \infty$. Jeśli dla pewnych stałych $a, b > 0$ spełnione jest*

$$\|\partial_M^2 f(x,y)\|_{\mathrm{op}} \leqslant a \quad \textit{oraz} \quad \|\partial_M^2 f(x,y)\|_{\mathrm{HS}} \leqslant b$$

dla dowolnych $(x,y) \in \mathbb{R}^n \times \mathbb{R}^n$, wówczas dla każdego $t \geqslant 0$,

$$\mathbb{P}\Big(|f(X,Y) - \Pi_1 f(X,Y)| \geqslant t\Big) \leqslant C \exp\Big(-c \min\Big(\frac{t}{a}, \frac{t^2}{b^2}\Big)\Big),$$

gdzie $C, c > 0$ są pewnymi stałymi numerycznymi.

Zanim przejdziemy do dowodu, skomentujmy optymalność powyższej nierówności. Weźmy $f(x,y) = x^t A y$, gdzie $A = (a_{ij})$ jest macierzą rozmiaru $n \times n$. Wówczas $\Pi_1 f(X,Y) \equiv 0$ p.n., zatem nierówność z Twierdzenia 6.1 daje oszacowanie górne na ogon zmiennej losowej $Z = X^t A Y = \sum_{i,j=1}^n a_{ij} g_i \bar{g}_j$, gdzie $(g_1, \ldots, g_n)^t = X$, $(\bar{g}_1, \ldots, \bar{g}_n)^t = Y$, czyli $g_i, \bar{g}_j$ są niezależnymi kopiami zmiennej losowej o rozkładzie $\mathcal{N}(0,1)$. Zmienna losowa Z jest „uniezależnionym" (ang. *decoupled*) chaosem gaussowskim rzędu 2. Oszacowanie ogonów (oraz momentów) takich zmiennych jest znane w literaturze (patrz np. [18]): dla każdego $t \geqslant 0$ oraz $p \geqslant 1$,

$$C_1 \exp\Big(-c_1 \min\Big(\frac{t}{a}, \frac{t^2}{b^2}\Big)\Big) \leqslant \mathbb{P}\Big(|X^t A Y| \geqslant t\Big) \leqslant C_2 \exp\Big(-c_2 \min\Big(\frac{t}{a}, \frac{t^2}{b^2}\Big)\Big), \tag{6.4}$$

$$c \max(pa, \sqrt{p} b) \leqslant \|X^t A Y\|_p \leqslant C \max(pa, \sqrt{p} b), \tag{6.5}$$

gdzie $a = \|A\|_{\mathrm{op}}$, $b = \|A\|_{\mathrm{HS}}$, zaś $C, c, C_1, c_1, C_2, c_2 > 0$ są pewnymi stałymi uniwersalnymi. Tym samym, Twierdzenie 6.1 można interpretować jako uogólnienie powyższych oszacowań z przypadku dwuliniowego na przypadek nieliniowy.

Ponieważ w dowodzie Twierdzenia 6.1 będziemy wykorzystywać oszacowanie momentów zmiennych losowych Z, dla kompletności podamy teraz prosty dowód górnych oszacowań z (6.4) i (6.5), korzystający z gaussowskiej nierówności koncentracyjnej dla funkcji lipschitzowskich (por. 4.7). Wykorzystując fakt, że $\mathcal{L}\Big(\sum b_i g_i\Big) = \mathcal{L}\Big((\sum b_i^2)^{1/2} g_1\Big)$, mamy

$$\begin{aligned} \mathbb{E}|Z|^p &= \mathbb{E}\,\mathbb{E}\Big[\Big|\sum_i \Big(\sum_j a_{ij}\bar{g}_j\Big) g_i\Big|^p \Big| \bar{X}\Big] \\ &= \mathbb{E}\,\mathbb{E}\Big[\Big|\Big(\sum_i \Big(\sum_j a_{ij}\bar{g}_j\Big)^2\Big)^{1/2} g_1\Big|^p \Big| \bar{X}\Big] = \|g_1\|_p^p \,\mathbb{E}\left|A\bar{X}\right|^p. \end{aligned} \tag{6.6}$$

Ponieważ funkcja $\mathbb{R}^n \ni x \mapsto |Ax|$ jest lipschitzowska ze stałą $a = \|A\|_{\mathrm{op}}$, to na mocy (4.7), dla każdego $t \geqslant 0$,

$$\mathbb{P}\left(\Big||A\bar{X}| - \mathbb{E}|A\bar{X}|\Big| \geqslant t\right) \leqslant 2e^{-t^2/(2a^2)}.$$

Wykorzystując wzór na całkowanie przez części dostajemy

$$
\begin{aligned}
\mathbb{E}\Big||A\bar{X}| - \mathbb{E}|A\bar{X}|\Big|^p &= \int_0^\infty pt^{p-1}\mathbb{P}\left(\Big||A\bar{X}| - \mathbb{E}|A\bar{X}|\Big| \geqslant t\right)\, dt \\
&\leqslant 2p \int_0^\infty t^{p-1} e^{-t^2/(2a^2)}\, dt \leqslant C^p p^{p/2} a^p,
\end{aligned}
$$

dla pewnej stałej numerycznej $C > 0$. Korzystając teraz z nierówności Minkowskiego dla normy L^p, otrzymujemy

$$
\|A\bar{X}\|_p \leqslant \mathbb{E}|A\bar{X}| + C\sqrt{p}a. \tag{6.7}
$$

Pozostaje oszacować wyrażenie $\mathbb{E}|A\bar{X}|$. Korzystając z monotoniczności norm L^p (czyli z nierówności Höldera) i przyjmując oznaczenie $(b_{ij})_{i,j\leqslant n} = A^tA$,

$$
\begin{aligned}
\Big(\mathbb{E}|A\bar{X}|\Big)^2 &\leqslant \mathbb{E}|A\bar{X}|^2 = \mathbb{E}\langle A\bar{X}, A\bar{X}\rangle = \mathbb{E}\langle \bar{X}, A^tA\bar{X}\rangle \\
&= \mathbb{E}\Big(\sum_{i,j=1}^n \bar{g}_i b_{ij} \bar{g}_j\Big) = \sum_{i=1}^n b_{ii}^2 = \mathrm{tr}(A^tA) = \|A\|_{\mathrm{HS}}^2 = b^2.
\end{aligned}
$$

Z powyższego oraz z (6.6) i (6.7) ostatecznie otrzymujemy

$$
\|Z\|_p \leqslant C(\sqrt{p}b + pa) \leqslant C' \max(pa, \sqrt{p}b).
$$

Górne oszacowanie z (6.4) otrzymamy stosując nierówność Czebyszewa (6.2) i zauważając, że $t = C\max(pa, \sqrt{p}b)$ wtedy i tylko wtedy, gdy $p = \min(t/(Ca), t^2/(Cb)^2)$.

Twierdzenie 6.1 jest prostym wnioskiem z następującej nierówności typu Sobolewa, będącej modyfikacją znanej nierówności Maureya-Pisiera [39, Theorem 2.2] (patrz także (4.2)):

Twierdzenie 6.2. *Niech $f\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ będzie klasy $\mathcal{C}^2$, $\mathbb{E}|f(X,Y)| < \infty$ oraz $\Psi\colon \mathbb{R} \to \mathbb{R}$ będzie funkcją wypukłą. Wówczas*

$$
\mathbb{E}\,\Psi\Big(f(X,Y) - \Pi_1 f(X,Y)\Big) \leqslant \mathbb{E}\,\Psi\Big((\pi/2)^2\left(\partial_M^2 f(X,Y)\right)(\bar{X},\bar{Y})\Big).
$$

Dowód. Wystarczy naśladować idee dowodu z pracy [39].

Dla $\theta_1, \theta_2 \in [0, \pi/2]$ wprowadźmy wektory gaussowskie

$$
\begin{aligned}
X(\theta_1) &= X\sin\theta_1 + \bar{X}\cos\theta_1, \\
Y(\theta_2) &= Y\sin\theta_2 + \bar{Y}\cos\theta_2.
\end{aligned}
$$

Przez $X'(\theta_1)$ będziemy rozumieć pochodną funkcji $\theta_1 \mapsto X(\theta_1)$, podobnie $Y'(\theta_2)$. Z faktu, że zmienne losowe gaussowskie nieskorelowane o łącznym rozkładzie gaussowskim są niezależne, wynika następująca kluczowa własność: dla dowolnych $\theta_1, \theta_2 \in [0, \pi/2]$,

$$
\begin{aligned}
\mathcal{L}\Big((X(\theta_1), X'(\theta_1))\Big) &= \mathcal{L}\Big((X, \bar{X})\Big), \\
\mathcal{L}\Big((Y(\theta_2), Y'(\theta_2))\Big) &= \mathcal{L}\Big((Y, \bar{Y})\Big).
\end{aligned} \tag{6.8}
$$

Założenie o gładkości f pozwala nam napisać

$$\begin{aligned}
&f(X,Y) - f(\bar{X},Y) - f(X,\bar{Y}) + f(\bar{X},\bar{Y}) \\
&= \int_0^{\pi/2}\int_0^{\pi/2} \frac{\partial^2 f(X(\theta_1),Y(\theta_2))}{\partial\theta_1\partial\theta_2}\, d\theta_1 d\theta_2 \\
&= \int_0^{\pi/2}\int_0^{\pi/2} \Big(\partial^2_M f(X(\theta_1),Y(\theta_2))\Big)(X'(\theta_1),Y'(\theta_2))\, d\theta_1 d\theta_2 \quad \text{p.n.}
\end{aligned}$$

Obkładając obie strony nierówności funkcją Ψ i stosując nierówność Jensena otrzymujemy

$$\begin{aligned}
&\Psi\Big(f(X,Y) - f(\bar{X},Y) - f(X,\bar{Y}) + f(\bar{X},\bar{Y})\Big) \\
&\leqslant \int_0^{\pi/2}\int_0^{\pi/2} \Psi\Big((\pi/2)^2\Big(\partial^2_M f(X(\theta_1),Y(\theta_2))\Big)(X'(\theta_1),Y'(\theta_2))\Big)(2/\pi)^2\, d\theta_1 d\theta_2 \quad \text{p.n.}
\end{aligned}$$

Przykładając wartość oczekiwaną, stosując do prawej strony twierdzenie Fubiniego oraz (6.8), dostajemy

$$\begin{aligned}
&\mathbb{E}\,\Psi\Big(f(X,Y) - f(\bar{X},Y) - f(X,\bar{Y}) + f(\bar{X},\bar{Y})\Big) \\
&\leqslant \int_0^{\pi/2}\int_0^{\pi/2} \mathbb{E}\,\Psi\Big((\pi/2)^2\left(\partial^2_M f(X,Y)\right)(\bar{X},\bar{Y})\Big)(2/\pi)^2\, d\theta_1 d\theta_2,
\end{aligned} \tag{6.9}$$

a tym samym całkowanie względem θ_1 i θ_2 można pominąć. Dowód kończymy stosując do lewej strony (6.9) nierówność Jensena, warunkowo względem $\sigma(X,Y)$:

$$\begin{aligned}
&\mathbb{E}\,\mathbb{E}\Big[\Psi\Big(f(X,Y) - f(\bar{X},Y) - f(X,\bar{Y}) + f(\bar{X},\bar{Y})\Big)\Big| X,Y\Big] \\
&\geqslant \mathbb{E}\,\Psi\Big(f(X,Y) - \Pi_1 f(X,Y)\Big).
\end{aligned}$$

□

Uwaga 6.3. W powyższym twierdzeniu nie było istotne to, że X i Y mają rozkład *kanoniczny* gaussowski, ani to, że ich rozkłady są na przestrzeni tego samego wymiaru, tj. na $\mathbb{R}^n$. Wystarczy bowiem założyć, że X i Y są niezależnymi wektorami gaussowskimi, na $\mathbb{R}^{n_1}$ i $\mathbb{R}^{n_2}$ (odpowiednio), o średniej zero. Gdy $\bar{X}$ i $\bar{Y}$ to niezależne kopie X i Y (odpowiednio), zależność (6.8) jest spełniona. W przypadku oryginalnej nierówności Maureya-Pisiera obserwacja ta pozwala bezpośrednio uzyskać koncentrację dla funkcji lipschitzowskich określonych na przestrzeni Banacha wyposażonej w dowolną scentrowaną miarę gaussowską.

Dowód Twierdzenia 6.1. Zastosujmy Twierdzenie 6.2 do $\Psi(u) = |u|^p$:

$$\|f(X,Y) - \Pi_1 f(X,Y)\|_p \leqslant \Big(\mathbb{E}\,\mathbb{E}\Big[\Big|(\pi/2)^2\left(\partial^2_M f(X,Y)\right)(\bar{X},\bar{Y})\Big|^p\Big|X,Y\Big]\Big)^{1/p}.$$

Warunkowa wartość oczekiwana występującą po prawej stronie zawiera chaos gaussowski rzędu 2, $Z = \bar{X}^t A\bar{Y}$, gdzie $A = \partial^2_M f(X,Y)$. Na mocy (6.5) otrzymujemy więc

$$\begin{aligned}
&\mathbb{E}\Big[\Big|(\pi/2)^2\left(\partial^2_M f(X,Y)\right)(\bar{X},\bar{Y})\Big|^p\Big|X,Y\Big] \\
&\leqslant C^p \max\Big(p\,\|\partial^2_M f(X,Y)\|_{\mathrm{op}}, \sqrt{p}\,\|\partial^2_M f(X,Y)\|_{\mathrm{HS}}\Big)^p \\
&\leqslant C^p \max\left(pa, \sqrt{p}b\right)^p \quad \text{p.n.,}
\end{aligned}$$

skąd

$$\|f(X,Y) - \Pi_1 f(X,Y)\|_p \leqslant C \max\left(pa, \sqrt{p}b\right).$$

Dowód kończymy w taki sam sposób, w jaki z (6.5) uzyskaliśmy górne oszacowanie z (6.4). $\square$

6.4. Uogólnienie na inne rozkłady

W oryginalnej nierówności Maureya-Pisiera,

$$\mathbb{E}\,\Psi\Big(f(X) - \mathbb{E}f(X)\Big) \leqslant \mathbb{E}\,\Psi\Big((\pi/2)\langle\nabla f(X), \bar{X}\rangle\Big),$$

można zastąpić wektor losowy X o rozkładzie gaussowskim, wektorem $W = S(X)$, gdzie S jest przekształceniem lipchitzowskim. Dokładniej, spełniona jest nierówność

$$\mathbb{E}\,\Psi\Big(f(W) - \mathbb{E}f(W)\Big) \leqslant \mathbb{E}\,\Psi\Big(L(\pi/2)\langle\nabla f(W), \bar{X}\rangle\Big),$$

gdzie L jest stałą Lipschitza przekształcenia S. Uwagę na ten temat można znaleźć w pracy [39, str. 181].

Pokażemy teraz, że w taki sam sposób można uogólnić także nierówność z Twierdzenia 6.2.

Wniosek 6.4. *Niech $S, T\colon \mathbb{R}^n \to \mathbb{R}^n$ będą przekształceniami lipschitzowskimi, klasy $\mathcal{C}^2$, ze stałymi Lipschitza odpowiednio L_1 i L_2. Wprowadźmy wektory losowe $W = S(X), Z = T(Y)$, tzn. W ma rozkład $\gamma_n \circ S^{-1}$, Z ma rozkład $\gamma_n \circ T^{-1}$ oraz wszystkie wektory losowe $W, Z, \bar{X}, \bar{Y}$ są niezależne ($\bar{X}$ i $\bar{Y}$ mają rozkład γ_n). Niech $f\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ będzie klasy $\mathcal{C}^2$ oraz spełnia $\mathbb{E}|f(W,Z)| < \infty$. Wówczas dla dowolnej funkcji wypukłej $\Psi\colon \mathbb{R} \to \mathbb{R}$,*

$$\mathbb{E}\,\Psi\Big(f(W,Z) - \Pi_1 f(W,Z)\Big) \leqslant \mathbb{E}\,\Psi\Big(L_1 L_2 (\pi/2)^2 \left(\partial^2_M f(W,Z)\right)(\bar{X}, \bar{Y})\Big). \tag{6.10}$$

Dowód. Zdefiniujmy funkcję $g\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ wzorem $g(x,y) = f(S(x), T(y))$. Oczywiście $\mathcal{L}(f(W,Z)) = \mathcal{L}(g(X,Y))$, skąd

$$\mathbb{E}\,\Psi\Big(f(W,Z) - \Pi_1 f(W,Z)\Big) = \mathbb{E}\,\Psi\Big(g(X,Y) - \Pi_1 g(X,Y)\Big).$$

Na mocy założeń funkcja g jest klasy $\mathcal{C}^2$, zatem stosując do niej Twierdzenie 6.2 otrzymujemy

$$\begin{aligned}
\mathbb{E}\,\Psi\Big(g(X,Y) - \Pi_1 g(X,Y)\Big) &\leqslant \mathbb{E}\,\Psi\Big((\pi/2)^2\Big(\partial^2_M g(X,Y)\Big)(\bar{X}, \bar{Y})\Big) \\
&= \mathbb{E}\,\Psi\Big((\pi/2)^2\Big(\partial^2_M f(S(X), T(Y))\Big)(\mathrm{D}S(X)\cdot\bar{X}, \mathrm{D}T(Y)\cdot\bar{Y})\Big) \\
&= \mathbb{E}\,\Psi\Big((\pi/2)^2\Big(\partial^2_M f(W,Z)\Big)(\mathrm{D}S(X)\cdot\bar{X}, \mathrm{D}T(Y)\cdot\bar{Y})\Big).
\end{aligned}$$

Na mocy niezależności zmiennych X, Y, W, Z od zmiennych $\bar{X}, \bar{Y}$, do ostatniego wyrażenia możemy zastosować twierdzenie Fubiniego otrzymując, że

$$\begin{aligned} &\mathbb{E}\,\Psi\Big(f(W,Z) - \Pi_1 f(W,Z)\Big) \\ &\qquad \leqslant \mathbb{E}_{X,Y}\mathbb{E}_{\bar{X},\bar{Y}}\,\Psi\Big((\pi/2)^2\Big(\partial_M^2 f(W,Z)\Big)(\mathrm{D}S(X)\cdot\bar{X}, \mathrm{D}T(Y)\cdot\bar{Y})\Big), \end{aligned} \tag{6.11}$$

gdzie $\mathbb{E}_{X,Y}$ oznacza całkowanie względem zmiennych X, Y, W, Z, zaś $\mathbb{E}_{\bar{X},\bar{Y}}$ — całkowanie względem zmiennych $\bar{X}, \bar{Y}$. Dla ustalonych X, Y (a więc i W, Z) oszacujemy wewnętrzną całkę iterowaną z ostatniego wyrażenia. Na mocy twierdzenia Minkowskiego, dla ustalonych $x, y \in \mathbb{R}^n$, endomorfizmy $\mathrm{D}S(x)/L_1$ oraz $\mathrm{D}T(y)/L_2$ możemy przedstawić jako skończone kombinacje wypukłe przekształceń ortogonalnych (te ostatnie są bowiem punktami ekstremalnymi kuli jednostkowej w przestrzeni endomorfizmów na $\mathbb{R}^n$ wyposażoną w normą operatorową):

$$\mathrm{D}S(x) = L_1 \sum_i \alpha_i U_i, \qquad \mathrm{D}T(y) = L_2 \sum_j \beta_j V_j,$$

gdzie $\alpha_i, \beta_j > 0$, $\sum_i \alpha_i = 1$, $\sum_j \beta_j = 1$, $U_i, V_j \in O(n)$. Wobec tego, przyjmując $w = S(x), z = T(y)$ i korzystając z dwuliniowości $(u, v) \mapsto \partial_M^2 f(w,z)(u,v)$, dla dowolnych wektorów $\bar{x}, \bar{y} \in \mathbb{R}^n$ dostajemy

$$\begin{aligned} &\Big(\partial_M^2 f(w,z)\Big)(\mathrm{D}S(x)\cdot\bar{x}, \mathrm{D}T(y)\cdot\bar{y}) \\ &= \Big(\partial_M^2 f(w,z)\Big)\left(L_1\left(\sum_i \alpha_i U_i\right)\cdot\bar{x}, L_2\left(\sum_j \beta_j V_j\right)\cdot\bar{y}\right) \\ &= L_1 L_2 \sum_{i,j} \alpha_i\beta_j \Big(\partial_M^2 f(w,z)\Big)(U_i\cdot\bar{x}, V_j\cdot\bar{y}). \end{aligned}$$

Przykładając funkcję wypukłą Ψ, zastępując $\bar{x}, \bar{y}$ przez zmienne losowe $\bar{X}, \bar{Y}$ i całkując względem tych ostatnich otrzymujemy

$$\begin{aligned} &\mathbb{E}_{\bar{X},\bar{Y}}\,\Psi\Big((\pi/2)^2\Big(\partial_M^2 f(w,z)\Big)(\mathrm{D}S(x)\cdot\bar{X}, \mathrm{D}T(y)\cdot\bar{Y})\Big) \\ &= \mathbb{E}_{\bar{X},\bar{Y}}\,\Psi\left(L_1L_2(\pi/2)^2\sum_{i,j}\alpha_i\beta_j\Big(\partial_M^2 f(w,z)\Big)(U_i\cdot\bar{X}, V_j\cdot\bar{Y})\right) \\ &\leqslant \sum_{i,j}\alpha_i\beta_j\mathbb{E}_{\bar{X},\bar{Y}}\,\Psi\left(L_1L_2(\pi/2)^2\Big(\partial_M^2 f(w,z)\Big)(U_i\cdot\bar{X}, V_j\cdot\bar{Y})\right), \end{aligned}$$

gdzie na końcu zastosowaliśmy nierówność Jensena. Ponieważ dla dowolnych ustalonych i, j, wektory $U_i\bar{X}$ i $V_j\bar{Y}$ są niezależne i mają rozkład kanoniczny gaussowski, to poprzednią nierówność możemy przepisać następująco:

$$\begin{aligned} &\mathbb{E}_{\bar{X},\bar{Y}}\,\Psi\Big((\pi/2)^2\Big(\partial_M^2 f(w,z)\Big)(\mathrm{D}S(x)\cdot\bar{X}, \mathrm{D}T(y)\cdot\bar{Y})\Big) \\ &\leqslant \mathbb{E}_{\bar{X},\bar{Y}}\,\Psi\left(L_1L_2(\pi/2)^2\Big(\partial_M^2 f(w,z)\Big)(\bar{X},\bar{Y})\right). \end{aligned}$$

Teraz wystarczy zastąpić w powyższej nierówności x, y, w, z przez X, Y, W, Z, scałkować względem $\mathbb{E}_{X,Y}$, by w połączeniu z (6.11) i z twierdzeniem Fubiniego otrzymać żądaną nierówność. $\square$

Z powyższego wniosku łatwo dostajemy uogólnienie nierówności koncentracyjnej z Twierdzenia 6.1, gdyż oszacowanie z góry prawej strony nierówności (6.10) tak jak poprzednio, dla $\Psi(u) = |u|^p$, sprowadza się do oszacowań momentów chaosów gaussowskich rzędu 2 (nierówność (6.5)): dla $W = S(X), Z = T(Y)$, gdzie S i T są lipschitzowskie ze stałymi L_1 i L_2, odpowiednio,

$$\mathbb{P}\Big(|f(W,Z) - \Pi_1 f(W,Z)| \geqslant t\Big) \leqslant C \exp\Big(-c \min\Big(\frac{t}{L_1 L_2 a}, \frac{t^2}{(L_1 L_2 b)^2}\Big)\Big).$$

6.5. Uogólnienie na d zmiennych wektorowych

Twierdzenie 6.2 łatwo jest uogólnić na przypadek d zmiennych wektorowych, zastępując m.in. parę wektorów losowych X, Y, przez ciąg $X_1, \ldots, X_d$. W celu ścisłego sformułowania tego uogólnienia, wprowadźmy jeszcze nieco notacji. Niech $E_j = \mathbb{R}^{n_j}$ dla $j = 1, \ldots, d$. Będziemy rozważać funkcje $f\colon E_1 \times \cdots \times E_d \to \mathbb{R}$, klasy $\mathcal{C}^d$ oraz działający na nich operator pochodnych mieszanych rzędu d, $\partial_M^d f$, który dla ustalonych $(x^{(1)}, \ldots, x^{(d)}) \in E_1 \times \cdots \times E_d$ jest formą d-liniową na $E_1 \times \cdots \times E_d$, zadaną wzorem

$$\Big(\partial_M^d f(x^{(1)}, \ldots, x^{(d)})\Big)(v^{(1)}, \ldots, v^{(d)}) := \Big(\mathrm{D}^d f(x^{(1)}, \ldots, x^{(d)})\Big)(\imath_1(v^{(1)}), \ldots, \imath_d(v^{(d)})),$$

gdzie $v^{(1)}, \ldots, v^{(d)}$ są dowolnymi wektorami z $E_1, \ldots, E_d$ (odpowiednio), zaś $\imath_j\colon E_j \to E_1 \times \cdots \times E_d$ jest kanonicznym włożeniem E_j w j-ty czynnik produktu kartezjańskiego $E_1 \times \cdots \times E_d$.

Niech $X_1, \bar{X}_1, \ldots, X_d, \bar{X}_d$ będą niezależnymi wektorami gaussowskimi, przy czym $X_j, \bar{X}_j$ (dla $j = 1, \ldots, d$) mają rozkład kanoniczny gaussowski na E_i. Ponadto, podobnie jak w Punkcie 6.2, powyższe wektory losowe określone są na przestrzeni probabilistycznej $(\Omega, \mathcal{F}, \mathbb{P})$ z wyróżnionymi σ-ciałami $\mathcal{F}_1, \ldots, \mathcal{F}_d$, które są niezależne i spełniają

$$\mathcal{F} = \sigma(\mathcal{F}_1, \ldots, \mathcal{F}_d).$$

Będziemy zakładać, że wektory losowe $X_j, \bar{X}_j$ są $\mathcal{F}_j$-mierzalne, dla $j = 1, \ldots, d$. Przyjmijmy jeszcze oznaczenie: dla $K \subseteq \{1, \ldots, d\}$,

$$\mathbb{E}_K(\cdot) = \mathbb{E}\Big[\,\cdot\, \Big|\sigma\big(\big\{\mathcal{F}_j \;\big|\; j \notin K\big\}\big)\Big].$$

Dla dowolnej całkowalnej zmiennej losowej $Z\colon (\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$ zdefiniujmy[1]

$$\Pi_{d-1} Z = \sum_{\emptyset \neq K \subseteq \{1,\ldots,d\}} (-1)^{\#K - 1} \mathbb{E}_K Z.$$

Naśladując dowód Twierdzenia 6.2 można pokazać

[1] Operator Π_{d-1} ma związek z tzw. *rozkładem Hoeffdinga*, znanym z teorii U-statystyk.

Twierdzenie 6.5. *Niech* $f\colon E_1 \times \dots \times E_d \to \mathbb{R}$ *będzie klasy* $\mathcal{C}^d$, $\mathbb{E}|f(X_1,\dots,X_d)| < \infty$ *oraz* $\Psi\colon \mathbb{R} \to \mathbb{R}$ *będzie funkcją wypukłą. Wówczas*

$$\mathbb{E}\,\Psi\Big(f(X_1,\dots,X_d) - \Pi_1 f(X_1,\dots,X_d)\Big)$$
$$\leqslant \mathbb{E}\,\Psi\Big((\pi/2)^d \left(\partial_M^d f(X_1,\dots,X_d)\right)(\bar{X}_1,\dots,\bar{X}_d)\Big).$$

W podobny sposób jak poprzednio, można z powyższego twierdzenia wyprowadzić odpowiednik Twierdzenia 6.1 dla d-zmiennych wektorowych. Wymaga to jednak znajomości oszacowań momentów chaosów gaussowskich rzędu d. Dokładne (dwustronne) tego typu oszacowania, dla dowolnego $d \geqslant 2$, zostały znalezione przez R. Latałę [26], co stanowi nader nietrywialne uogólnienie nierówności (6.4) i (6.5).

W związku z powyższym, przytoczymy teraz notację używaną w pracy [26]. Potrzebna będzie do określenia, co rozumiemy przez ograniczoność pochodnej mieszanej $\partial_M^d f$. Dla ustalonych $(x^{(1)},\dots,x^{(d)}) \in E_1 \times \dots \times E_d$, forma d-liniowa $A = \partial_M^d f(x^{(1)},\dots,x^{(d)})$ może być reprezentowana przez macierz $(a_{\boldsymbol{i}})$ z multi-indeksem rzędu d:

$$\boldsymbol{i} = (i_1,\dots,i_d) \in \Pi_{j=1}^d \{1,\dots,n_j\}\,.$$

Oznaczając przez $(e_i^{(j)})_{i=1,\dots,n_j}$ standardową bazę w E_j ($j = 1,\dots,d$), będziemy mieli

$$a_{\boldsymbol{i}} = A(e_{i_1}^{(1)},\dots,e_{i_d}^{(d)}).$$

Ponadto wprowadźmy obcięcie multi-indeksu $\boldsymbol{i} = (i_j)_{j\in\{1,\dots,d\}}$ do podzbioru $I \subseteq \{1,\dots,d\}$:

$$i_I := (i_j)_{j\in I}.$$

Dalej, dla dowolnego podziału $(I_1,\dots,I_k)$ zbioru $\{1,\dots,d\}$ na k niepustych podzbiorów ($k \leqslant d$) definiujemy normę $\|A\|_{I_1,\dots,I_k}$ jako

$$\|A\|_{I_1,\dots,I_k} = \sup\left\{\sum_{\boldsymbol{i}} a_{\boldsymbol{i}} x_{i_{I_1}}^{(1)} \cdots x_{i_{I_k}}^{(k)} \;\Big|\; \sum_{i_{I_1}} (x_{i_{I_1}}^{(1)})^2 \leqslant 1,\dots,\sum_{i_{I_k}} (x_{i_{I_k}}^{(k)})^2 \leqslant 1\right\}.$$

(W przypadku $d = 2$, $\boldsymbol{i} = (i_1, i_2)$, mamy więc do czynienia z dwiema normami: $\|A\|_{\{1\}\{2\}}$, która jest normą operatorową macierzy (a_{i_1,i_2}), oraz normą $\|A\|_{\{1,2\}}$, która jest normą Hilberta-Schmidta). Wreszcie, przez $\|A\|_{(k)}$ oznaczmy maksimum z liczb $\|A\|_{I_1,\dots,I_k}$, gdzie $(I_1,\dots,I_k)$ przebiega wszystkie podziały zbioru $\{1,\dots,d\}$ na k (niepustych) podzbiorów.

Twierdzenie 6.6. *Niech* $f\colon E_1 \times \dots \times E_d \to \mathbb{R}$ *będzie klasy* $\mathcal{C}^d$ *oraz* $\mathbb{E}|f(X_1,\dots,X_d)| < \infty$*. Jeśli dla pewnych stałych* $a_1,\dots,a_d > 0$ *spełnione jest*

$$\|\partial_M^d f(x^{(1)},\dots,x^{(d)})\|_{(k)} \leqslant a_k$$

dla dowolnych $(x^{(1)},\dots,x^{(d)}) \in E_1 \times \dots \times E_d$ *i* $k = 1,\dots,d$*, wówczas dla każdego* $t \geqslant 0$,

$$\mathbb{P}\Big(|f(X_1,\dots,X_d) - \Pi_1 f(X_1,\dots,X_d)| \geqslant t\Big) \leqslant C(d)\exp\Big(-c(d)\min_{1\leqslant k\leqslant d}\Big((t/a_k)^{2/k}\Big)\Big),$$

gdzie $C(d), c(d) > 0$ *są pewnymi stałymi, zależącymi tylko od* d*.*

Bibliografia

[1] S. Aida i D. Stroock, *Moment estimates derived from Poincaré and logarithmic Sobolev inequalities*, Math. Res. Lett. 1 (1994), no. 1, 75–86.

[2] D. Bakry, *L'hypercontractivité et son utilisation en théorie des semigroupes*, w: Lectures on probability theory (Saint-Flour, 1992), Lecture Notes in Math. 1581, Springer, Berlin, 1994, 1–114.

[3] D. Bakry i M. Émery, *Diffusions hypercontractives*, w: Séminaire de probabilités, XIX, 1983/84, Lecture Notes in Math., 1123, Springer, Berlin, 1985, 177–206.

[4] W. Beckner, *Inequalities in Fourier analysis*, Ann. of Math. (2) 102 (1975), 159–182.

[5] W. Beckner, *A generalized Poincaré inequality for Gaussian measures*, Proc. Amer. Math. Soc. 105 (1989), 397–400.

[6] S.G. Bobkov, *An isoperimetric inequality on the discrete cube, and an elementary proof of the isoperimetric inequality in Gauss space*, Ann. Probab. 25 (1997), 206–214.

[7] S.G. Bobkov, M. Ledoux, *Poincaré's inequalities and Talagrand's concentration phenomenon for the exponential distribution*, Probab. Theory Related Fields 107 (1997), 383–400.

[8] A. Bonami, *Étude des coefficients de Fourier des fonctions de $L^p(G)$*, Ann. Inst. Fourier (Grenoble) 20 (1970), no. 2, 335–402.

[9] C. Borell, *The Brunn-Minkowski inequality in Gauss space*, Invent. Math., 30 (1975), 207—216.

[10] C. Borell, *On the integrability of Banach space valued Walsh polynomials*, w: Séminaire de Probabilités XIII, Lecture Notes in Math. 721, Springer, 1979, 1–3.

[11] C. Borell, *On polynomial chaos and integrability*, Probab. Math. Statist. 3 (1984), 191–203.

[12] S. Boucheron, O. Bousquet, G. Lugosi i P. Massart, *Moment inequalities for functions of independent random variables*, Ann. Probab. 33 (2005), 514–560.

[13] E.B. Davies i B. Simon, *Ultracontractivity and the heat kernel for Schrödinger operators and Dirichlet Laplacians*, J. Funct. Anal. 59 (1984), 335–395.

[14] P. Diaconis i L. Saloff-Coste, *Logarithmic Sobolev inequalities for finite Markov chains*, Ann. Appl. Prob. 6 (1996), 695–750.

[15] I. Dinur, E. Friedgut i O. Regev, *Independent sets in graph powers are almost contained in juntas* (2006), przyjęta do Geom. Funct. Anal.

[16] L. Gross, *Logarithmic Sobolev inequalities*, Amer. J. of Math. 97 (1975), 1061–1083.

[17] A. Guionnet i B. Zegarlinski, *Lectures on logarithmic Sobolev inequalities*, w: Séminaire de Probabilités, XXXVI, Lecture Notes in Math. 1801, Springer, Berlin, 2003, 1–134.

[18] D.L. Hanson i F.T. Wright, *A bound on tail probabilities for quadratic forms in independent random variables*, Ann. Math. Statist. 42 (1971), 1079–1083.

[19] R.C. James, *Orthogonality and linear functionals in normed linear spaces*, Trans. Amer. Math. Soc. 61 (1947), 265–292.

[20] J. Kahn, G. Kalai, N. Linial, *The influence of variables on Boolean functions*, w: Proc 29th IEEE Symp. on Foundations of Computer Science, 1988, 68–80.

[21] W. Krakowiak i J. Szulga, *Hypercontraction principle and random multilinear forms*, Probab. Theory Related Fields 77 (1988), 325–342.

[22] S. Kwapień i J. Szulga, *Hypercontraction methods in moment inequalities for series of independent random variables in normed spaces*, Ann. Probab. 19 (1991), 369–379.

[23] S. Kwapień i W. A. Woyczyński, *Random Series and Stochastic Integrals: Single and Multiple*, Probability and its Applications, Birkhäuser, New York, 1992.

[24] R. Latała, *Hiperkontraktywne zmienne losowe albo kilka uwag o momentach*, Praca magisterska, Uniwersytet Warszawski, 1994.

[25] R. Latała, *Estimation of moments of sums of independent real random variables*, Ann. Probab. 25 (1997), 1502–1513.

[26] R. Latała, *Estimates of moments and tails of Gaussian chaoses*, Ann. Probab. 34 (2006), 2315–2331.

[27] R. Latała i K. Oleszkiewicz, *Between Sobolev and Poincaré*, w: Geometric Aspects of Functional Analysis, Lecture Notes in Math. 1745, Springer, Berlin, 2000, 147–168.

[28] M. Ledoux, *On Talagrand's deviation inequalities for product measures*, ESAIM Probab. Statist. 1 (1997), 63–87.

[29] M. Ledoux, *The concentration of measure phenomenon*, Mathematical Surveys and Monographs, 89, American Mathematical Society, Providence, RI, 2001.

[30] M. Ledoux i M. Talagrand, *Probability in Banach spaces. Isoperimetry and Processes*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3), 23, Springer-Verlag, Berlin, 1991.

[31] J. Lindenstrauss i L. Tzafriri, *Classical Banach spaces II. Function spaces*, Springer-Verlag, 1979.

[32] B. Maurey, *Some deviation inequalities*, Geom. Funct. Anal. 1 (1991), 188–197.

[33] S. Mazur i W. Orlicz, *Grundlegende Eigenschaften der polynomischen Operationen. Erste Mitteilung*, Studia Math. 5 (1934), 50–68.

[34] E. Nelson, *A quartic interaction in two dimensions*, w: Mathematical Theory of Elementary Particles (R. Goodman i I. Segal, edytorzy.), MIT Press, Cambridge, MA, 1966, 69–73.

[35] E. Nelson, *The free Markov field*, J. Funct. Anal. 12 (1973), 211–227.

[36] G. Nordlander, *The modulus of convexity in normed linear spaces*, Ark. Mat. 4 (1960), 15–17.

[37] K. Oleszkiewicz, *Uogólnienie metody hiperkontrakcji i nierówności Chinczyna-Kahane'a*, Praca magisterska, Uniwersytet Warszawski, 1994.

[38] K. Oleszkiewicz, *On a nonsymmetric version of the Khinchine-Kahane inequality*, w: Stochastic Inequalities and Applications, Progr. Probab. 56, Birkhäuser, Basel, 2003, 157–168.

[39] G. Pisier, *Probabilistic methods in the geometry of Banach spaces*, w: Probability and analysis (Verenna, 1985), Lecture Notes in Math. 1206, Springer, 1986, 167–241.

[40] O. Rothaus, *Logarithmic Sobolev inequalities and the spectrum of Sturm-Liouville operators*, J. Funct. Anal. 39 (1980), 42–56.

[41] I.E. Segal, *Construction of nonlinear local quantum processes*, Ann. of Math. (2) 92 (1970), 462–481.

[42] V.N. Sudakov i B.S. Tsirel'son, *Extremal properties of half-spaces for spherically invariant measures* (w j. rosyjskim), Zap. Nauchn. Sem. L.O.M.I. 41 (1974), 14-–24.

[43] J. Szulga, *A note on hypercontractivity of stable random variables*, Ann. Probab. 18 (1990), 1746–1758.

[44] M. Talagrand, *A new isoperimetric inequality and the concentration of measure phenomenon*, w: Geometric aspects of functional analysis (1989–90), Lecture Notes in Math., 1469, Springer, 1991, 94–124.

[45] M. Talagrand, *On Russo's approximate zero-one law*, Ann. Probab. 22 (1994), 1576–1587.

[46] P. Wolff, *Hypercontractivity of simple random variables*, Studia Math. 180 (2007), 219–236.

[47] P. Wolff, *Some remarks on functionals with the tensorization property*, Bull. Pol. Acad. Sci. Math. 55 (2007), 279–291.