

Praca doktorska

Operatory Markowa w ujęciu topologicznym

Bartosz Frej

Promotor:

dr hab. Tomasz Downarowicz,
prof. nadzw. PWr.

Instytut Matematyki
Wydział Podstawowych Problemów Techniki
Politechniki Wrocławskiej

10 czerwca 2003 roku

Spis treści

1	Wstęp	2
1.1	Podstawowe pojęcia	2
1.2	Zawartość pracy	7
1.3	Oznaczenia	9
2	Entropia operatorowa	11
2.1	Aksjomatyczna definicja miarowa	11
2.1.1	Aksjomaty	11
2.1.2	Asymptotyczna stabilność kratowa	14
2.1.3	Wyróżnione rozbicia	16
2.1.4	Dowód głównego twierdzenia	18
2.2	Jawna definicja miarowa	20
2.2.1	Definicja	20
2.2.2	Podaddytywność i istnienie granicy	23
2.2.3	Własności entropii	29
2.3	Entropia topologiczna	35
2.3.1	Nowa definicja entropii	35
2.3.2	Entropia Langevina-Walczaka	43
2.4	Relacja między entropią miarową a topologiczną	45
3	Rodziny operatorów Markowa	48
3.1	Układy przejścia. Trajektorie	48
3.2	Niezgodność komutowania	51
3.3	Pewien przykład trójki operatorów Markowa	55

Rozdział 1

Wstęp

1.1 Podstawowe pojęcia

W centrum zainteresowania teorii ergodycznej i dynamiki topologicznej znajdują się układy dynamiczne złożone z przestrzeni mierzalnej i transformacji mierzalnej, bądź z przestrzeni topologicznej i przekształcenia ciągłego. W dziedzinach tych szczególną wagę przykładają się do badania tych własności wspomnianych odwzorowań, które decydują o ewolucji układu zadanej przez kolejne iteracje przekształcenia. Jednym z pierwszych ważnych rezultatów dotyczących tego typu zagadnień było twierdzenie Poincarégo o powracaniu; warto też wspomnieć słynne twierdzenia ergodyczne Birkhoffa i von Neumanna, a w przypadku wielowymiarowym twierdzenia o wielopowracaniu Birkhoffa oraz Furstenberga-Szemeriediego. Inne istotne pytania, studiowane nie tylko w tych działach matematyki, dotyczą problemu izomorfizmu, czyli rozstrzygania, które układy możemy traktować jako nierozróżnialne z punktu widzenia danej teorii. W oba nurty badań dobrze wpisuje się teoria entropii, oferując, zarówno w przypadku miarowym, jak i topologicznym, wielkości mierzące stopień złożoności działania i będące jednocześnie niezmiennikami naturalnych izomorfizmów.

Uogólnienie klasycznych układów dynamicznych można otrzymać przez zastąpienie deterministycznej transformacji punktowej operacją, której wynik jest losowy. Stosowny formalizm matematyczny zapewnia następująca definicja.

Definicja 1.1.1 *Prawdopodobieństwem przejścia* na przestrzeni mierzalnej (X, Σ) nazywamy funkcję $P : X \times \Sigma \rightarrow [0, 1]$ spełniającą warunki:

- (i) dla każdego $x \in X$ funkcja zbioru $P(x, \cdot)$ jest miarą probabilistyczną,
- (ii) dla każdego $A \in \Sigma$ funkcja $P(\cdot, A)$ jest mierzalna.

Zauważmy, że mając dane prawdopodobieństwa przejścia P_1 i P_2 , możemy zdefiniować ich superpozycję, będącą także prawdopodobieństwem przejścia, kładąc:

$$P_1 P_2(x, A) = \int P_2(y, A) P_1(x, dy).$$

Prawdopodobieństwa przejścia związane są z pewnymi specjalnymi operatorami na przestrzeniach funkcyjnych, zwanymi zwykle operatorami stochastycznymi lub operatorami (bądź procesami) Markowa. Teoria takich operatorów, której podstawy opisał E. Hopf w swej klasycznej pracy [H], a którą rozwijali wcześniej m. in. J. L. Doob, S. Kakutani, K. Yosida, okazała się być bogatą i wieloaspektową dziedziną badań. Zajmowano się zarówno przenoszeniem na przypadek operatorowy znanych własności klasycznych układów dynamicznych (twierdzenia ergodyczne, istnienie miar niezmienniczych, twierdzenia o powracaniu), jak i studiowaniem własności zyskujących sens dopiero dla operatorów znacznie odbiegających od transformacji punktowych (prawo 0-2, aproksymacja operatora przez transformacje, układy ściśle niepunkto-we). Stosunkowo niedawno podjęto też próby definiowania pojęcia entropii dla operatorów Markowa.

Bogactwo i różnorodność tematyki zmusiły autorów do elastycznego posługiwania się definicjami operatora Markowa i operatora stochastycznego (por. definicje w [F], [H], [I1], [I3], [N], [S]). W niniejszej pracy będziemy rozumieć je w następujący sposób. Niech (X, Σ) będzie przestrzenią mierzalną. Załóżmy, że μ jest miarą probabilistyczną na (X, Σ) . Przez $L^p(\mu)$, $1 \leq p < \infty$, będziemy oznaczać przestrzeń całkowalnych w p -tej potędze funkcji (właściwie klas funkcji) rzeczywistych na X , a przez $L^\infty(\mu)$ przestrzeń rzeczywistych funkcji ograniczonych μ prawie wszędzie.

Definicja 1.1.2 *Operatorem stochastycznym* będziemy nazywać operator liniowy $S : L^1(\mu) \rightarrow L^1(\mu)$ spełniający warunki

(S1) $Sf \geq 0$ dla $f \geq 0$ (tzn. S jest dodatni),

(S2) $S^*1 = 1$, gdzie S^* oznacza operator sprzężony do S , a 1 jest funkcją stale równą 1.

Operator $T = S^* : L^\infty(\mu) \rightarrow L^\infty(\mu)$ będziemy nazywać *operatorem *-stochastycznym*.

Nietrudno zauważyć, że tak operator stochastyczny, jak i *-stochastyczny są ograniczone; S jest w rzeczywistości kontrakcją w $L^1(\mu)$. Ponadto warunek (S2) można zastąpić następującym warunkiem nie odwołującym się do własności operatora sprzężonego:

(S2') $\int \mathbf{S}u \, d\mu = \int u \, d\mu$ dla każdej funkcji $u \in L^1(\mu)$.

Operator *-stochastyczny $\mathbf{T}$ można równoważnie zdefiniować jako operator liniowy $L^\infty(\mu) \rightarrow L^\infty(\mu)$ spełniający warunki:

(S1*) $\mathbf{T}f \geq 0$ dla $f \geq 0$,

(S2*) $\mathbf{T}\mathbf{1} = \mathbf{1}$,

(S3*) jeśli $f_n \in L^\infty(\mu)$, $f_n \searrow 0$ μ -p.w., to $\mathbf{T}f_n \searrow 0$ μ -p.w.

Utożsamiając przestrzeń $L^1(\mu)$ ze zbiorem wszystkich miar absolutnie ciągłych względem μ , możemy, mając dany operator *-stochastyczny $\mathbf{T}$, zdefiniować $\mathbf{S}$, dla którego $\mathbf{S}^* = \mathbf{T}$, wzorem $\mathbf{S}\nu(A) = \int \mathbf{T}\mathbf{1}_A \, d\nu$. Każdy operator *-stochastyczny $\mathbf{T}$ (a więc i stochastyczny) określa prawdopodobieństwo przejścia $P(x, A) = \mathbf{T}\mathbf{1}_A(x)$; warunek (S3*) odpowiada za przeliczalną addytywność miary $P(x, \cdot)$. Na odwrót, każde prawdopodobieństwo przejścia zadaje operator stochastyczny na $L^1(\mu)$ wzorem

$$\mathbf{S}\nu(A) = \int P(x, A) \nu(dx)$$

oraz operator *-stochastyczny na $L^\infty(\mu)$ wzorem

$$\mathbf{T}f(x) = \int f(y)P(x, dy). \quad (1.1)$$

Łatwo sprawdzić, że operator *-stochastyczny można przedłużyć do kontrakcji w $L^1(\mu)$ (jednakże to przedłużenie nie musi być operatorem stochastycznym). Co więcej, korzystając z reprezentacji operatora *-stochastycznego jako całki względem pewnego prawdopodobieństwa przejścia i z nierówności Jensena, można wykazać, że odpowiednie przedłużenia na przestrzenie $L^p(\mu)$ są też kontrakcjami w normach $L^p(\mu)$.

Definicja 1.1.3 *Operatorem podwójnie stochastycznym* nazywamy operator stochastyczny $\mathbf{S}$ spełniający dodatkowo warunek

(S3) $\mathbf{S}\mathbf{1} = \mathbf{1}$, lub równoważnie

(S3') $\int \mathbf{S}^*f \, d\mu = \int f \, d\mu$ dla każdej funkcji f .

Zauważmy, że operator sprzężony do podwójnie stochastycznego (określony na $L^\infty(\mu)$) można przedłużyć do operatora podwójnie stochastycznego na $L^1(\mu)$. Będziemy swobodnie używać tej nazwy w obu przypadkach — jest ona zresztą zgodna ze spotykanym w literaturze określeniem operatora podwójnie stochastycznego jako liniowego, dodatniego i zachowującego całkę

funkcji operatora na $L^1(\mu)$ ([B]) lub $L^\infty(\mu)$ ([K-SL]), dla którego funkcja stale równa 1 jest funkcją niezmienniczą.

W przypadku topologicznym, tzn. gdy X jest zwartą przestrzenią Hausdorffa, a Σ jej σ -ciałem zbiorów borelowskich, przyjmujemy następującą definicję.

Definicja 1.1.4 *Operatorem Markowa* będziemy nazywać operator liniowy $\mathbf{T} : C(X) \rightarrow C(X)$ spełniający warunki

$$(M1) \quad \mathbf{T}f \geq 0 \text{ dla } f \geq 0,$$

$$(M2) \quad \mathbf{T}1 = 1.$$

Zauważmy, że warunki te implikują ograniczoność operatora $\mathbf{T}$ (w normie supremum).

Przykład 1.1.5 Niech X będzie przestrzenią n -elementową z σ -ciałem wszystkich podzbiorów. Niech μ oznacza miarę (wektor probabilistyczny) na X . Elementy u przestrzeni $L^1(\mu)$ będziemy zapisywać w postaci wektorów wierszowych, zaś elementy f przestrzeni $L^\infty(\mu)$ (lub $C(X)$) jako wektory kolumnowe. Wówczas macierz kwadratowa $A = [a_{ij}]$ stopnia n , dla której $\sum_i a_{ij} = 1$, definiuje operator stochastyczny na $L^1(\mu)$ wzorem $S\nu = \nu A$ (gdzie ν jest miarą absolutnie ciąglą względem μ) oraz operator $*$ -stochastyczny na $L^\infty(\mu)$ (Markowa na $C(X)$) wzorem $\mathbf{T}f = Af$. Elementy macierzy są wartościami związanego z $\mathbf{S}$ i $\mathbf{T}$ prawdopodobieństwa przejścia. Takie macierze noszą zwykle w literaturze nazwę *macierzy stochastycznych* lub *macierzy Markowa*. Jeśli dodatkowo $\mu A = \mu$, to zdefiniowane powyżej $\mathbf{S}$ i $\mathbf{T}$ są operatorami podwójnie stochastycznymi. $\square$

Niech $T : X \rightarrow X$ będzie transformacją punktową na przestrzeni (X, Σ) .

Definicja 1.1.6 Operator liniowy (na dowolnej przestrzeni funkcyjnej) określony wzorem $\mathbf{T}f(x) = f(Tx)$ nazywamy *operatorem punktowo generowanym*.

Jeśli na X dana jest miara probabilistyczna μ , a T jest nieosobliwą transformacją mierzalną, to $\mathbf{T}$ jest operatorem $*$ -stochastycznym (odpowiedni operator na $L^1(\mu)$ nazywa się wtedy operatorem Frobeniusa-Perrona); jeśli zachowuje miarę μ , to $\mathbf{T}$ jest podwójnie stochastyczny. Gdy X jest zwartą przestrzenią Hausdorffa, a T jest ciągłą, $\mathbf{T}$ jest operatorem Markowa na $C(X)$.

Rozważmy teraz przypadek, gdy (X, Σ) jest zwartą przestrzenią Hausdorffa z σ -ciałem borelowskim, a $\mathbf{T}$ jest operatorem Markowa. Wiadomo,

że wtedy każdy taki operator jest również generowany przez prawdopodobieństwo przejścia $P(x, A) = \mathbf{T}^* \delta_x(A)$, gdzie δ_x oznacza miarę probabilistyczną skupioną w punkcie x , a operator $\mathbf{T}^*$, sprzężony do $\mathbf{T}$, działa na przestrzeni znakowanych miar Radona na X (dualnej do $C(X)$). Takie prawdopodobieństwo przejścia, będące ciągłą funkcją z przestrzeni X w zbiór $\mathcal{M}(X)$ miar probabilistycznych z topologią $*$ -słabą, nazywamy *fellerowskim*. Podzbiór $\mathcal{M}(X)$ złożony z wszystkich miar $\mathbf{T}^*$ -niezmienniczych (czyli spełniających $\int \mathbf{T}f d\mu = \int f d\mu$ dla każdej f ciągłej) jest niepusty, wypukły i $*$ -słabo zwarty. Dla każdej takiej miary μ , operator $\mathbf{T}$ można traktować jako operator podwójnie stochastyczny, przyjmując (1.1) jako definicję działania operatora na funkcjach nieciągłych.

Wprowadzimy jeszcze pojęcie układu przejścia, które będzie wykorzystywane w rozdziale 3. Założmy, że X jest zwarta i metryzowalna. Oznaczmy przez 2^X zbiór wszystkich niepustych domkniętych (zatem zwartych) podzbiorów przestrzeni X . Prawdopodobieństwo przejścia $P(x, \cdot)$ zapewnia pełny opis możliwej ewolucji procesu znajdującego się w punkcie x . W przypadku topologicznym możemy rozważać opis uboższy, podający jedynie najmniejszy (domknięty) zbiór możliwych przyszłych stanów. Ściśle mówiąc, możemy ograniczyć się do badania przekształcenia $\varphi_P : X \rightarrow 2^X$, przypisującego każdemu x nośnik topologiczny miary $P(x, \cdot)$. To przekształcenie jest dolnie półciągłe, tzn. dla każdego zbioru otwartego $U \subset X$ zbiór $\{x : \varphi_P(x) \cap U \neq \emptyset\}$ jest otwarty w X . Przyjmujemy następującą definicję.

Definicja 1.1.7 *Układem przejścia* nazywamy dolnie półciągłe przekształcenie $\varphi : X \rightarrow 2^X$.

Prostym przykładem układu przejścia jest odwzorowanie $\varphi(x) = \{Tx\}$, gdzie T jest ciągłym przekształceniem przestrzeni X . Powiemy, że układ przejścia φ jest *wpisany* w inny układ przejścia ψ ($\varphi \prec \psi$), gdy $\varphi(x) \subset \psi(x)$ dla każdego $x \in X$. Dla danego układu przejścia φ istnienie prawdopodobieństwa przejścia P (zatem i operatora Markowa), którego układ φ_P jest wpisany w φ , wynika z klasycznego twierdzenia Michaela o ciągłych selektorach. Co więcej, stosując jeden z subtelniejszych wariantów tego twierdzenia (tw. 3.1''' w [Mi]), daje się uzyskać istnienie prawdopodobieństwa przejścia, dla którego zachodzi równość $\varphi_P = \varphi$.

Zauważmy, że układ $\varphi : X \rightarrow 2^X$ możemy przedłużyć do przekształcenia $2^X \rightarrow 2^X$, dla każdego domkniętego zbioru F kładąc

$$\overline{\varphi}(F) = \overline{\bigcup_{x \in F} \varphi(x)}. \quad (1.2)$$

Pozwala to na określenie superpozycji układów przejścia wzorem $(\varphi \circ \psi)(x) = \overline{\varphi}(\psi(x))$. Nietrudno sprawdzić, że dla prawdopodobieństw przejścia P_1, P_2

zachodzi

$$\varphi_{P_1 P_2} = \varphi_{P_2} \circ \varphi_{P_1}. \quad (1.3)$$

1.2 Zawartość pracy

Rozprawa składa się z dwóch części o nieco odmiennej tematyce. Rozdział drugi traktuje o entropii operatorów na przestrzeniach funkcyjnych. W literaturze można znaleźć kilka prób określenia entropii operatorowej. W przypadku miarowym, najdokładniej zbadana jest chyba definicja entropii kwantowej zaproponowana przez R. Alickiego, J. Andriesa, M. Fannesa i P. Tuylsa w [A-A-F-T] (obszerną monografią na temat kwantowych układów dynamicznych jest [A-F]). Entropia ta bazuje na zdefiniowanej przez J. von Neumanna entropii macierzy gęstości opisującej stan układu kwantowego. Podobne pojęcie, ale w sformułowaniu dla operatorów podwójnie stochastycznych (zwanych tam operatorami Markowa) na przestrzeni funkcji całkowalnych, badał I. I. Makarow w pracy [M]. Zupełnie inne podejście zostało zaprezentowane przez E. Ghysa, R. Langevina i P. Walczaka w [G-L-W]. Ich entropię studiowali niedawno B. Kamiński i J. de Sam Lazaro ([K-SL]). Jedyna znana autorowi definicja entropii topologicznej pochodzi od R. Langevina i P. Walczaka ([L-W]). Wszystkie powyższe definicje uogólniają pojęcia klasyczne, t.j. dla operatorów punktowo generowanych zgadzają się z entropią Kołmogorowa-Sinaja lub entropią topologiczną Adlera-Konheima-McAndrew generującej transformacji punktowej. Jednakże w przypadku operatorów nie punktowych relacje między tymi wielkościami nie były dotychczas znane.

W podrozdziale 2.1 zaprezentowane zostało aksjomatyczne podejście do miarowej entropii operatorowej. W celu uzyskania pojęcia posiadającego pewne, jak się wydaje, nieodzowne własności, wyróżnionych zostało pięć kroków konstrukcyjnych i cztery aksjomaty entropii. Głównym rezultatem jest twierdzenie mówiące, że wszystkie pojęcia spełniające podane warunki zgadzają się na zbiorze wszystkich operatorów podwójnie stochastycznych. W szczególności dotyczy to wyżej wspomnianych znanych definicji entropii miarowej, co pociąga ich równość dla każdego operatora podwójnie stochastycznego. Autor pragnie nadmienić, że jego wkład w wyniki dotyczące aksjomatycznej teorii entropii jest niewielki. Ponieważ jednak rezultaty te są istotne dla całości przedstawionej teorii, zamieszczenie ich w niniejszej rozprawie wydawało się niezbędne.

W podrozdziale 2.2 wprowadzono nową jawną definicję entropii miarowej, również spełniającą aksjomaty entropii. Definicja ta posiada pewne pożądane własności, których nie stwierdzono w przypadku innych znanych pojęć. Najważniejszą cechą wydaje się być fakt istnienia granicy w jednym z typowych

kroków konstrukcyjnych. Dowód tej własności opiera się na zastosowaniu twierdzenia o reprezentacji całkowitej pewnych operatorów funkcyjnych, udowodnionego przez A. Iwanika w pracy [I2]. W omawianym fragmencie można też znaleźć kilka elementarnych własności nowej entropii, między innymi fakt, że entropia operatorowa jest niezmiennikiem izomorfizmu oraz wzór na entropię potęgi operatora. Uzasadnia się również stwierdzenie, że entropia operatorowa stanowi istotne uogólnienie entropii transformacji punktowej.

Dla przypadku topologicznego podano, w podrozdziale 2.3, trzy naturalne metody określania entropii operatora Markowa, prowadzące, jak się okazuje, do tej samej wielkości. Zgodnie z oczekiwaniami, nowe pojęcie jest uogólnieniem klasycznej entropii topologicznej. Podano przykłady dyskwalifikujące dwie inne naturalne próby zdefiniowania entropii, jako dające zawyżone wyniki. Zamieszczono też analizę entropii wprowadzonej w [L-W], choć pytanie o relacje między tym pojęciem a nową definicją pozostaje otwarte.

W podrozdziale 2.4 sformułowano i udowodniono (przy wykorzystaniu definicji entropii topologicznej z niniejszej pracy) odpowiednik słynnego twierdzenia Goodwyna, stanowiącego część zasady wariacyjnej (patrz np. [D-G-S], tw. (18.4) lub [W], tw. 8.6): entropia topologiczna operatora Markowa na $C(X)$ ogranicza z góry jego entropię miarową (jako operatora podwójnie stochastycznego) względem każdej niezmienniczej miary probabilistycznej Radona. Pozostaje pytanie, czy operatorowa entropia topologiczna jest realizowana jako kres górny entropii miarowych.

Wyniki rozdziału 2 zostały wysłane do renomowanego czasopisma, jako praca [D-F].

Rozdział 3 jest inspirowany zagadnieniem wielopowracania dla operatorów Markowa na $C(X)$, gdzie X jest zwartą przestrzenią metryczną. Przypomnijmy, że punkt x jest *powracający* względem transformacji ciągłej $T : X \rightarrow X$, gdy istnieje rosnący ciąg liczb naturalnych (n_k) taki, że dla każdego otoczenia U punktu x zachodzi $T^{n_k}x \in U$, o ile k jest dostatecznie duże. Punkt x jest *wielopowracający* dla rodziny transformacji $\{T_l\}_{l=1}^L$, gdy jest powracający względem każdej z tych transformacji pod wspólnym ciągiem (n_k) . Odpowiednie sformułowania definicji dla operatorów Markowa uzyskujemy zastępując warunek $T^{n_k}x \in U$ przez $T^{n_k}\mathbf{1}_U(x) > 0$ (lub równoważnie $P^{n_k}(x, U) > 0$, gdzie P^{n_k} jest prawdopodobieństwem przejścia odpowiadającym T^{n_k}). Twierdzenie Birkhoffa o wielopowracaniu udziela pozytywnej odpowiedzi na pytanie o istnienie punktów wielopowracających dla rodziny komutujących transformacji ciągłych. Blisko związane, ale sformułowane w przypadku miarowym, jest też słynne twierdzenie Furstenberga-Szemeriediego. Przypadkiem operatorowym zajmowano się w pracach [I3] i [D-I]. W pierwszej z nich udowodniono, między innymi, że dla każdej skończonej rodziny złożonej z iteracji pojedynczego operatora Markowa istnieją

punkty wielopowracające oraz że para komutujących operatorów Markowa posiada takie punkty, o ile spełnione jest pewne dodatkowe założenie gwarantujące domkniętość zbioru trajektorii; w drugiej, przeprowadzono dowód tego drugiego faktu bez użycia wspomnianego założenia. W [I3] używano jako narzędzia układów przejścia związanych z komutującymi operatorami Markowa. Dowód przedstawiony w [D-I] oparty jest na konstrukcji procesu stacjonarnego, generowanego przez komutujące operatory Markowa na wzór twierdzenia Ionescu-Tulcea'ów. Niemniej, jak pokazujemy w rozdziale 3, miara odpowiadająca temu procesowi w istocie jest skupiona na zbiorze trajektorii pewnych układów przejścia.

Podrozdział 3.1 rozpoczyna się dyskusją na temat układów przejścia. Szczególny nacisk położony jest na zagadnienia związane z istnieniem trajektorii; w szczególności dowodzi się wspomnianego wyżej faktu dotyczącego układów pochodzących od komutujących operatorów Markowa.

W podrozdziale 3.2 uzasadnia się, że komutowanie układów przejścia może nie implikować komutowania operatorów, od których dane układy pochodzą. W istocie pokazano nieco więcej: komutowanie może być wykluczone nawet dla operatorów Markowa związanych z układami wpisanymi. Wyniki tych podrozdziałów (w przypadku 3.2 tylko skrócona wersja przykładu) opublikowane zostały w pracy [Fr2].

Rozprawę kończy przykład trójki operatorów Markowa, których żadne iteracje nie dopuszczają trajektorii (podrozdział 3.3). Jest on kontynuacją przykładu, który podano w [D-I] celem wskazania trudności w przenoszeniu konstrukcji procesu stochastycznego związanego z trzema operatorami Markowa. Tamten przykład dopuszczał jednak trajektorie już dla kwadratów zaproponowanych operatorów. Podany w bieżącej pracy uzasadnia, że metoda użyta w dowodzie twierdzenia o istnieniu punktów wielopowracających dla pary operatorów Markowa w istocie nie może być przeniesiona na przypadki liczniejszych rodzin. Przykład ten znalazł się w pracy [Fr1].

1.3 Oznaczenia

Poza nielicznymi wyjątkami, w oznaczeniach będziemy się kierować następującymi regułami:

przestrzeń, na której działamy będzie oznaczana przez X , a σ -ciało przez Σ ,

funkcje rzeczywiste na rozważanej przestrzeni mierzalnej lub topologicznej będziemy oznaczać literami f lub g (z ewentualnymi indeksami),

skończone rodziny takich funkcji (podstawowe narzędzie w definicjach entropii) — $\mathcal{F}, \mathcal{G}, \mathcal{H}$

transformacje punktowe na X — S, T ,

operatory (a w podrozdziale 3.3 także związane z nimi prawdopodobieństwa przejścia) — $\mathbf{S}, \mathbf{T}, \mathbf{U}$,

układy przejścia — φ, ψ ,

rozbicia mierzalne przestrzeni X — $\mathcal{A}, \mathcal{B}$,

rozbicia na odcinki przedziału $[0, 1]$ — α ,

pokrycia otwarte przestrzeni topologicznej X — $\mathcal{U}, \mathcal{W}$,

pokrycia otwarte przedziału $[0, 1]$ — $\mathcal{V}$.

Oprócz tego, przyjmujemy standardowo:

$\mathbb{N}$ — zbiór liczb naturalnych,

$\mathbb{N}_0 = \mathbb{N} \cup \{0\}$,

$\mathbb{R}$ — zbiór liczb rzeczywistych,

$B(X)$ — przestrzeń rzeczywistych funkcji ograniczonych na X ,

$C(X)$ — przestrzeń rzeczywistych funkcji ciągłych na X ,

$L^\infty(\mu)$ lub $L^\infty(X, \mu)$ — przestrzeń rzeczywistych funkcji ograniczonych μ prawie wszędzie na X ,

$L^1(\mu)$ lub $L^1(X, \mu)$ — przestrzeń rzeczywistych funkcji całkowalnych na X względem miary μ .

* * *

Chciałbym wyrazić głęboką wdzięczność dwóm osobom. Pierwszą z nich jest promotor rozprawy, profesor Tomasz Downarowicz. Czas, który mi poświęcił i możliwość wspólnej pracy były dla mnie bezcenne. Drugą osobą jest moja żona, Paulina. Na jej pomoc mogłem liczyć i przy organizacji życia w domu, i przy redakcji pracy. Ich wiara w moje umiejętności miała dla mnie ogromne znaczenie.

Dziękuję.

Rozdział 2

Entropia operatorowa

2.1 Aksjomatyczna definicja miarowa

2.1.1 Aksjomaty

Niech (X, Σ) będzie przestrzenią mierzalną, a L pewnym zbiorem funkcji mierzalnych na X przyjmujących wartości z odcinka $[0, 1]$. Rozważmy operator $\mathbf{T} : L \rightarrow L$ (niekoniecznie liniowy) i miarę probabilistyczną na Σ . Wydaje się, że chcąc zdefiniować jakiekolwiek pojęcie entropii $h_\mu(\mathbf{T})$ operatora $\mathbf{T}$, należy postępować według ustalonego schematu:

- (1) określić zbiór $\mathbb{F}$ złożony z wybranych skończonych rodzin $\mathcal{F}$ funkcji należących do L ; zbiór ten powinien być $\mathbf{T}$ -niezmienniczy, tzn. dla $\mathcal{F} \in \mathbb{F}$ żądamy, by $\mathbf{T}\mathcal{F} \in \mathbb{F}$, gdzie $\mathbf{T}\mathcal{F}$ jest zbiorem $\{\mathbf{T}f : f \in \mathcal{F}\}$;
- (2) zdefiniować operację *połączenia* rodzin $\sqcup$ taką, że zbiór $\mathbb{F}$ jest zamknięty na działanie $\sqcup$; będziemy zakładać, że działanie to jest łączne i przemienne oraz, że liczność rodziny $\mathcal{F} \sqcup \mathcal{G}$ zależy jedynie od liczności rodzin $\mathcal{F}$ i $\mathcal{G}$;
- (3) zdefiniować entropię „statyczną” $H_\mu(\mathcal{F})$ rodziny $\mathcal{F} \in \mathbb{F}$ względem miary μ ;
- (4) położyć

$$h_\mu(\mathbf{T}, \mathcal{F}) = \limsup_{n \rightarrow \infty} \frac{1}{n} H_\mu(\mathcal{F}^n),$$

gdzie

$$\mathcal{F}^n = \bigsqcup_{k=0}^{n-1} \mathbf{T}^k \mathcal{F};$$

(5) przyjąć

$$h_\mu(\mathbf{T}) = \sup_{\mathcal{F} \in \mathbb{F}} h_\mu(\mathbf{T}, \mathcal{F}).$$

Przykładowo, w klasycznej definicji Kołmogorowa-Sinaja entropii transformacji mierzalnej zbiór $\mathbb{F}$ składa się z rodzin funkcji charakterystycznych odpowiadających mierzalnemu rozbiciom przestrzeni, a połączenie rodzin może być realizowane przez mnożenie przez siebie elementów rodzin lub, równoważnie, przez operację minimum. Niektóre ogólniejsze definicje ([A-F], [G-L-W]) przyjmują za $\mathbb{F}$ zbiór mierzalnych *rozkładów jedności*, tzn. rodzin $\mathcal{F} = \{f_i : 1 \leq i \leq r\}$ składających się z mierzalnych funkcji nieujemnych, spełniających warunek $\sum_i f_i = 1$ (lub $\sum_i f_i^2 = 1$ w przypadku [A-F]). Połączenie określa się wtedy, jako zbiór odpowiednich iloczynów funkcji. W podanej dalej nowej definicji entropii (patrz podrozdział 2.2), $\mathbb{F}$ jest zbiorem wszystkich skończonych rodzin funkcji mierzalnych o wartościach z odcinka $[0, 1]$, a połączenie rodzin jest ich sumą mnogościową.

Powyższej standardowej konstrukcji towarzyszy zwykle definicja entropii warunkowej, przydatnej w dowodach własności entropii. Na tak ogólnym poziomie rozważań przyjmujemy, że entropia warunkowa jest zadana wzorem

$$H_\mu(\mathcal{F}|\mathcal{G}) = H_\mu(\mathcal{F} \sqcup \mathcal{G}) - H_\mu(\mathcal{G}).$$

Każde pojęcie entropii powinno uogólniać klasyczną definicję Kołmogorowa-Sinaja, zachowując jej elementarne własności w przypadku ogólnym. Postulat ten prowadzi do sformułowania następujących warunków, które nazwiemy *aksjomatami entropii*.

(A) AKSJOMAT MONOTONICZNOŚCI I PODADDYTYWNOŚCI

Dla $\mathcal{F}$, $\mathcal{G}$ i $\mathcal{H}$ należących do $\mathbb{F}$ zachodzą nierówności

$$0 \leq H_\mu(\mathcal{F}|\mathcal{H}) \leq H_\mu(\mathcal{F} \sqcup \mathcal{G}|\mathcal{H}) \leq H_\mu(\mathcal{F}|\mathcal{H}) + H_\mu(\mathcal{G}|\mathcal{H}),$$

przy czym przyjmujemy konwencję, że $H_\mu(\mathcal{F}|\mathcal{H}) = H_\mu(\mathcal{F})$, gdy $\mathcal{H}$ jest rodziną pustą.

Z aksjomatu tego wynika nierówność

$$H_\mu\left(\bigcup_{k=1}^n \mathcal{F}_k \left| \bigcup_{k=1}^n \mathcal{G}_k \right.\right) \leq \sum_{k=1}^n H_\mu(\mathcal{F}_k|\mathcal{G}_k), \quad (2.1)$$

której standardowy dowód pomijamy (por. np. [D-S]). Ponadto, dla każdego $n \geq 1$,

$$h_\mu(\mathbf{T}, \mathbf{T}^n \mathcal{F}) = h_\mu(\mathbf{T}, \mathcal{F}). \quad (2.2)$$

(B) AKSJOMAT CIĄGŁOŚCI (W L^1)

Odległość (w L^1) dwóch rodzin $\mathcal{F} = \{f_i : 1 \leq i \leq r\}$ i $\mathcal{G} = \{g_i : 1 \leq i \leq r'\}$, $r' \leq r$, definiujemy wzorem

$$\text{dist}(\mathcal{F}, \mathcal{G}) = \min_{\pi} \left\{ \max_{1 \leq i \leq r} \int |f_i - g_{\pi(i)}| d\mu \right\},$$

gdzie π przebiega zbiór wszystkich permutacji zbioru $\{1, 2, \dots, r\}$, a $\mathcal{G}$ traktujemy jako rodzinę r -elementową przyjmując $g_i \equiv 0$ dla $r' < i \leq r$. W aksjomacie tym zakładamy, że dla dowolnych $r \in \mathbb{N}$ i $\varepsilon > 0$ istnieje $\delta_\varepsilon > 0$ taka, że jeśli liczności $\mathcal{F}$, $\mathcal{G}$ i $\mathcal{H}$ nie przekraczają r , a $\text{dist}(\mathcal{F}, \mathcal{G}) < \delta_\varepsilon$ to

$$\text{dist}(\mathcal{F} \sqcup \mathcal{H}, \mathcal{G} \sqcup \mathcal{H}) < \varepsilon$$

oraz

$$|H_\mu(\mathcal{F}) - H_\mu(\mathcal{G})| < \varepsilon. \quad (2.3)$$

W szczególności, z aksjomatu tego wynika, że dla każdego $r \in \mathbb{N}$ i $\varepsilon > 0$ istnieje $\delta > 0$, dla której $\text{dist}(\mathcal{F}, \mathcal{G}) < \delta$ pociąga

$$|H_\mu(\mathcal{H}|\mathcal{F}) - H_\mu(\mathcal{H}|\mathcal{G})| < \varepsilon, \quad (2.4)$$

o ile wszystkie występujące we wzorze rodziny mają nie więcej niż r elementów.

(C) AKSJOMAT O ROZBICIACH

Dla mierzalnego rozbitcia skończonego $\mathcal{A}$ przestrzeni X oznaczmy przez $\mathbf{1}_{\mathcal{A}} = \{\mathbf{1}_A : A \in \mathcal{A}\}$ rodzinę funkcji charakterystycznych elementów $\mathcal{A}$. Aksjomat jest spełniony, gdy funkcja charakterystyczna każdego zbioru mierzalnego należy do L , dla każdego rozbitcia mierzalnego $\mathcal{A}$ przestrzeni X rodzina $\mathbf{1}_{\mathcal{A}}$ jest elementem $\mathbb{F}$, a połączenie i entropia H_μ dla takich rodzin zgadzają się z pojęciami klasycznymi:

$$\mathbf{1}_{\mathcal{A}} \sqcup \mathbf{1}_{\mathcal{B}} = \mathbf{1}_{\mathcal{A} \vee \mathcal{B}}, \quad H_\mu(\mathbf{1}_{\mathcal{A}}) = - \sum_{A \in \mathcal{A}} \mu(A) \log \mu(A).$$

Dowód faktu, że dane pojęcie entropii uogólnia entropię Kołmogorowa-Sinaja, musi opierać się na oszacowaniu z góry entropii rodziny funkcji $\mathcal{F}$ przez entropię pewnego rozbitcia. Takie rozbitcie otrzymuje się zazwyczaj rozważając przeciwobrazy (przez funkcje $f \in \mathcal{F}$) dostatecznie drobnego rozbitcia odcinka $[0, 1]$. Poniższy aksjomat, w połączeniu z aksjomatem (C), gwarantuje zgodność z entropią Kołmogorowa-Sinaja dla transformacji zachowujących miarę.

(D) AKSJOMAT O OSZACOWANIU

Dla dowolnych $r \in \mathbb{N}$ i $\varepsilon > 0$ istnieje $\gamma > 0$ taka, że dla każdej rodziny $\mathcal{F} = \{f_i : 1 \leq i \leq r\}$ i każdego rozbitcia α odcinka $[0, 1]$ na skończoną liczbę przedziałów o długości nie przekraczającej γ zachodzi

$$H_\mu(\mathcal{F} | 1_{V_i f_i^{-1}(\alpha)} \sqcup \bar{\alpha}) < \varepsilon,$$

gdzie $\bar{\alpha}$ jest rodziną funkcji zależną tylko od α i spełniającą warunek $\lim_n \frac{1}{n} H_\mu(\bigsqcup_{k=1}^n \bar{\alpha}) = 0$ (zazwyczaj $\bar{\alpha}$ jest rodziną pustą, co oznacza, że należy ją zaniedbać, lub składa się z funkcji stałych).

Oba wspomniane pojęcia entropii spełniają powyższe aksjomaty (dla entropii kwantowej, aksjomat (A) wynika z lematów 11.1 i 11.4 w [A-F], zaś aksjomat (D) pojawia się tamże w dowodzie twierdzenia 11.2; dla entropii [G-L-W] aksjomat (A) można wywnioskować z wyników pracy [K-SL], a aksjomat (D) został wykorzystany w dowodzie głównego twierdzenia w pracy [G-L-W]; prawdziwość pozostałych aksjomatów jest w obu przypadkach oczywista). Pozostałą część tego podrozdziału zajmuje dowód faktu, że wszystkie pojęcia entropii spełniające nasze aksjomaty zgadzają się nie tylko na zbiorze transformacji mierzalnych, ale też na zbiorze wszystkich operatorów podwójnie stochastycznych. Innymi słowy, dowodzimy następującego twierdzenia:

Twierdzenie 2.1.1 *Jeśli T jest operatorem podwójnie stochastycznym, to aksjomaty (A)–(D) (razem z krokami konstrukcyjnymi (1)–(5)) jednoznacznie określają wartość entropii $h_\mu(T)$.*

Dowód jest oparty na dwóch obserwacjach dotyczących operatorów podwójnie stochastycznych: kolejne iteracje operatora zachowują od pewnego miejsca operacje kratowe z zadaną z góry dokładnością i (w konsekwencji) przekształcają pewne specjalne funkcje charakterystyczne na funkcje bliskie charakterystycznym.

2.1.2 Asymptotyczna stabilność kratowa

Przez resztę podrozdziału 2.1 zakładamy, że T jest operatorem podwójnie stochastycznym.

Lemat 2.1.2 *Niech f i g będą ograniczonymi funkcjami mierzalnymi na X . Dla każdej liczby dodatniej δ istnieje $N \in \mathbb{N}$ takie, że dla dowolnych $k \in \mathbb{N}$, $l \geq N$ zachodzi nierówność*

$$\int \left| T^k(T^l f \vee T^l g) - (T^{k+l} f \vee T^{k+l} g) \right| d\mu < \delta$$

oraz

$$\int \left| \mathbf{T}^k(\mathbf{T}^l f \wedge \mathbf{T}^l g) - (\mathbf{T}^{k+l} f \wedge \mathbf{T}^{k+l} g) \right| d\mu < \delta.$$

Dowód Jest jasne, że $\mathbf{T}(f \vee g) \geq \mathbf{T}f$ i $\mathbf{T}(f \vee g) \geq \mathbf{T}g$, co pociąga

$$\begin{aligned} \mathbf{T}(f \vee g) &\geq \mathbf{T}f \vee \mathbf{T}g \\ (\text{i analogicznie } \mathbf{T}(f \wedge g) &\leq \mathbf{T}f \wedge \mathbf{T}g). \end{aligned} \quad (2.5)$$

Ponieważ $\mathbf{T}$ zachowuje całkę, dla każdego $n \in \mathbb{N}$ mamy

$$\int \mathbf{T}^n f \vee \mathbf{T}^n g d\mu = \int \mathbf{T}(\mathbf{T}^n f \vee \mathbf{T}^n g) d\mu \geq \int \mathbf{T}^{n+1} f \vee \mathbf{T}^{n+1} g d\mu,$$

więc ciąg całek $\int \mathbf{T}^n f \vee \mathbf{T}^n g d\mu$ jest zbieżny. Dla zadanej $\delta > 0$ można znaleźć $N \in \mathbb{N}$ tak duże, by dla każdego $l \geq N$ i każdego $k \in \mathbb{N}$

$$0 \leq \int \mathbf{T}^l f \vee \mathbf{T}^l g d\mu - \int \mathbf{T}^{k+l} f \vee \mathbf{T}^{k+l} g d\mu \leq \delta. \quad (2.6)$$

Ponieważ $\mathbf{T}^k$ zachowuje miarę, a wynikająca z (2.5) nierówność $\mathbf{T}^k(\mathbf{T}^l f(x) \vee \mathbf{T}^l g(x)) \leq \mathbf{T}^{k+l} f(x) \vee \mathbf{T}^{k+l} g(x)$ zachodzi dla wszystkich x , powyższa różnica jest równa odległości funkcji w $L^1(\mu)$. Dowód dla operacji minimum $\wedge$ jest analogiczny. $\square$

Definicja 2.1.3 Dla funkcji f i stałych $a < b$ oznaczmy $f_a^b = (f \vee a) \wedge b$. Mówimy, że f ma własność CZ(δ), gdy

$$\int \left| \mathbf{T}^n(f_a^b) - (\mathbf{T}^n f)_a^b \right| d\mu < \delta,$$

dla każdego $n \geq 0$ i każdej pary $a < b$.

Lemat 2.1.4 Dla każdej funkcji ograniczonej f i każdej liczby dodatniej δ istnieje $l \in \mathbb{N}$, dla którego $\mathbf{T}^l f$ ma własność CZ(δ).

Dowód Wybierzmy stałe a, b . Zgodnie z lematem 2.1.2 istnieje l' takie, że dla każdego n odległość w $L^1(\mu)$ między $\mathbf{T}^n((\mathbf{T}^{l'} f) \vee a)$ a $(\mathbf{T}^{n+l'} f) \vee a$ jest mniejsza niż $\frac{\delta}{9}$. Stosując ten sam lemat do funkcji $g = (\mathbf{T}^{l'} f) \vee a$ i stałej b , znajdziemy l'' , dla którego prawdziwa jest przy dowolnym n nierówność

$$\int \left| \mathbf{T}^n((\mathbf{T}^{l''} g) \wedge b) - (\mathbf{T}^{n+l''} g) \wedge b \right| d\mu < \frac{\delta}{9}.$$

Zarówno operator $\mathbf{T}$, jak i przekształcenie $g \mapsto g \wedge b$ są kontrakcjami w $L^1(\mu)$, zatem

$$\int \left| \mathbf{T}^n((\mathbf{T}^{l'+l''} f)_a^b) - (\mathbf{T}^{n+l'+l''} f)_a^b \right| d\mu < \frac{\delta}{3}.$$

Niech l będzie największą z liczb $l' + l''$ uzyskanych dla skończenie wielu par stałych a, b rozmieszczonych co $\frac{\delta}{3}$. Modyfikując każdą ze stałych a, b o liczbę mniejszą niż $\frac{\delta}{3}$, otrzymujemy funkcję f_a^b , która różni się od poprzedniej (w normie L^1) o mniej niż $\frac{\delta}{3}$. Zatem ostatnia nierówność jest spełniona dla każdego a i b , gdy zastąpimy $l' + l''$ przez l , a ułamek $\frac{\delta}{3}$ przez δ . $\square$

Następny wynik zostanie wykorzystany dopiero w podrozdziale 2.2, jednakże technika dowodu, taka jak w poprzednim lemacie, uzasadnia jego wcześniejsze przytoczenie.

Wielomianem kratowym (Boole'a) nazywamy dowolne wyrażenie θ zawierające skończenie wiele argumentów będących elementami ustalonej kraty (lub algebry Boole'a) oraz operacje kratowe $\vee$ i $\wedge$ (a także nawiasy) (podstawowe informacje o wielomianach kratowych można znaleźć np. w [R]). Będziemy posługiwać się takimi wielomianami na algebrze Boole'a funkcji z L . Przykładem wielomianu kratowego trzech funkcji jest wyrażenie $\theta(\cdot, \cdot, \cdot)$ określone wzorem $\theta(f, g, h) = (f \vee g) \wedge h$. Pisząc $\theta(\mathcal{F})$, będziemy zakładać, że elementy $\mathcal{F}$ zostały w pewien sposób ponumerowane.

Definicja 2.1.5 Mówimy, że rodzina $\mathcal{F} = \{f_i : 1 \leq i \leq r\}$ jest *kratowo δ -stabilna względem $\mathbf{T}$* , gdy dla każdego r -argumentowego wielomianu kratowego θ i każdego $n \in \mathbb{N}$ prawdziwa jest nierówność

$$\int |\mathbf{T}^n(\theta(\mathcal{F})) - \theta(\mathbf{T}^n \mathcal{F})| d\mu < \delta. \quad (2.7)$$

Stwierdzenie 2.1.6 Dla każdej rodziny $\mathcal{F}$ funkcji ograniczonych i każdej liczby $\delta > 0$ istnieje $l \geq 0$, dla którego rodzina $\mathbf{T}^l \mathcal{F}$ jest kratowo δ -stabilna.

Dowód Dla rodzin dwuelementowych powyższe stwierdzenie wynika z lematu 2.1.2. Dla większych rodzin (i bardziej skomplikowanych wielomianów kratowych) przeprowadzamy dowód indukcyjny. Załóżmy, że wielomian kratowy $\theta(\mathcal{F})$ jest postaci $\theta_1(\mathcal{F}_1) \vee \theta_2(\mathcal{F}_2)$ lub $\theta_1(\mathcal{F}_1) \wedge \theta_2(\mathcal{F}_2)$, gdzie θ_i i $\mathcal{F}_i$ są takie, że dla pewnego l' zachodzi wzór (2.7) przy θ_i , $\mathbf{T}^{l'} \mathcal{F}_i$ i $\frac{\delta}{6}$ w miejsce θ , $\mathcal{F}$, δ ($i = 1, 2$). Następnie dla $\frac{\delta}{3}$ oraz funkcji $\theta_1(\mathbf{T}^{l'} \mathcal{F}_1)$ i $\theta_2(\mathbf{T}^{l'} \mathcal{F}_2)$ znajdziemy l'' zgodnie z lematem 2.1.2. Wówczas suma $l = l' + l''$ zapewnia δ -stabilność wielomianu θ na rodzinie $\mathbf{T}^l(\mathcal{F})$.

Ponieważ każdy wielomian składa się ze skończenie wielu operacji $\vee$ i $\wedge$, a dla ustalonej liczności r istnieje tylko skończenie wiele różnych wielomianów kratowych, można znaleźć l spełniające tezę dowodzonego stwierdzenia. $\square$

2.1.3 Wyróżnione rozbicia

Lemat 2.1.7 Niech $\mathcal{F} \subset \mathbb{F}$ składa się z r funkcji o wartościach w $[0, 1]$, spełniających warunek CZ(δ^3). Wtedy można znaleźć skończenie wiele punktów ξ_j

rozmieszczonych na odcinku $[0, 1]$ w odległościach nie przekraczających $2r\delta$ i takich, że dla każdego $n \in \mathbb{N}$ i każdej funkcji $f \in \mathcal{F}$ zachodzi

$$\int \left| \mathbf{T}^n \mathbf{1}_{\{x: f(x) \geq \xi_j\}} - \mathbf{1}_{\{x: \mathbf{T}^n f(x) \geq \xi_j\}} \right| d\mu < 4\delta.$$

Dowód Pokryjmy przedział $[0, 1]$ odcinkami o długości $2\delta^2$ tak, by ich liczba nie przekraczała $\frac{1}{2\delta^2} + 1$. Dla ustalonej funkcji $f \in \mathcal{F}$, co najwyżej $\left(\frac{1}{\delta} - 1\right)$ przeciwobrazów tych odcinków ma miary większe niż δ . Zatem łączna ilość przeciwobrazów (przez dowolne $f \in \mathcal{F}$) o miarach przekraczających δ wynosi co najwyżej $r\left(\frac{1}{\delta} - 1\right)$. Każdy przedział długości $2r\delta$ musi więc zawierać punkt ξ_j będący środkiem pewnego odcinka, którego przeciwobraz przez dowolną funkcję $f \in \mathcal{F}$ jest miary nie większej niż δ , tzn.

$$\mu\{x : \xi_j - \delta^2 \leq f(x) < \xi_j + \delta^2\} \leq \delta.$$

Jest prawdą, że

$$F_f := \frac{f_{\xi_j}^{\xi_j + \delta^2} - \xi_j}{\delta^2} \leq \mathbf{1}_{\{x: f(x) \geq \xi_j\}} \leq \frac{f_{\xi_j - \delta^2}^{\xi_j} - (\xi_j - \delta^2)}{\delta^2} =: G_f.$$

Zauważmy, że

$$0 \leq G_f - F_f \leq \mathbf{1}_{\{x: \xi_j - \delta^2 \leq f(x) < \xi_j + \delta^2\}}.$$

Stąd $\int |G_f - F_f| d\mu \leq \delta$. Nakładając $\mathbf{T}^n$, otrzymujemy

$$\mathbf{T}^n F_f \leq \mathbf{T}^n (\mathbf{1}_{\{x: f(x) \geq \xi_j\}}) \leq \mathbf{T}^n G_f,$$

przy czym ponownie odległość (w L^1) między skrajnymi wyrazami jest mniejsza niż δ , gdyż $\mathbf{T}$ jest kontrakcją w $L^1(\mu)$. Z drugiej strony,

$$F_{\mathbf{T}^n f} = \frac{(\mathbf{T}^n f)_{\xi_j}^{\xi_j + \delta^2} - \xi_j}{\delta^2} \leq \mathbf{1}_{\{x: \mathbf{T}^n f(x) \geq \xi_j\}} \leq \frac{(\mathbf{T}^n f)_{\xi_j - \delta^2}^{\xi_j} - (\xi_j - \delta^2)}{\delta^2} = G_{\mathbf{T}^n f},$$

ale tym razem tracimy oszacowanie odległości. Korzystając jednak z własności CZ(δ^3), możemy stwierdzić, że odległość między $F_{\mathbf{T}^n f}$ a $\mathbf{T}^n F_f$ (podobnie między $G_{\mathbf{T}^n f}$ a $\mathbf{T}^n G_f$) jest mniejsza niż $\frac{\delta^3}{\delta^2} = \delta$. Zatem odległość (w L^1) między $\mathbf{1}_{\{x: \mathbf{T}^n f(x) \geq \xi_j\}}$ a $\mathbf{T}^n (\mathbf{1}_{\{x: f(x) \geq \xi_j\}})$ jest mniejsza niż 4δ . $\square$

Lemat 2.1.8 Niech rodzina $\mathcal{F} = \{f_i : 1 \leq i \leq r\} \subset L$ składa się z funkcji mających własność CZ(δ^3) i niech α będzie rozbiciem $[0, 1]$ na m odcinków: $A_0 = [0, \xi_1)$, $A_j = [\xi_j, \xi_{j+1})$ ($j = 1, \dots, m-2$), $A_{m-1} = [\xi_{m-1}, 1]$, przy czym punkty ξ_j spełniają tęzę lematu 2.1.7. Wówczas

$$\text{dist} \left(\mathbf{T}^n (\mathbf{1}_{\bigvee_i f_i^{-1}(\alpha)}), \mathbf{1}_{\bigvee_i \mathbf{T}^n f_i^{-1}(\alpha)} \right) < 8rm^r \delta,$$

dla każdego $n \geq 0$.

Dowód Rodzina $\mathbf{1}_{\bigvee_i f_i^{-1}(\alpha)}$ składa się z funkcji postaci

$$\mathbf{1}_{\bigcap_i \{x: f_i(x) \in A_{j_i}\}} = \inf_i \mathbf{1}_{\{x: f_i(x) \in A_{j_i}\}},$$

o łącznej sumie (przebiegającej wszystkie $i = 1, \dots, r$ oraz wszystkie ciągi $(j_i) \in \{0, 1, \dots, m-1\}^r$) będącej funkcją stale równą 1. Zgodnie z (2.5), dla każdego n zachodzi nierówność

$$\mathbf{T}^n \left(\mathbf{1}_{\bigcap_i \{x: f_i(x) \in A_{j_i}\}} \right) \leq \inf_i \mathbf{T}^n \left(\mathbf{1}_{\{x: f_i(x) \in A_{j_i}\}} \right), \quad (2.8)$$

a łączna suma funkcji występujących z lewej strony nierówności ponownie jest stale równa 1. Ponieważ

$$\mathbf{1}_{\{x: f_i(x) \in A_{j_i}\}} = \mathbf{1}_{\{x: f_i(x) \geq \xi_{j_i}\}} - \mathbf{1}_{\{x: f_i(x) \geq \xi_{j_i+1}\}},$$

z lematu 2.1.7 wynika, że odległość (w L^1) między funkcjami $\mathbf{T}^n(\mathbf{1}_{\{x: f_i(x) \in A_{j_i}\}})$ a $\mathbf{1}_{\{x: \mathbf{T}^n f_i(x) \in A_{j_i}\}}$ jest mniejsza niż 8δ . Zatem

$$\int \left| \inf_i \mathbf{T}^n(\mathbf{1}_{\{x: f_i(x) \in A_{j_i}\}}) - \inf_i \mathbf{1}_{\{x: \mathbf{T}^n f_i(x) \in A_{j_i}\}} \right| d\mu < 8r\delta. \quad (2.9)$$

Zauważmy, że ostatnią funkcję można zapisać jako $\mathbf{1}_{\bigcap_i \{x: \mathbf{T}^n f_i(x) \in A_{j_i}\}}$. Ponieważ wszystkie takie funkcje składają się na rodzinę $\mathbf{1}_{\bigvee_i \mathbf{T}^n f_i^{-1}(\alpha)}$, ich sumą jest funkcja stale równa 1. Korzystając z (2.8) i (2.9), otrzymujemy

$$\mathbf{T}^n \left(\mathbf{1}_{\bigcap_i \{x: f_i(x) \in A_{j_i}\}} \right) \leq \mathbf{1}_{\bigcap_i \{x: \mathbf{T}^n f_i(x) \in A_{j_i}\}} + g,$$

gdzie g jest pewną funkcją nieujemną o całce nie przekraczającej $8r\delta$. Ponieważ porównujemy rodziny o jednakowych sumach, największa z różnic

$$\left| \mathbf{1}_{\bigcap_i \{x: \mathbf{T}^n f_i(x) \in A_{j_i}\}} - \mathbf{T}^n \left(\mathbf{1}_{\bigcap_i \{x: f_i(x) \in A_{j_i}\}} \right) \right|$$

może mieć całkę co najwyżej $8r\delta$ razy większą od liczności rodziny funkcji $\mathbf{1}_{\bigcap_i \{x: \mathbf{T}^n f_i(x) \in A_{j_i}\}}$ (która z kolei jest mniejsza lub równa m^r). To kończy dowód. $\square$

2.1.4 Dowód głównego twierdzenia

Dowód twierdzenia 2.1.1 Rozważmy pojęcie entropii zdefiniowane w krokach (1)–(5) i spełniające aksjomaty (A)–(D). Ustalmy $\varepsilon > 0$ i wybierzmy rodzinę $\mathcal{F} = \{f_i : 0 \leq i \leq r\} \in \mathbb{F}$ spełniającą

$$h_\mu(\mathbf{T}) < h_\mu(\mathbf{T}, \mathcal{F}) + \varepsilon$$

(por. punkt (5)). Niech liczba γ spełnia aksjomat (D) o oszacowaniu dla r i ε . Wybierzmy m między $\frac{1}{\gamma}$ a $\frac{2}{\gamma}$. Niech δ będzie taką liczbą dodatnią, że nierówność (2.4) jest spełniona dla rodzin o liczności m^r i ε . Mając na uwadze własność (2.2), zastąpmy rodzinę $\mathcal{F}$ przez $\mathbf{T}^l \mathcal{F}$ (dalej oznaczaną już przez $\mathcal{F}$), gdzie l jest dobrane tak, by każda funkcja z nowej rodziny miała własność CZ($(\frac{\delta}{8rm^r})^3$) (patrz lemat 2.1.4). Liczba $\frac{\delta}{4m^r}$ ogranicza z góry odległości między punktami ξ_j z lematu 2.1.7 (w miejsce δ przyjmujemy $\frac{\delta}{8rm^r}$), a ponieważ jest ona znacznie mniejsza od γ (w szczególności, mniejsza niż $\frac{\gamma}{2}$), możemy spośród ξ_j wybrać $(m-1)$ punktów zadających rozbięcie α odcinka $[0, 1]$ na m przedziałów o długościach mniejszych niż γ . Dla każdego k , aksjomat (D), zapisany dla $\mathbf{T}^k \mathcal{F}$, daje oszacowanie

$$H_\mu(\mathbf{T}^k \mathcal{F} | \mathbf{1}_{V_i \mathbf{T}^k f_i^{-1}(\alpha)} \sqcup \bar{\alpha}) < \varepsilon.$$

Stosując wzór (2.1), otrzymujemy dla każdego n nierówność

$$H_\mu\left(\mathcal{F}^n \left| \bigsqcup_{k=0}^{n-1} (\mathbf{1}_{V_i \mathbf{T}^k f_i^{-1}(\alpha)} \sqcup \bar{\alpha}) \right.\right) < n\varepsilon.$$

Ponieważ $\lim_n \frac{1}{n} H_\mu(\bigsqcup_{k=0}^{n-1} \bar{\alpha}) = 0$, z aksjomatu (A) i kroku konstrukcyjnego (4) wynika nierówność

$$h_\mu(\mathbf{T}, \mathcal{F}) \leq \limsup_{n \rightarrow \infty} \frac{1}{n} H_\mu\left(\bigsqcup_{k=0}^{n-1} \mathbf{1}_{V_i \mathbf{T}^k f_i^{-1}(\alpha)}\right) + \varepsilon.$$

Oznaczmy powyższą granicę górną przez $h_\mu(\mathbf{T}, \mathcal{F}, \alpha)$ i zauważmy, że w rzeczywistości zależy ona jedynie od wartości entropii na pewnych rozbiciach, zatem jest jednoznacznie określona przez aksjomat (C). Z lematu 2.1.8 otrzymujemy

$$\text{dist}(\mathbf{1}_{V_i \mathbf{T}^k f_i^{-1}(\alpha)}, \mathbf{T}^k(\mathbf{1}_{V_i f_i^{-1}(\alpha)})) < \delta,$$

co, w połączeniu z (2.4), pociąga

$$H_\mu(\mathbf{1}_{V_i \mathbf{T}^k f_i^{-1}(\alpha)} | \mathbf{T}^k(\mathbf{1}_{V_i f_i^{-1}(\alpha)})) < H_\mu(\mathbf{1}_{V_i \mathbf{T}^k f_i^{-1}(\alpha)} | \mathbf{1}_{V_i \mathbf{T}^k f_i^{-1}(\alpha)}) + \varepsilon.$$

Zgodnie z aksjomatem (C) o rozbiciach, entropia warunkowa, występująca po prawej stronie nierówności, znika. Ponownie stosując (2.1) i krok (4) konstrukcji, wnioskujemy, że

$$h_\mu(\mathbf{T}, \mathcal{F}, \alpha) \leq h_\mu(\mathbf{T}, \mathbf{1}_{V_i f_i^{-1}(\alpha)}) + \varepsilon \leq h_\mu(\mathbf{T}) + \varepsilon.$$

Udowodniliśmy, że

$$h_\mu(\mathbf{T}, \mathcal{F}, \alpha) - \varepsilon \leq h_\mu(\mathbf{T}) \leq h_\mu(\mathbf{T}, \mathcal{F}, \alpha) + 2\varepsilon,$$

więc wartość entropii $h_\mu(\mathbf{T})$ jest jednoznacznie określona przez aksjomaty (A)–(D). $\square$

2.2 Jawna definicja miarowa

2.2.1 Definicja

Jedną z możliwych jawnych definicji entropii operatora podwójnie stochastycznego jest wielkość $h_\mu(\mathbf{T}, \mathcal{F}, \alpha)$ użyta w dowodzie twierdzenia 2.1.1. Jej wadą jest jednak obecność niewygodnych technicznych założeń dotyczących rozbicia α (specjalnie dobierane punkty podziału ξ_j). Opuszczenie tych założeń, jak pokazujemy w przykładzie 2.2.14, istotnie zmienia własności rozważanego pojęcia. Proponujemy zatem inną, naszym zdaniem naturalną, definicję entropii. Definicja ta jest poprawna dla dowolnych przekształceń $\mathbf{T} : L \rightarrow L$, gdzie L jest zbiorem funkcji mierzalnych przyjmujących wartości z odcinka $[0, 1]$. Postępując według metody opisanej w poprzednim podrozdziale, wykonujemy następujące kroki konstrukcyjne.

(1°) Za $\mathbb{F}$ przyjmujemy rodzinę wszystkich skończonych rodzin $\mathcal{F} \subset L$.

(2°) $\mathcal{F} \sqcup \mathcal{G} \stackrel{\text{def}}{=} \mathcal{F} \cup \mathcal{G}$

(3°) Dla $f \in L$ definiujemy

$$A_f = \{(x, t) \in X \times [0, 1] : t \leq f(x)\}$$

i oznaczamy przez $\mathcal{A}_f$ rozbitcie produktu $X \times [0, 1]$ składające się z A_f i jego dopełnienia A_f^c . Dla zadanej rodziny $\mathcal{F}$ kładziemy

$$\mathcal{A}_{\mathcal{F}} = \bigvee_{f \in \mathcal{F}} \mathcal{A}_f.$$

Entropię rodziny $\mathcal{F}$ definiujemy wzorem

$$H_\mu(\mathcal{F}) = H_{\hat{\mu}}(\mathcal{A}_{\mathcal{F}}) = - \sum_{A \in \mathcal{A}_{\mathcal{F}}} \hat{\mu}(A) \log \hat{\mu}(A),$$

gdzie $\hat{\mu}$ oznacza produkt miary μ z miarą Lebesgue'a na odcinku $[0, 1]$.

(4°) $h_\mu(\mathbf{T}, \mathcal{F}) = \limsup_{n \rightarrow \infty} \frac{1}{n} H_\mu(\mathcal{F}^n)$

(5°) $h_\mu(\mathbf{T}) = \sup_{\mathcal{F} \in \mathbb{F}} h_\mu(\mathbf{T}, \mathcal{F})$

Uwaga 2.2.1

(i) Dla dowolnych $\mathcal{F}, \mathcal{G} \in \mathbb{F}$ zachodzi równość $\mathcal{A}_{\mathcal{F} \cup \mathcal{G}} = \mathcal{A}_{\mathcal{F}} \vee \mathcal{A}_{\mathcal{G}}$.

- (ii) Dla każdej skończonej rodziny $\mathcal{F}$ można skonstruować skończoną (ale zazwyczaj liczniejszą) rodzinę $\Theta(\mathcal{F})$, której elementy tworzą rosnący ciąg funkcji i która generuje rozbitcie $\mathcal{A}_{\Theta(\mathcal{F})} = \mathcal{A}_{\mathcal{F}}$.

Dowód Niech r oznacza liczbę elementów rodziny $\mathcal{F}$. Na zbiorze $\{0, 1\}^r$ określamy relację porządku liniowego kładąc $\alpha = (\alpha_i) \preceq \beta = (\beta_i)$, gdy $\sum_i \alpha_i > \sum_i \beta_i$ lub gdy sumy są równe i pierwsza jedynka występuje w (α_i) na tej samej lub wcześniejszej pozycji niż w (β_i) . Niech

$$\theta_\beta(\mathcal{F}) = \sup_{\alpha \preceq \beta} \inf \{f_i : \alpha_i = 1\}.$$

Zbiór $\{\theta_\beta(\mathcal{F})\}_\beta$ z dołączonymi funkcjami stałymi 0 (jako najmniejszy element) i 1 (jako największy) jest szukaną rodziną $\Theta(\mathcal{F})$. $\square$

Zauważmy, że $\Theta(\mathcal{F})$ zależy od kolejności ponumerowania funkcji w $\mathcal{F}$, co jednak nie ma znaczenia dla dalszych rozważań. Jest jasne, że entropie „stacyjne” obu rodzin są równe, co więcej,

$$H_\mu(\mathcal{F}) = H_\mu(\Theta(\mathcal{F})) = - \sum_g \left(\int g d\mu \right) \log \left(\int g d\mu \right), \quad (2.10)$$

gdzie g przebiega wszystkie różnice następujących po sobie elementów $\Theta(\mathcal{F})$ (w kontekście innych, znanych definicji, ciekawym jest spostrzeżenie, że funkcje g tworzą rozkład jedności). Jednakże w definicji $h_\mu(\mathbf{T}, \mathcal{F})$ nie można zastąpić $\mathcal{F}$ przez $\Theta(\mathcal{F})$, gdyż $\Theta(\mathbf{T}(\mathcal{F})) \neq \mathbf{T}(\Theta(\mathcal{F}))$.

Łatwo można sprawdzić, że zdefiniowana przez nas entropia jest zgodna z definicją aksjomatyczną podaną w poprzednim podrozdziale. Aksjomaty (A)–(C) wynikają ze znanych własności entropii Kołmogorowa-Sinaja. Aby uzasadnić (D), przybliżmy każdą funkcję $f_i \in \mathcal{F}$ funkcją prostą s_i o wartościach w punktach podziału ξ_j rozbitcia α . Wobec punktu (3°) naszej definicji oraz wzoru (2.1), entropia $H_\mu(\mathcal{F}|\{s_i\})$ jest mała, zaś rozbitcie generowane przez rodzinę $\mathbf{1}_{\bigvee_i f_i^{-1}(\alpha)} \cup \bar{\alpha}$, gdzie $\bar{\alpha}$ jest rodziną funkcji stałych o wartościach ξ_j , jest drobniejsze od rozbitcia zadanego przez $\{s_i\}$.

Nasza definicja posiada też następującą własność, zgodną z intuicją opartą na interpretacji entropii w kontekście teorii informacji:

$$H_\mu(\mathcal{F}|\mathcal{F}) = 0.$$

Równość ta nie jest spełniona ani przez entropię [G-L-W] (patrz [K-SL], wniosek z Proposition 1), ani przez entropię kwantową (wystarczy rozważyć przestrzeń dwupunktową z miarą równo rozłożoną $\{\frac{1}{2}, \frac{1}{2}\}$ oraz rozkład jedności w sensie [A-F], składający się z funkcji $(\frac{1}{\sqrt{2}}, 0)$ i $(\frac{1}{\sqrt{2}}, 1)$). Ze wzoru (2.4)

oraz z powyższej równości wynika istnienie dla każdego $r \in \mathbb{N}$ i $\varepsilon > 0$ pewnej liczby $\delta > 0$, dla której

$$\text{dist}(\mathcal{F}, \mathcal{G}) < \delta \implies H_\mu(\mathcal{F}|\mathcal{G}) < \varepsilon, \quad (2.11)$$

o ile liczności rodzin są równe r . Uwzględniając (2.1), otrzymujemy dla dowolnej kontrakcji $\mathbf{T}$ w $L^1(\mu)$ implikację

$$\text{dist}(\mathcal{F}, \mathcal{G}) < \delta \implies |h_\mu(\mathbf{T}, \mathcal{F}) - h_\mu(\mathbf{T}, \mathcal{G})| < \varepsilon. \quad (2.12)$$

Ponadto, dla każdej rodziny $\mathcal{F}_\phi$ funkcji niezmienniczych względem $\mathbf{T}$ zachodzi

$$h_\mu(\mathbf{T}, \mathcal{F}_\phi) = 0. \quad (2.13)$$

Przypomnijmy, że w przypadku zachowującej miarę transformacji punktowej entropia rozbitcia i entropia jego przeciwobrazu są równe, co oznacza, że dla operatora punktowo generowanego zachodzi $H_\mu(\mathbf{T}\mathcal{F}) = H_\mu(\mathcal{F})$. Niestety, równość ta nie jest prawdziwa dla dowolnych operatorów podwójnie stochastycznych. Co gorsza, jak pokazujemy w poniższym przykładzie, entropia rodziny $\mathcal{F}$ może wzrosnąć pod działaniem $\mathbf{T}$. Okazuje się jednak, że entropia stabilizuje się dla dalekich iteracji operatora podwójnie stochastycznego. Własność tę będziemy nazywać *asymptotyczną niezmienniczością* (patrz lemat 2.2.3).

Przykład 2.2.2 Niech $\mathbf{T}$ będzie operatorem podwójnie stochastycznym, określonym na odcinku $[0, 1]$ z miarą Lebesgue'a wzorem

$$\mathbf{T}f(x) = \frac{1}{2}(f(x) + f(1-x)).$$

Niech $\mathcal{F}$ składa się z funkcji charakterystycznych $\mathbf{1}_{[0, \frac{1}{4}]}$ i $\mathbf{1}_{(\frac{1}{4}, 1]}$. Wówczas $H_\mu(\mathcal{F}) = 2 \log 2 - \frac{3}{4} \log 3 < \log 2$, zaś $H_\mu(\mathbf{T}\mathcal{F}) = \frac{3}{2} \log 2$. Zauważmy, że skuteczny niekiedy zabieg, polegający na wzbogaceniu rodziny $\mathcal{F}$ pewną liczbą dość gęsto rozmieszczonych funkcji stałych, w tym przypadku nie zmienia relacji między entropią $\mathcal{F}$ a entropią $\mathbf{T}\mathcal{F}$. $\square$

Lemat 2.2.3 (Asymptotyczna niezmienniczość entropii) *Niech $\mathbf{T}$ będzie operatorem podwójnie stochastycznym, a $\mathcal{F}$ pewną rodziną funkcji z $\mathbb{F}$. Dla każdego $\varepsilon > 0$ istnieje liczba $N \in \mathbb{N}$ taka, że dla każdego naturalnego k zachodzi nierówność*

$$|H_\mu(\mathbf{T}^{k+N}\mathcal{F}) - H_\mu(\mathbf{T}^N\mathcal{F})| < \varepsilon.$$

Dowód Skorzystamy z operacji $\Theta(\cdot)$ opisanej w uwadze 2.2.1. Ustalmy rodzinę $\mathcal{F}$ o liczności r i liczbę dodatnią ε . Dobierzmy $\delta > 0$ zgodnie z aksjomatem ciągłości (B). Na mocy stwierdzenia 2.1.6, można znaleźć liczbę naturalną N , dla której rodzina $\mathbf{T}^N \mathcal{F}$ jest kratowo δ -stabilna względem $\mathbf{T}$. Wtedy $\text{dist}(\Theta(\mathbf{T}^{N+k} \mathcal{F}), \mathbf{T}^k \Theta(\mathbf{T}^N \mathcal{F})) < \delta$ dla każdej liczby naturalnej k , więc (2.3) implikuje

$$|H_\mu(\Theta(\mathbf{T}^{k+N} \mathcal{F})) - H_\mu(\mathbf{T}^k \Theta(\mathbf{T}^N \mathcal{F}))| < \varepsilon.$$

Wobec (2.10) i faktu, że $\mathbf{T}^k$ zachowuje całkę, mamy $H_\mu(\mathbf{T}^k \Theta(\mathbf{T}^N \mathcal{F})) = H_\mu(\Theta(\mathbf{T}^N \mathcal{F}))$, a stąd wynika teza. $\square$

2.2.2 Podaddytywność i istnienie granicy

Najważniejszym wynikiem bieżącej części jest twierdzenie 2.2.6, mówiące, że granicę górną w punkcie (4°) naszej definicji można zastąpić granicą. Warto nadmienić, że analogiczne rezultaty dla pozostałych entropii operatorowych nie są znane. Podstawową trudność w dowodzie stanowi brak podaddytywności ciągu $H_\mu(\mathcal{F}^n)$, która to własność jest używana w klasycznych dowodach dla transformacji punktowych. Co prawda, zgodnie z aksjomatem (A), zachodzi nierówność

$$H_\mu(\mathcal{F}^{n+m}) \leq H_\mu(\mathcal{F}^n) + H_\mu(\mathbf{T}^n \mathcal{F}^m),$$

ale, jak już wiemy, w drugim składniku sumy nie można zaniedbać $\mathbf{T}^n$. Nie wystarczy też powołać się na asymptotyczną niezmienniczość gwarantowaną przez lemat 2.2.3, gdyż zwiększając liczbę m , zmieniamy wyjściową rodzinę funkcji i, chcąc zachować to samo N , tracimy kontrolę nad ε .

Odpowiednim dla naszych potrzeb wzmocnieniem asymptotycznej stabilności jest lemat 2.2.4. Kluczowe w jego dowodzie jest twierdzenie o reprezentacji całkowitej pewnych operatorów funkcyjnych, udowodnione przez A. Iwanika w [I2]. Stwierdza ono, że w przypadku przestrzeni standardowej każdy operator $\mathbf{T}$, zadany wzorem (1.1) przez pewne prawdopodobieństwo przejścia na zbiorze wszystkich ograniczonych funkcji mierzalnych, jest postaci

$$\mathbf{T}f(x) = \int_{\Omega} f(\phi_{\omega}(x)) d\lambda(\omega),$$

gdzie (Ω, λ) oznacza odcinek $[0, 1]$ z miarą Lebesgue'a, a $(\omega, x) \mapsto \phi_{\omega}(x)$ jest mierzalnym odwzorowaniem produktu $\Omega \times X$ w X . Rozważmy operator Φ na przestrzeni $B(\Omega \times X)$, punktowo generowany przez transformację $(\omega, x) \mapsto (\omega, \phi_{\omega}(x))$. Dla zadanej funkcji f , mierzalnej na X , oznaczmy przez $\bar{f}$ funkcję $(\omega, x) \mapsto f(x)$. Wówczas

$$\mathbf{T}f(x) = \int \Phi \bar{f}(\omega, x) d\lambda(\omega). \quad (2.14)$$

Mimo, że operator Φ nie jest podwójnie stochastyczny (bo nie zachowuje całki względem miary produktowej), z twierdzenia Fubiniego uzyskujemy równość

$$\iint \Phi \bar{f} d\lambda d\mu = \int \mathbf{T} f d\mu = \int f d\mu. \quad (2.15)$$

Ponieważ każda iteracja $\mathbf{T}$ jest zadana przez pewne prawdopodobieństwo przejścia, możemy zdefiniować punktowo generowany operator Φ_k odpowiadający $\mathbf{T}^k$. Zauważmy, że Φ_k nie musi być równy iteracji Φ^k .

Lemat 2.2.4 *Niech $\mathbf{T}$ będzie operatorem podwójnie stochastycznym, a $\mathcal{F}$ pewną rodziną funkcji z $\mathbb{F}$. Dla każdego $\varepsilon > 0$ istnieje liczba $N \in \mathbb{N}$ taka, że dla każdego $k, m \in \mathbb{N}$ zachodzi nierówność*

$$|\mathbf{H}_\mu(\mathbf{T}^{k+N}\mathcal{F}^m) - \mathbf{H}_\mu(\mathbf{T}^N\mathcal{F}^m)| < m\varepsilon.$$

Dowód Najpierw, przeprowadzimy dowód zakładając, że X jest przestrzenią standardową, a $\mathbf{T}$ spełnia założenia twierdzenia Iwanika.

Niech ν oznacza miarę produktową $\lambda \times \mu$. Udowodnimy, że

$$\begin{aligned} &\text{dla dowolnej funkcji } f \in L \text{ i dostatecznie dużej liczby natu-} \\ &\text{ralnej } l \text{ odległość (w } L^1(\nu)) \text{ między funkcjami } \overline{\mathbf{T}^{k+l}f} \text{ i } \Phi_k \overline{\mathbf{T}^l f} \\ &\text{jest mała dla każdego } k \geq 1. \end{aligned} \quad (2.16)$$

Ustalmy $M \in \mathbb{N}$ i $\delta > 0$. Ponieważ operator Φ_k jest punktowo generowany, dla każdej funkcji ograniczonej g zachodzi

$$\int |\Phi_k \bar{g}| d\nu = \int \Phi_k |\bar{g}| d\nu \stackrel{(2.15)}{=} \int |g| d\mu.$$

Ponownie korzystając z (2.15) dostajemy równość

$$\int (\Phi_k \bar{g})^+ d\nu = \int g^+ d\mu. \quad (2.17)$$

Ponieważ $(\mathbf{T}^l f)^+ = \mathbf{T}^l f \vee \mathbf{T}^l 0$, na mocy (2.6) możemy znaleźć liczbę naturalną N , taką że dla wszystkich $l \geq N$ i $k \geq 1$ zachodzi

$$0 < \int (\mathbf{T}^l f)^+ d\mu - \int (\mathbf{T}^{k+l} f)^+ d\mu < \delta.$$

Podstawiając $g = \mathbf{T}^l f$ we wzorze (2.17), otrzymujemy

$$0 \leq \int (\Phi_k \overline{\mathbf{T}^l f})^+ d\nu - \int (\mathbf{T}^{k+l} f)^+ d\mu < \delta. \quad (2.18)$$

Aby udowodnić (2.16), wystarczy sprawdzić, że miara zbioru wszystkich punktów (ω, x) , dla których $|\Phi_k \overline{\mathbf{T}^l f}(\omega, x) - \mathbf{T}^{k+l} f(x)| \geq \frac{2}{M}$, jest mała. Położmy

$$B(f) = \left\{ (\omega, x) : \Phi_k \overline{\mathbf{T}^l f}(\omega, x) > \frac{1}{M}, \quad \mathbf{T}^{k+l} f(x) \leq 0 \right\}.$$

Miarę $B(f)$ szacujemy w następujący sposób.

$$\begin{aligned} \frac{1}{M} \nu(B(f)) &\leq \int_{B(f)} \Phi_k \overline{\mathbf{T}^l f} d\nu \\ &\leq \int_{\{(\omega, x) : \mathbf{T}^{k+l} f(x) \leq 0\}} (\Phi_k \overline{\mathbf{T}^l f})^+ d\nu \\ &= \int (\Phi_k \overline{\mathbf{T}^l f})^+ d\nu - \int_{\{(\omega, x) : \mathbf{T}^{k+l} f(x) > 0\}} (\Phi_k \overline{\mathbf{T}^l f})^+ d\nu \\ &\leq \int (\Phi_k \overline{\mathbf{T}^l f})^+ d\nu - \int_{\{(\omega, x) : \mathbf{T}^{k+l} f(x) > 0\}} \Phi_k \overline{\mathbf{T}^l f} d\nu \\ &\stackrel{(2.14)}{=} \int (\Phi_k \overline{\mathbf{T}^l f})^+ d\nu - \int_{\{x : \mathbf{T}^{k+l} f(x) > 0\}} \mathbf{T}^{k+l} f d\mu \\ &\stackrel{(2.18)}{<} \delta. \end{aligned}$$

Zatem $\nu(B(f)) < M\delta$. Oczywiście, liczbę N można wybrać tak, by powyższe rachunki dało się przeprowadzić dla (skończenie wielu) funkcji postaci $f - \frac{i}{M}$, gdzie $i = 0, \dots, M-1$. Otrzymujemy więc oszacowanie:

$$\begin{aligned} \nu \left(\left\{ (\omega, x) : \Phi_k \overline{\mathbf{T}^l f}(\omega, x) - \mathbf{T}^{k+l} f(x) > \frac{2}{M} \right\} \right) \\ \leq \sum_{i=0}^{M-1} \nu \left(B(f - \frac{i}{M}) \right) < M^2 \delta. \end{aligned}$$

Podstawiając $(-f)$ w miejsce f mamy też

$$\nu \left(\left\{ (\omega, x) : \Phi_k \overline{\mathbf{T}^l f}(\omega, x) - \mathbf{T}^{k+l} f(x) < -\frac{2}{M} \right\} \right) < M^2 \delta,$$

a stąd

$$\int |\Phi_k \overline{\mathbf{T}^l f} - \overline{\mathbf{T}^{k+l} f}| d\mu < \frac{2}{M} + 2M^2 \delta.$$

Dobierając odpowiednio M i δ kończymy dowód tezy (2.16).

Ustalmy $\varepsilon > 0$. Wybierzmy N tak duże, by dla każdego $k \geq 1$ i $l \geq N$ odległość (w $L^1(\nu)$) między rodzinami $\overline{\mathbf{T}^{k+l} \mathcal{F}}$ a $\Phi_k \overline{\mathbf{T}^l \mathcal{F}}$ była tak mała, że, zgodnie z (2.11)

$$H_\nu(\overline{\mathbf{T}^{k+l} f} | \Phi_k \overline{\mathbf{T}^l f}) < \frac{\varepsilon}{r} \quad \text{oraz} \quad H_\nu(\Phi_k \overline{\mathbf{T}^l f} | \overline{\mathbf{T}^{k+l} f}) < \frac{\varepsilon}{r}. \quad (2.19)$$

Weźmy $m \geq 1$ i niech $\overline{T^N \mathcal{F}^m}$ oznacza rodzinę funkcji $\overline{T^l f}$, gdzie $f \in \mathcal{F}$, $l = N, \dots, N + m - 1$. Wtedy

$$\begin{aligned} H_\mu(\mathbf{T}^{k+N} \mathcal{F}^m) &= H_\nu(\overline{\mathbf{T}^{k+N} \mathcal{F}^m}) \\ &\leq H_\nu(\Phi_k \overline{\mathbf{T}^N \mathcal{F}^m}) + H_\nu(\overline{\mathbf{T}^{k+N} \mathcal{F}^m} | \Phi_k \overline{\mathbf{T}^N \mathcal{F}^m}). \end{aligned}$$

Na mocy uwagi 2.2.1(ii),

$$\begin{aligned} H_\nu(\Phi_k \overline{\mathbf{T}^N \mathcal{F}^m}) &= H_\nu(\Theta(\Phi_k \overline{\mathbf{T}^N \mathcal{F}^m})) = H_\nu(\Phi_k \overline{\Theta(\mathbf{T}^N \mathcal{F}^m)}) \\ &\stackrel{(2.15)}{=} H_\mu(\Theta(\mathbf{T}^N \mathcal{F}^m)) = H_\mu(\mathbf{T}^N \mathcal{F}^m), \end{aligned}$$

zatem

$$\begin{aligned} H_\mu(\mathbf{T}^{k+N} \mathcal{F}^m) &\leq H_\mu(\mathbf{T}^N \mathcal{F}^m) + \sum_{l=N}^{N+m-1} \sum_{f \in \mathcal{F}} H_\nu(\overline{\mathbf{T}^{k+l} f} | \Phi_k \overline{\mathbf{T}^l f}) \\ &\stackrel{(2.19)}{<} H_\mu(\mathbf{T}^N \mathcal{F}^m) + m\varepsilon. \end{aligned}$$

Ponadto,

$$\begin{aligned} H_\mu(\mathbf{T}^N \mathcal{F}^m) &= H_\nu(\Phi_k \overline{\mathbf{T}^N \mathcal{F}^m}) \\ &\leq H_\nu(\overline{\mathbf{T}^{k+N} \mathcal{F}^m}) + H_\nu(\Phi_k \overline{\mathbf{T}^N \mathcal{F}^m} | \overline{\mathbf{T}^{k+N} \mathcal{F}^m}) \\ &< H_\mu(\mathbf{T}^{k+N} \mathcal{F}^m) + m\varepsilon, \end{aligned}$$

co kończy dowód w sytuacji, gdy spełnione są założenia twierdzenia Iwanika.

Rozważmy teraz przypadek ogólny. Niech (X, μ) będzie przestrzenią probabilistyczną, a $\mathbf{T}$ dowolnym operatorem podwójnie stochastycznym. Zauważmy, że wszystkie funkcje pojawiające się w powyższej argumentacji, czyli rodzina $\mathcal{F}$, jej obrazy względem iteracji $\mathbf{T}$ i obrazy rodzin $\mathbf{T}^N \mathcal{F}^m$ przez operację $\Theta(\cdot)$, tworzą zbiór przeliczalny. Oznaczmy przez $\mathcal{L}$ zespoloną podalgebrę algebry (również zespolonej) $L^\infty(\mu)$, generowaną przez te funkcje. Wiadomo, że $\mathcal{L}$ jest izometrycznie izomorficzna z algebrą (zespolonych) funkcji ciągłych na pewnej zwartej przestrzeni Hausdorffa $\mathcal{X}$. Izomorfizm ten (oznaczymy go przez τ) zachowuje sprzężenie, więc przeprowadza funkcje o wartościach rzeczywistych na funkcje o wartościach rzeczywistych. Ponadto można dowieść, że τ jest izomorfizmem krat, zatem obrazy funkcji nieujemnych są funkcjami nieujemnymi. Miara μ traktowana jako funkcjonał na $L^\infty(\mu)$ indukuje funkcjonał liniowy na przestrzeni zespolonych funkcji ciągłych, reprezentowany przez pewną borelowską miarę probabilistyczną na $\mathcal{X}$. Całki z funkcji, będących swoimi obrazami przez τ względem odpowiednich miar, są równe. Na

mocy wniosku 10.10 z [Z], przestrzeń $\mathcal{X}$ jest metryzowalna (zatem standardowa), ponieważ zbiór generatorów algebry $\mathcal{L}$ jest przeliczalny. Jest jasne, że operator $\mathcal{T} = \tau \mathbf{T} \tau^{-1}$ określony na (rzeczywistej) podalgebrze wszystkich rzeczywistych funkcji ciągłych $C(\mathcal{X})$ jest operatorem Markowa, zatem pochodzi od pewnego prawdopodobieństwa przejścia. Z pierwszej części dowodu otrzymujemy więc prawdziwość tezy twierdzenia dla operatora $\mathcal{T}$. Na mocy uwagi 2.2.1(ii) i faktu, że τ jest izomorfizmem krat, entropie rodzin $\mathbf{T}^N \mathcal{F}^m$ i $\mathbf{T}^{k+N} \mathcal{F}^m$ są równe entropiom odpowiadających im rodzin w $C(\mathcal{X})$, co kończy dowód. $\square$

Poniższa własność, nazwana przez nas *quasi-podaddytywnością*, jest operatorowym odpowiednikiem podaddytywności klasycznej entropii.

Stwierdzenie 2.2.5 (Quasi-podaddytywność entropii) *Niech $\mathbf{T}$ będzie operatorem podwójnie stochastycznym. Dla każdego $\varepsilon > 0$ istnieją: $N \in \mathbb{N}$ oraz liczba rzeczywista c , takie, że dla dowolnych $k \in \mathbb{N}$ i $m \geq N$ zachodzi*

$$H_\mu(\mathcal{F}^{k+m}) \leq H_\mu(\mathcal{F}^k) + H_\mu(\mathcal{F}^m) + c + m\varepsilon.$$

Dowód Ustalmy $\varepsilon > 0$ i wybierzmy N zgodnie z lematem 2.2.4. Korzystając z faktu, że $H_\mu(\mathcal{F}) \leq r \log 2$, gdzie r jest licznością rodziny $\mathcal{F}$, otrzymujemy dla $k > 0$ i $m > N$ nierówności

$$\begin{aligned} H_\mu(\mathcal{F}^{k+m}) &\leq H_\mu(\mathcal{F}^k) + H_\mu(\mathbf{T}^k \mathcal{F}^N) + H_\mu(\mathbf{T}^{k+N} \mathcal{F}^{m-N}) \\ &\leq H_\mu(\mathcal{F}^k) + Nr \log 2 + H_\mu(\mathbf{T}^N \mathcal{F}^{m-N}) + (m-N)\varepsilon \\ &\leq H_\mu(\mathcal{F}^k) + H_\mu(\mathcal{F}^m) + Nr \log 2 + m\varepsilon. \end{aligned}$$

$\square$

Twierdzenie 2.2.6 *Jeśli $\mathbf{T}$ jest operatorem podwójnie stochastycznym, to ciąg $\frac{1}{n} H_\mu(\mathcal{F}^n)$, wykorzystany w punkcie (4°) definicji entropii, jest zbieżny.*

Dowód Wobec quasi-podaddytywności ciągu $H_\mu(\mathcal{F}^n)$, wystarczy jedynie zmodyfikować dowód klasycznego lematu orzekającego o zbieżności ciągu $(\frac{1}{n} a_n)$ w przypadku (a_n) podaddytywnego (patrz np. [D-G-S], Proposition (10.7) lub [W], Theorem 4.9).

Oznaczmy $a_n = H_\mu(\mathcal{F}^n)$. Ustalmy $\varepsilon > 0$ i dobierzmy N oraz c zgodnie ze stwierdzeniem 2.2.5. Dla $n > N$, $n = Nd + r$, gdzie $d \in \mathbb{N}$, $0 \leq r < N$, zachodzi

$$a_n \leq a_{Nd} + a_r + c + r\varepsilon \leq a_{Nd} + cd + n\varepsilon + a_r.$$

Dzieląc przez n i obliczając granicę górną względem n , otrzymujemy

$$\limsup_{n \rightarrow \infty} \frac{a_n}{n} \leq \frac{a_N}{N} + \frac{c}{N} + \varepsilon,$$

skąd, po nałożeniu granicy dolnej względem N , wynika nierówność

$$\limsup_n \frac{a_n}{n} \leq \liminf_n \frac{a_n}{n}.$$

□

W przeciwieństwie do przypadku klasycznego, granica w twierdzeniu 2.2.6 nie musi być równa kresowi dolnemu ciągu. Pełną podaddytywność, niezbędną dla tej równości, można otrzymać przez wprowadzenie w definicji entropii dodatkowego kroku konstrukcyjnego. Proponujemy dwie takie modyfikacje.

Definicja 2.2.7

$H'_\mu(\mathcal{F}) = \lim_{k \rightarrow \infty} H_\mu(\mathbf{T}^k \mathcal{F})$ (istnienie granicy zostało zapewnione w Lemacie 2.2.3)

$H''_\mu(\mathcal{F}) = \sup_{k \in \mathbb{N}} H_\mu(\mathbf{T}^k \mathcal{F})$

Lemat 2.2.8 Ciągi $H'_\mu(\mathcal{F}^n)$ i $H''_\mu(\mathcal{F}^n)$ są podaddytywne.

Dowód Jest oczywiste, że

$$H'_\mu(\mathbf{T}^m \mathcal{F}) = \lim_{k \rightarrow \infty} H_\mu(\mathbf{T}^{m+k} \mathcal{F}) = H'_\mu(\mathcal{F}).$$

Stąd

$$H'_\mu(\mathcal{F}^{m+n}) \leq H'_\mu(\mathcal{F}^m) + H'_\mu(\mathbf{T}^m \mathcal{F}^n) = H'_\mu(\mathcal{F}^m) + H'_\mu(\mathcal{F}^n).$$

Dowód podaddytywności ciągu $H''_\mu(\mathcal{F}^n)$ jest prawie identyczny. Należy jedynie zaznaczyć, że $H''_\mu(\mathbf{T}^m \mathcal{F})$ jest nie większa niż $H''_\mu(\mathcal{F})$. □

Oznaczmy przez $h'_\mu(\mathbf{T}, \mathcal{F})$ i $h''_\mu(\mathbf{T}, \mathcal{F})$ modyfikacje wielkości $h_\mu(\mathbf{T}, \mathcal{F})$ otrzymane przez podstawienie w miejsce H_μ (w punkcie (4°) konstrukcji entropii) odpowiednio H'_μ lub H''_μ (zgodnie z powyższym lematem, granicę górną można w obu przypadkach zastąpić granicą).

Stwierdzenie 2.2.9 Dla każdego operatora stochastycznego $\mathbf{T}$ zachodzą równości

$$h_\mu(\mathbf{T}, \mathcal{F}) = h'_\mu(\mathbf{T}, \mathcal{F}) = h''_\mu(\mathbf{T}, \mathcal{F}).$$

Dowód Oczywiście, $h_\mu(\mathbf{T}, \mathcal{F}) \leq h_\mu''(\mathbf{T}, \mathcal{F})$ oraz $h_\mu'(\mathbf{T}, \mathcal{F}) \leq h_\mu''(\mathbf{T}, \mathcal{F})$.

Udowodnimy, że entropia $h_\mu''(\mathbf{T})$ jest szacowana z góry przez $h_\mu'(\mathbf{T})$. Ustalmy liczbę naturalną n . Dla rodziny $\mathcal{F}^n$ i $\varepsilon = 1$ znajdziemy N spełniające tezę Lematu 2.2.3. Niech m będzie liczbą naturalną większą niż $n + N$ i dobierzmy $p \in \mathbb{N}$ tak, by $(p-1)n \leq m \leq pn$. Następnie znajdziemy k , dla którego różnica między $H_\mu(\mathbf{T}^k \mathcal{F}^m)$ a $H_\mu''(\mathcal{F}^m)$ jest mniejsza niż 1. Niech K będzie większą z liczb N i k .

$$\begin{aligned} H_\mu''(\mathcal{F}^m) &\leq H_\mu(\mathbf{T}^k \mathcal{F}^m) + 1 \\ &\leq H_\mu(\mathcal{F}^N) + H_\mu\left(\bigcup_{i=0}^{p-1} \mathbf{T}^{K+in} \mathcal{F}^n\right) + 1 \\ &\leq H_\mu(\mathcal{F}^N) + pH_\mu(\mathbf{T}^N \mathcal{F}^n) + p + 1. \end{aligned}$$

Dzieląc skrajne strony nierówności przez m i obliczając granicę górną względem m otrzymujemy

$$h_\mu''(\mathbf{T}, \mathcal{F}) \leq \frac{1}{n} H_\mu(\mathbf{T}^N \mathcal{F}^n) + \frac{1}{n}.$$

Obliczając granicę względem N , a następnie względem n uzyskujemy nierówność $h_\mu''(\mathbf{T}, \mathcal{F}) \leq h_\mu'(\mathbf{T}, \mathcal{F})$.

Aby dowieść relacji $h_\mu(\mathbf{T}, \mathcal{F}) \geq h_\mu'(\mathbf{T}, \mathcal{F})$, mając dany $\varepsilon > 0$, znajdujemy, na mocy lematu 2.2.4, liczbę N taką, że dla dowolnych l, m naturalnych zachodzi

$$H_\mu(\mathbf{T}^{l+N} \mathcal{F}^m) \leq H_\mu(\mathbf{T}^N \mathcal{F}^m) + m\varepsilon.$$

Gdy l dąży do nieskończoności, otrzymujemy

$$H_\mu'(\mathcal{F}^m) \leq H_\mu(\mathbf{T}^N \mathcal{F}^m) + m\varepsilon,$$

skąd wynika $h_\mu'(\mathbf{T}, \mathcal{F}) \leq h_\mu(\mathbf{T}, \mathbf{T}^N \mathcal{F}) + \varepsilon \stackrel{(2.2)}{=} h_\mu(\mathbf{T}, \mathcal{F}) + \varepsilon$ dla dowolnych $\varepsilon > 0$. $\square$

Niestety, wielkości H_μ' i H_μ'' nie spełniają aksjomatu o rozbiciach, zaś prawdziwość aksjomatu o oszacowaniu trudno jest zweryfikować — przypuszczalnie również nie jest on spełniony.

2.2.3 Własności entropii

Jak wykazemy poniżej, niektóre znane w klasycznej teorii entropii własności h_μ pozostają prawdziwe w przypadku naszej entropii. Na początek sformułujmy pewne standardowe (dla transformacji) definicje w przypadku operatorowym.

Definicja 2.2.10 Niech $\mathbf{T}_1, \mathbf{T}_2$ będą operatorami podwójnie stochastycznymi odpowiednio na $L^\infty(X_1, \mu_1)$ i $L^\infty(X_2, \mu_2)$ (lub na przestrzeniach funkcji całkowalnych).

- (i) $\mathbf{T}_1$ jest *faktorem* $\mathbf{T}_2$, gdy istnieje zachowujące miarę przekształcenie π z X_2 na X_1 , spełniające dla każdej funkcji $f \in L^\infty(X_1, \mu_1)$ (lub $L^1(\mu)$) warunek

$$(\mathbf{T}_1 f) \circ \pi = \mathbf{T}_2(f \circ \pi).$$

- (ii) $\mathbf{T}_1$ i $\mathbf{T}_2$ są *izomorficzne*, gdy zdefiniowane powyżej π jest przekształceniem odwracalnym i jego odwrotność π^{-1} jest transformacją mierzalną (zatem także zachowuje miarę).

Stwierdzenie 2.2.11 Niech $\mathbf{T}_1$ i $\mathbf{T}_2$ będą operatorami podwójnie stochastycznymi.

- (i) Jeśli $\mathbf{T}_1$ jest faktorem $\mathbf{T}_2$, to $h_{\mu_1}(\mathbf{T}_1) \leq h_{\mu_2}(\mathbf{T}_2)$.
- (ii) Jeśli $\mathbf{T}_1$ i $\mathbf{T}_2$ są izomorficzne, to ich entropie są równe.

Dowód Niech $\pi : X_2 \rightarrow X_1$ będzie odwzorowaniem faktorującym, a $\mathcal{F}$ ustaloną rodziną funkcji. Połóżmy

$$\pi\mathcal{F} = \{f \circ \pi : f \in \mathcal{F}\}.$$

Prawdą jest, że $h_{\mu_1}(\mathbf{T}_1, \mathcal{F}) = h_{\mu_2}(\mathbf{T}_2, \pi\mathcal{F})$, zatem (i) jest konsekwencją kroku konstrukcyjnego (5°), implikując jednocześnie (ii). $\square$

Stwierdzenie 2.2.12 Jeśli $\mathbf{T}$ jest dowolnym operatorem na zbiorze L złożonym z funkcji mierzalnych przyjmujących wartości z odcinka $[0, 1]$, to dla każdej liczby naturalnej k zachodzi

$$h_\mu(\mathbf{T}^k) = k h_\mu(\mathbf{T}).$$

Dowód Upewnijmy się, że standardowa argumentacja (patrz np. [D-G-S], Proposition (10.12)(c) lub [W], Theorem 4.13) zachowuje poprawność w przypadku operatorowym. Zauważmy, że wobec Lematu 2.2.1

$$\begin{aligned} h_\mu(\mathbf{T}^k, \mathcal{F}) &= \limsup_{n \rightarrow \infty} \frac{1}{n} H_\mu \left(\bigcup_{i=0}^{n-1} \mathbf{T}^{ki} \mathcal{F} \right) \\ &\leq k \limsup_{n \rightarrow \infty} \frac{1}{kn} H_\mu \left(\bigcup_{i=0}^{kn-1} \mathbf{T}^i \mathcal{F} \right) \\ &\leq k h_\mu(\mathbf{T}, \mathcal{F}). \end{aligned}$$

Na odwrót, ustalmy k i n , i niech $m(n) = k \cdot d(n)$ będzie najmniejszą wielokrotnością k większą niż n . Wtedy,

$$\begin{aligned} h_\mu(\mathbf{T}, \mathcal{F}) &= \limsup_{n \rightarrow \infty} \frac{1}{n} H_\mu \left(\bigcup_{i=0}^{n-1} \mathbf{T}^i \mathcal{F} \right) \\ &\leq \limsup_{n \rightarrow \infty} \frac{1}{n} H_\mu \left(\bigcup_{i=0}^{m(n)-1} \mathbf{T}^i \mathcal{F} \right) \\ &= \frac{1}{k} \limsup_{n \rightarrow \infty} \frac{m(n)}{n} \cdot \frac{1}{d(n)} \cdot H_\mu \left(\bigcup_{i=0}^{d(n)-1} \bigcup_{j=0}^{k-1} \mathbf{T}^{ki+j} \mathcal{F} \right). \end{aligned}$$

Ponieważ $\lim_{n \rightarrow \infty} \frac{m(n)}{n} = 1$, a $d(n)$ przebiega wszystkie wartości naturalne, gdy n dąży do nieskończoności, otrzymujemy

$$\begin{aligned} h_\mu(\mathbf{T}, \mathcal{F}) &\leq \frac{1}{k} \limsup_{n \rightarrow \infty} \frac{1}{n} H_\mu \left(\bigcup_{i=0}^{n-1} \mathbf{T}^{ki} \left(\bigcup_{j=0}^{k-1} \mathbf{T}^j \mathcal{F} \right) \right) \\ &= \frac{1}{k} h_\mu \left(\mathbf{T}^k, \bigcup_{j=0}^{k-1} \mathbf{T}^j \mathcal{F} \right). \end{aligned}$$

□

Następne stwierdzenie dotyczy istoty zachowań chaotycznych rozróżnianych przez entropię. Jego treścią jest naturalna, jak się wydaje, własność naszego pojęcia: w przypadku trywialnej, w pewnym sensie, dynamiki układu entropia wynosi zero.

Stwierdzenie 2.2.13 *Niech $\mathbf{T}$ będzie L^1 -kontrakcją na zbiorze L złożonym z funkcji mierzalnych przyjmujących wartości z odcinka $[0, 1]$ i takim, że dla każdej funkcji $f \in L$ ciąg $\mathbf{T}^n f$ zbiega w $L^1(\mu)$ do funkcji niezmienniczej ϕ_f . Wówczas $h_\mu(\mathbf{T}) = 0$.*

Dowód Ustalmy rodzinę $\mathcal{F}$ o liczności r i liczbę $\varepsilon > 0$. Dobierzmy $\delta > 0$ tak, by prawdziwy był wzór (2.12). Niech $\mathcal{F}_\phi$ oznacza rodzinę r funkcji niezmienniczych będących granicami (w $L^1(\mu)$) ciągów funkcyjnych $\mathbf{T}^n f$, $f \in \mathcal{F}$. Wówczas istnieje N , dla którego $\text{dist}(\mathbf{T}^N \mathcal{F}, \mathcal{F}_\phi) < \delta$ oraz

$$h_\mu(\mathbf{T}, \mathcal{F}) \stackrel{(2.2)}{=} h_\mu(\mathbf{T}, \mathbf{T}^N \mathcal{F}) \stackrel{(2.12)}{<} h_\mu(\mathbf{T}, \mathcal{F}_\phi) + \varepsilon \stackrel{(2.13)}{=} \varepsilon,$$

skąd, wobec dowolności $\mathcal{F}$ i ε , wynika teza. □

Przypomnijmy, że z powodu ograniczeń technicznych dotyczących wyboru rozbitcia α , zrezygnowaliśmy z używania wielkości $h_\mu(\mathbf{T}, \mathcal{F}, \alpha)$ jako jawnej

definicji entropii. Przedstawimy teraz wspomniany na początku działu przykład, ilustrujący niemożność opuszczenia niewygodnych założeń.

Przykład 2.2.14 Niech (X, Σ, μ) będzie przestrzenią wszystkich ciągów o elementach 0 lub 1, z σ -ciałem produktowym i równomiernie rozłożoną miarę produktową. Zdefiniujmy operator podwójnie stochastyczny $\mathbf{T}$ wzorem

$$\mathbf{T}f(x) = \frac{1}{2} \left(f(\sigma x) + \int_X f(x) d\mu \right),$$

gdzie σ jest przesunięciem $(x_n) \mapsto (x_{n+1})$ na X . Niech $\mathcal{F}$ składa się z jednej tylko funkcji $f(x) = x_0$, zaś α będzie rozbiem odcinka $[0, 1]$ na przedziały $[0, \frac{1}{2}]$ i $(\frac{1}{2}, 1]$. Zauważmy, że f jest funkcją charakterystyczną cylindra

$$C = \{x \in X : x_0 = 1\}.$$

Zatem

$$\begin{aligned} \mathbf{T}^n f(x) &= \frac{1}{2} + \frac{1}{2^n} \left(\mathbf{1}_{\sigma^{-n}(C)}(x) - \frac{1}{2} \right) \\ &= \begin{cases} \frac{1}{2} - \frac{1}{2^{n+1}} & \text{dla } x_n = 0 \\ \frac{1}{2} + \frac{1}{2^{n+1}} & \text{dla } x_n = 1 \end{cases} \end{aligned}$$

zbiega w $L^1(\mu)$ do funkcji stałe równej $\frac{1}{2}$. Wobec stwierdzenia 2.2.13, entropia h_μ operatora $\mathbf{T}$ wynosi zero. Jednakże $H_\mu \left(\bigsqcup_{k=0}^{n-1} \mathbf{1}_{\mathbf{T}^k f^{-1}(\alpha)} \right)$ jest entropią rozbita przestrzeni X na cylindry o ustalonych n współrzędnych, więc $h_\mu(\mathbf{T}, \mathcal{F}, \alpha)$ jest równa $\log 2$. $\square$

Entropia operatorowa stanowi istotne rozszerzenie definicji Kołmogorowa-Sinaja — jest wrażliwa nie tylko na zjawiska punktowe, ale też na „dynamikę operatorową”. Fakt ten wyrażamy następującym twierdzeniem.

Twierdzenie 2.2.15 *Istnieje operator podwójnie stochastyczny o entropii dodatniej, którego jedynym faktorem punktowo generowanym jest faktor trywialny.*

Dowód Niech (X, Σ, μ) będzie taka jak w poprzednim przykładzie. Oznaczmy przez ν rozkład geometryczny na $\mathbb{N}$, tzn. $\nu(k) = 2^{-k}$ dla $k \in \mathbb{N}$. Element (x, k) przestrzeni produktowej $X \times \mathbb{N}$ będziemy traktować jako ciąg $(x_1, x_2, \dots, x_k^*, x_{k+1}, \dots)$ o wyrazach 0 lub 1, z umieszczonym na pozycji k znacznikiem (gwiazdka).

Dla każdego bloku $B = (b_1, b_2, \dots, b_k) \in \{0, 1\}^k$ definiujemy przekształcenie $\sigma_B : X \rightarrow X$ kładąc

$$(\sigma_B x)_n = \begin{cases} b_n & \text{dla } n \leq k \\ x_{n+1} & \text{dla } n > k \end{cases}.$$

Następnie określamy operator $\mathbf{T}$ na $L^\infty(\mu \times \nu)$ wzorem:

$$\mathbf{T}f(x, k) = f(\sigma x, k-1) \quad \text{dla } k > 1,$$

$$\mathbf{T}f(x, 1) = \sum_{k=1}^{\infty} 2^{-k} \sum_{B \in \{0,1\}^k} 2^{-k} f(\sigma_B x, k).$$

Odpowiadające $\mathbf{T}$ prawdopodobieństwo przejścia $P((x, k), \cdot)$ możemy opisać jako operację przesunięcia ciągu o jedną pozycję w lewo (przesuwającą także znacznik), o ile znacznik nie znajduje się na pozycji pierwszej. Jeśli znacznik jest na pozycji pierwszej, to ciąg zostaje przesunięty w lewo, a następnie w miejsce początkowego bloku o długości k (wybranej zgodnie z prawdopodobieństwem geometrycznym) podstawia się losowo wybrany blok zer i jedynek, o tej samej długości. Znacznik zapisujemy wtedy na pozycji k .

Uzasadnimy najpierw, że $\mathbf{T}$ jest podwójnie stochastyczny względem miary produktowej $\mu \times \nu$. W tym celu rozważmy funkcję charakterystyczną $f = \mathbf{1}_C$ cylindra postaci

$$C = C(y_1, y_2, \dots, y_k^*, y_{k+1}, \dots, y_n) = \{(x, k) : x_i = y_i, i = 1, \dots, n\}.$$

Jest jasne, że całka z funkcji f wynosi $2^{-(k+n)}$. Każdy punkt należący do któregoś z dwóch cylindrów $C(x_0, y_1, y_2, \dots, y_k^*, y_{k+1}, \dots, y_n)$ ($x_0 \in \{0, 1\}$) przechodzi z prawdopodobieństwem 1 na punkt należący do C . Zatem na obu tych cylindrach $\mathbf{T}f = 1$, a całka z $\mathbf{T}f$ po tej części przestrzeni jest równa $2^{-(k+1)} \cdot 2 \cdot 2^{-(n+1)} = 2^{-(k+n+1)}$ (bo znacznik znajduje się na pozycji $k+1$). Ponadto, z prawdopodobieństwem 2^{-2k} do C wpadają też punkty z cylindrów postaci $C(x_0^*, x_1, \dots, x_k, y_{k+1}, \dots, y_n)$, gdzie $x_0, \dots, x_k$ są dowolnie wybrane (bowiem z takim prawdopodobieństwem wybieramy przekształcenie σ_B dla $B = (y_1, y_2, \dots, y_k)$). Dla takich punktów $\mathbf{T}f(x) = P((x, 1), C) = 2^{-2k}$. Całka na tej części przestrzeni wynosi $2^{-2k} \cdot 2^{k+1} \cdot 2^{-1} \cdot 2^{-(n+1)} = 2^{-(k+n+1)}$. Suma obu tych składników jest równa wyjściowej całce z funkcji f , więc $\mathbf{T}$ jest podwójnie stochastyczny względem $\mu \times \nu$.

Każdy punktowo generowany faktor operatora $\mathbf{T}$ odpowiada takiemu σ -ciału $\Sigma' \subset \Sigma$, że $P((x, k), A) = 0$ lub 1 dla każdego $A \in \Sigma'$ i prawie każdego x . Rozważmy parę punktów (x, k) , (x', k') o jednakowych ogonach, tzn. takich, że $x_n = x'_n$, o ile n jest większe od pewnego k'' . Możemy założyć, że $k'' > k$ i $k'' > k'$. Rozważmy punkt (x'', k'') taki, że $x_n = x'_n = x''_n$ dla $n > k''$.

Prawdopodobieństwa przejścia z punktu $(0x, 1) = 0^*x$ do każdego z punktów (x, k) i (x'', k'') są dodatnie. To oznacza, że faktor utożsamia (x, k) i (x'', k'') (o ile 0^*x jest poza pewnym zbiorem miary zero). Podobnie, faktor ten utożsamia (x', k') i (x'', k'') , a zatem także (x, k) i (x', k') . Pokazaliśmy więc, że faktor punktowy utożsamia (prawie wszystkie) punkty o jednakowych ogonach, niezależnie od pozycji występowania znacznika. Wobec tego Σ' jest (z dokładnością do miary μ) zawarte w σ -ciele ogonowym przestrzeni $\{0, 1\}^{\mathbb{N}}$. Na mocy zero-jedynkowego prawa Kołmogorowa, takie σ -ciało jest trywialne względem miary μ . Zatem $\mathbf{T}$ nie posiada nietrywialnych (miarowych) faktorów punktowych.

Kończymy dowód uzasadniając, że $h_\mu(\mathbf{T}) \geq \log 2$. Rozważmy funkcję $f_i(x, k) = x_i$. Sprawdźmy, że

$$\mathbf{T}^n f_i(x, k) = \alpha \frac{1}{2} + (1 - \alpha)x_{i+n},$$

gdzie $\alpha \leq 2^{-(i-1)}$. Jest to oczywiście prawdą dla $n < k$ (wówczas $\alpha = 0$), gdyż wtedy $\mathbf{T}^n$ działa deterministycznie w punkcie (x, k) . Rozważmy teraz $n = k$. Jest to pierwszy moment, w którym prawdopodobieństwo $P^n((x, k), \cdot)$ nie jest deterministyczne. Zauważmy, że rozkład pozycji znacznika w obrazie (x, k) (czyli nośniku $P^n((x, k), \cdot)$) jest dany przez prawdopodobieństwo geometryczne ν . W następnym, $(n + 1)$ -szym kroku, z prawdopodobieństwem $\frac{1}{2}$ znaczniki zostaną przesunięte (dając połowę tego samego rozkładu geometrycznego), a z prawdopodobieństwem $\frac{1}{2}$ rozmieszczone losowo (znów zgodnie z tym samym rozkładem). W rezultacie, rozkład pozycji znaczników względem $P^{n+1}((x, k), \cdot)$ pozostaje ten sam. Podobnie jest w następnych krokach. Dowiedliśmy więc, że dla $n \geq k$ pozycja znacznika względem $P^n((x, k), \cdot)$ ma rozkład geometryczny. Punkty należące do nośnika miary $P^n((x, k), \cdot)$, których znaczniki znajdują się na pozycjach wcześniejszych niż pewne i , mają x_{i+n} na pozycji i -tej, gdyż do tego momentu wartość x_{i+n} była przekształcana deterministycznie (przesuwana). Prawdopodobieństwo zdarzenia, że w losowo wybranym punkcie z nośnika znacznik występuje na pozycji wcześniejszej niż i wynosi

$$\sum_{j=1}^{i-1} 2^{-j} = 1 - 2^{-(i-1)},$$

skąd otrzymujemy składnik $(1 - 2^{-(i-1)})x_{i+n}$ we wzorze na $\mathbf{T}^n f_i(x, k)$. Pozostałe punkty z n -tego obrazu (x, k) przyjmują na pozycji i -tej wartość 0 lub 1 z równym prawdopodobieństwem, co daje składnik $2^{-(i-1)}\frac{1}{2}$. Zatem otrzymujemy żądany wzór w przypadku $n \geq k$ (α wynosi $2^{-(i-1)}$).

Z powyższych obliczeń wynika, że n -ty obraz funkcji f_i różni się od funkcji f_{i+n} tym, że amplituda skoków funkcji $\mathbf{T}^n f_i$ przy przechodzeniu między

cylindrami wynosi przynajmniej $1 - 2^{-(i-1)}$, zamiast być stale równa 1 (jak w przypadku f_{i+n}). Jednakże pozwala to otrzymać entropię dowolnie bliską $\log 2$ poprzez rozpatrywanie odpowiednio dużych i . $\square$

2.3 Entropia topologiczna

2.3.1 Nowa definicja entropii

Niech $C_I(X)$ oznacza zbiór wszystkich rzeczywistych funkcji ciągłych $f : X \rightarrow [0, 1]$, gdzie X jest zwartą przestrzenią Hausdorffa. Niech T będzie operatorem Markowa działającym na $C(X)$. Poniżej podamy trzy równoważne sposoby definiowania entropii topologicznej. Pierwszy z nich jest najbardziej zbliżony do podanej w poprzednim podrozdziale definicji entropii miarowej: kluczową rolę odgrywają pokrycia $\mathcal{U}_f^\varepsilon$, otrzymane przez „pogrubienie” elementów rozbięcia $\mathcal{A}_f$. Druga metoda polega na przeniesieniu pokrycia z odcinka $[0, 1]$ do przestrzeni X przy pomocy zadanej rodziny funkcji ciągłych. W trzeciej, wprowadzamy w X pewną pseudo-metrykę indukowaną przez skończoną rodzinę funkcji ciągłych. Prowadzi to do definicji przypominającej znaną definicję Bowena. Dla skrócenia zapisu, skończone pokrycie przestrzeni X zbiorami otwartymi będziemy nazywać *pokryciem otwartym*.

Dla dowolnej funkcji ciągłej f zdefiniujemy

$$\begin{aligned} U_{<f}^\varepsilon &= \{(x, t) \in X \times [0, 1] : t < f(x) + \varepsilon\} \\ U_{>f}^\varepsilon &= \{(x, t) \in X \times [0, 1] : t > f(x) - \varepsilon\} \\ \mathcal{U}_f^\varepsilon &= \{U_{<f}^\varepsilon, U_{>f}^\varepsilon\}. \end{aligned}$$

Mając daną skończoną rodzinę $\mathcal{F} \subset C_I(X)$, określamy pokrycie otwarte produktu $X \times [0, 1]$, kładąc

$$\mathcal{U}_\mathcal{F}^\varepsilon = \bigvee_{f \in \mathcal{F}} \mathcal{U}_f^\varepsilon.$$

Dla pokrycia otwartego $\mathcal{V}$ odcinka $[0, 1]$ położymy

$$\mathcal{F}^{-1}(\mathcal{V}) = \bigvee_{f \in \mathcal{F}} f^{-1}(\mathcal{V}).$$

Następujący lemat zawiera podstawowe własności zdefiniowanych pokryć.

Lemat 2.3.1 *Niech $\mathcal{F}$ i $\mathcal{G}$ będą skończonymi podzbiorami $C_I(X)$, $\mathcal{V}$ pokryciem otwartym odcinka $[0, 1]$, zaś ε dowolną liczbą dodatnią. Wówczas*

$$(i) \quad \mathcal{U}_{\mathcal{F} \cup \mathcal{G}}^\varepsilon = \mathcal{U}_\mathcal{F}^\varepsilon \vee \mathcal{U}_\mathcal{G}^\varepsilon,$$

$$(ii) (\mathcal{F} \cup \mathcal{G})^{-1}(\mathcal{V}) = \mathcal{F}^{-1}(\mathcal{V}) \vee \mathcal{G}^{-1}(\mathcal{V}),$$

(iii) jeśli $\mathbf{T}f = f \circ T$ dla pewnej transformacji ciągłej $T : X \rightarrow X$, to

$$\mathcal{U}_{\mathcal{F}^n}^\varepsilon = \bigvee_{i=0}^{n-1} (T \times \text{Id})^{-i}(\mathcal{U}_{\mathcal{F}}^\varepsilon)$$

oraz

$$(\mathcal{F}^n)^{-1}(\mathcal{V}) = \bigvee_{i=0}^{n-1} T^{-i}(\mathcal{F}^{-1}(\mathcal{V})).$$

Dowód Wynika z definicji połączenia pokryw i własności przeciwobrazu. $\square$

Przypomnijmy, że symbol $N(\mathcal{U})$, gdzie $\mathcal{U}$ jest dowolnym pokryciem otwartym, oznacza minimalną liczbę podpokrycia wybranego z $\mathcal{U}$ (tzn. podzbioru $\mathcal{U}$ będącego pokryciem). W poniższych definicjach, $\mathcal{F}$ jest skończonym podzbiorem $C_I(X)$.

Definicja 2.3.2 Niech $\varepsilon > 0$. Definiujemy

$$(i) H_1(\mathcal{F}, \varepsilon) = \log N(\mathcal{U}_{\mathcal{F}}^\varepsilon),$$

$$(ii) h_1(\mathbf{T}, \mathcal{F}, \varepsilon) = \limsup_{n \rightarrow \infty} \frac{1}{n} H_1(\mathcal{F}^n, \varepsilon),$$

$$(iii) h_1(\mathbf{T}) = \sup_{\mathcal{F}} \sup_{\varepsilon} h_1(\mathbf{T}, \mathcal{F}, \varepsilon).$$

Definicja 2.3.3 Niech $\mathcal{V}$ będzie pokryciem otwartym odcinka $[0, 1]$.

$$(i) H_2(\mathcal{F}, \mathcal{V}) = \log N(\mathcal{F}^{-1}(\mathcal{V})),$$

$$(ii) h_2(\mathbf{T}, \mathcal{F}, \mathcal{V}) = \limsup_{n \rightarrow \infty} \frac{1}{n} H_2(\mathcal{F}^n, \mathcal{V}),$$

$$(iii) h_2(\mathbf{T}) = \sup_{\mathcal{F}} \sup_{\mathcal{V}} h_2(\mathbf{T}, \mathcal{F}, \mathcal{V}).$$

Wspomnianą pseudometrykę na X określamy wzorem

$$d_{\mathcal{F}}(x, y) = \sup_{f \in \mathcal{F}} |f(x) - f(y)|.$$

Mówimy, że pewien podzbiór przestrzeni X jest $(d_{\mathcal{F}}, \varepsilon)$ -rozdzielony, gdy jest ε -rozdzielony w pseudo-metryce $d_{\mathcal{F}}$. Jako że przestrzeń X jest zwarta, można znaleźć w X skończony $(d_{\mathcal{F}}, \varepsilon)$ -rozdzielony podzbiór o maksymalnej liczności. Liczność tę będziemy oznaczać symbolem $s(d_{\mathcal{F}}, \varepsilon)$.

Definicja 2.3.4 Dla $\varepsilon > 0$ definiujemy

- (i) $H_3(\mathcal{F}, \varepsilon) = \log s(d_{\mathcal{F}}, \varepsilon)$;
- (ii) $h_3(\mathbf{T}, \mathcal{F}, \varepsilon) = \limsup_{n \rightarrow \infty} \frac{1}{n} H_3(\mathcal{F}^n, \varepsilon)$;
- (iii) $h_3(\mathbf{T}) = \sup_{\mathcal{F}} \sup_{\varepsilon} h_3(\mathbf{T}, \mathcal{F}, \varepsilon)$.

Twierdzenie 2.3.5 Dla każdego operatora Markowa $\mathbf{T}$ zachodzą równości

$$h_1(\mathbf{T}) = h_2(\mathbf{T}) = h_3(\mathbf{T}).$$

Dowód Na początek, udowodnimy, że $h_1(\mathbf{T}) \leq h_2(\mathbf{T})$. Wybierzmy $\varepsilon > 0$ i pokrycie otwarte $\mathcal{V}$ odcinka $[0, 1]$ składające się ze zbiorów o średnicach nie większych niż ε . Twierdzimy, że pokrycie produktu $X \times [0, 1]$ zdefiniowane wzorem

$$\mathcal{W}_n = \{U \times V : U \in (\mathcal{F}^n)^{-1}(\mathcal{V}), V \in \mathcal{V}\}$$

jest wpisane w $\mathcal{U}_{\mathcal{F}^n}^\varepsilon$. Istotnie, dla każdego $U \times V \in \mathcal{W}_n$ kładąc

$$\mathcal{F}' = \{f \in \mathcal{F}^n : (\forall_{x \in U}) f(x) \geq \inf V\}.$$

otrzymujemy

$$U \times V \subset \bigcap_{f \in \mathcal{F}'} U_{<f}^\varepsilon \cap \bigcap_{f \in \mathcal{F}^n \setminus \mathcal{F}'} U_{>f}^\varepsilon \in \mathcal{U}_{\mathcal{F}^n}^\varepsilon.$$

Zatem

$$N(\mathcal{U}_{\mathcal{F}^n}^\varepsilon) \leq N(\mathcal{W}_n) \leq N((\mathcal{F}^n)^{-1}(\mathcal{V})) \cdot N(\mathcal{V}),$$

a skoro $N(\mathcal{V})$ nie zależy od n , wnioskujemy, że

$$h_1(\mathbf{T}, \mathcal{F}, \varepsilon) \leq h_2(\mathbf{T}, \mathcal{F}, \mathcal{V}).$$

Kończymy tę część dowodu, obliczając odpowiednie kresy górne.

Następnie udowodnimy, że $h_2(\mathbf{T}) \leq h_3(\mathbf{T})$. Niech $\mathcal{V}$ będzie pokryciem skończonym odcinka $[0, 1]$, mającym liczbę Lebesgue'a δ . Przez E oznaczmy $(d_{\mathcal{F}^n}, \frac{\delta}{2})$ -rozdzielony podzbiór przestrzeni X o maksymalnej liczności. Z maksymalności E wynika, że zbiór $\{B^{\frac{\delta}{2}}(x) : x \in E\}$ kul w pseudo-metryce $d_{\mathcal{F}^n}$ jest pokryciem otwartym przestrzeni X . Dla każdej funkcji $f \in \mathcal{F}^n$ i dowolnego $x \in E$ przedział $(f(x) - \frac{\delta}{2}, f(x) + \frac{\delta}{2})$ zawiera się w pewnym elemencie $V_f(x)$ pokrycia $\mathcal{V}$. Stąd

$$\begin{aligned} B^{\frac{\delta}{2}}(x) &= \bigcap_{f \in \mathcal{F}^n} f^{-1} \left(f(x) - \frac{\delta}{2}, f(x) + \frac{\delta}{2} \right) \\ &\subset \bigcap_{f \in \mathcal{F}^n} f^{-1}(V_f(x)) \in (\mathcal{F}^n)^{-1}(\mathcal{V}) \end{aligned}$$

oraz

$$N((\mathcal{F}^n)^{-1}(\mathcal{V})) \leq \text{card}\{B^{\frac{\varepsilon}{2}} : x \in E\} = s(d_{\mathcal{F}^n}, \frac{\varepsilon}{2}),$$

co pociąga $h_2(\mathbf{T}) \leq h_3(\mathbf{T})$.

Kończymy dowód uzasadnieniem nierówności $h_3(\mathbf{T}) \leq h_1(\mathbf{T})$. Niech $D \subset X$ będzie $(d_{\mathcal{F}}, \varepsilon)$ -rozdzielonym zbiorem o maksymalnej liczności. Połóżmy $\gamma = \frac{\varepsilon}{6}$ oraz

$$\tilde{\mathcal{F}} = \left\{ \frac{1}{2}f + i\gamma : f \in \mathcal{F}, i \in \mathbb{Z}, 0 \leq i \leq \frac{1}{2\gamma} \right\}.$$

Udowodnimy, że pokrycie $\mathcal{U}_{\tilde{\mathcal{F}}}^{\gamma}$ rozdziela punkty zbioru $D \times \left\{ \frac{1}{2} \right\}$ w tym sensie, że każdy element pokrycia $\mathcal{U}_{\tilde{\mathcal{F}}}^{\gamma}$ zawiera co najwyżej jeden punkt należący do $D \times \left\{ \frac{1}{2} \right\}$. Rozważmy $x, y \in D$. Istnieje funkcja $f \in \mathcal{F}$, dla której $|f(x) - f(y)| \geq \varepsilon$. Bez straty ogólności możemy założyć, że $f(x) + \varepsilon \leq f(y)$. Wybierzmy indeks i taki, że funkcja $\tilde{f} = \frac{1}{2}f + i\gamma$ należy do $\tilde{\mathcal{F}}$ i spełnia warunek

$$\gamma \leq \frac{1}{2} - \tilde{f}(x) \leq 2\gamma.$$

Wówczas

$$\begin{aligned} \tilde{f}(y) - \frac{1}{2} &= \frac{1}{2}f(y) + i\gamma - \frac{1}{2} \\ &\geq \frac{1}{2}f(x) + \frac{\varepsilon}{2} + i\gamma - \frac{1}{2} \\ &= \tilde{f}(x) - \frac{1}{2} + \frac{\varepsilon}{2} \geq \gamma. \end{aligned}$$

Zatem $(y, \frac{1}{2})$ należy do $U_{<\tilde{f}}^{\gamma}$, lecz nie należy do $U_{>\tilde{f}}^{\gamma}$, podczas gdy $(x, \frac{1}{2})$ jest elementem $U_{>\tilde{f}}^{\gamma}$, ale nie $U_{<\tilde{f}}^{\gamma}$. Ponieważ każdy zbiór z pokrycia $\mathcal{U}_{\tilde{\mathcal{F}}}^{\gamma}$ zawiera się w $U_{<\tilde{f}}^{\gamma}$ lub w $U_{>\tilde{f}}^{\gamma}$, pokrycie to rozdziela punkty zbioru $D \times \left\{ \frac{1}{2} \right\}$. Co więcej, każde pokrycie wybrane z $\mathcal{U}_{\tilde{\mathcal{F}}}^{\gamma}$ posiada tę samą własność, zatem

$$s(d_{\mathcal{F}}, \varepsilon) \leq N(\mathcal{U}_{\tilde{\mathcal{F}}}^{\gamma}).$$

Przypomnijmy, że $\mathbf{T}$, jako operator Markowa, jest liniowy i zachowuje funkcje stałe. Wobec tego $\tilde{\mathcal{F}}^n = (\tilde{\mathcal{F}})^n$, co pozwala nam zastąpić w powyższych rozważaniach $\mathcal{F}$ przez $\mathcal{F}^n$, by otrzymać

$$s(d_{\mathcal{F}^n}, \varepsilon) \leq N(\mathcal{U}_{(\tilde{\mathcal{F}})^n}^{\gamma}).$$

Dowód kończymy obliczając odpowiednie granice górne i kresy. □

W dalszej części pracy będziemy pisać h_{top} mając na myśli wspólną wartość wielkości h_1 , h_2 i h_3 . Następne twierdzenie wskazuje, że zbieżność z klasycznym oznaczeniem entropii dla ciągłych transformacji jest celowa.

Twierdzenie 2.3.6 *Jeśli $Tf = f \circ T$ jest punktowo generowanym operatorem Markowa, to entropia $h_{\text{top}}(T)$ jest równa entropii topologicznej transformacji T .*

Dowód twierdzenia poprzedzimy następującym lematem.

Lemat 2.3.7 *Niech $\mathcal{W} = \{W_1, \dots, W_r\}$ będzie pokryciem otwartym przestrzeni X , o minimalnej liczności. Istnieje rozkład jedności $\mathcal{F} = \{f_1, \dots, f_r\}$ składający się z funkcji ciągłych na X i taki, że $f_i(x) = 0$ dla każdego $x \in W_i^c$ ($i = 1, \dots, r$). (Wtedy również $f_i(x) = 1$, o ile $x \in \bigcap_{j \neq i} (W_j^c)$, tj. W_i jest jedynym zbiorem pokrywającym x .)*

Dowód Oznaczmy przez $W_{\geq i}$ sumę $\bigcup_{j=i}^r W_j$. Na mocy lematu Urysohna, istnieje ciągła funkcja $f_1 : X \rightarrow [0, 1]$, stale równa 1 na zbiorze $(W_{\geq 2})^c$, a znikająca na W_1^c . Załóżmy dla dowodu indukcyjnego, że wybraliśmy już nieujemne funkcje ciągłe $f_1, \dots, f_{i-1}$, dla których spełnione są warunki:

- (a) $f_j(x) = 0$ dla $x \in W_j^c$ ($j = 1, \dots, i-1$),
- (b) $\sum_{j=1}^{i-1} f_j(x) \leq 1$ dla każdego $x \in X$,
- (c) $f_{i-1}(x) = 1 - \sum_{j=1}^{i-2} f_j(x)$ dla każdego $x \in (W_{\geq i})^c$.

Zdefiniujmy funkcję pomocniczą $u_i : (W_i^c \cup (W_{\geq i+1})^c) \rightarrow [0, 1]$ wzorem

$$u_i(x) = \begin{cases} 1 - \sum_{j=1}^{i-1} f_j(x) & \text{dla } x \in (W_{\geq i+1})^c \\ 0 & \text{dla } x \in W_i^c \end{cases}.$$

Poprawność definicji u_i wynika z faktu, że dla $x \in W_i^c \cap (W_{\geq i+1})^c = (W_{\geq i})^c$ warunek (c) implikuje

$$1 - \sum_{j=1}^{i-1} f_j(x) = \left(1 - \sum_{j=1}^{i-2} f_j(x)\right) - f_{i-1}(x) = 0.$$

Zatem u_i jest dobrze określoną funkcją ciągłą na domkniętym podzbiorze przestrzeni X . Na mocy twierdzenia Tietzego-Urysohna, istnieje przedłużenie u_i do ciągłej funkcji $\bar{u}_i : X \rightarrow [0, 1]$. Połóżmy

$$f_i(x) = \min \left(\bar{u}_i(x), \left(1 - \sum_{j=1}^{i-1} f_j(x) \right) \right).$$

Nietrudno przekonać się, że warunki (a), (b) oraz (c) pozostają spełnione, gdy w miejsce $i - 1$ podstawimy i . Na koniec, niech

$$f_r = 1 - \sum_{j=1}^{r-1} f_j.$$

Funkcja f_r jest, oczywiście, ciągła. Wobec warunku (b) dla $i = r$ oraz faktu, że f_i są funkcjami dodatnimi dla $i < r$, mamy $0 \leq f_r \leq 1$. Zatem $\mathcal{F} = \{f_1, \dots, f_r\}$ jest rozkładem jedności złożonym z funkcji ciągłych. Jako że $W_r^c = (W_{\geq r})^c$, a f_{r-1} spełnia warunek (c), otrzymujemy $f_r(x) = 1 - \sum_{j=1}^{r-1} f_j(x) = 0$, o ile $x \in W_r^c$. $\square$

Dowód twierdzenia 2.3.6 Z lematu 2.3.1(iii) wynika, że $h_2(\mathbf{T})$ jest mniejsza lub równa entropii T .

W dowodzie odwrotnej nierówności posłużymy się definicją h_1 . Niech $\mathcal{W}$ będzie pokryciem otwartym przestrzeni X . Spośród jego elementów wybierzmy podpokrycie $\{W_1, \dots, W_r\}$ o minimalnej liczności. Niech $\mathcal{F} = \{f_1, \dots, f_r\}$ będzie rodziną spełniającą tezę poprzedniego lematu. Ustalmy $0 < \varepsilon < \frac{1}{2r}$. Każdy zbiór z pokrycia $\mathcal{U}_{\mathcal{F}^n}^\varepsilon$ jest postaci

$$\bigcap_{g \in \mathcal{F}^n} U_g = \bigcap_{k=0}^{n-1} \bigcap_{g \in \mathbf{T}^k \mathcal{F}} U_g,$$

gdzie $U_g \in \mathcal{U}_g^\varepsilon$. Udowodnimy, że każde podpokrycie $\mathcal{U}'$ wybrane z $\mathcal{U}_{\mathcal{F}^n}^\varepsilon$ zadaje wybór podpokrycia z $\mathcal{W}^n$, o co najwyżej takiej samej liczności. Każdy zbiór z pokrycia $\mathcal{U}'$ spełniający warunek

$$(\forall_{k < n}) (\exists_{g_k \in \mathbf{T}^k \mathcal{F}}) U_{g_k} = U_{<g_k}^\varepsilon \quad (2.20)$$

wyznacza pewien element pokrycia $\mathcal{W}^n$. Element ten ma postać $\bigcap_{k=0}^{n-1} \widetilde{W}_k$, gdzie $\widetilde{W}_k \in T^{-k} \mathcal{W}$ jest taki, że funkcja g_k jest stale równa 1 na zbiorze pokrytym wyłącznie przez $\widetilde{W}_k$, zaś znika na $\widetilde{W}_k^c$. Wybierzmy $x \in X$. Ponieważ dla każdego $k = 0, \dots, n-1$ można znaleźć funkcję $g_k \in T^k \mathcal{F}$, dla której wartość $g_k(x)$ jest nie mniejsza niż $\frac{1}{r}$, punkt $(x, \frac{1}{2r})$ nie należy do $\bigcap_{g \in \mathbf{T}^k \mathcal{F}} U_{>g}^\varepsilon$ dla

żadnego k . Zatem zbiór $U \in \mathcal{U}'$ zawierający $(x, \frac{1}{2r})$ spełnia warunek (2.20) i wyznacza pewien element pokrycia $\mathcal{W}^n$. Ponieważ dla $k = 0, \dots, n-1$ funkcja $g_k(x)$ jest dodatnia (nie mniejsza niż $\frac{1}{r}$), element ten zawiera x . Z dowolności x wynika, że zbiory postaci $\bigcap_{k=0}^{n-1} \widehat{W}_k$ tworzą pokrycie przestrzeni X , zatem

$$N(\mathcal{W}^n) \leq N(\mathcal{U}_{\mathcal{F}^n}^\varepsilon).$$

□

Poniższy wynik jest odpowiednikiem stwierdzenia 2.2.13.

Stwierdzenie 2.3.8 *Jeśli dla każdej funkcji f istnieje funkcja niezmiennicza ϕ_f taka, że*

$$\lim_{n \rightarrow \infty} \sup_{x \in X} |\mathbf{T}^n f(x) - \phi_f(x)| = 0,$$

to $h_{\text{top}}(\mathbf{T}) = 0$.

Dowód Ustalmy $\varepsilon > 0$ oraz rodzinę $\mathcal{F}$ funkcji ciągłych, i połóżmy

$$\Phi(\mathcal{F}) = \{\phi_f : f \in \mathcal{F}\}.$$

Istnieje wówczas liczba naturalna N taka, że dla każdej funkcji $f \in \mathcal{F}$ i wszystkich $n \geq N$ zachodzi

$$\sup_{x \in X} |\mathbf{T}^n f(x) - \phi_f(x)| < \frac{\varepsilon}{3}.$$

Wtedy dla dowolnej pary x, y elementów przestrzeni X mamy

$$\begin{aligned} |\mathbf{T}^n f(x) - \mathbf{T}^n f(y)| &\leq |\mathbf{T}^n f(x) - \phi_f(x)| + |\phi_f(x) - \phi_f(y)| \\ &\quad + |\phi_f(y) - \mathbf{T}^n f(y)| \\ &< \frac{2\varepsilon}{3} + |\phi_f(x) - \phi_f(y)|. \end{aligned}$$

Zatem dla każdego $n \geq N$ prawdziwa jest nierówność

$$s(d_{\mathcal{F}^n}, \varepsilon) \leq s(d_{\mathcal{F}^N \cup \Phi(\mathcal{F})}, \frac{\varepsilon}{3}),$$

przy czym prawa strona nierówności nie zależy od n . □

Wzorując się na miarowej definicji 2.2.10, możemy określić faktor topologiczny operatora Markowa, zastępując we wspomnianej definicji żądanie zachowywania miary ciągłością odwzorowania faktorującego. Operatory Markowa są izomorficzne, gdy odwzorowanie to jest homeomorfizmem. Powiemy, że zwarty zbiór Y jest *niezmienniczny* względem $\mathbf{T}$, jeśli $P_{\mathbf{T}}(x, Y^c) = 0$ dla

każdego $x \in Y$ ($P_{\mathbf{T}}$ oznacza prawdopodobieństwo przejścia odpowiadające $\mathbf{T}$). Korzystając ze wzoru (1.1) określamy operator Markowa na $C(Y)$, który możemy taktować jako „obcięcie” operatora $\mathbf{T}$ (ponieważ Y jest domknięty, każda funkcja $f \in C(Y)$ przedłuża się do funkcji ciągłej na X). Pozwolimy sobie pominąć standardowy dowód poniższych faktów.

Stwierdzenie 2.3.9

- (i) *Entropia fatora operatora Markowa $\mathbf{T}$ jest nie większa niż entropia $\mathbf{T}$.*
- (ii) *Operatory izomorficzne mają jednakowe entropie.*
- (iii) *Jeśli Y jest zwartym podzbiorem niezmienniczym X to entropia obcięcia operatora Markowa $\mathbf{T}$ do przestrzeni $C(Y)$ jest nie większa niż entropia $\mathbf{T}$ na $C(X)$.*
- (iv) *Dla każdego $k \in \mathbb{N}$ zachodzi $h_{\text{top}}(\mathbf{T}^k) = k h_{\text{top}}(\mathbf{T})$.*

Przedstawimy teraz dwa inne sposoby definiowania entropii operatora Markowa na $C(X)$. W obu przypadkach, jako entropię operatora przyjmuje się klasyczną entropię topologiczną pewnej transformacji ciągłej, w naturalny sposób związanej z operatorem: entropię „shift”-u na przestrzeni trajektorii lub entropię operatora sprzężonego na przestrzeni $\mathcal{M}(X)$ wszystkich probabilistycznych miar Radona. Poniższe przykłady wykazują, że żadna z tych metod nie spełnia naszych oczekiwań.

Przykład 2.3.10 Niech X będzie zbiorem n elementowym. Oznaczmy przez p miarę probabilistyczną równo rozłożoną na X i określmy operator Markowa $\mathbf{T}$ wzorem

$$\mathbf{T}f(x) = \int_X f dp = \frac{1}{n} \sum_{y \in X} f(y).$$

Obraz każdej funkcji jest funkcją stałą, więc oczekujemy, że entropia $\mathbf{T}$ będzie równa zero. Istotnie, stwierdzenie 2.3.8 gwarantuje, że $h_{\text{top}}(\mathbf{T}) = 0$. Z drugiej strony, układ symboliczny zdefiniowany na zbiorze trajektorii tego operatora jest pełnym układem symbolicznym (*full shift*) nad alfabetem n -elementowym, więc jego entropia wynosi $\log n$. $\square$

Przykład 2.3.11 Rozważmy jednostronny pełny układ symboliczny (X, T) nad alfabetem $\{0, 1\}$. Niech $(a_n)_{n \in \mathbb{N}}$ będzie dowolnym ciągiem o wyrazach z odcinka $[0, 1]$. Dla każdej liczby naturalnej n określmy miarę $P_{a_n} = \{a_n, 1 - a_n\}$ na $\{0, 1\}$ i niech $\mu_{(a_n)} = \prod P_{a_n}$ będzie miarą produktową na X . Nietrudno sprawdzić, że

$$T^* \mu_{(a_n)} = \mu_{(a_{n+1})}.$$

Jako że $(a_n)_{n \in \mathbb{N}}$ może być dowolnym ciągiem liczb z odcinka $[0, 1]$, układ $(\mathcal{M}(X), T^*)$ zawiera podukład izomorficzny z pełnym układem symbolicznym nad nieskończonym alfabetem $[0, 1]$, czyli ma entropię nieskończoną. Wobec tego entropia topologiczna transformacji T^* nie jest równa klasycznej entropii topologicznej w przypadku punktowym. $\square$

W rzeczywistości, z rezultatów otrzymanych przez E. Glasnera i B. Weissa w [G-W] wynika, że dodatniość entropii T wymusza nieskończoną entropię sprzężenia T^* (ci sami autorzy udowodnili wcześniej, że entropia T^* znika, gdy T ma entropię zerową). Naturalne jest pytanie, czy entropia dowolnego operatora Markowa $\mathbf{T}$ jest realizowana przez obcięcie transformacji sprzężonej $\mathbf{T}^*$ do najmniejszego zwartego zbioru niezmienniczego zawierającego wszystkie miary Diraca.

2.3.2 Entropia Langevina-Walczaka

W pracy [L-W] znaleźć można inną definicję entropii topologicznej, bazującą na rozkładach jedności złożonych z nieujemnych funkcji ciągłych. Dla takiego rozkładu $\mathcal{F}$ wprowadza się oznaczenie

$$N_{[\text{L-W}]}(\mathcal{F}) = \sum_{f \in \mathcal{F}} \sup_{x \in X} f(x),$$

zaś jako połączenie rozkładów $\mathcal{F}, \mathcal{G}$ przyjmuje się rozkład jedności $\mathcal{F} \cdot \mathcal{G} = \{f \cdot g : f \in \mathcal{F}, g \in \mathcal{G}\}$. Następnie definiuje się

- (i) $H_{[\text{L-W}]}(\mathcal{F}) = \log N_{[\text{L-W}]}(\mathcal{F})$,
- (ii) $h_{[\text{L-W}]}(\mathbf{T}, \mathcal{F}) = \limsup_{n \rightarrow \infty} \frac{1}{n} H_{[\text{L-W}]}(\mathcal{F}^{n\bullet})$, gdzie $\mathcal{F}^{n\bullet}$ oznacza połączenie (przez mnożenie funkcji) rozkładów $\mathbf{T}^k \mathcal{F}$ przy $k = 0, 1, \dots, n-1$ ($\mathbf{T} \mathcal{F}$ oznacza jak zwykle rodzinę obrazów $\{\mathbf{T}f : f \in \mathcal{F}\}$, która w tym przypadku jest też rozkładem jedności).

W pracy [L-W] entropia operatora $\mathbf{T}$ została zdefiniowana jako kres górny, po wszystkich rozkładach jedności $\mathcal{F}$, wyrażenia $\liminf_{k \rightarrow \infty} \frac{1}{k} h_{[\text{L-W}]}(\mathbf{T}^k, \mathcal{F})$. Wydaje się jednak, że jest to definicja niepoprawna: liczba $N_{[\text{L-W}]}(\mathcal{F})$ jest zawsze szacowana z góry przez liczbę $\mathcal{F}$, zatem $h_{[\text{L-W}]}(\mathbf{T}^k, \mathcal{F})$ nie przekracza logarytmu z tej liczby, bez względu na wartość k . Granica dolna da więc w wyniku zero, czego nie zmieni już obliczany na końcu kres górny. Spośród możliwych prób poprawienia definicji wybieramy najprostszą, opuszczenie $\liminf_k \frac{1}{k}$, przyjmując

- (iii) $h_{[\text{L-W}]}(\mathbf{T}) = \sup_{\mathcal{F}} h_{[\text{L-W}]}(\mathbf{T}, \mathcal{F})$.

Modyfikując argumentację przeprowadzoną w [L-W], można uogólnić twierdzenie, sformułowane tamże dla polihedronów, na przypadek zwartej przestrzeni Hausdorffa.

Twierdzenie 2.3.12 *Niech X będzie zwartą przestrzenią Hausdorffa, a $\mathbf{T}$ operatorem Markowa na $C(X)$ punktowo generowanym przez transformację ciągłą T . Wówczas*

$$h_{[L-W]}(\mathbf{T}) = h_{\text{top}}(T).$$

Dowód Uzasadnimy najpierw, że $h_{[L-W]}(\mathbf{T}) \geq h_{\text{top}}(T)$. Wykorzystamy w tym celu wielkość h_3 z poprzedniego rozdziału. Niech $\mathcal{G}$ będzie ustaloną rodziną funkcji ciągłych, a $\mathcal{V}$ pokryciem odcinka $[0, 1]$ przedziałami otwartymi o średnicach mniejszych od ε . Oznaczmy przez $\mathcal{W}$ pokrycie wybrane z $\mathcal{G}^{-1}(\mathcal{V})$ o minimalnej liczności i dobierzmy do $\mathcal{W}$ rozkład jedności $\mathcal{F}$ spełniający tezę lematu 2.3.7. Niech D będzie zbiorem $(d_{\mathcal{G}^n}, \varepsilon)$ -rozdzielonym o maksymalnej liczności ($\mathcal{G}^n$ oznacza tu połączenie przez sumy mnogościowe rodzin). Ustalmy $x \in D$ i znajdziemy funkcję $f = f_{i_1} \cdot (f_{i_2} \circ T) \cdot \dots \cdot (f_{i_n} \circ T^{n-1})$ należącą do $\mathcal{F}^{n\bullet}$, dla której $f(x) > 0$. Dla każdego innego punktu $y \in D$, przynajmniej jedna z wartości $f_{i_j}(T^{j-1}y)$ musi być równa 0, więc $f(y) = 0$. Zastępując $\sup f$ przez wartość f w punkcie $x \in D$, w którym ta funkcja jest dodatnia (lub przez zero, gdy f jest stale równa zero na D), otrzymujemy

$$N_{[L-W]}(\mathcal{F}^{n\bullet}) \geq \sum_{x \in D} \left(\sum_{\{f \in \mathcal{F}^{n\bullet} : f(x) > 0\}} f(x) \right).$$

Ponieważ $\mathcal{F}^{n\bullet}$ jest rozkładem jedności, suma wewnątrz nawiasu wynosi 1, zatem prawa strona jest równa $s(d_{\mathcal{G}^n}, \varepsilon)$, co pociąga $h_{[L-W]}(\mathbf{T}) \geq h_3(\mathbf{T}) = h_{\text{top}}(T)$.

Dla dowodu odwrotnej nierówności ustalmy rozkład jedności $\mathcal{F} = \{f_i\}_{i=1}^r$, liczbę dodatnią ε i pokrycie $\mathcal{V}$ odcinka $[0, 1]$ przedziałami o średnicach mniejszych niż $\frac{\varepsilon}{r}$. Ustalmy n i rozważmy podpokrycie $\mathcal{U} \subset (\mathcal{F}^{n\bullet})^{-1}(\mathcal{V})$ o minimalnej liczności. Oznaczmy przez $\mathcal{U}(x)$ liczbę tych elementów U , do których należy x . Dla funkcji $f = f_{i_1} \dots (f_{i_n} \circ T^{n-1}) \in \mathcal{F}^{n\bullet}$ i $U \in \mathcal{U}$ połóżmy

$$\begin{aligned} \bar{f}_U(x) &= \frac{\mathbf{1}_U(x)}{\mathcal{U}(x)} \sup_{y \in U} f(y), \\ \bar{f}^\Sigma(x) &= \sum_{f \in \mathcal{F}^{n\bullet}} \sum_{U \in \mathcal{U}} \bar{f}_U(x), \\ \tilde{f}_U(x) &= \frac{\bar{f}_U(x)}{\bar{f}^\Sigma(x)}. \end{aligned}$$

Oznaczmy przez $\overline{\mathcal{F}^{n\bullet}}$ i $\widetilde{\mathcal{F}^{n\bullet}}$ rodziny złożone odpowiednio z funkcji $\bar{f}_U$ i $\tilde{f}_U$. Dla każdego x z ustalonego zbioru $U \in \mathcal{U}$ zachodzi

$$\begin{aligned}\bar{f}_U(x) &= \frac{1}{\mathcal{U}(x)} \sup_{y \in U} f(y) \\ &\leq \frac{1}{\mathcal{U}(x)} \sup_{y \in U} f_{i_1}(y) \dots \sup_{y \in T^{-n+1}U} f_{i_n}(T^{n-1}y) \\ &\leq \frac{1}{\mathcal{U}(x)} \left(f_{i_1}(x) + \frac{\varepsilon}{r} \right) \dots \left(f_{i_n}(T^{n-1}x) + \frac{\varepsilon}{r} \right).\end{aligned}$$

Zatem

$$\bar{f}^\Sigma(x) \leq \sum_{U \in \mathcal{U}} (1 + \varepsilon)^n \cdot \mathbf{1}_U(x) \cdot \frac{1}{\mathcal{U}(x)} = (1 + \varepsilon)^n,$$

co pociąga $N_{[\text{L-w}]}(\overline{\mathcal{F}^{n\bullet}}) \leq (1 + \varepsilon)^n \cdot N_{[\text{L-w}]}(\widetilde{\mathcal{F}^{n\bullet}})$. Zauważmy dalej, że funkcja $\tilde{f}_U$ osiąga największą wartość na zbiorze punktów pokrytych wyłącznie przez U . Wybierając z każdego U dowolny punkt x_U , otrzymujemy

$$N_{[\text{L-w}]}(\widetilde{\mathcal{F}^{n\bullet}}) = \sum_{U \in \mathcal{U}} \sum_{f \in \mathcal{F}^{n\bullet}} \tilde{f}_U(x_U).$$

Łatwo widać, że sumując najpierw po f , a potem po U , otrzymujemy licznosc pokrycia $\mathcal{U}$, czyli $N((\mathcal{F}^{n\bullet})^{-1}(\mathcal{V}))$. Jest jasne, że $N_{[\text{L-w}]}(\mathcal{F}^{n\bullet}) \leq N_{[\text{L-w}]}(\overline{\mathcal{F}^{n\bullet}})$, więc

$$N_{[\text{L-w}]}(\mathcal{F}^{n\bullet}) \leq (1 + \varepsilon)^n N((\mathcal{F}^{n\bullet})^{-1}(\mathcal{V})).$$

Dowód kończymy biorąc odpowiednie granice górne i suprema. $\square$

Zwróćmy uwagę, że obie części dowodu korzystają w istotny sposób z faktu, że $\mathbf{T}$ jest punktowo generowany. Dlatego, mimo użycia w jednej z nich definicji h_3 , pytanie o relacje między $h_{[\text{L-w}]}$ a entropią topologiczną podaną w poprzednim podrozdziale pozostaje na razie otwarte.

2.4 Relacja między entropią miarową a topologiczną

Przyjmując definicję entropii topologicznej z podrozdziału 2.3.1, jesteśmy w stanie udowodnić następujące twierdzenie. Dowód jest stosowną modyfikacją znanego dowodu zasady wariacyjnej podanego przez Misiurewicza.

Twierdzenie 2.4.1 *Niech X będzie zwartą przestrzenią Hausdorffa, a $\mathbf{T}$ operatorem Markowa na przestrzeni $C(X)$. Dla każdej niezmienniczej miary probabilistycznej Radona μ na X zachodzi nierówność*

$$h_\mu(\mathbf{T}) \leq h_{\text{top}}(\mathbf{T}).$$

Dowód Będziemy korzystać z jawnej postaci entropii, zdefiniowanej w podrozdziale 2.2. Wobec gęstości zbioru wszystkich funkcji ciągłych w $L^\infty(X, \mu)$ oraz na mocy wzoru (2.12), entropia miarowa jest realizowana na rodzinach $\mathcal{F}$ składających się z funkcji ciągłych. Niech $\mathcal{F}$ będzie rodziną funkcji należących do $C_I(X)$ o liczności r . Ustalmy dodatnią liczbę ε , dla której spełniona jest nierówność $2r\varepsilon \log(2r\varepsilon) < \frac{1}{2^r}$. Dla każdego zbioru $A \in \mathcal{A}_{\mathcal{F}}$ oznaczmy przez $\mathcal{F}_{\geq A}$ rodzinę tych elementów $\mathcal{F}$, dla których $f(x) \geq t$, o ile $(x, t) \in A$. Analogicznie, $\mathcal{F}_{< A}$ jest zbiorem wszystkich funkcji z $\mathcal{F}$, dla których $f(x) < t$, pod warunkiem, że $(x, t) \in A$. Nietrudno zauważyć, że suma mnogościowa $\mathcal{F}_{< A}$ i $\mathcal{F}_{\geq A}$ stanowi całą rodzinę $\mathcal{F}$. Niech $\mathcal{B}_{\mathcal{F}}^\varepsilon$ będzie rozbiem przestrzeni $X \times [0, 1]$ złożonym ze zwartych zbiorów postaci

$$B_A = B_{A, \mathcal{F}}^\varepsilon = \left\{ (x, t) : (\forall f \in \mathcal{F}_{\geq A}) t \leq f(x) - \varepsilon \right\} \cap \left\{ (x, t) : (\forall f \in \mathcal{F}_{< A}) t \geq f(x) + \varepsilon \right\},$$

gdzie A należy do $\mathcal{A}_{\mathcal{F}}$, oraz zbioru otwartego

$$\tilde{B} = \tilde{B}_{\mathcal{F}}^\varepsilon = \bigcap_{A \in \mathcal{A}_{\mathcal{F}}} B_A^c = \{ (x, t) : (\exists f \in \mathcal{F}) |f(x) - t| < \varepsilon \}.$$

Zauważmy, że $A \cap B_{A'} = B_A$ wtedy i tylko wtedy, gdy $A = A'$. W przeciwnym wypadku, przekrój $A \cap B_{A'}$ jest pusty. Zatem,

$$\begin{aligned} H_\mu(\mathcal{A}_{\mathcal{F}} | \mathcal{B}_{\mathcal{F}}^\varepsilon) &= H_\mu(\mathcal{A}_{\mathcal{F}} \vee \mathcal{B}_{\mathcal{F}}^\varepsilon) - H_\mu(\mathcal{B}_{\mathcal{F}}^\varepsilon) \\ &= - \sum_{A \in \mathcal{A}_{\mathcal{F}}} \hat{\mu}(A \cap \tilde{B}) \log(\hat{\mu}(A \cap \tilde{B})) \end{aligned}$$

(przypomnijmy, że $\hat{\mu}$ oznacza produkt miary μ z miarą Lebesgue'a). Jako że $\hat{\mu}(\tilde{B}) < 2r\varepsilon$, a $\mathcal{A}_{\mathcal{F}}$ zawiera co najwyżej 2^r elementów, otrzymujemy

$$H_\mu(\mathcal{A}_{\mathcal{F}} | \mathcal{B}_{\mathcal{F}}^\varepsilon) < 1. \quad (2.21)$$

Podobnie, dla każdego k zachodzi

$$H_\mu(\mathcal{A}_{\mathbf{T}^k \mathcal{F}} | \mathcal{B}_{\mathbf{T}^k \mathcal{F}}^\varepsilon) < 1, \quad (2.22)$$

gdyż ε zależy jedynie od liczności rodziny $\mathcal{F}$. Aby skrócić zapis, połączenie rozbić $\bigvee_{k=0}^{n-1} \mathcal{B}_{\mathbf{T}^k \mathcal{F}}^\varepsilon$ będziemy oznaczać symbolem $\mathcal{B}^n$ (warto zwrócić uwagę, że to rozbiecie jest różne od $\mathcal{B}_{\mathcal{F}^n}^\varepsilon$). Korzystając ze wzorów (2.21) i (2.22) możemy napisać:

$$H_\mu(\mathcal{A}_{\mathcal{F}^n} | \mathcal{B}^n) \leq \sum_{k=0}^{n-1} H_\mu(\mathcal{A}_{\mathbf{T}^k \mathcal{F}} | \mathcal{B}_{\mathbf{T}^k \mathcal{F}}^\varepsilon) < n.$$

Stąd

$$H_{\hat{\mu}}(\mathcal{A}_{\mathcal{F}^n}) < H_{\hat{\mu}}(\mathcal{B}^n) + n.$$

Ponieważ entropia rozbitcia nie przekracza nigdy logarytmu z jego liczności, wystarczy oszacować liczbę elementów $\mathcal{B}^n$.

Niech $\mathcal{U}'$ będzie minimalnym podpokryciem wybranym z $\mathcal{U}_{\mathcal{F}^n}^{\varepsilon}$. Jest jasne, że

$$\text{card } \mathcal{B}^n \leq \sum_{U \in \mathcal{U}'} \text{card } \{B \in \mathcal{B}^n : B \cap U \neq \emptyset\}.$$

Każdy zbiór $U \in \mathcal{U}_{\mathcal{F}^n}^{\varepsilon}$ jest postaci $\bigcap_{k=0}^{n-1} U_k$, gdzie $U_k \in \mathcal{U}_{\mathbf{T}^k \mathcal{F}}^{\varepsilon}$. Ponadto, dla każdego U_k istnieje $A \in \mathcal{A}_{\mathbf{T}^k \mathcal{F}}$, dla którego suma zbiorów $B_{A, \mathbf{T}^k \mathcal{F}}^{\varepsilon}$ i $\tilde{B}_{\mathbf{T}^k \mathcal{F}}^{\varepsilon}$ zawiera U_k . Zatem

$$\text{card } \{B \in \mathcal{B}^n : B \cap U \neq \emptyset\} \leq 2^n,$$

co pociąga

$$\text{card } \mathcal{B}^n \leq 2^n \cdot N(\mathcal{U}_{\mathcal{F}^n}^{\varepsilon}).$$

W rezultacie mamy

$$H_{\hat{\mu}}(\mathcal{A}_{\mathcal{F}^n}) < \log N(\mathcal{U}_{\mathcal{F}^n}^{\varepsilon}) + n \log 2 + n,$$

czyli

$$h_{\mu}(\mathbf{T}) \leq h_{\text{top}}(\mathbf{T}) + \log 2 + 1.$$

Stosując taką samą argumentację dla $\mathbf{T}^k$ w miejsce $\mathbf{T}$ oraz korzystając ze stwierdzeń 2.2.12 i 2.3.9(iv), otrzymujemy

$$h_{\mu}(\mathbf{T}) = \frac{1}{k} h_{\mu}(\mathbf{T}^k) \leq \frac{1}{k} (h_{\text{top}}(\mathbf{T}^k) + \log 2 + 1) \leq h_{\text{top}}(\mathbf{T}) + \frac{\log 2 + 1}{k}$$

dla dowolnej liczby naturalnej k . □

Brak pełnej podaddytywności ciągu $H_{\mu}(\mathcal{F}^n)$ powoduje liczne trudności przy próbach dowodzenia nierówności odwrotnej. W klasycznych dowodach konstruuje się niezmienniczą miarę μ , realizującą entropię bliską entropii topologicznej, jako granicę ciągu miar μ_n , dla których wartości $H_{\mu_n}(\mathcal{F}^n)$ rosną wraz ze wzrostem n . Dzięki podaddytywności, wartości $H_{\mu_n}(\mathcal{F}^n)$ pozostają duże dla dalekich indeksów $m > n$. Ponieważ H_{μ} jest ciągłą (lub górną półciągłą) funkcją miary, wszystkie wyrazy ciągu $H_{\mu}(\mathcal{F}^n)$ są również odpowiednio duże. W naszej sytuacji indeks N , powyżej którego ciąg $H_{\mu_n}(\mathcal{F}^k)$ jest „prawie podaddytywny”, zależy od μ_n i może być znacznie większy niż n . Alternatywne funkcje $H'_{\mu}(\mathcal{F}^n)$ i $H''_{\mu}(\mathcal{F}^n)$ są podaddytywne, ale nie muszą być górną półciągłe.

Rozdział 3

Rodziny operatorów Markowa

3.1 Układy przejścia. Trajektorie

Aby nie mnożyć oznaczeń, przyjmujemy w całym rozdziale konwencję oznaczania tą samą literą danego operatora Markowa i jego prawdopodobieństwa przejścia.

Niech X będzie zwartą przestrzenią metryczną. Rozważmy układy przejścia $\varphi, \psi : X \rightarrow 2^X$. Przez $\varphi \cup \psi$ będziemy oznaczać układ przejścia dany wzorem

$$(\varphi \cup \psi)(x) = \varphi(x) \cup \psi(x).$$

Oczywiście, $\varphi \prec \varphi \cup \psi$ i $\psi \prec \varphi \cup \psi$. Ponadto, dla wypukłej kombinacji prawdopodobieństw przejścia $P = \alpha P_1 + (1 - \alpha)P_2$, $0 < \alpha < 1$, mamy $\varphi_P = \varphi_{P_1} \cup \varphi_{P_2}$. Gdy φ i ψ są układami przejścia indukowanymi przez ciągłe przekształcenia $S, T : X \rightarrow X$, będziemy pisać $S \cup T$ zamiast $\varphi \cup \psi$.

Interesującą klasę układów przejścia stanowią układy, dla których przekształcenie $\varphi : X \rightarrow 2^X$ jest ciągłe. Wówczas odwzorowanie $\bar{\varphi} : 2^X \rightarrow 2^X$ jest również ciągłe, więc para $(2^X, \bar{\varphi})$ jest układem dynamicznym w klasycznym sensie (układy takie były rozważane w [D]). Dzięki ciągłości można też opuścić domknięcie w definicji $\bar{\varphi}$ (suma $\bigcup_{x \in F} \varphi(x)$ we wzorze (1.2) jest już zbiorem domkniętym) — jak zobaczymy w przykładzie 3.1.4 własność komutowania układów przejścia zyskuje wtedy istotne znaczenie.

W przypadku przestrzeni skończonych badanie układów przejścia sprowadza się do studiowania macierzy, podobnie jak w przypadku prawdopodobieństw przejścia. Z każdym układem przejścia na przestrzeni n -elementowej jest bowiem skojarzona macierz kwadratowa $[\varphi_{ij}]_{n \times n}$ o elementach 0 i 1, dla której $\varphi_{ij} = 1$ wtedy i tylko wtedy, gdy element o numerze j (w pewnej ustalonej numeracji) należy do obrazu elementu i -tego. Zatem macierz takiego dyskretnego układu przejścia otrzymuje się z macierzy stochastycznej

przez zastąpienie dodatnich elementów jedynkami. Składanie układów przejścia odpowiada następującemu „mnożeniu” macierzy:

$$(\varphi \odot \psi)_{ij} = \text{sgn}((\varphi\psi)_{ij}).$$

Zauważmy też, że wtedy $(\varphi \cup \psi)_{ij} = \varphi_{ij} \vee \psi_{ij}$.

Przykład 3.1.1 Niech C_n będzie grupą cykliczną liczb całkowitych modulo n . Oznaczmy przez $I : C_n \rightarrow C_n$ przekształcenie identycznościowe, a przez $R : C_n \rightarrow C_n$ operację dodawania jedynki, t.j. $R(p) = p + 1 \pmod{n}$. Gdy $n = 3$, układy $I \cup R$, $I \cup R^2$, $R \cup R^2$ odpowiadają macierzom

$$\begin{aligned} I \cup R &\sim \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} & I \cup R^2 &\sim \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \\ R \cup R^2 &\sim \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \end{aligned}$$

a $\odot$ -iloczyn dowolnych dwóch spośród nich jest macierzą złożoną z samych jedynek. $\square$

Definicja 3.1.2 *Trajektorią układu przejścia φ nazywamy ciąg $(x^n) \in X^{\mathbb{N}_0}$, dla którego $x^{n+1} \in \varphi(x^n)$.*

Trajektorią skończonej rodziny układów przejścia $\varphi_1, \dots, \varphi_m$ nazywamy macierz $[x^{n_1, \dots, n_m}] \in X^{\mathbb{N}_0 \times \dots \times \mathbb{N}_0}$ wymiaru $m \times \infty$, dla której zachodzi:

$$\begin{aligned} x^{n_1+1, n_2, \dots, n_m} &\in \varphi_1(x^{n_1, \dots, n_m}) \\ x^{n_1, n_2+1, \dots, n_m} &\in \varphi_2(x^{n_1, \dots, n_m}) \\ &\vdots \\ x^{n_1, n_2, \dots, n_m+1} &\in \varphi_m(x^{n_1, \dots, n_m}). \end{aligned}$$

Konstrukcję zbioru trajektorii dla komutującej pary ciągłych układów przejścia można znaleźć w [I3] (zbiór Ω_∞ w dowodzie twierdzenia 3). Dla skończonej rodziny operatorów Markowa, przez trajektorię będziemy rozumieć trajektorię indukowanej rodziny układów przejścia. Jak wiadomo z [D-I], jeśli $\mathbf{S}$ i $\mathbf{T}$ są komutującymi prawdopodobieństwami przejścia, a P_0 miarą probabilistyczną na X , to istnieje miara probabilistyczna μ na $X^{\mathbb{N}_0 \times \mathbb{N}_0}$ określająca proces zadany przez $\mathbf{S}$ i $\mathbf{T}$. Proces ten (miara) jest stacjonarny, gdy P_0 jest miarą niezmienniczą względem operatorów $\mathbf{S}$ i $\mathbf{T}$.

Twierdzenie 3.1.3 *Miara μ zbioru wszystkich trajektorii pary S, T jest równa 1.*

Dowód Dla ustalonych $k_0, l_0 \in \mathbb{N}_0$ rozważmy zbiór $C_h(k_0, l_0)$ wszystkich punktów $(x^{k,l}) \in X^{\mathbb{N}_0 \times \mathbb{N}_0}$ spełniających warunek $x^{k_0+1, l_0} \in \varphi_T(x^{k_0, l_0})$. Zgodnie z twierdzeniem 1 z pracy [D-I] miara tego zbioru wynosi

$$\begin{aligned} \int_X \int_X \dots \int_X \int_{\varphi_T(x^{k_0, l_0})} & T(x^{k_0, l_0}, dx^{k_0+1, l_0}) \dots T(x^{0, l_0}, dx^{1, l_0}) \\ & \times S(x^{0, l_0-1}, dx^{0, l_0}) \dots S(x^{0, 0}, dx^{0, 1}) \\ & \times P_0(dx^{0, 0}), \end{aligned}$$

więc jest równa 1, bo $\varphi_T(x^{k_0, l_0})$ jest nośnikiem miary $T(x^{k_0, l_0}, \cdot)$. Podobnie, miara zbioru $C_v(k_0, l_0) = \{(x^{k,l}) : x^{k_0, l_0+1} \in \varphi_S(x^{k_0, l_0})\}$ także jest równa 1. Zbiór wszystkich trajektorii pary S, T można zapisać jako przeliczalny przekrój

$$\bigcap_{k \in \mathbb{N}_0} \bigcap_{l \in \mathbb{N}_0} [C_h(k, l) \cap C_v(k, l)],$$

zatem ma on miarę 1. $\square$

Jak wskazuje poniższy przykład, komutująca (w sensie mnożenia $\odot$) para dowolnych układów przejścia może nie posiadać trajektorii.

Przykład 3.1.4 Rozważmy przestrzeń

$$Y = \{p\} \cup \{a, a_1, a_2, \dots\} \cup [0, \pi],$$

gdzie $p, a, a_1, a_2, \dots$ są różnymi liczbami z odcinka $[0, \pi]$, p jest punktem izolowanym, zaś ciąg $a_1, a_2, \dots$ zbiega do punktu a . Oczywiście, Y jest zbiorem zwartym. Zdefiniujmy układy przejścia φ i ψ na Y w następujący sposób:

$$\begin{aligned} \varphi(y) &= \begin{cases} \{a, a_1, a_2, \dots\} & \text{dla } y = p \\ \{\frac{k}{n} : k \in \mathbb{N}\} \cap [0, \pi] & \text{dla } y = a_n \\ \{0\} & \text{dla } y = a \\ \{p\} & \text{dla } y \in [0, \pi] \end{cases} \\ \psi(y) &= \begin{cases} \{a, a_1, a_2, \dots\} & \text{dla } y = p \\ \{\pi - \frac{k}{n} : k \in \mathbb{N}\} \cap [0, \pi] & \text{dla } y = a_n \\ \{\pi\} & \text{dla } y = a \\ \{p\} & \text{dla } y \in [0, \pi] \end{cases}. \end{aligned}$$

Przekształcenia φ i ψ są dolnie półciągłe (jedynym punktem nieciągłości jest punkt a). Ponadto φ i ψ komutują, gdyż

$$(\varphi \circ \psi)(y) = (\psi \circ \varphi)(y) = \begin{cases} [0, \pi], & \text{dla } y = p \\ \{p\}, & \text{dla } y \in \{a, a_1, a_2, \dots\} \\ \{a, a_1, a_2, \dots\}, & \text{dla } y \in [0, \pi] \end{cases}.$$

Udowodnimy, że para φ, ψ nie posiada trajektorii. Ponieważ dla każdego $y \in Y$ jeden z obrazów $\varphi(y), \varphi^2(y), \varphi^3(y)$ jest równy $\{p\}$, wystarczy ograniczyć się do trajektorii o początku w p . Zauważmy, że dla $x^{0,0} = p$ otrzymujemy $x^{1,0} \in \varphi(x^{0,0}) = \{a, a_1, a_2, \dots\}$ oraz $x^{0,1} \in \psi(x^{0,0}) = \{a, a_1, a_2, \dots\}$. Zatem, jednocześnie $x^{1,1}$ ma należeć do zbioru $\varphi(x^{0,1})$, złożonego z liczb wymiernych, i do $\psi(x^{1,0})$ zawierającego wyłącznie liczby niewymierne. Sprzeczność. $\square$

3.2 Niezgodność komutowania

Rozważmy rodzinę $T_1, \dots, T_m$ komutujących operatorów Markowa. Zgodnie ze wzorem (1.3), odpowiadające im układy przejścia również komutują. Odwrotne stwierdzenie na ogół nie jest prawdziwe. Istotnie, wobec twierdzenia 3.1.3 jest jasne, że para φ, ψ skonstruowana w przykładzie 3.1.4 nie jest związana z żadną parą komutujących operatorów Markowa. Poniżej podajemy mocniejszy przykład *ciągłych* komutujących układów przejścia nie związanych z komutującymi operatorami Markowa. Co więcej, w przykładzie tym żadne układy wpisane nie pochodzą od komutujących operatorów Markowa. Wobec tego, twierdzenia wyrażane w języku układów przejścia są istotnie ogólniejsze niż formułowane dla operatorów Markowa. (Na odwrót, kontrprzykłady polegające na konstrukcji operatorów Markowa są automatycznie kontrprzykładami w teorii układów przejścia.)

Przykład 3.2.1 Niech $X = C_3 \times C_3$. Zdefiniujmy układ przejścia na X wzorem:

$$\varphi = \begin{cases} I \times (I \cup R) & \text{na } \{0\} \times C_3 \\ I \times (I \cup R^2) & \text{na } \{1\} \times C_3 \\ I \times (R \cup R^2) & \text{na } \{2\} \times C_3 \end{cases}$$

$$\psi = \begin{cases} R \times (I \cup R) & \text{na } \{0\} \times C_3 \\ R \times (I \cup R^2) & \text{na } \{1\} \times C_3 \\ R \times (R \cup R^2) & \text{na } \{2\} \times C_3 \end{cases}.$$

Ponieważ I i R komutują, a złożenie dowolnych dwóch układów spośród $I \cup R, I \cup R^2, R \cup R^2$ jest równe $I \cup R \cup R^2$, układy φ i ψ również komutują. Udowodnimy, że żadne macierze stochastyczne odpowiadające układom wpisanym w φ i ψ nie mogą komutować.

Przypuśćmy, że takie macierze istnieją i oznaczmy je przez T_φ i T_ψ . Ponumerujmy elementy przestrzeni X tak, by punkt (p, q) miał numer $3p + q$. Wówczas dla pewnych $a, b, c, \dots, z \in [0, 1]$ spełniających $a + a' = 1, b + b' =$

$1, \dots, z + z' = 1$, możemy zapisać:

$$\mathbf{T}_\varphi = \left[\begin{array}{ccc|ccc|ccc} a & a' & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & b & b' & 0 & 0 & 0 & 0 & 0 & 0 \\ c' & 0 & c & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & d & 0 & d' & 0 & 0 & 0 \\ 0 & 0 & 0 & e' & e & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & f' & f & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & g & g' \\ 0 & 0 & 0 & 0 & 0 & 0 & h' & 0 & h \\ 0 & 0 & 0 & 0 & 0 & 0 & i & i' & 0 \end{array} \right],$$

$$\mathbf{T}_\psi = \left[\begin{array}{ccc|ccc|ccc} 0 & 0 & 0 & r & r' & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & s & s' & 0 & 0 & 0 \\ 0 & 0 & 0 & t' & 0 & t & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & u & 0 & u' \\ 0 & 0 & 0 & 0 & 0 & 0 & v' & v & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & w' & w \\ \hline 0 & x & x' & 0 & 0 & 0 & 0 & 0 & 0 \\ y' & 0 & y & 0 & 0 & 0 & 0 & 0 & 0 \\ z & z' & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right].$$

Aby macierze $\mathbf{T}_\varphi$ i $\mathbf{T}_\psi$ komutowały, niezbędne jest, by:

$$\begin{bmatrix} a & a' & 0 \\ 0 & b & b' \\ c' & 0 & c \end{bmatrix} \cdot \begin{bmatrix} r & r' & 0 \\ 0 & s & s' \\ t' & 0 & t \end{bmatrix} = \begin{bmatrix} r & r' & 0 \\ 0 & s & s' \\ t' & 0 & t \end{bmatrix} \cdot \begin{bmatrix} d & 0 & d' \\ e' & e & 0 \\ 0 & f' & f \end{bmatrix},$$

czyli by

$$\begin{bmatrix} ar & ar' + a's & a's' \\ b't' & bs & bs' + b't \\ c'r + ct' & c'r' & ct \end{bmatrix} = \begin{bmatrix} rd + r'e' & r'e & rd' \\ se' & se + s'f' & s'f \\ t'd & tf' & tf + t'd' \end{bmatrix}. \quad (3.1)$$

Otrzymujemy równości

$$\begin{aligned} ar &= rd + r'e' & ar' + a's &= r'e \\ bs &= se + s'f' & bs' + b't &= s'f \\ ct &= tf + t'd' & ct' + c'r &= t'd, \end{aligned}$$

a następnie

$$\begin{aligned} r(a - d) &= r'e' & r'(e - a) &= a's \\ s(b - e) &= s'f' & s'(f - b) &= b't \\ t(c - f) &= t'd' & t'(d - c) &= c'r. \end{aligned} \quad (3.2)$$

Przypuśćmy, że $r, s, t \in (0, 1)$. Ponieważ wszystkie współczynniki są nieujemne, mamy $a \geq d$, $b \geq e$, $c \geq f$ oraz $e \geq a$, $f \geq b$, $d \geq c$. Zatem, $a \geq d \geq c \geq f \geq b \geq e \geq a$, co pociąga równość współczynników $a, \dots, f$. Wobec tego, lewe strony nierówności (3.2) są zerami, więc musi zachodzić $a = b = \dots = f = 1$. Wtedy

$$\begin{aligned} \begin{bmatrix} u & 0 & u' \\ v' & v & 0 \\ 0 & w' & w \end{bmatrix} &= \begin{bmatrix} u & 0 & u' \\ v' & v & 0 \\ 0 & w' & w \end{bmatrix} \cdot \begin{bmatrix} 0 & g & g' \\ h' & 0 & h \\ i & i' & 0 \end{bmatrix} \\ &= \begin{bmatrix} u'i & ug + u'i' & ug' \\ vh' & v'g & v'g' + vh \\ w'h' + wi & wi' & w'h \end{bmatrix}. \end{aligned}$$

Ponieważ, w szczególności,

$$0 = ug + u'i', \quad (3.3)$$

musimy mieć $u = 0$ lub $g = 0$. Ale jeśli $u = 0$, to również $i' = 0$, skąd $0 = u = u'i = 1$. Sprzeczność. Z drugiej strony, jeśli $g = 0$, to $v = v'g = 0$ i $1 = v' = vh' = 0$ — ponownie otrzymujemy sprzeczność.

Udowodniliśmy zatem, że macierze Markowa indukujące układy przejścia φ, ψ nie mogą komutować. Aby uzasadnić, że podobnie jest dla układów wpisanych w φ, ψ , wystarczy wykazać, że żaden ze współczynników r, s, t nie może należeć do $\{0, 1\}$. Przypuśćmy, że $r = 0$ (przy dowolnych pozostałych współczynnikach). Wobec (3.2), mamy

$$0 = r(a - d) = r'e' = e',$$

co pociąga $a's = a'$. Zatem $a = 1$ lub $s = 1$. Dla komutowania $\mathbf{T}_\varphi$ i $\mathbf{T}_\psi$ potrzeba, by

$$\begin{bmatrix} du & d'w' & du' + d'w \\ e'u + ev' & ev & e'u' \\ f'v' & f'v + fw' & fw \end{bmatrix} = \begin{bmatrix} u'i & ug + u'i' & ug' \\ vh' & v'g & v'g' + vh \\ w'h' + wi & wi' & w'h \end{bmatrix} \quad (3.4)$$

oraz

$$\begin{bmatrix} gy' + g'z & g'z' & gy \\ hz & h'x + hz' & h'x' \\ i'y' & ix & ix' + i'y \end{bmatrix} = \begin{bmatrix} x'c' & xb & xb' + x'c \\ y'a + yc' & y'a' & yc \\ za & za' + z'b & z'b' \end{bmatrix}. \quad (3.5)$$

Gdy we wzorze (3.5) podstawimy $a = 1$, otrzymamy $h'x + hz' = 0$. Wtedy zachodzi jedna z następujących możliwości:

- (i) $h = 0$ i $x = 0$,

$$(ii) \ h = 1 \text{ i } z = 1,$$

$$(iii) \ x = 0 \text{ i } z = 1.$$

Jeśli zachodzi (i), to z warunku $yc = h'x' = 1$ otrzymujemy $c = y = 1$. Ponadto,

$$\begin{aligned} z &= za = i'y' = 0 \\ b &= za' + z'b = ix = 0 \\ g &= gy = xb' + x'c = 1 \end{aligned}$$

Po podstawieniu do (3.4) mamy:

$$\begin{aligned} v &= ev = v'g = v' \Rightarrow v = \frac{1}{2}, \\ fw &= w'h = 0 \Rightarrow w = 0 \text{ or } f = 0. \end{aligned}$$

Ale jeśli $w = 0$, to $1 = w'h' + wi = f'v' = \frac{1}{2}f'$, co nie jest możliwe. Natomiast gdy $w \neq 0$, wzór $du' + d'w = ug' = 0$ pociąga $d = u = 1$ — sprzeczność z $du = u'i$.

Zakładając prawdziwość (ii), dochodzimy do sprzeczności w (3.4), gdyż jednocześnie $v' = e'u + ev' = vh' = 0$ i $v = ev = v'g = 0$.

Ostatni z warunków, (iii), implikuje $i'y' = za = 1$, a następnie $i = y = 0$, $c' = x'c' = gy' + g'z = 1$ oraz $h = hz = y'a + yc' = 1$, co prowadzi do rozważanego już (ii). Zatem $a \neq 1$.

Rozważmy teraz przypadek $s = 1$. Jako że $s(b - e) = s'f'$ (patrz (3.2)), mamy $b = e = 1$. Z (3.5) otrzymujemy $ix' + i'y = z'b' = 0$, więc ponownie musimy rozważyć trzy przypadki:

$$(i) \ i = 0 \text{ i } y = 0,$$

$$(ii) \ i = 1 \text{ i } x = 1,$$

$$(iii) \ x = 1 \text{ i } y = 0.$$

Zakładając (i) i korzystając z (3.5), uzyskujemy

$$za = i'y' = 1 \Rightarrow a = z = 1,$$

co stoi w sprzeczności z udowodnionym uprzednio faktem, że $a \neq 1$. Jeśli zachodzi (ii), ponownie równość (3.5) pociąga $g'z' = xb = 1$, więc $g = z = 0$. Porównując ten wynik z (3.4), dostajemy jednocześnie $v = ev = v'g = 0$ i $v' = e'u + ev' = vh' = 0$. Sprzeczność. Zauważmy, że (iii) prowadzi do $g = z = 0$ i $i = ix = za' + z'b = 1$, które implikują (ii).

Dowiedliśmy zatem, że $r \neq 0$. Jest jasne, że to samo jest prawdą dla s i t . Pozostaje tylko wykluczyć możliwość, że którykolwiek ze współczynników r, s, t wynosi 1. Przypuśćmy, że to $r = 1$. Wtedy (3.2) pociąga $a = d$ i $a's = 0$. Ponieważ $s \neq 0$, musi zachodzić $a = 1$. Następnie otrzymujemy:

$$\begin{aligned} d &= 1 \\ c' &= c'r = t'(d - c) = t'c' \Rightarrow c't = 0 \stackrel{t \neq 0}{\Rightarrow} c = 1 \\ tf' &= t(c - f) = t'd' = 0 \stackrel{t \neq 0}{\Rightarrow} f = 1. \end{aligned}$$

Jak jednak wynika z (3.4), $ug + u'i' = d'w' = 0$, więc otrzymujemy sprzeczność tak jak w przypadku warunku (3.3).

Zatem udowodniliśmy, że macierze $\mathbf{T}_\varphi$ i $\mathbf{T}_\psi$ nie mogą komutować. $\square$

Zwróćmy uwagę na następujący związek układów przejścia z teorią układów symbolicznych skończonego typu. Wiadomo, że zbiór trajektorii pojedynczego układu przejścia φ na przestrzeni skończonej można utożsamiać z układem symbolicznym skończonego typu. Z drugiej strony, każdy układ skończonego typu można przekodować (przy pomocy tzw. *higher block codes*; patrz np. [L-M]) uzyskując układ symboliczny na trajektoriach układu przejścia. Jak wiemy, taki układ jest nośnikiem miary probabilistycznej μ odpowiadającej procesowi stacjonarnemu zadanemu przez prawdopodobieństwo przejścia $\mathbf{T}$ spełniające $\varphi_{\mathbf{T}} = \varphi$. Zbiór trajektorii skończonej rodziny układów przejścia można utożsamiać z wielowymiarowym układem skończonego typu (o ile zbiór ten nie jest pusty). Powyższy przykład wskazuje, że wielowymiarowy układ skończonego typu może nie posiadać miary niezmienniczej odpowiadającej procesowi Markowa w sensie [D-I].

3.3 Pewien przykład trójki operatorów Markowa

Jak wiadomo z [D-I], w przypadku większych niż dwuelementowe rodzin operatorów Markowa, zbiór trajektorii może być pusty. Przykład komutujących operatorów Markowa $\mathbf{T}$, $\mathbf{S}$, $\mathbf{U}$, jaki podamy poniżej, ma nawet mocniejszą własność: $\mathbf{T}^k$, $\mathbf{S}^k$ i $\mathbf{U}^k$ nie posiadają trajektorii dla żadnej liczby naturalnej k (przykład skonstruowany w [D-I] dopuszczał trajektorie już dla $\mathbf{T}^2$, $\mathbf{S}^2$ i $\mathbf{U}^2$). Operatory te są zadane przez prawdopodobieństwa przejścia związane, jak łatwo sprawdzić, z ciągłymi układami przejścia.

Niech $(X_k)_{k=1}^\infty$ będzie ciągiem zwartych przestrzeni metrycznych. Załóżmy, że na każdej przestrzeni X_k określone jest fellerowskie prawdopodobieństwo przejścia $\mathbf{T}_k$. Wówczas przez $\mathbf{T} = \prod_{k=1}^\infty \mathbf{T}_k$ rozumiemy prawdopo-

dobieństwo przejścia zdefiniowane na przestrzeni produktowej $X = \prod_{k=1}^{\infty} X_k$ tak, że dla każdego $x = (x_k) \in X$ prawdopodobieństwo $\mathbf{T}(x, \cdot)$ jest miarą produktową $\mathbf{T}_1(x_1, \cdot) \times \mathbf{T}_2(x_2, \cdot) \times \dots$

Lemat 3.3.1 *Niech $(X_k)_{k=1}^{\infty}$ będzie ciągiem zwartych przestrzeni metrycznych. Dla każdego $k \in \mathbb{N}$ niech $\mathbf{T}_k, \mathbf{S}_k, \mathbf{U}_k$ będą komutującymi fellerowskimi prawdopodobieństwami przejścia na X_k , dla których trójka $\mathbf{T}_k^k, \mathbf{S}_k^k, \mathbf{U}_k^k$ nie posiada trajektorii. Oznaczmy $X = \prod_{k=1}^{\infty} X_k$ oraz*

$$\mathbf{T} = \prod_{k=1}^{\infty} \mathbf{T}_k, \quad \mathbf{S} = \prod_{k=1}^{\infty} \mathbf{S}_k, \quad \mathbf{U} = \prod_{k=1}^{\infty} \mathbf{U}_k.$$

Wówczas $\mathbf{T}, \mathbf{S}$ i $\mathbf{U}$ są komutującymi fellerowskimi prawdopodobieństwami przejścia zaś trójka $\mathbf{T}^k, \mathbf{S}^k, \mathbf{U}^k$ nie posiada trajektorii dla żadnego $k \in \mathbb{N}$.

Dowód Na początek, zauważmy, że przekształcenie $X \rightarrow \prod_{k=1}^{\infty} \mathcal{M}(X_k)$ (produkt przestrzeni miar probabilistycznych) przypisujące punktowi $(x_1, x_2, \dots)$ ciąg miar $(\mathbf{T}_1(x_1, \cdot), \mathbf{T}_2(x_2, \cdot), \dots)$ jest ciągłe. Korzystając z faktu, że operacja tworzenia miary produktowej z zadanego ciągu miar jest ciągła, otrzymujemy fellerowskość prawdopodobieństwa przejścia $\mathbf{T}$. Podobnie, $\mathbf{S}$ i $\mathbf{U}$ są fellerowskie. Fakt komutowania prawdopodobieństw $\mathbf{T}, \mathbf{S}$ i $\mathbf{U}$ wynika z twierdzenia Fubiniego.

Jest oczywiste, że zbiór trajektorii jest pusty. Wystarczy bowiem zauważyć, że mając trajektorię trójki $\mathbf{T}^k, \mathbf{S}^k, \mathbf{U}^k$, uzyskalibyśmy trajektorie dla $\mathbf{T}_k^k, \mathbf{S}_k^k, \mathbf{U}_k^k$ przez rzutowanie na k -tą oś produktu. $\square$

Lemat 3.3.2 *Niech X będzie zwartą przestrzenią metryczną, a $\mathbf{T}, \mathbf{S}, \mathbf{U}$ trójką komutujących fellerowskich prawdopodobieństw przejścia na X . Dla $k \in \mathbb{N}$ oznaczmy przez $+_k$ operację dodawania modulo k . Załóżmy, że*

1. $X = \bigcup_{i=0}^{k-1} X_i$, gdzie $X_0, \dots, X_{k-1}$ są rozłącznymi domkniętymi podzbiórmi X takimi, że dla każdego $P \in \{\mathbf{T}, \mathbf{S}, \mathbf{U}\}$ układ przejścia φ_P spełnia jeden z warunków:
 - (a) dla każdego $i = 0, \dots, k$ zachodzi $\varphi_P(X_i) \subset X_{i+_k 1}$, lub
 - (b) dla każdego $i = 0, \dots, k$ zachodzi $\varphi_P(X_{i+_k 1}) \subset X_i$;
2. $\mathbf{T}^k, \mathbf{S}^k, \mathbf{U}^k$ nie posiadają trajektorii o punkcie początkowym należącym do X_0 .

Wtedy istnieje zwarta przestrzeń metryzowalna $\widehat{X}$ oraz komutujące prawdopodobieństwa przejścia $\widehat{\mathbf{T}}, \widehat{\mathbf{S}}, \widehat{\mathbf{U}}$ na $\widehat{X}$, dla których trójka $\widehat{\mathbf{T}}^k, \widehat{\mathbf{S}}^k, \widehat{\mathbf{U}}^k$ nie posiada trajektorii.

Dowód Oznaczmy przez $\widehat{X}$ podzbiór przestrzeni X^k złożony z tych punktów $(x_1, \dots, x_k)$, dla których istnieje $m \in [0, k)$ takie, że $x_j \in X_{m+kj}$ dla każdego $j \in [1, k]$. Jest jasne, że $\widehat{X}$ jest zwartą przestrzenią metryzowalną. Niech $\widehat{\mathbf{T}} = \mathbf{T} \times \dots \times \mathbf{T}$, $\widehat{\mathbf{S}} = \mathbf{S} \times \dots \times \mathbf{S}$, $\widehat{\mathbf{U}} = \mathbf{U} \times \dots \times \mathbf{U}$. Z 1(a) i 1(b) wynika, że $\varphi_{\widehat{\mathbf{T}}}(\widehat{X}) \subset \widehat{X}$, $\varphi_{\widehat{\mathbf{S}}}(\widehat{X}) \subset \widehat{X}$, $\varphi_{\widehat{\mathbf{U}}}(\widehat{X}) \subset \widehat{X}$, zatem $\widehat{\mathbf{T}}, \widehat{\mathbf{S}}$ i $\widehat{\mathbf{U}}$ są dobrze określonymi komutującymi fellerowskimi prawdopodobieństwami przejścia na $\widehat{X}$. Przypuśćmy, że trójka $\widehat{\mathbf{T}}^k, \widehat{\mathbf{S}}^k, \widehat{\mathbf{U}}^k$ posiada trajektorię na $\widehat{X}$, i oznaczmy punkt początkowy tej trajektorii przez $x = (x_1, \dots, x_k)$. Jedna ze współrzędnych punktu x , powiedzmy x_j , musi należeć do X_0 . Wtedy jednak, rzut tej trajektorii na j -tą oś jest trajektorią trójki $\mathbf{T}^k, \mathbf{S}^k, \mathbf{U}^k$ o początku w X_0 , co prowadzi do sprzeczności. $\square$

Dowód poniższego twierdzenia zawiera konstrukcję przykładu zapowiedzianego na początku podrozdziału.

Twierdzenie 3.3.3 *Istnieją komutujące fellerowskie operatory Markowa $\mathbf{T}, \mathbf{S}$ i $\mathbf{U}$, dla których żadna trójka iteracji $\mathbf{T}^k, \mathbf{S}^k, \mathbf{U}^k$, $k \in \mathbb{N}$, nie posiada trajektorii.*

Dowód Wobec lematów 3.3.1 i 3.3.2, wystarczy dla każdej liczby naturalnej k skonstruować trójkę komutujących fellerowskich prawdopodobieństw przejścia spełniających założenia lematu 3.3.2.

Ustalmy $k \in \mathbb{N}$. Zdefiniujemy macierze pomocnicze stopnia $k+1$:

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ & & & \ddots & \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}, \quad \mathbf{R} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ & & & \ddots & \\ 1 & 0 & 0 & \dots & 0 \end{bmatrix},$$

$$\mathbf{M} = \sum_{i=0}^k \frac{1}{k+1} \mathbf{R}^i = \begin{bmatrix} \frac{1}{k+1} & \frac{1}{k+1} & \dots & \frac{1}{k+1} \\ \frac{1}{k+1} & \frac{1}{k+1} & \dots & \frac{1}{k+1} \\ & & \ddots & \\ \frac{1}{k+1} & \frac{1}{k+1} & \dots & \frac{1}{k+1} \end{bmatrix}.$$

Oznaczmy przez $\mathbf{0}$ macierz kwadratową stopnia $k+1$ złożoną z samych zer.

[illegible]

$$\mathbf{U} = \begin{bmatrix}
 \begin{array}{cccc|cccc|cccc|cccc}
 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & \ddots & 0 & \mathbf{M} \\
 \mathbf{I} & 0 & \ddots & 0 & 0 & 0 & \ddots & 0 & 0 & 0 & \ddots & 0 & 0 & \ddots & 0 & 0 \\
 0 & \mathbf{I} & \ddots & \ddots & 0 & 0 & \ddots & \ddots & 0 & 0 & \ddots & \ddots & 0 & \ddots & \ddots & 0 \\
 \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots
 \end{array} &
 \begin{array}{cccc|cccc|cccc|cccc}
 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & \ddots & 0 & 0 \\
 0 & 0 & \ddots & 0 & \mathbf{M} & 0 & \ddots & 0 & 0 & 0 & \ddots & 0 & 0 & \ddots & 0 & 0 \\
 0 & 0 & \ddots & \ddots & 0 & \mathbf{M} & \ddots & \ddots & 0 & 0 & \ddots & \ddots & 0 & \ddots & \ddots & 0 \\
 \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots
 \end{array} &
 \begin{array}{cccc|cccc|cccc|cccc}
 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & \ddots & 0 & 0 \\
 0 & 0 & \ddots & 0 & 0 & 0 & \ddots & 0 & \mathbf{I} & 0 & \ddots & 0 & 0 & \ddots & 0 & 0 \\
 0 & 0 & \ddots & \ddots & 0 & 0 & \ddots & \ddots & 0 & \mathbf{I} & \ddots & \ddots & 0 & \ddots & \ddots & 0 \\
 \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots
 \end{array} &
 \begin{array}{cccc|cccc|cccc|cccc}
 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & \ddots & 0 & 0 \\
 0 & 0 & \ddots & 0 & 0 & 0 & \ddots & 0 & 0 & 0 & \ddots & 0 & \mathbf{M} & \ddots & 0 & 0 \\
 0 & 0 & \ddots & \ddots & 0 & 0 & \ddots & \ddots & 0 & 0 & \ddots & \ddots & 0 & \ddots & \ddots & 0 \\
 \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots
 \end{array} \\
 \begin{array}{cccc|cccc|cccc|cccc}
 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & \ddots & 0 & 0 \\
 0 & 0 & \ddots & 0 & 0 & 0 & \ddots & 0 & 0 & 0 & \ddots & 0 & \mathbf{M} & \ddots & 0 & 0 \\
 \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\
 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & \dots & \mathbf{M} & 0
 \end{array}
 \end{bmatrix}$$

przy czym każda klatka ma wymiar $k \times k$, więc $\mathbf{T}, \mathbf{S}$ i $\mathbf{U}$ są wymiaru $4k(k+1) \times 4k(k+1)$. Możemy rozważać działanie macierzy $\mathbf{T}, \mathbf{S}, \mathbf{U}$ na przestrzeni $C_{4k} \times C_{k+1}$, numerując elementy tej przestrzeni tak, by $(p, q) \in C_{4k} \times C_{k+1}$ odpowiadało elementowi $i = p(k+1) + q + 1$ ($i \in \{1, 2, \dots, 4k(k+1)\}$). Oczywiście, $\mathbf{T}, \mathbf{S}$ i $\mathbf{U}$ są fellerowskie, gdyż działają na przestrzeni dyskretniej. Mnożąc odpowiednie macierze i korzystając z równości $\mathbf{RM} = \mathbf{MR} = \mathbf{IM} = \mathbf{MI} = \mathbf{MM} = \mathbf{M}$, łatwo można sprawdzić, że $\mathbf{T}, \mathbf{S}$ i $\mathbf{U}$ komutują. Udowodnimy, że spełniają one założenia lematu 3.3.2. Dla $i = 0, \dots, k-1$ połączmy

$$X_i = \{i, i+k, i+2k, i+3k\} \times C_{k+1}.$$

Wówczas $X_0, \dots, X_{k-1}$ są rozłącznymi zbiorami domkniętymi, których suma jest całą przestrzenią X . Obrazem zbioru X_i pod działaniem $\mathbf{T}$ bądź $\mathbf{S}$ jest X_{i+k+1} , podczas gdy $\mathbf{U}X_{i+k+1} = X_i$. Jest jasne, że X_0 jest zbiorem niezmienniczym względem $\mathbf{T}^k, \mathbf{S}^k$ i $\mathbf{U}^k$, a obciążenia tych iteracji do X_0 są opisane macierzami:

$$\mathbf{T}^k = \begin{bmatrix} 0 & \mathbf{M} & 0 & 0 \\ 0 & 0 & \mathbf{R}^{-1} & 0 \\ 0 & 0 & 0 & \mathbf{M} \\ \mathbf{I} & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{S}^k = \begin{bmatrix} 0 & \mathbf{M} & 0 & 0 \\ 0 & 0 & \mathbf{I} & 0 \\ 0 & 0 & 0 & \mathbf{M} \\ \mathbf{I} & 0 & 0 & 0 \end{bmatrix},$$

$$U^k = \begin{bmatrix} 0 & 0 & 0 & M \\ I & 0 & 0 & 0 \\ 0 & M & 0 & 0 \\ 0 & 0 & I & 0 \end{bmatrix}$$

(gdyż $I^k = I$, $R^k = R^{-1}$ i $M^k = M$).

Dowód faktu, że trójka T^k, S^k, U^k nie posiada trajektorii przebiega tak jak w [D-I]. Zauważmy najpierw, że każda trajektoria musi zawierać punkt postaci $(0, q)$, gdzie $q \in C_{k+1}$. Dlatego wystarczy uzasadnić, że żadna trajektoria nie może rozpoczynać się w takim punkcie. Przypuśćmy, że istnieje trajektoria (x^{n_1, n_2, n_3}) , $n_1, n_2, n_3 \in \mathbb{N}_0$, o początku w $x^{0,0,0} = (0, 0)$. Wtedy $x^{1,0,0} \in \varphi_{T^k}(x^{0,0,0}) = \{k\} \times C_{k+1}$. Niech $x^{1,0,0} = (k, p)$. Z postaci macierzy S^k i U^k wynika, że $x^{1,1,0} \in \varphi_{S^k}(k, p) = \{(2k, p)\}$ oraz $x^{1,0,1} \in \varphi_{U^k}(k, p) = \{(0, p)\}$. Ponadto, patrząc na „przeciwbrazy” operatorów (przez „przeciwbraz” rozumiemy zbiór punktów, jakie mogą poprzedzać w trajektorii bieżący stan), wnioskujemy, że $x^{0,0,1} = (3k, p)$, $x^{0,1,1} = (0, p)$, $x^{0,1,0} = (k, p)$, a wreszcie $x^{1,1,0} = (2k, p')$, gdzie $p' + k1 = p$. Sprzeczność. Jest jasne, że rozpoczynając rozumowanie od dowolnego innego punktu $(0, q)$ ($q \neq 0$), dochodzimy do sprzeczności w podobny sposób, co kończy dowód. $\square$

Bibliografia

- [A-A-F-T] R. Alicki, J. Andries, M. Fannes, P. Tuyls *An algebraic approach to the Kolmogorov-Sinai entropy*, Rev. Math. Phys., 8, 167–184
- [A-F] R. Alicki, M. Fannes, *Quantum Dynamical Systems*, Oxford University Press, Oxford 2001
- [B] J. Brown, *Ergodic Theory and Topological Dynamics*, Academic Press, New York 1976
- [D-G-S] M. Denker, C. Grillenberger, K. Sigmund, *Ergodic Theory on Compact Spaces*, Lecture Notes in Mathematics 527, Springer-Verlag, New York 1976
- [D] T. Downarowicz, *Strictly nonpointwise Markov operators and weak mixing*, Ergodic Theory and its Connections with Harmonic Analysis, Proc. of the 1993 Alexandria Conference, Lecture Note Series 205, London 1995, 259–272
- [D-F] T. Downarowicz, B. Frej, *Measure-theoretic and topological entropy of operators on function spaces*, preprint
- [D-I] T. Downarowicz, A. Iwanik, *Multiple recurrence for discrete time Markov processes II*, Colloq. Math. 55 (1988), pp. 311–316.
- [D-S] T. Downarowicz, J. Serafin, *Fiber entropy and conditional variational principles in compact non-metrizable spaces*, Fund. Math. 172, 2002, 217–247
- [E] R. Engelking, *Topologia ogólna*, PWN, Warszawa 1977
- [F] S. R. Foguel, *The Ergodic Theory of Markov Processes*, Van Nostrand Reinhold Company, New York 1969

- [Fr1] B. Frej, *Transition probabilities and transition systems*, I Konferencja dla Młodych Matematyków: Zastosowania Matematyki-Karpacz 2000, Prace Naukowe Instytutu Matematyki Politechniki Wrocławskiej, 2001, pp. 45–53
- [Fr2] B. Frej, *A note on Markov operators and transition systems*, Colloq. Math. 91 (2002), pp. 183–190
- [G-L-W] E. Ghys, R. Langevin, P. G. Walczak, *Entropie mesurée et partitions de l'unité*, C. R. Acad. Sci. Paris, Sér. I, 303 (1986), 251–254
- [G-W] E. Glasner, B. Weiss, *Quasifactors of ergodic maps with positive entropy*, preprint
- [H] E. Hopf, *The general temporally discrete Markov processes*, J. Rational Mech. Anal., 3 (1954)
- [I1] A. Iwanik, *Approximation theorems for stochastic operators*, Indiana Univ. Math. J., 29, 1980, 415–425
- [I2] A. Iwanik, *Integral representations of stochastic kernels*, in: Aspects of Positivity in Functional Analysis (Tübingen, 1985), Elsevier Sci. Publ. (North-Holland), 1986, 223–230
- [I3] A. Iwanik, *Multiple recurrence for discrete time Markov processes*, Colloq. Math 51 (1987), 165–174.
- [K-SL] B. Kamiński, J. de Sam Lazaro, *A note on the entropy of a doubly stochastic operator*, Colloq. Math. 84/85 (2000), 1, 245–254
- [L-M] D. Lind, B. Marcus, *An introduction to symbolic dynamics and coding*, Cambridge University Press, 1995.
- [L-W] R. Langevin, P. G. Walczak, *Entropy, transverse entropy and partitions of unity*, Ergod. Th. & Dynam. Sys. (1994), 14, 551–563
- [M] I. I. Makarov, *Dynamical entropy for Markov operators*, J. Dyn. Control Syst. 6 (2000), No. 1, 1–11
- [Mi] E. Michael, *Continuous selections*, Ann. Math. 63 (1956), pp. 361–382.
- [N] J. Neveu, *Mathematical foundations of the calculus of probability*, Holden-Day, Inc., 1965.

- [R] D. E. Rutherford, *Introduction to Lattice Theory*, Oliver & Boyd Ltd., Edinburgh, 1965
- [S] R. Sine, *Geometric theory of a single Markov operator*, Pacific J. Math., 27, 1968(1), 155–166
- [W] P. Walters, *An Introduction to Ergodic Theory*, Springer-Verlag, Berlin 1982
- [Z] W. Żelazko, *Algebry Banacha*, PWN, Warszawa 1968

