Uniwersytet w Białymstoku

Wydział Matematyki

Sonia Gazeau

Promień spektralny operatora Ruelle'a-Perrona-Frobeniusa

Praca dyplomowa napisana
pod kierunkiem
dr hab. Bartosza K. Kwaśniewskiego

Białystok 2022

Składam serdeczne podziękowania
dr hab. Bartoszowi Kwaśniewskiemu
za cenne wskazówki podczas pisania pracy
oraz za okazaną cierpliwość i wyrozumiałość.

Sonia Gazeau

Spis treści

Wstęp

Podstawowym celem niniejszej pracy jest zaprezentowanie Zasady Wariacyjnej wyrażającej promień spektralny operatora Ruelle'a-Perrona-Frobeniusa, który jest operatorem $\mathcal{L} : C(X) \to C(X)$ postaci

$$\mathcal{L}(f)(x) := \sum_{y \in \varphi^{-1}(x)} c(y) f(y), \qquad x \in X, f \in C(X).$$

W powyższym wzorze $\varphi : X \to X$ jest ciągłym odwzorowaniem rozszerzającym, natomiast $c : X \to [0, +\infty)$ jest ciągłym potencjałem określonym na zwartej przestrzeni metrycznej X. Operator Ruelle'a-Perrona-Frobeniusa często nazywany jest także operatorem przejścia. Operatory takie określają w jaki sposób odwzorowania zmieniają się podczas iteracji. Stąd też znajdują one zastosowanie w typowo fizycznych zagadnieniach, takich jak chaos kwantowy czy też mechanika statystyczna, gdzie jednym z głównych przedmiotów badań jest zmiana funkcji w czasie.

Zasada Wariacyjna pozwalająca obliczyć promień spektralny naszego operatora jest tak naprawdę produktem ubocznym Twierdzenia Ruelle'a-Perrona-Frobeniusa. Twierdzenie to głównie znajduje zastosowanie w konstrukcji i dowodzie jednoznaczności miary Gibbsa dla trójki (X, φ, c) przy pomocy własności spektralnych operatora $\mathcal{L}$. Wspomniana Zasada Wariacyjna przyjmuje następującą postać

$$\ln r(\mathcal{L}) = \max_{\mu \in \mathcal{M}(X,\varphi)} \left(\int_X \ln c \, d\mu + h_\mu(\varphi) \right) = \max_{\mu \in \mathrm{Erg}(X,\varphi)} \left(\int_X \ln c \, d\mu + h_\mu(\varphi) \right), \tag{0.1}$$

gdzie $\mathcal{M}(X, \varphi)$ jest zbiorem φ-niezmienniczych miar unormowanych na X, $\mathrm{Erg}(X, \varphi)$ jest zbiorem miar φ-ergodycznych, natomiast $h_\mu(\varphi)$ jest entropią Kołmogorowa Sinai'a. Równość (0.1) w skrócie mówi o tym, że logarytm naturalny z promienia spektralnego operatora $\mathcal{L}$ jest równy ciśnieniu topologicznemu dla trójki (X, φ, c).

Przy czym należy tu podkreślić, że w literaturze wzór (0.1) dowodzi się zazwyczaj przy dodatkowych założeniach, takich jak:

- c jest ściśle dodatnia i jej logarytm $\ln c$ jest hölderowsko ciągły
- odwzorowanie φ jest mieszające

Celem pracy jest wykazanie wzoru (0.1) bez tych założeń (dowód taki ukazał się dopiero w zeszłym roku w [4]) oraz szczegółowe omówienie i wprowadzenie wszystkich wyżej wymionych pojęć.

W pierwszym rozdziale opisane zostały pokrótce pojęcia zaczerpnięte z różnych dziedzin matematyki, w celu przypomnienia najważniejszych pojęć i treści, zawartych i wykorzystywanych w dalszej części pracy. Zostały tu także omówione pewne fakty (częściowo wraz dowodami) z Teorii Ergodycznej. Ta część wykracza poza ramy podstawowych zagadnień omawianych na studiach. Po wprowadzeniu bazowych informacji na temat operatorów i funkcjonałów na przestrzeniach Banacha skupiamy się głównie na omówieniu własności zbioru miar $\mathcal{M}(X)$ borelowskich oraz zbioru miar φ-niezmienniczych $\mathcal{M}(X, \varphi)$.

Drugi rozdział został poświęcony omówieniu operatorów przejścia $\mathcal{L} : C(X) \to C(X)$ zadanych na zwartej przestrzeni metrycznej X z odwzorowaniem ciągłym $\varphi : X \to X$. Ogólnie są to dodatnie operatory liniowe spełniające tożsamość homologiczną

$$\mathcal{L}((f \circ \varphi) \cdot g) = f\mathcal{L}g,$$

dla wszystkich $f, g \in C(X)$. Wykazujemy tu, że gdy odwzorowanie φ jest lokalnym homeomorfizmem, to operator przejścia sprowadza się do operatora Ruelle'a-Perrona-Frobeniusa. Natomiast, przy φ będącym homeomorfizmem operator przejścia staje się operatorem ważonego przesunięcia. W tym rozdziale zostały także podane i omówione dwie Zasady Wariacyjne, jedna pozwalająca na obliczenie promienia spektralnego ważonego operatora przejścia, natomiast druga, będąca wnioskiem z pierwszej pozwalająca na wyliczenie promienia spektralnego operatora ważonego przesunięcia. Pierwsza z nich jest postaci

$$\ln r(\mathcal{L}e^b) = \max_{\mu \in \mathcal{M}(X,\varphi)} \left(\int_X b d\mu + \tau_{\mathcal{L}}(\mu) \right).$$

gdzie $\tau_{\mathcal{L}}(\mu)$ jest t-entropią zdefiniowaną przez Antonevicha, Bakhtina i Lebedeva w 2011 roku, patrz [2], a b dowolną rzeczywistą funkcją ciągłą. Gdy φ jest homeomorfizmem zachowujący jedynkę operator przejścia w działaniu na funkcję ciągłą przyjmuje postać $T_\varphi f(x) = f(\varphi(x))$ i w zasadzie powyższy wzór sprowadza się do

$$\ln r(aT_\varphi) = \max_{\mu \in \mathrm{Erg}(X,\varphi)} \int_X \ln |a(x)| d\mu,$$

gdzie $a \in C(X)$ jest dowolną funkcją ciągłą. Ten drugi wzór jest dobrze znany i został wykazany w 1979 roku przez Kitovera i Lebedeva [13], [19].

Trzeci, najdłuższy rozdział jest dość szczegółowym wprowadzeniem do pojęcia entropii - jednego z głównych pojęć teorii ergodycznej. Głównie interesuje nas tu entropia metryczna, czyli entropia Kołmogorowa Sinai'a. Jednakże, omawiamy tu także pojęcie entropii topologicznej w celu zaprezentowania kolejnej Zasady Wariacyjnej mówiącej o tym, że na przestrzeni zwartej entropia topologiczna jest równa supremum po entropiach metrycznych. Materiał w tym rozdziale oparty jest o klasyczną książkę Waltersa [30], ale prezentacja i wybór treści jest autorski. Większość faktów, w tym Zasada Wariacyjna, przedstawiona jest z pełnymi dowodami. Omawiamy tu też szereg konkretnych przykładów.

W czwartym rozdziale wprowadzamy pojęcie ciśnienia topologicznego jako uogólnienie entropii topologicznej. Omawiamy tu najważniejsze elementy formalizmu termodynamicznego, w tym miary Gibbsa oraz klasyczne Twierdzenie Ruelle'a-Perrona Frobeniusa. Tu też zrealizowany został główny cel pracy - dowód Zasady Wariacyjnej (0.1) na promień spektralny operatora Ruelle'a-Perrona-Frobeniusa, gdzie $\varphi : X \to X$ jest dowolnym odwzorowaniem rozszerzającym, a $c : X \to [0, +\infty)$ dowolną funkcją ciągłą.

Rozdział 1

Preliminaria

W tym rozdziale, dla wygody czytelnika, przypomnimy podstawowe pojęcia z topologii, analizy funkcjonalnej, teorii miary oraz teorii ergodycznej, skupiając się na pojęciach i twierdzeniach wykorzystywanych w dalszej części pracy. W większości są to fakty dobrze znane każdemu studentowi matematyki. Jedynie omówione tu elementy teorii ergodycznej wykraczają poza standardowy program studiów.

1.1 Topologia

Przestrzenią topologiczną nazywamy zbiór X z wyróżnioną rodziną zbiorów τ, która zawiera zbiór pusty i całą przestrzeń, jest zamknięta na dowolne sumy i skończone zbiory. Wtedy zbiory z τ nazywamy *zbiorami otwartymi*, natomiast ich dopełnienia *zbiorami domkniętymi. Domknięciem zbioru* nazywamy najmniejszy zbiór domknięty zawierający dany zbiór. Zbiór nazywamy *gęstym*, jeżeli jego domknięcie jest całą przestrzenią. Jeśli w przestrzeni X istnieje zbiór przeliczalny gęsty, to przestrzeń tę nazywamy *przestrzenią ośrodkową*. Mówimy, że przestrzeń topologiczna X posiada *przeliczalną bazę*, jeśli istnieje ciąg $\{U_n\}_{n=1}^{\infty}$ zbiorów otwartych taki, że każdy zbiór otwarty $U \subseteq X$ jest sumą zbiorów z $\{U_n\}_{n=1}^{\infty}$. Każda przestrzeń z przeliczalną bazą jest ośrodkowa.

Mówimy, że przestrzeń topologiczna X jest *przestrzenią zwartą*, jeśli z każdego pokrycia zbiorami otwartymi tej przestrzeni jesteśmy w stanie wybrać podpokrycie skończone. Natomiast X jest *przestrzenią Hausdorffa*, jeśli dla dwóch różnych punktów z tej przestrzeni istnieją otoczenia otwarte, których część wspólna jest zbiorem pustym. Każda przestrzeń metryczna jest przestrzenią Hausdorffa.

Przestrzenią metryczną nazywamy zbiór X z zadaną na nim *metryką*, to znaczy funkcją $d : X \times X \to [0, \infty)$ określającą odległość między punktami. Formalnie zakładamy, że $d(x, y) = 0$ tylko wtedy, gdy $x = y$. Natomiast $d(x, y) = d(y, x)$ oraz $d(x, y) \leqslant d(x, z) + d(z, y)$ dla wszystkich $x, y, z \in X$. Zbiory otwarte w przestrzeni metrycznej zdefiniowane są jako sumy kul otwartych, tj. zbiorów postaci $B(x, r) = \{y \in X : d(x, y) < r\}$, gdzie $x \in X$, $r > 0$. Mówimy, że przestrzeń topologiczna jest *metryzowalna*, jeśli istnieje na X metryka zadająca daną topologię. Na przykład, zwarta X przestrzeń Hausdorffa jest metryzowalna wtedy i tylko wtedy, gdy posiada ona przeliczalną bazę [10]. Ciąg $\{x_n\}_{n=1}^{\infty} \subseteq X$ w przestrzeni metrycznej (X, d) jest *zbieżny* do $x \in X$, jeżeli $\lim_{n\to\infty} d(x_n, x) = 0$. Ciąg $\{x_n\}_{n=1}^{\infty} \subseteq X$ nazywamy *ciągiem Cauchy*, jeśli $\lim_{n,m\to\infty} d(x_n, x_m) = 0$. Każdy ciąg zbieżny jest ciągiem Cauchy. Jeśli w danej przestrzeni zachodzi implikacja odwrotna, to mówimy, że przestrzeń jest *zupełna*. Przestrzeń metryczna jest zwarta wtedy i tylko wtedy, gdy każdy ciąg w tej przestrzeni posiada podciąg zbieżny. W szczególności, każda zwarta przestrzeń metryczna jest zupełna.

Funkcję $f : X \to Y$ określoną między dwiema przestrzeniami topologicznymi X i Y nazywamy *funkcją ciągłą* jeśli przeciwobraz $f^{-1}(V)$ każdego zbioru otwartego $V \subseteq Y$ jest zbiorem otwartym w X. Mówimy, że f jest ciągła w punkcie x_0, jeżeli dla każdego zbioru otwartego $V \ni f(x_0)$ istnieje zbiór otwarty $U \ni x_0$ taki, że $f(U) \subseteq V$. Funkcja jest ciągła wtedy i tylko wtedy, gdy jest ciągła w każdym punkcie swojej dziedziny. W przypadku, gdy (X, d_X) i (Y, d_Y) są przestrzeniami metrycznymi, to kryterium Heinego ciągłości funkcji w punkcie mówi, że f jest ciągła w punkcie $x_0 \in X$, jeśli dla każdego ciągu $\{x_n\}_{n=1}^{\infty}$ zbieżnego do x_0 ciąg $\{f(x_n)\}_{n=1}^{\infty}$ zbiega do $f(x_0)$.

1.2 Analiza funkcjonalna

Przestrzenie liniowe (wektorowe) będziemy rozpatrywać głównie nad ciałem liczb zespolonych $\mathbb{C}$, ale czasem także nad ciałem liczb rzeczywistych $\mathbb{R}$. Niech E będzie *przestrzenią liniową* nad ciałem $\mathbb{F} = \mathbb{R}, \mathbb{C}$. To znaczy zbiór E jest wyposażony w operację dodawania $E \times E(f, g) \mapsto f + g \in E$ oraz mnożenia przez skalary $\mathbb{F} \times E \ni (\lambda, f) \mapsto \lambda f \in E$ spełniające standardowe aksjomaty rozdzielności mnożenia względem dodawania, łączności, itd. Normą na E nazywamy funkcję $\|\cdot\| : E \to [0, +\infty)$ spełniającą:

(1) $\|f\| = 0 \iff f = 0,$ (*niezdegenerowanie*)

(2) $\|\lambda f\| = |\lambda|\|f\|,$ (*jednorodność*)

(3) $\|f + g\| \leqslant \|f\| + \|g\|,$ (*warunek trójkąta*)

dla wszystkich $f, g \in E$ oraz $\lambda \in \mathbb{F}$. Wtedy wzór $d(f, g) = \|f - g\|$ definiuje metryką na E. Zatem ciąg $\{f_n\}_{n=1}^{\infty}$ jest zbieżny w przestrzeni unormowanej E do wektora $f_0 \in E$, jeżeli $\lim_{n\to\infty} \|f_n - f_0\| = 0$. Natomiast $\{f_n\}_{n=1}^{\infty} \subseteq E$ jest ciągiem Cauchy, jeżeli $\lim_{n,m\to\infty} \|f_n - f_m\| = 0$. Niech dany będzie ciąg $\{f_n\}_{n=1}^{\infty}$. Przestrzeń nazywamy *przestrzenią Banacha*, jeśli jest unormowana (posiada normę) oraz każdy ciąg Cauchy'ego posiada w niej granicę (jest zupełna).

Najważniejszym przykładem przestrzeni Banacha w tej pracy jest przestrzeń funkcji ciągłych na przestrzeni zwartej.

Przykład 1.1 (Przestrzeń funkcji ciągłych)**.** Niech $C(X) := \{f : X \to \mathbb{C} \text{ funkcja ciągła}\}$ będzie zbiorem wszystkich zespolonych funkcji ciągłych na przestrzeni zwartej X. Jest to przestrzeń Banacha z działaniami określonymi punktowo, tzn.

$$(f + g)(x) = f(x) + g(x), \qquad (\lambda f)(x) = \lambda f(x),$$

oraz normą zadaną wzorem
$$\|f\| = \sup_{t\in X} |f(t)|.$$
Skończoność normy w powyższym wzorze wynika z faktu, że funkcje ciągłe na zbiorze zwartym osiągają swoje kresy, stąd $\|f\| = \max_{t\in X}|f(t)| < \infty$. Zbieżność ciągu funkcji w tej normie jest *zbieżnością jednostajną*.

Innym ważnym przykładem jest przestrzeń funkcji całkowalnych w p-tej potędze. Przy czym chodzi tu o całkę Lebesgue'a względem miary. Przypomnijmy podstawowe składniki potrzebne do zdefiniowania takiej całki.

Przestrzenią mierzalną nazywamy zbiór X wyposażony w *σ-ciało*, to jest niepustą rodzinę Σ podzbiorów X taką, że $A \in \Sigma \to A' \in \Sigma$, gdzie $A' = X\backslash A$, oraz $\bigcup_{n=1}^{\infty} A_n \in \Sigma$ dla każdego ciągu $\{A_n\}_{n=1}^{\infty} \subseteq \Sigma$. *Miarą* na zbiorze X, a właściwie na przestrzeni mierzalnej (X,Σ), nazywamy funkcję $\mu : \Sigma \to [0,+\infty]$ taką, że $\mu(\emptyset) = 0$, oraz $\mu(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty}\mu(A_n)$ dla parami rozłącznych zbiorów $A_1, A_2, ... \in \Sigma$. W tej pracy skupimy się na miarach skończonych, tzn. takich, że $\mu(X) < \infty$, a w zasadzie na miarach probabilistycznych, tzn. takich, że $\mu(X) = 1$. Jeśli X jest przestrzenią topologiczną, to będziemy ją rozpatrywać jako przestrzeń mierzalną wyposażoną w *σ-ciało zbiorów borelowskich*, to jest najmniejsze σ-ciało zawierające wszystkie zbiory otwarte, które będziemy oznaczać $\mathcal{B}(X)$. Miary określone na $\mathcal{B}(X)$ nazywamy *miarami borelowskimi*.

Niech (X,Σ,μ) będzie przestrzenią z miarą μ. Funkcję $f : X \to \mathbb{R}$ nazywamy *mierzalną*, jeżeli przeciwobraz każdego zbioru borelowskiego na $\mathbb{R}$ jest mierzalny w X:
$$\forall_{A\in\mathcal{B}(\mathbb{R})} \quad f^{-1}(A) \in \Sigma,$$
Całkę z funkcji mierzalnej $f : X \to \mathbb{R}$ definiuje się w trzech krokach. Dla prostoty załóżmy, że μ jest miarą skończoną. Jeśli f jest funkcją *prostą*, to znaczy jest postaci $f = \sum_{i=1}^{n} a_i \mathbf{1}_{A_i}$ gdzie $a_i \in \mathbb{R}$ są liczbami, a $\mathbf{1}_{A_i}$ są funkcjami charakterystycznymi zbiorów mierzalnych $A_i \in \Sigma$, $i = 1, ..., n$, to całkę z funkcji f definiujemy wzorem
$$\int_X f d\mu := \sum_{i=1}^{n} a_i\mu(A_i). \tag{1.1}$$
Całkę z nieujemnej funkcji mierzalnej $f : X \to \mathbb{R}$ definiuje się wzorem
$$\int_X f d\mu := \sup\left\{\int_X s d\mu : 0 \leqslant s \leqslant f,\ s \text{ jest funkcją prostą}\right\}.$$
W tym wypadku funkcję f nazywamy całkowalną, jeśli $\int_X f d\mu < \infty$. Ogólnie mówimy, że funkcja mierzalna jest całkowalna, jeśli całkowalne są funkcje $f^+ := \max\{f, 0\}$ i $f^- := -\min\{f, 0\}$, i wtedy całkę definiujemy wzorem
$$\int_X f d\mu = \int_X f^+ d\mu - \int_X f^- d\mu.$$
Niech $f : X \to \mathbb{C}$ będzie funkcją zespoloną, tzn. $f = u+iv$, gdzie u jest częścią rzeczywistą funkcji f, a v jej częścią urojoną. Wtedy całkę z funkcji f można przedstawić w następujący sposób
$$\int_X f d\mu = \int_X u d\mu + i\int_X v d\mu.$$

Przykład 1.2 (Przestrzenie L^p)**.** Niech $L^p_\mu(X) = \{f : X \to \mathbb{C} : \int_X |f(t)|^p d\mu < \infty\}$, gdzie f jest funkcją mierzalną. Tak zadana przestrzeń jest przestrzenią Banacha z działaniami określonymi punktowo i normą

$$\|f\|_p := \left(\int_X |f(t)|^p d\mu\right)^{\frac{1}{p}},$$

gdzie $1 \leqslant p < \infty$. W rzeczywistości powyższy wzór określa tylko półnormę, gdyż warunek $\|f\|_p = 0$ implikuje jedynie, że f jest równe zero μ-prawie wszędzie, to znaczy $\mu(x \in X : f(x) = 0) = 0$. Dlatego formalnie elementy przestrzeni $L^p_\mu(X)$ definiuje się nie jako funkcje, lecz jako klasy równoważności funkcji równych sobie μ-prawie wszędzie.

Niech E i F będą przestrzeniami Banacha nad ciałem $\mathbb{F} = \mathbb{R}, \mathbb{C}$ i niech $T : E \to F$ będzie *operatorem liniowym*, czyli odwzorowaniem takim, że $T(f + g) = Tf + Tg$, oraz $T(\lambda f) = \lambda Tf$, gdzie $\lambda \in \mathbb{F}$ i $f, g \in E$. Mówimy, że operator T jest ograniczony, jeśli istnieje stała $C > 0$, taka że $\|Tf\| \leqslant C\|f\|$, dla każdego $f \in E$ (normy w E i F oznaczamy tym samym symbolem, gdyż nie prowadzi to do nieporozumień). Ograniczoność operatora liniowego jest równoważna jego ciągłości. Najmniejszą ze stałych C w nierówności powyżej nazywamy normą operatora ograniczonego i oznaczamy $\|T\|$. Normę operatora wyraża się też wzorem $\|T\| = \sup_{\|x\|=1} \|Tx\|$. W przypadku, gdy $F = \mathbb{F}$ jest ciałem skalarów, to operator $T : E \to \mathbb{F}$ nazywamy *funkcjonałem*.

Zbiór ograniczonych operatorów liniowych $B(E, F)$ wraz z działaniami określonymi punktowo i normą operatorową jest przestrzenią Banacha. Przestrzeń ograniczonych funkcjonałów liniowych $B(E, \mathbb{F})$ nazywa się *przestrzenią dualną* do E i oznacza się przez E^*. Na przestrzeni $B(E) := B(E, E)$ operacja składania odwzorowań zadaje iloczyn, wraz z którym $B(E)$ tworzy algebrę Banacha. Rozważmy *grupę operatorów odwracalnych* $\mathrm{Inv}(E) = \{T \in B(E) :\ T \text{ jest odwracalny w } B(E)\}$, w której elementem neutralnym jest operator identycznościowy $I : E \to E$. Na mocy twierdzenia Banacha o odwzorowaniu otwartym $\mathrm{Inv}(E) = \{T \in B(E) :\ T : E \to E \text{ bijekcja}\}$. Grupa operatorów odwracalnych jest zbiorem otwartym, patrz [29, Sekcja 1.2, strona 16].

Widmem operatora $T \in B(E)$ nazywamy zbiór

$$\sigma(T) := \{\lambda \in \mathbb{F} :\ T - \lambda I \notin \mathrm{Inv}(E)\}.$$

Widmo operatora jest zbiorem zwartym, a także niepustym, jeżeli $\mathbb{F} = \mathbb{C}$, patrz [29, Twierdzenie 1.7 i Twierdzenie 1.9]. *Promień spektralny operatora* $T \in B(X)$ jest to promień najmniejszej kuli o środku w zerze, zawierającej widmo, tzn.

$$r(T) := \max_{\lambda \in \sigma(T)} |\lambda|.$$

Promień spektralny można również wyrazić *wzorem Gelfanda*, nazwanym tak na cześć matematyka Israela M. Gelfanda, który pozwala obliczyć promień spektralny dla danego operatora w terminach jego normy:

$$r(T) = \inf_{n \in \mathbb{N}} \|T^n\|^{\frac{1}{n}} = \lim_{n \to \infty} \|T^n\|^{\frac{1}{n}},$$

patrz [29, Twierdzenie 1.9].

1.3 Miary borelowskie na przestrzeniach zwartych

W tej pracy X będzie zazwyczaj metryzowalną przestrzenią zwartą. Miary borelowskie na takich przestrzeniach posiadają dobre własności i odpowiadają funkcjonałom liniowym na przestrzeni $C(X)$. Metryzowalność X odpowiada ośrodkowości przestrzeni $C(X)$, patrz [10]. Przypomnijmy, że $\mathcal{B}(X)$ oznacza σ-ciało zbiorów borelowskich. Natomiast *zbiór borelowskich miar unormowanych* (probabilistycznych) będziemy oznaczać przez $\mathcal{M}(X)$. Każdy punkt $x \in X$ określa element δ_x (miarę Diraca) zbioru $\mathcal{M}(X)$ wzorem:

$$\delta_x(A) = \begin{cases} 1, & \text{gdy } x \in A \\ 0, & \text{gdy } x \notin A \end{cases}$$

Odwzorowanie $x \to \delta_x$ zanurza X w $\mathcal{M}(X)$.

Twierdzenie 1.3. *Każda miara borelowska $m : \mathcal{B}(X) \to [0, +\infty)$ na przestrzeni metrycznej X jest regularna, tzn. dla każdego $A \in \mathcal{B}(X)$ mamy*

$$m(A) = \inf\{m(U) : A \subseteq U,\ U \textit{ zbiór otwarty}\} = \sup\{m(K), K \subseteq A,\ K \textit{ zbiór domknięty}\},$$

Dowód. Niech $\mathcal{R}$ będzie rodziną zbiorów, które spełniają warunek regularności, tzn. $\mathcal{R} = \{A \in \mathcal{R} : \forall_{\varepsilon>0}\, \exists$ otwarty U_ε oraz domknięty C_ε, takie, że $C_\varepsilon \subseteq A \subseteq U_\varepsilon$ oraz $m(U_\varepsilon \backslash C_\varepsilon) < \varepsilon\}$.

Żeby otrzymać tezę trzeba pokazać, że $\mathcal{B}(X) \subseteq \mathcal{R}$. W tym celu wystarczy pokazać, że $\mathcal{R}$ jest σ-algebrą oraz że $\mathcal{R}$ zawiera rodzinę zbiorów generujących σ-algebrę $\mathcal{B}(X)$. Taką rodziną generującą $\mathcal{B}(X)$ mogą być na przykład wszystkie zbiory domknięte.

Pokażmy zatem, że dowolny zbiór domknięty C należy do $\mathcal{R}$. Niech $\varepsilon > 0$. Zdefiniujmy ciąg zbiorów $U_n := \{x \in X : d(C, x) < \frac{1}{n}\}$, $n \in \mathbb{N}$. Jest to zstępujący ciąg

$$U_1 \supseteq U_2 \supseteq ... \supseteq U_n ... \tag{1.2}$$

zbiorów otwartych taki, że

$$\bigcap_{i=1}^{\infty} U_i = C. \tag{1.3}$$

W szczególności (1.3) implikuje, że

$$m(U_k) \geqslant m(C) \quad \text{i} \quad m(U_k \backslash C) \geqslant 0, \tag{1.4}$$

dla każdego k. Ponadto, skoro m jest miarą σ-addytywną (a więc jest ciągła), to z (1.2) i (1.3) mamy, że

$$\lim_{n\to\infty} \mu(U_n) = m(C). \tag{1.5}$$

Teraz z (1.4) i (1.5)

$$\lim_{k\to\infty} m(U_k \backslash C) = 0.$$

Zatem możemy tak dobrać k tak, aby $m(U_k \backslash C) < \varepsilon$. Kładąc $U_\varepsilon = U_k$ i $C_\varepsilon = C$ widzimy, że warunek definiujący $\mathcal{R}$ jest spełniony. Czyli $C \in \mathcal{R}$.

Wystarczy teraz pokazać, że $\mathcal{R}$ jest σ-algebrą. Oczywistym jest, że $X \in \mathcal{R}$, bo X jest zbiorem domkniętym i otwartym. Załóżmy, że $A \in \mathcal{R}$. Chcemy pokazać, że $X \setminus A \in \mathcal{R}$. Dla $\varepsilon > 0$ istnieje zbiór otwarty U_ε oraz zbiór domknięty C_ε, takie, że $C_\varepsilon \subseteq A \subseteq U_\varepsilon$ oraz $m(U_\varepsilon \setminus C_\varepsilon) < \varepsilon$. Stąd $X \setminus U_\varepsilon \subseteq X \setminus A \subseteq X \setminus C_\varepsilon$ i $(X \setminus C_\varepsilon) \setminus (X \setminus U_\varepsilon) = U_\varepsilon \setminus C_\varepsilon$, więc

$$m((X \setminus C_\varepsilon) \setminus (X \setminus U_\varepsilon)) = m(U_\varepsilon \setminus C_\varepsilon) < \varepsilon.$$

Stąd $A \setminus X \in \mathcal{R}$.

Niech teraz $\{A_n\}_{n=1}^{\infty} \in \mathcal{R}$. Potrzebujemy pokazać, że $A := \bigcup_{n=1}^{\infty} A_n \in \mathcal{R}$ (czyli, że rodzina $\mathcal{R}$ jest zamknięta ze względu na przeliczalne sumy). Niech $\varepsilon > 0$. Z założenia, dla każdego $n \in \mathbb{N}$ istnieją zbiory $U_{\varepsilon,n}$ i $C_{\varepsilon,n}$, takie, że $U_{\varepsilon,n}$ jest otwarty, $C_{\varepsilon,n}$ jest domknięty, $C_{\varepsilon,n} \subseteq A_n \subseteq U_{\varepsilon,n}$ oraz $m(U_{\varepsilon,n} \setminus C_{\varepsilon,n}) < \frac{\varepsilon}{3^n}$. Połóżmy

$$\tilde{C}_\varepsilon := \bigcup_{n=1}^{\infty} C_{\varepsilon,n}.$$

Powyższy zbiór jest borelowski, natomiast zbiór $U_\varepsilon := \bigcup_{n=1}^{\infty} U_{\varepsilon,n}$ jest otwarty i $\tilde{C}_\varepsilon \subset U_\varepsilon$, a z tego $m(\tilde{C}_\varepsilon) \leqslant m(U_\varepsilon) < \infty$. Teraz, jeśli dla każdego k położymy

$$C_\varepsilon^k := \bigcup_{n=1}^{k} C_{\varepsilon,n},$$

to mamy, że

$$C_\varepsilon^1 \subseteq C_\varepsilon^2 \subseteq ... \subseteq C_\varepsilon^k \subseteq ... \quad \text{i} \quad \bigcup_{k=1}^{\infty} C_\varepsilon^k = \tilde{C}_\varepsilon.$$

Tak jak poprzednio z ciągłości miary m mamy, że

$$\lim_{k \to \infty} m(\tilde{C}_\varepsilon \setminus C_\varepsilon^k) = 0.$$

Więc możemy wybrać k takie, że $m(\tilde{C}_\varepsilon \setminus \bigcup_{n=1}^{k} C_{\varepsilon,n}) < \frac{\varepsilon}{2}$. Zbiór $C_\varepsilon := \bigcup_{n=1}^{k} C_{\varepsilon,n}$ jest domknięty jako skończona suma domkniętych zbiorów $C_{\varepsilon,n}$. Mamy, że $C_\varepsilon \subseteq A \subseteq U_\varepsilon$. Ponadto

$$\begin{aligned} m(U_\varepsilon \setminus C_\varepsilon) &\leqslant m(U_\varepsilon \setminus \tilde{C}_\varepsilon) + m(\tilde{C}_\varepsilon \setminus C_\varepsilon) \\ &\leqslant \sum_{n=1}^{\infty} m(U_{\varepsilon,n} \setminus C_{\varepsilon,n}) + m(\tilde{C}_\varepsilon \setminus C_\varepsilon) \leqslant \sum_{n=1}^{\infty} \frac{\varepsilon}{3^n} + \frac{\varepsilon}{2} = \epsilon. \end{aligned}$$

To dowodzi, że $A = \bigcup_{n=1}^{\infty} A_n \in \mathcal{R}$. Zatem $\mathcal{R}$ jest σ-algebrą. $\square$

Każda miara borelowska $\mu : \mathcal{B}(X) \to [0, +\infty)$ zadaje poprzez całkę ograniczony funkcjonał liniowy J na $C(X)$, którego norma jest równa $\mu(X)$ (przestrzeń takich funkcjonałów oznaczamy przez $C(X)^*$ i nazywamy przestrzenią dualną do $C(X)$). Rzeczywiście, z liniowości całki otrzymujemy, że wzór

$$J(f) := \int_X f(t) d\mu(t), \qquad f \in C(X),$$

definiuje funkcjonał liniowy $J : C(X) \to \mathbb{C}$. Ponadto,

$$|J(f)| = \left|\int_X f(t)d\mu(t)\right| \leqslant \int_X |f(t)|d\mu(t) \leqslant \int_X \sup_{s\in X} |f(s)|d\mu(t) = \|f\|_\infty \cdot \mu(X).$$

Zatem $\|J\| \leqslant \mu(X)$. Z drugiej strony, $J(1) = \mu(X)$, skąd $\|J\| \geqslant 1$, a zatem $\|J\| = 1$. Zauważmy, że dodatniość miary μ implikuje, iż funkcjonał J jest *dodatni*, tzn.

$$f \in C(X) \text{ oraz } f \geqslant 0 \implies J(f) \geqslant 0.$$

Poniższe twierdzenie mówi o tym, że każdy funkcjonał liniowy z przestrzeni $C(X)^*$, który jest dodatni można przedstawić za pomocą całki z pewnej miary dodatniej.

Twierdzenie 1.4 (Twierdzenie Riesza o reprezentacji funkcjonałów na $C(X)$)**.** *Niech X będzie zwartą przestrzenią metryczną, a $J : C(X) \to \mathbb{C}$ ciągłym, liniowym odwzorowaniem dodatnim (tzn. $J(f) \geqslant 0$) dla $f \geqslant 0$). Wtedy istnieje dokładnie jedna miara $\mu \in \mathcal{M}(X)$ taka, że $J(f) = \int_X f d\mu$ dla wszystkich f w $C(X)$.*

Dowód. Dowód istnienia miary jest dość długi i żmudny, patrz [26]. Dlatego wykażemy tu jedynie jednoznaczność miary μ. Pokażemy, że jeśli $m(f) := \int_X f dm = \int_X f d\mu := \mu(f)$ dla każdego $f \in C(X)$, wtedy $m = \mu$. Na mocy Twierdzenia 1.3 wystarczy pokazać, że $m(C) = \mu(C)$ dla wszystkich zbiorów domkniętych $C \subseteq X$, gdyż takie wartości determinują miary regularne. Załóżmy, że C jest domknięty. Niech $\varepsilon > 0$. Z regularności m mamy, że istnieje zbiór otwarty U taki, że $C \subseteq U$ oraz $m(U \backslash C) < \epsilon$. Zdefiniujmy $f : X \to \mathbb{R}$

$$f(x) = \begin{cases} 0 & \text{gdy } x \notin U, \\ \frac{d(x,X\backslash U)}{d(x,X\backslash U)+d(x,C)} & \text{gdy } x \in U. \end{cases}$$

Jest to poprawnie zdefiniowane, o ile mianownik jest niezerowy. Ponadto f jest funkcją ciągłą, f przyjmuje zerowe wartości na zbiorze $X\backslash U$, $f = 1$ na zbiorze C oraz $0 < f(x) < 1$ dla każdego $x \in X$. Stąd

$$\mu(C) \leqslant \int_X f d\mu = \int_X f dm \leqslant m(U) < m(U) + \varepsilon.$$

Zatem $\mu(C) < m(C) + \varepsilon$ dla dowolnie małego $\varepsilon > 0$, więc $\mu(C) \leqslant m(C)$. Z symetryczności otrzymujemy $m(C) \leqslant \mu(C)$. □

Z Twierdzenia Riesza wynika, że odwzorowanie $\mu \to J$ jest bijekcją pomiędzy $\mathcal{M}(X)$, a zbiorem unormowanych, dodatnich funkcjonałów liniowych na $C(X)$. Oznaczmy J_μ jako obraz μ. Zauważmy, że ta bijekcja jest odwzorowaniem *afinicznym*, tzn. zachowuje kombinacje wypukłe:

$$J_{p\mu+(1-p)m} = pJ_\mu + (1-p)J_m, \qquad p \in [0,1],\ m, \mu \in M(X).$$

W szczególności, $\mathcal{M}(X)$ zanurza się jako zbiór wypukły w kulę jednostkową w $C(X)^*$. Przy tym zanurzeniu, $\mathcal{M}(X)$ dziedziczy *-słabą topologię z przestrzeni $C(X)^*$, którą można opisać w następujący sposób.

Definicja 1.5. *Topologia *-słaba* na $\mathcal{M}(X)$ jest najmniejszą topologią w której dla każdego $f \in C(X)$ funkcjonał $\mathcal{M}(X) \ni \mu \to \int_X f d\mu \in \mathbb{R}$ jest ciągły. Baza tej topologii jest dana przez rodzinę zbiorów postaci

$$V_\mu(f_1, ..., f_k, \varepsilon) = \left\{ m \in \mathcal{M}(X) : \left| \int f_i dm - \int f_i d\mu \right| < \varepsilon, 1 \leqslant i \leqslant k \right\},$$

gdzie $\mu \in \mathcal{M}(X)$, $k \geqslant 1$, $f_i \in C(X)$ i $\varepsilon > 0$. Oczywistym jest, że ta topologia nie zależy od metryki zadanej na X.

Twierdzenie 1.6. *Jeśli X jest zwartą przestrzenią metryzowalną, to *-słaba topologia na $\mathcal{M}(X)$ jest metryzowalna. Dokładniej, jeśli $\{f_n\}_{n=1}^\infty$ jest gęstym podzbiorem w przestrzeni $C(X)$ (taki ciąg zawsze istnieje, gdyż $C(X)$ jest ośrodkowa), to*

$$D(m, \mu) = \sum_{n=1}^{\infty} \frac{|\int f_n dm - \int f_n d\mu|}{2^n \|f_n\|},$$

*jest metryką na $\mathcal{M}(X)$ zadającą *-słabą topologię.*

Dowód. Jasne jest, że funkcja $D : \mathcal{M}(X) \times \mathcal{M}(X) \to [0, +\infty)$ zdefiniowana powyżej jest metryką. Rozważmy przestrzeń metryczną $(\mathcal{M}(X), D)$. Dla każdego ustalonego i odwzorowanie $\mu \to \int f_i d\mu$ jest ciągłe na (X, D), ponieważ $|\int f_i dm - \int f_i d\mu| \leqslant 2^i \|f_i\| D(m, \mu)$. Biorąc pod uwagę, że $\{f_i\}_{n=1}^\infty$ jest podzbiorem gęstym w $C(X)$ mamy, że dla każdego $f \in C(X)$ odwzorowanie $\mu \to \int f d\mu$ jest ciągłe na $(\mathcal{M}(X), D)$. Zatem każdy zbiór otwarty w *-słabej topologii jest otwarty w przestrzeni metrycznej $(\mathcal{M}(X), D)$. Aby przeprowadzić dowód w drugą stronę wystarczy wykazać, że każda kula $\{m \in \mathcal{M}(X) : D(m, \mu) < \varepsilon\}$ w $(\mathcal{M}(X), D)$ zawiera zbiór $V_\mu(g_1, ..., g_k; \delta)$, gdzie $k \geqslant 1$, $g_i \in C(X)$, $1 \leqslant i \leqslant k$, oraz $\delta > 0$. Jeśli $\mu \in \mathcal{M}(X)$ i $\varepsilon > 0$ są dane, należy dobrać N tak, aby

$$\sum_{n=N+1}^{\infty} \frac{2}{2^n} < \frac{\varepsilon}{2}.$$

Niech

$$\delta = \frac{\varepsilon}{2} \left(\sum_{n=1}^{N} \frac{1}{2^n \|f_n\|} \right)^{-1}.$$

Wtedy $V_\mu(f_1, ..., f_N; \delta) \subseteq \{m \in \mathcal{M}(X) \colon D(m, \mu) < \varepsilon\}$. □

Uwaga (1.5). (1) Ciąg miar $\{\mu_n\} \subseteq \mathcal{M}(X)$ zbiega w *-słabej topologii do $\mu \in \mathcal{M}(X)$ wtedy i tylko wtedy, gdy $\int f d\mu_n \to \int f d\mu$ dla każdej funkcji ciągłej $f \in C(X)$.

(2) Zanurzenie $X \to M(X)$ dane przez $x \to \delta_x$ jest ciągłe.

(3) Jeśli $\mu_n, \mu \in M(X)$, $n \geqslant 1$, to następujące warunki są równoważne:

(i) $\mu_n \to \mu$ w *-słabej topologii,

(ii) Dla każdego domkniętego podzbioru F zbioru X, $\lim_{n\to\infty}\sup \mu_n(F) \leqslant \mu(F)$,

(iii) Dla każdego otwartego podzbioru U zbioru X, $\lim_{n\to\infty}\inf \mu_n(U) \geqslant \mu(U)$,

(iv) Dla każdego borelowskiego zbioru A z $\mu(\partial A) = 0$, $\mu_n(A) \to \mu(A)$.

Dowód. Pokażemy, że (i)⇒(ii), a następnie, że (ii)⇒(iii) i (iv). Dowody pozostałych implikacji znajdują się w [23].

Niech F będzie domkniętym podzbiorem X i dla $k \geqslant 1$ niech $U_k = \{x \in X : d(x, F) < 1/k\}$. Zbiory U_k są otwarte oraz zbiegają do F. Stąd $\mu(U_k) \to \mu(F)$. Z Lematu Urysohn'a (patrz [26, Lemat 1.12]), wybieramy $f_k \in C(X)$ takie, że $0 \leqslant f_k \leqslant 1$, $f_k = 1$ na F oraz $f_k = 0$ na $X\backslash U_k$. Wtedy

$$\lim_{n\to\infty}\sup \mu_n(F) \leqslant \lim_{n\to\infty}\sup \int f_k d\mu_n = \int f_k d\mu \leqslant \mu(U_k),$$

więc $\lim_{n\to\infty}\sup \mu_n(F) \leqslant \mu(F)$. Więc mamy (i)⇒(ii). Załóżmy prawdziwość (ii) i niech U będzie otwartym podzbiorem X. Wtedy

$$\lim_{n\to\infty}\sup \mu_n(X\backslash U) \leqslant \mu(X\backslash U) \quad \text{więc} \quad \lim_{n\to\infty}\inf \mu_n(U) \geqslant \mu(U).$$

To dowodzi implikacji (ii)⇒(iii). Załóżmy prawdziwość (ii) oraz, że $\mu(\partial A) = 0$. Wtedy $\mu(\mathrm{int}(A)) = \mu(A) = \mu(\overline{A})$ oraz $\lim_{n\to\infty}\sup \mu_n(\overline{A}) \leqslant \mu(\overline{A}) = \mu(A)$ i $\lim_{n\to\infty}\inf \mu_n(\mathrm{int}(A)) \geqslant \mu(\mathrm{int}(A)) = \mu(A)$. Więc $\mu_n(A) \to \mu(A)$. Zatem zostało wykazane, że (i)⇒(iv). □

Kolejnym ważnym faktem jest, że przestrzeń $\mathcal{M}(X)$ jest zwarta w $*$-słabej topologii. Zazwyczaj pokazuje się to korzystając ze zwartości kuli jednostkowej w przestrzeni dualnej $C(X)^*$, co wynika z Twierdzenia Banacha-Alaoglu. W poniższym twierdzeniu udowodnimy ten fakt wprost, nie korzystając z Twierdzenia Banacha-Alaoglu.

Twierdzenie 1.7. *Jeśli X jest zwartą przestrzenią, wtedy $\mathcal{M}(X)$ jest zwarta w *-słabej topologii.*

Dowód. Wprowadźmy notację $\mu(f) := \int f d\mu$. Niech $\{\mu_n\}_{n=1}^{\infty}$ będzie ciągiem w $\mathcal{M}(X)$. Należy pokazać, że posiada on podciąg zbieżny, co dowiedzie zwartości $\mathcal{M}(X)$. Wybierzmy ciąg funkcji $f_1, f_2, ...$ tworzący zbiór gęsty w $C(X)$. Rozważmy ciąg liczb zespolonych $\{\mu_n(f_1)\}$. Jest on ograniczony przez $\|f_1\|$, więc posiada on podciąg zbieżny $\{\mu_n^{(1)}(f_1)\}$. Teraz rozważmy ciąg liczbowy $\{\mu_n^{(1)}(f_2)\}$. On także jest ograniczony i ma podciąg zbieżny $\{\mu_n^{(2)}(f_2)\}$. Zauważmy, że $\{\mu_n^{(2)}(f_1)\}$ także jest zbieżny. Kontynuujemy ten proces i dla każdego $i \geqslant 1$ konstruujemy podciąg $\{\mu_n^{(i)}\}$ ciągu $\{\mu_n\}$ taki, że $\{\mu_n^{(i)}\} \subseteq \{\mu_n^{(i-1)}\} \subseteq ... \subseteq \{\mu_n^{(1)}\} \subseteq \{\mu_n\}$. Łatwo zauważyć,że $\{\mu_n^{(i)}(f_i)\}$ zbiega dla $f = f_1, f_2, ..., f_i$. Rozważmy ciąg diagonalny $\{\mu_n^{(n)}\}$. Ciąg $\{\mu_n^{(n)}(f_i)\}$ jest zbieżny dla wszystkich i. Stąd $\{\mu_n^{(n)}(f)\}$ jest zbieżny dla wszystkich $f \in C(X)$. Niech $J(f) = \lim_{n\to\infty} \mu_n^{(n)}(f)$. Oczywistym jest, że $J : C(X) \to \mathbb{C}$ jest liniowe i ograniczone, gdyż $|J(f)| \leqslant \|f\|$. Także $J(1) = 1$ i jeśli $f \geqslant 0$, to $J(f) \geqslant 0$. Z Twierdzenie 1.4 istnieje borelowska miara probabilistyczna μ na X taka, że $J(f) = \int_X f d\mu$ dla wszystkich $f \in C(X)$. Stąd,

$$\int_X f d\mu_n^{(n)} \to \int_X f d\mu,$$

tzn. $\mu_n^{(n)} \to \mu$ w *-słabej topologii (patrz Uwaga 1.5 (1)). □

Zatem $\mathcal{M}(X)$ jest zwartą, wypukłą i metryzowalną przestrzenią co daje nam możliwość stosowania twierdzeń dla odwzorowań określonych na takich przestrzeniach.

1.4 Dyskretne układy dynamiczne

Przez *dyskretny układ dynamiczny* rozumiemy zbiór X z zadaną pewną strukturą oraz odwzorowanie $\varphi : X \to X$ zachowujące strukturę na X. Interesować nas będą głównie przypadki, gdy na X jest zadana topologia oraz miara.

Definicja 1.8. *Miarowym układem dynamicznym* nazywamy układ dynamiczny (X, φ), gdzie (X, Σ, μ) jest przestrzenią z miarą unormowaną, a $\varphi : X \to X$ jest odwzorowaniem mierzalnym takim, że

$$\forall_{A\in\Sigma} \quad \mu(\varphi^{-1}(A)) = \mu(A). \tag{1.6}$$

Jeśli zachodzi (1.6), to mówimy, że φ *zachowuje miarę* μ lub że miara μ jest φ*-niezmiennicza.*

Definicja 1.9. *Topologicznym układem dynamicznym* nazywamy parę (X,φ), gdzie X jest zwartą przestrzenią metryczną, natomiast $\varphi : X \to X$ jest odwzorowaniem ciągłym. Zbiór wszystkich unormowanych borelowskich miar φ-niezmienniczych oznaczamy $\mathcal{M}(X, \varphi)$.

Różne własności układów dynamicznych, a w szczególności ich własności spektralne można badać za pomocą operatorów kompozycji, które te układy zadają. Jednym z pierwszych, który tę technikę zastosował był Koopman [16]. Dlatego też, w kontekście układów dynamicznych operatory kompozycji nazywa się często operatorami Koopmana.

Stwierdzenie 1.10. *Jeśli (X, φ) jest miarowym układem dynamicznym na przestrzeni z miarą (X, Σ, μ), to dla dowolnego $1 \leqslant p < \infty$ operator kompozycji z odwzorowaniem $\varphi : X \to X$.*

$$T_\varphi f(x) := f(\varphi(x))$$

jest poprawnie określoną izometrią $T_\varphi : L^p_\mu(X) \to L^p_\mu(X)$.

Dowód. Wystarczy pokazać, że zachodzi to dla funkcji charakterystycznych zbiorów z Σ. Jest tak, ponieważ każdą całkę z dodatniej funkcji mierzalnej można przybliżać funkcjami prostymi, które zbiegają punktowo do tej funkcji. Zajdzie to również dla dowolnej funkcji ciągłej $f : X \to \mathbb{R}$ poprzez rozważenie jej dodatniej i ujemnej części. Dla $p \geqslant 1$ i dla każdego $A \in \Sigma$ mamy

$$\begin{aligned}\|T_\varphi \mathbf{1}_A\|_p^p = \|\mathbf{1}_A \circ \varphi\|_p^p = \int |\mathbf{1}_A \circ \varphi|^p d\mu = \int |\mathbf{1}_{\varphi^{-1}(A)}|^p d\mu \\ = \mu(\varphi^{-1}(A)) = \mu(A) = \int |\mathbf{1}_A|^p d\mu = \|\mathbf{1}_A\|_p^p.\end{aligned}$$

□

Ustalmy teraz topologiczny układ dynamiczny (X, φ) i zdefiniujmy odwzorowanie $\varphi_* : \mathcal{M}(X) \to \mathcal{M}(X)$ zadane wzorem

$$\forall_{A\in\Sigma} \quad \varphi_*\mu(A) := \mu(\varphi^{-1}(A)).$$

Lemat 1.11. *Przy powyższych założeniach zachodzi*

$$\forall_{f\in C(X)} \qquad \int f\circ\varphi\, d\mu = \int f d\varphi_*\mu.$$

Dowód. Wystarczy wykazać, że zachodzi to dla funkcji $f \in C(X)$ o rzeczywistym zbiorze wartości. Z definicji $\varphi_*\mu$ dla każdego $A \in \Sigma$ mamy $\int \mathbf{1}_A d\varphi_*\mu = \int \mathbf{1}_A \circ \varphi d\mu$, gdzie $\mathbf{1}_A$ jest funkcją charakterystyczną zbioru A. Stąd

$$\int h d\varphi_*\mu = \int h\circ\varphi d\mu,$$

jeśli h jest funkcją prostą. Zachodzi to także, gdy h jest nieujemną funkcją mierzalną, poprzez wybranie rosnącego ciągu funkcji prostych zbiegającego punktowo do h. Zatem powyższe równanie jest spełnione dla dowolnej funkcji ciągłej $f : X \to \mathbb{R}$ poprzez rozważenie dodatniej i ujemnej części tej funkcji. □

Stwierdzenie 1.12. *Jeśli (X, φ) jest topologicznym układem dynamicznym, to operator kompozycji jest poprawnie określoną kontrakcją $T_\varphi : C(X) \to C(X)$, która jest izometrią wtedy i tylko wtedy, gdy φ jest surjekcją. Miara $\mu \in \mathcal{M}(X)$ jest φ-niezmiennicza wtedy i tylko wtedy, gdy*

$$\int T_\varphi f d\mu = \int f d\mu \text{ dla wszystkich } f \in C(X).$$

Dowód. Najpierw pokażemy, że T_φ jest kontrakcją. Weźmy $f \in C(X)$, wtedy

$$\|T_\varphi f\| = \max_{x\in X} |T_\varphi f(x)| = \max_{x\in X} |f(\varphi(x))| \leqslant \max_{y\in X} |f(y)| = \|f\|. \tag{1.7}$$

Teraz pokażemy, że T_φ jest izometrią (tzn. $\|T_\varphi f\| = \|f\|$) wtedy i tylko wtedy, gdy φ jest surjekcją ($\varphi(X) = X$). Jeśli $\varphi(X) = X$, to $\max_{x\in X} |f(\varphi(x))| = \max_{y\in X} |f(y)|$, więc w (1.7) mamy wszędzie równości, co dowodzi, że T_φ jest izometrią. Z drugiej strony, jeśli $\varphi(X) \neq X$, to wtedy $\varphi(X) := Y$ jest zwartym podzbiorem X (obraz zbioru zwartego przy ciągłym odwzorowaniu także jest zwarty). Weźmy dowolny punkt $x_0 \in X\backslash Y$. Wtedy z Lematu Urysohn'a (patrz [26, Lemat 1.12]) istnieje funkcja $f \in C(X)$ taka, że

$$0 \leqslant f(x) \leqslant 1, \quad x \in X,$$

oraz

$$f(x) = \begin{cases} 1, & \text{gdy } x = x_0 \\ 0, & \text{gdy } x \in Y. \end{cases}$$

Dla tej funkcji mamy $\|f\| = \max_{x\in X} |f(x)| = |f(x_0)| = 1$ oraz

$$\|T_\varphi f\| = \max_{x\in X} |T_\varphi f(x)| = \max_{x\in X} |f(\varphi(x))| = \max_{x\in Y} |f(x)| = 0,$$

Zatem $\|T_\varphi f\| = 0 < \|f\| = 1$. Czyli T_φ nie jest izometrią.

Druga część twierdzenia wynika z Lematu 1.11 oraz z jednoznaczności w Twierdzeniu 1.4. □

Dla każdego topologicznego układu dynamicznego istnieje borelowska miara niezmiennicza. Jest to tak zwane Twierdzenie Bogoliubowa-Krylowa:

Twierdzenie 1.13 (Bogoliubow-Krylow)**.** *Dla każdego topologicznego układu dynamicznego (X, φ) zbiór $\mathcal{M}(X, \varphi)$ jest niepusty, tzn. istnieje borelowska miara φ-niezmiennicza.*

Dowód. Niech $\{\sigma_n\}_{n=1}^{\infty}$ będzie ciągiem z $\mathcal{M}(X)$. Może to być na przykład $\sigma_n = \delta_y$ dla każdego n, gdzie $y \in X$ dowolnie wybrany punkt. Rozważmy ciąg miar $\{\mu_n\}_{n=1}^{\infty}$ danych wzorem $\mu_n := \frac{1}{n}\sum_{i=0}^{n-1} \varphi_*^i \sigma_n$. Wtedy $\{\mu_n\}_{n=1}^{\infty} \subseteq \mathcal{M}(X)$, a więc ze zwartości zbioru $\mathcal{M}(X)$ (patrz Twierdzenie 1.7) otrzymujemy, że ciąg ten posiada podciąg μ_{n_j} zbieżny w *-słabej topologi do pewnej miary $\mu \in \mathcal{M}(X)$. Pokażemy, że μ należy do $\mathcal{M}(X, \varphi)$. Dla dowolnego $f \in C(X)$ mamy

$$\begin{aligned}
\left|\int f \circ \varphi d\mu - \int f d\mu\right| &= \lim_{j\to\infty}\left|\int f \circ \varphi d\mu_{n_j} - \int f d\mu_{n_j}\right| \\
&= \lim_{j\to\infty}\left|\frac{1}{n_j}\int \sum_{i=0}^{n_j-1}(f \circ \varphi^{i+1} - f \circ \varphi^i) d\sigma_{n_j}\right| \\
&= \lim_{j\to\infty}\left|\frac{1}{n_j}\int (f \circ \varphi^{n_j} - f) d\sigma_{n_j}\right| \\
&\leqslant \lim_{j\to\infty}\frac{2\|f\|}{n_j} = 0.
\end{aligned}$$

Zatem $\mu \in \mathcal{M}(X, \varphi)$ na mocy Stwierdzenia 1.12. □

Przykład 1.14 (Obroty okręgu)**.** Niech $X := S^1 = \{z \in \mathbb{C} : |z| = 1\}$ będzie okręgiem jednostkowym oraz niech $\varphi(z) := e^{2\pi i\omega}z$, gdzie $\omega \in [0, 1)$. Odwzorowanie to jest obrotem o kąt $2\pi\omega$. Zatem zachowuje ono długość łuku na S^1, czyli miarę Lebesgue'a m.

Przykład 1.15 (Endomorfizmy okręgu)**.** Niech, tak jak w poprzednim przykładzie, $X = S^1$ będzie okręgiem i niech $\varphi(z) = z^2$. Wtedy każdy punkt ma dokładnie dwa przeciwobrazy. Za pomocą odwzorowania $S^1 \ni e^{2\pi it} \leftrightarrow t \in [0, 1)$ możemy utożsamić okrąg z odcinkiem $[0, 1)$ modulo 1. Wtedy

$$\varphi(t) = \begin{cases} 2t & \text{gdy } t \in [0, \frac{1}{2}) \\ 2t - 1 & \text{gdy } x \in [\frac{1}{2}, 1) \end{cases}$$

Zatem $\varphi^{-1}((a, b)) = \left(\frac{a}{2}, \frac{b}{2}\right) \cup \left(\frac{a+1}{2}, \frac{b+1}{2}\right)$. Stąd oraz z własności miary Lebesgue'a m

$$m(\varphi^{-1}((a, b))) = m\left(\left(\frac{a}{2}, \frac{b}{2}\right)\right) + m\left(\left(\frac{a+1}{2}, \frac{b+1}{2}\right)\right) = \frac{1}{2}(b - a) + \frac{1}{2}(b - a) = m((a, b)).$$

Czyli odwzorowanie φ także zachowuje miarę Lebesgue'a m. Powyższe rozważania przenoszą się na wszystkie odwzorowania postaci $\varphi(z) = z^n$, dla pewnego $n > 1$. Zauważmy, że okrąg S^1 wraz z mnożeniem zespolonym jest grupą. Natomiast odwzorowanie φ jest endomorfizmem tej grupy - zachowuje iloczyn: $\varphi(z_1 \cdot z_2) = (z_1 \cdot z_2)^n = z_1^n \cdot z_2^n = \varphi(z_1)\varphi(z_2)$ dla $z_1, z_2 \in \mathbb{C}$.

Przykład 1.16 (Topologiczne przesunięcia)**.** Topologiczne przesunięcia nazywamy również "shiftami". Zadajmy przestrzeń

$$\Sigma_n := \{1, ..., n\}^{\mathbb{N}} = \{(x_1, x_2, ...) : x_i \in \{1, ..., n\}\}.$$

Wraz z topologią produktową tworzy ona zwartą przestrzeń metryzowalną. Wynika to z Twierdzenia Tichonowa, patrz [11]. Bazę topologii na Σ_n tworzą zbiory $A_{y_1,...,y_n} := \{(y_1, ..., y_k, x_{k+1}, ...) : x_j \in \{1, ..., n\}$ dla $j > k$, gdzie $y_1, ..., y_n \in \{1, ..., n\}$ są ustalone. *Topologicznym przesunięciem* nazywamy odwzorowanie $\sigma_n : \Sigma_n \to \Sigma_n$, gdzie

$$\sigma_n(x_1, x_2, ...) = (x_2, x_3, ...).$$

Podstawowym obiektem badań teorii ergodycznej są układy i miary ergodyczne:

Definicja 1.17. Miarowy układ dynamiczny (X, φ) na przestrzeni z miarą (probabilistyczną) (X, Σ, μ) nazywamy *ergodycznym* jeżeli zachodzi implikacja

$$\forall_{A\in\Sigma}\ \varphi^{-1}(A) = A \implies \mu(A) = 0 \text{ lub } \mu(A) = 1.$$

Mówimy wtedy też, że miara μ jest *φ-ergodyczna* lub że odwzorowanie $\varphi : X \to X$ jest *ergodyczne względem μ.*

Uwaga. Implikacja w powyższej definicji mówi, że układ ergodyczny (X, φ) jest nierozkładalny na mniejsze części (które byłyby istotne z punktu widzenia teorii miary). Rzeczywiście, jeżeli dla pewnego $A \in \Sigma$ mamy $\varphi^{-1}(A) = A$, to wtedy również $\varphi^{-1}(X\backslash A) = X\backslash A$, więc zamiast badać układ (X, φ) możemy badać jego "dwa podukłady" $\varphi : A \to A$ oraz $\varphi : X\backslash A \to X\backslash A$. Przy czym taka redukcja ma sens jedynie, jeśli $0 < \mu(A) < 1$.

Definicja 1.18. Jeżeli (X, φ) jest topologicznym układem dynamicznym, to zbiór borelowskich miar φ-ergodycznych oznaczamy przez

$$\mathrm{Erg}(X, \varphi) := \{\mu \in \mathcal{M}(X, \varphi) : \mu \text{ jest } \varphi\text{-ergodyczna}\}.$$

Przypomnijmy, że podzbiór A przestrzeni liniowej nazywamy *zbiorem wypukłym*, jeżeli wraz z każdymi dwoma swoimi punktami zawiera odcinek je łączący, tzn. jeśli $x_1, x_2 \in A$ to dla każdego $t \in [0, 1]$ kombinacja wypukła $tx_1 + (1 - t)x_2$ także należy do A. Punkty zbioru wypukłego A, które nie są kombinacjami wypukłymi dwóch różnych punktów (nie leżą na żadnym odcinku o końcach w A), nazywamy *punktami ekstremalnymi* tego zbioru.

Twierdzenie 1.19. *Niech (X, φ) będzie topologicznym układem dynamicznym. Zbiór miar niezmienniczych $\mathcal{M}(X, \varphi) \subseteq \mathcal{M}(X)$ jest zbiorem wypukłym i zwartym w *-słabej topologii. Miary ergodyczne* $\mathrm{Erg}(X, \varphi)$ *są punktami ekstremalnymi tego zbioru, tzn. jeśli $\mu \in \mathcal{M}(X, \varphi)$, to*

$$\mathrm{Erg}(X, \varphi) := \{\mu \in \mathcal{M}(X, \varphi) : \forall_{t\in[0,1]}\forall_{\mu_1,\mu_2\in\mathcal{M}(X,\varphi)}\ \mu = t\mu_1 + (1 - t)\mu_2 \Rightarrow \mu = \mu_1 = \mu_2\}.$$

Dowód. Aby wykazać zwartość $\mathcal{M}(X,\varphi)$ wystarczy pokazać, że jest to domknięty podzbiór zbioru $\mathcal{M}(X)$, który wiemy, że jest zwarty na mocy Twierdzenia 1.7. Niech $\{\mu_n\}_{n=1}^{\infty}$ będzie ciągiem z $\mathcal{M}(X,\varphi)$ i $\mu_n \to \mu$ w $\mathcal{M}(X)$. Wtedy

$$\int f d\varphi_* \mu = \int f \circ \varphi d\mu = \lim_{n\to\infty} \int f \circ \varphi d\mu_n = \lim_{n\to\infty} \int f d\mu_n = \int f d\mu,$$

zatem $\mu \in \mathcal{M}(X,\varphi)$, co dowodzi domkniętości zbioru $\mathcal{M}(X,\varphi)$. Zatem $\mathcal{M}(X,\varphi)$ jest zwarty, jako domknięty podzbiór zbioru zwartego $\mathcal{M}(X)$.

Wypukłość wynika z tego, że jeśli $\mu_1, \mu_2 \in \mathcal{M}(X,\varphi)$ i $\lambda \in [0,1]$ to $\lambda\mu_1 + (1-\lambda)\mu_2 \in \mathcal{M}(X,\varphi)$.

Przejdźmy teraz do dowodu ostatniej części twierdzenia. Załóżmy, że $\mu \in \mathcal{M}(X,\varphi)$ i μ nie jest miarą ergodyczną. Wtedy istnieje zbiór borelowski B taki, że $\varphi^{-1}(B) = B$ i $0 < \mu(B) < 1$. Zdefiniujmy miary μ_1 i μ_2 w następujący sposób

$$\mu_1(A) = \frac{\mu(A \cap B)}{\mu(B)} \quad \text{i} \quad \mu_2(A) = \frac{\mu(A \cap (X \setminus B))}{\mu(X \setminus B)},$$

gdzie A także jest zbiorem borelowskim. Łatwo zauważyć, że $\mu_1, \mu_2 \in \mathcal{M}(X,\varphi)$, ponieważ

$$\begin{aligned}\mu_1(\varphi^{-1}(A)) &= \frac{\mu(\varphi^{-1}(A) \cap B)}{\mu(B)} = \frac{\mu(\varphi^{-1}(A) \cap \varphi^{-1}(B))}{\mu(B)} = \frac{\mu(\varphi^{-1}(A \cap B))}{\mu(B)} \\ &= \frac{\mu(A \cap B)}{\mu(B)} = \mu_1(A).\end{aligned}$$

Analogicznie mamy dla $\mu_2(A)$. Ponadto

$$\mu(A) = \mu(B)\mu_1(A) + (1-\mu(B))\mu_2(A).$$

Zatem μ nie jest punktem ekstremalnym zbioru $\mathcal{M}(X,\varphi)$.

Z drugiej strony, niech $\mu \in \mathcal{M}(X,\varphi)$ będzie miarą ergodyczną i

$$\mu = p\mu_1 + (1-p)\mu_2,$$

gdzie $\mu_1, \mu_2 \in \mathcal{M}(X,\varphi)$ i $p \in (0,1)$. Miara μ_1 jest absolutnie ciągła względem miary μ, tzn. $\mu(A) = 0$ implikuje $\mu_1(A) = 0$. Zatem istnieje pochodna Radona-Nikodyma $\frac{d\mu_1}{d\mu}$, tzn. dla każdego zbioru borelowskiego B mamy

$$\mu_1(B) = \int_B \frac{d\mu_1(x)}{d\mu} d\mu(x),$$

patrz [30, Twierdzenie 0.10]. Mamy, że $\frac{d\mu_1}{d\mu} \geqslant 0$. Niech $B = \{x : \frac{d\mu_1}{d\mu}(x) < 1\}$. Wtedy

$$\begin{aligned}\int_{B\cap\varphi^{-1}(B)} \frac{d\mu_1}{d\mu} d\mu + \int_{B\setminus\varphi^{-1}(B)} \frac{d\mu_1}{d\mu} d\mu &= \mu_1(B) = \mu_1(\varphi^{-1}(B)) \\ &= \int_{B\cap\varphi^{-1}(B)} \frac{d\mu_1}{d\mu} d\mu + \int_{\varphi^{-1}(B)\setminus B} \frac{d\mu_1}{d\mu} d\mu.\end{aligned}$$

Zatem

$$\int_{B\setminus\varphi^{-1}(B)} \frac{d\mu_1}{d\mu} d\mu = \int_{\varphi^{-1}(B)\setminus B} \frac{d\mu_1}{d\mu} d\mu.$$

Skoro $\frac{d\mu_1}{d\mu} < 1$ na $B \setminus \varphi^{-1}(B)$ i $\frac{d\mu_1}{d\mu} \geqslant 1$ na $\varphi^{-1}(B) \setminus B$ i z tego, że $\mu(\varphi^{-1}(B) \setminus B) = \mu(\varphi^{-1}(B)) - \mu(\varphi^{-1}(B) \cap B) = \mu(B) - \mu(\varphi^{-1}(B) \cap B) = \mu(B \setminus \varphi^{-1}(B))$ mamy $\mu(B \setminus \varphi^{-1}(B)) = 0 = \mu(\varphi^{-1}(B) \setminus B)$. Stąd $\mu(\varphi^{-1}(B) \triangle B) = 0$, a zatem $\mu(B) = 0$ lub 1, patrz [30, Twierdzenie 1.5]. Jeśli $\mu(B) = 1$ to $\mu_1(X) = \int_B \frac{d\mu_1}{d\mu} d\mu < \mu(B) = 1$, czyli $\mu_1(X) = 1$. Zatem $\mu(B)$ jest równe zeru.

Kładąc $C = \{x : \frac{d\mu_1}{d\mu} > 1\}$ w analogiczny sposób jak powyżej otrzymujemy, że $\mu(C) = 0$. Zatem $\frac{d\mu_1}{d\mu} = 1$ i stąd $\mu_1 = \mu$. Czyli μ jest punktem ekstremalnym zbioru $M(X, \varphi)$. $\square$

Twierdzenie Krejna-Milmana mówi, że każdy zwarty zbiór wypukły (w lokalnie wypukłej przestrzeni liniowo-topologicznej) jest rozpięty przez swoje punkty ekstremalne (składa się z kombinacji wypukłych takich punktów). W szczególności każdy taki zbiór posiada punkty ekstremalne. Zatem łącząc powyższe twierdzenie z Twierdzeniem Bogoliubowa-Krylowa (patrz Twierdzenie 1.13) otrzymujemy:

Wniosek 1.20. *Dla każdego topologicznego układu dynamicznego (X, φ) zbiór $\mathrm{Erg}(X, \varphi)$ jest niepusty, tzn. istnieje borelowska miara φ-ergodyczna.*

Miary ergodyczne można również scharakateryzować w języku operatorów kompozycji, o czym mówi następujące twierdzenie:

Twierdzenie 1.21. *Niech $\varphi : X \to X$ będzie odwzorowaniem zachowującym miarę μ oraz niech $p \in [1, \infty)$. Wtedy następujące warunki są równoważne:*

1) Miara μ jest φ-ergodyczna.

2) Jeśli f jest funkcją mierzalną na X oraz $f \overset{\mu-p.w}{=} f \circ \varphi$ to $f \overset{\mu-p.w}{\equiv} const.$

3) Jedynka jest prostą własnością operatora kompozycji $T_\varphi : L^p_\mu(X) \to L^p_\mu(X)$, tzn. jeśli $f \in L^p_\mu(X)$ i $T_\varphi f = f$, to istnieje $\lambda \in \mathbb{C}$ takie, że $f = \lambda 1$ w $L^p_\mu(X)$.

Dowód. Przejście z 2) do 3) jest jasne.

3)⇒1). Niech $\varphi^{-1}(A) = A$ dla pewnego zbioru borelowskiego A. Wtedy funkcja charakterystyczna $\mathbf{1}_A$ należy do $L^p_\mu(X)$, ponieważ

$$\|\mathbf{1}_A\|_{L^p} = \left(\int \mathbf{1}^p_A d\mu\right)^{\frac{1}{p}} = \mu(A)^{\frac{1}{p}} < \infty,$$

oraz

$$T_\varphi \mathbf{1}_A = \mathbf{1}_A \circ \varphi = \mathbf{1}_{\varphi^{-1}(A)} = \mathbf{1}_A.$$

Zatem z 3) $\mathbf{1}_A$ jest μ-prawie wszędzie stała, więc jest μ-prawie wszędzie równa zero bądź jeden, a stąd $\mu(A) = 0$ lub 1.

1)⇒2). Niech $f = f \circ \varphi$ prawie wszędzie i niech φ będzie odwzorowaniem μ-ergodycznym.

Możemy założyć, że funkcja $f : X \to \mathbb{R}$ (przechodząc ewentualnie do części rzeczywistej bądź zespolonej funkcji f). Dla każdego $k \in \mathbb{Z}$ i $n > 0$ połóżmy

$$X_{k,n} = \left\{x \in X : \frac{k}{2^n} \leqslant f(x) < \frac{k+1}{2^n}\right\} = f^{-1}\left(\left[\frac{k}{2^n}, \frac{k+1}{2^n}\right)\right).$$

Wówczas

$$\varphi^{-1}(X_{k,n}) \triangle X_{k,n} \subseteq \{x : f(\varphi(x)) \neq f(x)\}.$$

Jeśli $x \in \varphi^{-1}(X_{k,n}) \setminus X_{k,n}$ to $\varphi(x) \in X_{k,n}$ i $x \notin X_{k,n}$. Natomiast, jeśli $x \in X_{k,n} \setminus \varphi^{-1}(X_{k,n})$ to $x \in X_{k,n}$ i $\varphi(x) \notin X_{k,n}$. Zatem $\mu(\varphi^{-1}(X_{k,n}) \triangle X_{k,n}) = 0$ i z [30, Twierdzenie 1.5] $\mu(X_{k,n}) = 0$ bądź $\mu(X_{k,n}) = 1$. Dla każdego n, $X = \bigsqcup_{k \in \mathbb{Z}} \mu(X_{k,n})$, zatem jeśli $\mu(X) = 1$ to istnieje tylko jedno takie k dla którego $\mu(X_{k,n}) = 1$. Połóżmy $Y = \bigcap_{n=1}^{\infty} X_{k_n,n}$. Wtedy $\mu(Y) = 1$, ponieważ

$$\mu(X \setminus Y) = \mu\left(\bigcup_{n=1}^{\infty} X \setminus X_{k_n,n}\right) \leqslant \sum_{n=1}^{\infty} \mu(X \setminus X_{k_n,n}) = \sum_{n=1}^{\infty} \mu(X) - \mu(X_{k_n,n}) = 1 - 1 = 0.$$

Ponadto, jeśli pewne x_1, x_2 należą do Y to dla każdego n, x_1, x_2 należą także do $X_{k_n,n}$. Zatem, dla każdego n, $|f(x_1) - f(x_2)| < \frac{1}{2^n} \overset{n\to\infty}{\to} 0$, a stąd $f(x_1) = f(x_2)$. Zatem funkcja f obcięta do Y jest stała i $\mu(X \setminus Y) = 0$. □

Na koniec przytoczymy fundamentalne twierdzenie Teorii Ergodycznej. Mówi ono, że dla układów ergodycznych średnia po czasie jest równa średniej po przestrzeni. Dowód powyższego twierdzenia można znaleźć w [30, Twierdzenie 1.14].

Twierdzenie 1.22 (Twierdzenie Ergodyczne Birkhoffa-Chinczyna). *Jeśli μ jest miarą φ-ergodyczną, to dla każdej funkcji całkowalnej $f \in L^1_\mu(X)$ mamy, że*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f \circ \varphi^i = \int_X f d\mu \quad \mu\text{-}p.w.$$

Jako ilustrację Twierdzenia Ergodycznego pokażemy jak można z niego otrzymać Mocne Prawo Wielkich Liczb Kołmogorowa. Niech $(\Omega, \mathcal{F}, P)$ będzie przestrzenią probabilistyczną z prawdopodobieństwem P i σ-algebrą $\mathcal{F}$. Połóżmy $X = \Omega^{\mathbb{N}}$. Niech $\Sigma = \mathcal{F} \bigotimes \mathcal{F} \bigotimes ...$ będzie σ-algebrą produktową i $\mu = P \bigotimes P \bigotimes ...$ miarą produktową. Wtedy (X, Σ, μ) także jest przestrzenią probabilistyczną.

Twierdzenie 1.23. *Jeśli (X, Σ, μ) jest przestrzenią probabilistyczną zadaną tak jak powyżej, to przesunięcie Bernoulliego $\sigma : X \to X$ zadane wzorem $\sigma(\omega_1, \omega_2, ...) = (\omega_2, \omega_3, ...)$ jest ergodyczne względem μ.*

Dowód. Patrz [30, Twierdzenie 1.12]. □

Wniosek 1.24 (Mocne Prawo Wielkich Liczb, Kołmogorow 1930). *Niech $\{\xi\}_{i=1}^{\infty}$ będzie ciągiem zmiennych losowych o tym samym rozkładzie $\xi_i \sim \xi$ posiadającym wartość oczekiwaną $E(\xi)$. Wtedy*

$$\frac{1}{n} \sum_{i=1}^{n} \xi_i \overset{p.w}{\to} E(\xi).$$

Dowód. Niech (Ω, Σ, μ)-przestrzeń probabilistyczna taka jak w Twierdzeniu 1.23, a ξ-zmienna losowa. Wtedy $\xi_i \sim \xi$ ciągiem niezależnych zmiennych losowych. Ponadto $\xi_i = \xi_1 \circ \sigma^{i-1}$, gdzie σ jest przesunięciem Bernoulliego. Z Twierdzenia 1.23 wiemy, że $\mu = P \bigotimes P \bigotimes ...$ jest σ-ergodyczna. Zatem, z Twierdzenia Ergodycznego $\frac{1}{n}\sum_{i=1}^{n}\xi_i = \frac{1}{n}\sum_{i=0}^{n-1}\xi_1\sigma^i \overset{\mu-p.w.}{\rightarrow} \int_X \xi_1 d\mu = E(\xi_1) = E(\xi)$. □

Rozdział 2

Operatory przejścia i ich promień spektralny

W tym rozdziale opiszemy ogólną postać operatorów przejścia dla topologicznych układów dynamicznych. Wyjaśnimy ich związek z operatorami Ruelle'a-Perrona-Frobeniusia, które są głównym obiektem w niniejszej pracy. Podamy również ogólną Zasadę Wariacyjną na promień spektralny operatorów przejścia otrzymaną niedawno przez Antonevicha, Bakthina i Lebedeva oraz pokażemy jak ją zastosować do obliczenia promienia spektralnego ważonych operatorów przesunięcia.

2.1 Operatory przejścia i dodatnie funkcjonały

Ten podrozdział został napisany w oparciu o publikacje [5, 18]. Niech (X, φ) będzie topologicznym układem dynamicznym. Czyli X jest zwartą przestrzenią metryczną, natomiast $\varphi : X \to X$ jest odwzorowaniem ciągłym.

Definicja 2.1. *Operatorem przejścia* dla układu (X, φ) nazywamy operator liniowy $\mathcal{L} : C(X) \to C(X)$, który

(1) jest *dodatni* tzn. $\mathcal{L}(f) \geqslant 0$ jest funkcją nieujemną dla każdej funkcji nieujemnej $f \geqslant 0$,

(2) spełnia *homologiczną tożsamość*

$$\mathcal{L}((f \circ \varphi) \cdot g) = f \cdot \mathcal{L}(g) \tag{2.1}$$

dla wszystkich $f, g \in C(X)$.

Uwaga. Tożsamość (2.1) można zapisać, korzystając z operatora kompozycji $T_\varphi : C(X) \to C(X)$, por. Stwierdzenie 1.12, jak następuje:

$$\mathcal{L}(T_\varphi(f) \cdot g) = f \cdot \mathcal{L}(g), \qquad f, g \in C(X).$$

Przy dodatkowym założeniu, że $\mathcal{L} : C(X) \to C(X)$ zachowuje jedynkę, tzn. $\mathcal{L}(1) = 1$, otrzymujemy, że operator $\mathcal{L}$ jest ledwo odwrotny do T_φ. Rzeczywiście mamy wtedy

$$\mathcal{L}(T_\varphi(f)) = \mathcal{L}(T_\varphi(f) \cdot 1) \overset{(2.1)}{=} f \cdot \mathcal{L}(1) = f, \qquad \text{dla } f \in C(X).$$

Czyli $\mathcal{L} \circ T_\varphi = id_{C(X)}$. Na ogół operatorów lewo odwrotnych do T_φ jest nieskończenie wiele. Jeżeli jednak φ jest homeomorfizmem, to operator T_φ jest odwracalny i wtedy $\mathcal{L} = (T_\varphi)^{-1} = T_{\varphi^{-1}}$ jest operatorem kompozycji z odwzorowaniem odwrotnym $\varphi^{-1} : X \to X$.

Przykład 2.2 (Ważone operatory przesunięcia)**.** Jeśli $\varphi : X \to X$ jest homeomorfizmem i $a \in C(X)$ funkcją ciągłą, to operator kompozycji z wagą a, zwany też *ważonym operatorem przesunięcia*, definiujemy wzorem

$$aT_{\varphi^{-1}}(f)(x) = a(x) f(\varphi^{-1}(x)).$$

Jeśli $a \leqslant 0$ jest funkcją nieujemną, to aT jest operatorem przejścia. Rzeczywiście jest on dodatni, bo a jest nieujemna i spełnia (2.1), ponieważ

$$\begin{aligned} aT_{\varphi^{-1}}(f \circ g)(x) &= a(x)(f \circ \varphi \circ \varphi^{-1})(x) g(\varphi^{-1}(x)) = f(x) a(x) g(\varphi^{-1}(x)) \\ &= f aT_{\varphi^{-1}} g(x). \end{aligned}$$

Poniżej wykażemy, że jeżeli $\varphi : X \to X$ jest homeomorfizmem, to każdy operator przejścia dla (X, φ) jest tej postaci.

Klasycznymi przykładami operatorów przejścia są operatory Ruelle'a-Perrona-Frobeniusa związane z lokalnymi homeomorfizmami. Przypomnijmy, że odwzorowanie $\varphi : X \to X$ jest *lokalnym homeomorfizmem*, jeśli dla każdego punktu $x \in X$ jesteśmy w stanie znaleźć takie otoczenie otwarte $U \subseteq X$, że zbiór $\varphi(U)$ jest otwarty w X oraz obcięcie $\varphi|_U : U \to \varphi(U)$ jest homeomorfizmem między podprzestrzeniami topologicznymi U i $\varphi(U)$. Oczywiście każdy homeomorfizm jest lokalnym homeomorfizmem. Odwzorowania z Przykładów 1.15, 1.16 są modelowymi przykładami lokalnych homeomorfizmów.

Definicja 2.3. *Operatorem Ruelle'a-Perrona-Frobeniusa* związanym z lokalnym homeomorfizmem $\varphi : X \to X$ oraz nieujemną funkcją ciągłą $c \in C(X)$, nazywamy operator $\mathcal{L}_c : C(X) \to C(X)$ dany wzorem

$$\mathcal{L}_c(f)(y) = \sum_{x \in \varphi^{-1}(y)} c(x) f(x), \qquad f \in C(X), y \in X.$$

Uwaga. Zauważmy, że dla dowolnej nieujemnej funkcji $c \in C(X)$ operator Ruelle'a-Perrona-Frobeniusa jest $\mathcal{L}_c$ jest poprawnie określony i jest operatorem przejścia. Rzeczywiście, skoro φ jest lokalnym homeomorfizmem, a X jest zwarty, to istnieje skończone pokrycie otwarte $\{U_i\}_{i=1}^n$ przestrzeni X takie, że $\varphi : U_i \to \varphi(U_i)$ jest homeomorfizmem dla każdego i. W szczególności, dla każdego $y \in X$ przeciwobraz $\varphi^{-1}(y)$ jest skończony, ma co najwyżej n elementów. Zatem suma $\sum_{x \in \varphi^{-1}(y)} c(x) f(x)$ jest skończona. Co więcej, przy ciągłej zmianie y suma ta zmienia się w sposób ciągły. Czyli $\mathcal{L}_c : C(X) \to C(X)$ jest poprawnie określonym operatorem. Liniowość jest jasna, a dodatniość wynika z nieujemności funkcji c. Tożsamość homologiczną weryfikuje prosty rachunek:

$$\begin{aligned}\mathcal{L}_c((f\circ\varphi)\cdot g)(y) &= \sum_{x\in\varphi^{-1}(y)} c(x)(f\circ\varphi)(x)g(x) = \sum_{x\in\varphi^{-1}(y)} c(x)f(\varphi(x))g(x)\\ &= \sum_{x\in\varphi^{-1}(y)} c(x)f(y)g(x) = f(y)\mathcal{L}_c(g)(y).\end{aligned}$$

Jak wyjaśnimy poniżej każdy operator przejścia związany z lokalnym homeomorfizmem jest operatorem Ruelle'a-Perrona-Frobeniusa dla pewnego $c \in C(X)$, patrz [18, Stwierdzenia 2.1, 2.6] (wziąć $\Delta = X$), [4, Podrodział 2.2].

Teraz omówimy ogólną postać operatorów przejścia. Zaczniemy od ważnego i często wykorzystywanego faktu, że operatory dodatnie są automatycznie ograniczone i ich norma realizuje się na jedynce:

Stwierdzenie 2.4. *Każdy operator liniowy dodatni $\mathcal{L} : C(X) \to C(X)$ jest ograniczony oraz $\|\mathcal{L}\| = \|\mathcal{L}(1)\|$.*

Dowód. Początkowo udowodnimy ograniczoność liniowego, dodatniego funkcjonału, powiedzmy $\psi : C(X) \to \mathbb{C}$. Dla funkcji $f \in C(X)$ o rzeczywistym zbiorze wartości i $\|f\| \leqslant 1$ mamy, że $1 \geqslant f \geqslant -1$, tzn. $1 - f \geqslant 0$ i $1 + f \geqslant 0$. Stąd oraz z liniowości i dodatniości ψ dla funkcji rzeczywistej dostajemy

$$\|f\| \leqslant 1 \Longrightarrow \psi(1-f) = \psi(1) - \psi(f) \geqslant 0 \quad \text{i} \quad \psi(1+f) = \psi(1) + \psi(f) \geqslant 0,$$

co oznacza, że

$$\|f\| \leqslant 1 \Longrightarrow |\psi(f)| \leqslant \psi(1). \tag{2.2}$$

Dowodzi to ograniczoności ψ dla funkcji rzeczywistych. Teraz przejdźmy do dowodu tej ograniczoności dla wszystkich ciągłych funkcji. Dowolną funkcję $f \in C(X)$ możemy zapisać jako kombinację liniową dwóch funkcji rzeczywistych:

$$f = f_1 + if_2, \quad \text{gdzie} \quad f_1 := \mathrm{Re} f, f_2 = \mathrm{Im}\, f,$$

i wtedy

$$|f_i| \leqslant |f|, \quad i = 1, 2.$$

Z powyższych nierówności oraz z (2.2) dostajemy

$$\|f\| \leqslant 1 \Rightarrow |\psi(f)| = |\psi(f_1 + if_2)| = |\psi(f_1) + i\psi(f_2)| \leqslant |\psi(f_1)| + |\psi(f_2)| \leqslant 2\psi(1),$$

a to dowodzi ograniczoności ψ.

Przejdźmy teraz do właściwej części dowodu. Zauważmy, że wystarczy jedynie pokazać, że $\|\mathcal{L}\| = \sup_{\|f\|=1} \|\mathcal{L}(f)\| = \|\mathcal{L}(1)\|$, gdyż stąd, że $\|\mathcal{L}(1)\| < \infty$ dowiedzie to również ograniczoności $\mathcal{L}$. Dla każdego $x \in X$ zdefiniujmy funkcjonał $\mu_x : C(X) \to \mathbb{C}$ wzorem

$$\mu_x(f) := \mathcal{L}(f)(x), \qquad f \in C(X)$$

Z liniowości i dodatniości operatora $\mathcal{L}$ wynika, że funkcjonał μ_x jest liniowy i dodatni. Zatem μ_x jest ograniczony na mocy pierwszej części dowodu. Stąd na mocy Twierdzenia Riesza (Twierdzenie 1.4) możemy utożsamić μ_x z pewną miarą borelowską na X i wtedy

$$\mu_x(f) = \int_X f(t)d\mu_x(t), \qquad f \in C(X).$$

Dalej mamy, że

$$\begin{aligned}\|\mathcal{L}\| &= \sup_{\|f\|=1} \|\mathcal{L}(f)\| = \sup_{\|f\|=1} \max_{x\in X} |\mathcal{L}(f)(x)| = \sup_{\|f\|=1} \max_{x\in X} |\mu_x(f)| = \sup_{\|f\|=1} \max_{x\in X} \Big|\int_X f(t)d\mu_x(t)\Big| \\ &\leqslant \sup_{\|f\|=1} \max_{x\in X} \int_X |f(t)|d\mu_x(t) \leqslant \max_{x\in X} \int_X 1d\mu_x(t) = \max_{x\in X} \alpha_x(1) = \|\mathcal{L}(1)\|.\end{aligned}$$

Czyli $\|\mathcal{L}\| \leqslant \|\mathcal{L}(1)\|$. Z drugiej strony $\|\mathcal{L}(1)\| \leqslant \|\mathcal{L}\| \cdot \|1\| = \|\mathcal{L}\|$. □

Ustalmy operator przejścia $\mathcal{L} : C(X) \to C(X)$ dla dowolnego odwzorowania ciągłego $\varphi : X \to X$. Tak jak w dowodzie Stwierdzenia 2.4, dla każdego punktu $x \in X$ zdefiniujmy funkcjonał $\mu_x : C(X) \to \mathbb{C}$ wzorem

$$\mu_x(f) := \mathcal{L}(f)(x), \quad f \in C(X), \tag{2.3}$$

i utożsammy go za pomocą Twierdzenia Riesza (Twierdzenie 1.4) z pewną miarą borelowską na X. Homologiczna tożsamość oznacza także, że dowolnej funkcji $f \in C(X)$ zachodzi

$$\mathcal{L}(f \circ \varphi)(x) = \mathcal{L}(f \circ \varphi \cdot 1)(x) = f(x) \cdot \mathcal{L}(1)(x).$$

Stąd

$$\frac{1}{\mathcal{L}(1)(x)} \mu_x(f \circ \varphi) = f(x),$$

a zatem

$$\operatorname{supp} \mu_x \subseteq \varphi^{-1}(x). \tag{2.4}$$

Dla punktu x możliwe są dwa przypadki:

(1) $\mathcal{L}(f)(x) = 0$. To znaczy, że $\mu_x(1) = 0$, dlatego $\mu_x = 0$ ponieważ μ_x jest dodatnim funkcjonałem,

(2) $\mathcal{L}(f)(x) \neq 0$. W tym przypadku $\mu_x \neq 0$ i z Twierdzenie Riesza (Twierdzenie 1.4) możemy mu przyporządkować miarę μ_x na X.

Jasnym jest, że odwzorowanie $x \to \mu_x$ jest ciągłe w $*$-słabej topologii.

Lemat 2.5. *Jeśli* $x \notin \varphi(X)$ *to* $\mathcal{L}(1)(x) = 0$.

Dowód. Załóżmy nie wprost, że $\mathcal{L}(1)(x) \neq 0$. Wtedy można wybrać taką funkcję $f \in C(X)$, aby $f|_{\varphi(X)} = 0$ i $f(x) = 1$. Stąd otrzymujemy sprzeczność, bo

$$0 = \frac{1}{\mathcal{L}(1)} \mathcal{L}(f \circ \varphi)(x) = f(x) = 1.$$

□

Rozważane powyżej obiekty dają pełny opis operatorów przejścia. Rzeczywiście dowolne ciągłe w $*$-słabej topologii odwzorowanie $x \to \mu_x$, gdzie μ_x są miarami borelowskimi, które spełniają:

(1) $\mu_x = 0$, gdy $x \notin \varphi(X)$,

(2) μ_x spełnia (2.4), gdy $x \in \varphi(X)$ (może tu być także $\mu_x = 0$),

zadaje operator przejścia (patrz wzór (2.5) poniżej), dla którego zachodzi (2.3). Z powyższych rozważań wynika następujące stwierdzenie:

Stwierdzenie 2.6. *Każdy operator przejścia $\mathcal{L}$ jest postaci*

$$\mathcal{L}(f)(x) = \int_{\varphi^{-1}(x)} f(y)d\mu_x(y) \qquad f \in C(X), x \in X, \tag{2.5}$$

gdzie nośnik miary μ_x jest zawarty w $\varphi^{-1}(x)$ oraz odwzorowanie

$$X \ni x \longmapsto \mu_x \in \mathcal{M}(X)$$

jest ciągłe w $*$-słabej topologii.

Dla każdego operatora $\mathcal{L}$ możemy zdefiniować funkcję $c : X \to [0, \infty)$ wzorem

$$c(x) := \mu_{\varphi(x)}(\{x\}) \quad x \in X. \tag{2.6}$$

W ogólności nie jest ona ciągła i nie określa operatora $\mathcal{L}$ jednoznacznie. Jeśli istnieje pokrycie otwarte $\{V_i\}_{i=1}^n$ zbioru X takie, że φ obcięte do V_i, $i = 1, ..., n$ jest różnowartościowe, to funkcja c jest ciągła. Faktycznie, jeśli weźmiemy $\{h_i\}_{i=1}^n \in C(X)$ jako rozkład jedynki dla tego pokrycia, patrz [26, Twierdzenie 2.13]. Wtedy

$$(\mathcal{L}(h_i) \circ \varphi)(x) = \mathcal{L}(h_i)(\varphi(x)) = c(x)h_i(x) \quad x \in V_i.$$

Stąd, dla każdego $i = 1, ..., n$ funkcja $c(x)h_i(x)$ jest ciągła (gdyż $\mathcal{L}$ przeprowadza funkcję ciągłą na funkcję ciągłą), a zatem funkcja $c = \sum_{i=1}^n c \cdot h_i$ także jest ciągła. Dodajmy jeszcze, że jeśli φ jest lokalnym homeomorfizmem to φ^{-1} jest skończony. Z tych rozważań wynika następujące stwierdzenie:

Stwierdzenie 2.7. *Jeśli φ jest lokalnym homeomorfizmem, to każdy operator przejścia $\mathcal{L}$ jest operatorem Ruelle'a-Perrona-Frobeniusa, tzn. jest zadany wzorem*

$$\mathcal{L}(f)(x) = \sum_{y \in \varphi^{-1}(x)} c(y)f(y), \quad x \in X, f \in C(X), \tag{2.7}$$

dla pewnej ciągłej nieujemnej funkcji $c \in C(X)$.

Wniosek 2.8. *Jeśli $\varphi : X \to X$ jest homeomorfizmem, to każdy operator przejścia $\mathcal{L}$ jest postaci*

$$\mathcal{L}(f)(x) = c(x)f(\varphi^{-1}(x))),$$

dla pewnej nieujemnej funkcji ciągłej $c \in C(X)$. To znaczy, $\mathcal{L}$ jest ważonym operatorem przesunięcia.

Dowód. Skoro φ jest homeomorfizmem to suma w (2.7) posiada tylko jeden składnik dla $y = \varphi^{-1}(x)$. $\square$

2.2 Promień spektralny operatora przejścia

W tym podrozdziale przedstawimy ogólną zasadę wariacyjną wyrażającą promień spektralny dowolnych operatorów przejścia otrzymaną przez Antonevicha, Bakhtina i Lebedeva w [2]. W otrzymanym wzorze pojawia się pewien funkcjonał zależny od operatora przejścia, którego autorzy [2] nazwali t-entropią operatora przejścia. W pracy [3] definicję t-entropii nieco uproszczono i tę definicję tu przytoczymy.

Ustalmy operator przejścia $\mathcal{L} : C(X) \to C(X)$ dla pewnego topologicznego układu dynamicznego (X, φ). Na mocy Stwierdzenia 2.7 operator $\mathcal{L}$ jest dany wzorem (2.5). Z takim operatorem możemy związać rodzinę operatorów $\mathcal{L}_{e^b} : C(X) \to C(X)$ zależnych od funkcjonalnego parametru $b \in C(X, \mathbb{R})$ (b ma wartości rzeczywiste), które działają na funkcję $f \in C(X)$ w następujący sposób

$$\mathcal{L}_{e^b} f := \mathcal{L}(e^b f).$$

Operatory tego typu są głównymi obiektami badanymi w teorii spektralnej układów dynamicznych, formalizmie termodynamicznym, teorii eliptycznej równań różniczkowych itp. Autorzy [2], [3] funkcjonał $\lambda : C(X, \mathbb{R}) \to \mathbb{R}$, gdzie $\lambda(b)$ jest logarytmem naturalnym z promienia spektralnego operatora $\mathcal{L}_{e^b}$ nazywają *spektralnym potencjałem* operatora przejścia $\mathcal{L}$. Zatem dla $b \in C(X, \mathbb{R})$, ze wzoru Gelfanda otrzymujemy, że

$$\lambda(b) = \ln r(\mathcal{L}_{e^b}) = \lim_{n\to\infty} \frac{1}{n} \ln \|\mathcal{L}^n_{e^b}\|.$$

Zauważmy, że $\lambda(0) = \ln r(\mathcal{L}_{e^0}) = \ln r(\mathcal{L})$. Ponadto z dodatniości operatora $\mathcal{L}^n_{e^b}$, patrz Stwierdzenie 2.4, mamy

$$\lambda(b) = \lim_{n\to\infty} \frac{1}{n} \ln \|\mathcal{L}^n_{e^b}(1)\|,$$

gdzie 1 oznacza funkcję tożsamościowo równą 1 na X. Dalej z homologicznej tożsamości, mamy, że

$$\mathcal{L}^n_{e^b}(1) = \mathcal{L}^n \left(e^{\sum_{k=0}^{n-1} b\circ\varphi^k} \right) = \mathcal{L}^n \left(\prod_{k=0}^{n-1} e^{b\circ\varphi^k} \right).$$

Stąd otrzymujemy skomplikowany wzór

$$\lambda(b) = \lim_{n\to\infty} \frac{1}{n} \max_{x_1\in X} \ln \left| \int_{\varphi^{-1}(x_1)} ... \int_{\varphi^{-1}(x_n)} \prod_{k=0}^{n-1} e^{b\circ\varphi^k} d\mu_{x_n} ... d\mu_{x_1} \right|.$$

W szczególności,

$$\ln r(\mathcal{L}) = \lambda(0) = \lim_{n\to\infty} \frac{1}{n} \max_{x_1\in X} \ln \left| \int_{\varphi^{-1}(x_1)} ... \int_{\varphi^{-1}(x_n)} 1 d\mu_{x_n} ... d\mu_{x_1} \right|.$$

Powyższe skomplikowane wzory autorzy [2], [3] zamienili na inny skomplikowany wzór:

Definicja 2.9 ([2], [3]). Niech $\mathcal{L} : C(X) \to C(X)$ będzie operatorem przejścia. *T-entropią* operatora $\mathcal{L}$ nazywamy funkcjonał $\tau_{\mathcal{L}} : \mathcal{M}(X, \varphi) \to [-\infty, \infty)$ określony na zbiorze wszystkich borelowskich miar φ-niezmienniczych dany wzorem

$$\tau_{\mathcal{L}}(\mu) = \inf_{n\in\mathbb{N}} \frac{1}{n} \inf_{G} \sum_{g_i\in G} \mu(g_i) \ln \frac{\mu(\mathcal{L}^n g_i)}{\mu(g_i)}, \tag{2.8}$$

gdzie drugie infinimum jest brane po wszystkich ciągłych rozkładach jedynki zbioru X, tzn. po rodzinach $G = \{g_1, ..., g_k\}$, gdzie $g_i \in C(X)$ są funkcjami nieujemnymi oraz $g_1 + ... + g_k \equiv 1$ na X. W (2.8) przyjmujemy konwencję, że $\ln(0) := -\infty$ oraz jeśli $\mu(g_i) = 0$, to przyjmujemy, że odpowiedni składnik w sumie jest zerem.

Jedynym z głównych wyników pracy [2] jest następujące twierdzenie:

Twierdzenie 2.10 (Zasada Wariacyjna dla spektralnego potencjału). *Niech $\mathcal{L} : C(X) \to C(X)$ będzie operatorem przejścia dla układu (X, φ). Dla dowolnego $b \in C(X, \mathbb{R})$ zachodzi*

$$\lambda(b) = \max_{\mu\in\mathcal{M}(X,\varphi)} \left(\int_X bd\mu + \tau_{\mathcal{L}}(\mu) \right). \tag{2.9}$$

Wniosek 2.11 (Zasada Wariacyjna dla promienia spektralnego operatora przejścia). *Logarytm promienia spektralnego operatora przejścia $\mathcal{L} : C(X) \to C(X)$ dla topologicznego układu dynamicznego (X, φ) jest równy maksymalnej t-entropii:*

$$\ln r(\mathcal{L}) = \lambda(0) = \max_{\mu\in\mathcal{M}(X,\varphi)} \tau_{\mathcal{L}}(\mu).$$

Dowód powyższego twierdzenia jest skomplikowany i opiera się na pojęciu transformacji Legendre'a, patrz [2]. Ponadto samo obliczenie t-entropii w konkretnych przypadkach wydaje się być trudne. Jedyny konkretny przypadek, jaki w pracach [2], [3] opisano, dotyczy sytuacji, gdy φ jest homeomorfizmem i $\mathcal{L} = T_{\varphi^{-1}}$ jest operatorem kompozycji:

Lemat 2.12. *Jeśli $\varphi : X \to X$ jest homeomorfizmem oraz $\mathcal{L} = T_{\varphi^{-1}}$, to $\tau_{\mathcal{L}}(\mu) = 0$.*

Dowód. Przy powyższych założeniach $\mathcal{L}^n = T_{\varphi^{-n}}$. Zatem dla dowolnej miary niezmienniczej $\mu \in \mathcal{M}(X, \varphi)$ mamy $\mu(\mathcal{L}^n(g)) = \mu(g)$, a stąd

$$\tau_{\mathcal{L}}(\mu) = \inf_{n\in\mathbb{N}} \frac{1}{n} \inf_{G} \sum_{g_i\in G} \mu(g_i) \ln \frac{\mu(\mathcal{L}^n g_i)}{\mu(g_i)} = \inf_{n\in\mathbb{N}} \frac{1}{n} \inf_{G} \sum_{g_i\in G} \mu(g_i) \ln 1 = 0.$$

□

Celem niniejszej pracy jest opisanie wzoru na promień spektralny operatora Ruelle'a-Perrona-Frobeniusa, tj. rozważenie przypadku, gdy φ jest lokalnym homeomorfizmem. W szczególności wyjaśnimy w tym przypadku związek t-entropii z klasyczną entropią Kołmogorowa-Sinai oraz rolą powyższych pojęć w formalizmie termodynamicznym. Zanim to nastąpi, zatrzymamy się jeszcze na chwilę w przypadku, gdy φ jest homeomorfizmem i żadna entropia się nie pojawia (por. Lemat 2.12).

2.3 Promienia spektralnego operatora ważonego przesunięcia

Jak widzieliśmy w poprzednim podrozdziale operatory przejścia dla homeomorfizmu $\varphi : X \to X$ są ważonymi operatorami przesunięcia z wagami dodatnimi. W tym podrozdziale jako ilustrację pokażemy jak z ogólnego wzóru (2.9) otrzymać zasadę wariacyjną na promień spektralny ważonych operatorów przesunięcia. Pierwotnie zasadę tę udowodnili niezależnie Kitover i Lebedev w 1979 roku [13], [19].

Ustalmy odwracalny topologiczny układ dynamiczny (X, φ). To znaczy, niech X będzie zwartą przestrzenią metryczną, a $\varphi : X \to X$ homeomorfizmem. Przypomnijmy, że operatory ważonego przesunięcia składają się z (są złożeniem) dwóch operatorów: *operatora przesunięcia* $T_\varphi : C(X) \to C(X)$ zadanego wzorem

$$T_\varphi f(x) = f(\varphi(x)), \qquad f \in C(X),$$

oraz *operatora wagi* (operatora mnożenia) $a : C(X) \to C(X)$, zadanego wzorem

$$af(x) = a(x)f(x), \qquad f \in C(X),$$

gdzie $X \ni x \to a(x) \in \mathbb{C}$ jest funkcją ciągłą. Dla operatora wagi będziemy używać tego samego symbolu, co dla funkcji która go zadaje. Zatem operatorami ważonego przesunięcia nazywamy operatory postaci $aT_\varphi : C(X) \to C(X)$, gdzie $a \in C(X)$. Wzór Gelfanda prowadzi do następującego wzoru na promień spektralny operatora ważonego przesunięcia:

Stwierdzenie 2.13. *Dla dowolnego $a \in C(X)$ promień spektralny ważonego operatora przesunięcia $aT_\varphi : C(X) \to C(X)$ wyraża się wzorem*

$$r(aT_\varphi) = \lim_{n\to\infty} \max_{x\in X} \left(\prod_{k=0}^{n-1} |a(\varphi^k(x)| \right)^{\frac{1}{n}}. \tag{2.10}$$

W szczególności $r(aT_\varphi) = r(|a|T_\varphi)$.

Dowód. Zauważmy najpierw, że norma ważonego operatora przesunięcia jest równa normie jego wagi. Rzeczywiście,

$$\begin{aligned} \|aT_\varphi\| &= \sup_{\|f\|_\infty \leqslant 1} \|aT_\varphi\|_\infty = \sup_{\|f\|_\infty \leqslant 1} \max_{x\in X} |a(x)f(\varphi(x))| \\ &\leqslant \sup_{\|f\|_\infty \leqslant 1} \max_{x\in X} |a(x)| \cdot \max_{x\in X} |f(\varphi(x))| \leqslant \max_{x\in X} |a(x)|, \end{aligned}$$

czyli $\|aT_\varphi\| \leqslant \|a\|_\infty$. Nierówność w drugą stronę otrzymujemy kładąc $f \equiv 1$. Zatem

$$\|aT_\varphi\| = \max_{x\in X} |a(x)| := \|a\|_\infty.$$

Dalej zauważmy, że $T_\varphi^n = T_{\varphi^n}$ oraz $T_\varphi af(x) = T_\varphi(af)(x) = a(\varphi(x))f(\varphi(x))$, czyli $T_\varphi a = (a \circ \varphi)T_\varphi$. Stąd

$$(aT_\varphi)^n = aT_\varphi \cdot aT_\varphi ... \cdot aT_\varphi = a \cdot (a \circ \varphi) \cdot (a \circ \varphi^{n-1})T_{\varphi^n},$$

jest ważonym operatorem przesunięcia z wagą $a\cdot(a\circ\varphi)\cdot(a\circ\varphi^{n-1})$ dla układu dynamicznego (X,φ^n). Stąd i z pierwszej uwagi otrzymujemy, że

$$\|aT_\varphi\| = \max_{x\in X}\prod_{k=0}^{n-1}|a(\varphi^k(x))|.$$

Zatem korzystając ze wzoru Gelfanda mamy

$$r(aT_\varphi) = \lim_{n\to\infty}\|(aT_\varphi)^n\|^{\frac{1}{n}} = \lim_{n\to\infty}\max_{x\in X}\left(\prod_{k=0}^{n-1}|a(\varphi^k(x)|\right)^{\frac{1}{n}}.$$

□

W niektórych przypadkach wzór (2.10) jest wystarczający, aby obliczyć promień spektralny odpowiednich operatorów. Drugim sposobem jest skorzystanie z poniższej zasady wariacyjnej, w którym kluczową rolę grają zdefiniowane wcześniej miary ergodyczne.

Twierdzenie 2.14 (Zasada wariacyjna na promień spektralny ważonego operatora przesunięcia)**.** *Niech $\varphi : X \to X$ będzie homeomorfizmem. Dla dowolnej funkcji $a \in C(X)$ mamy*

$$\ln r(aT_\varphi) = \max_{\mu\in\mathcal{M}(X,\varphi)}\int_X \ln|a(x)|d\mu = \max_{\mu\in\mathrm{Erg}(X,\varphi)}\int_X \ln|a(x)|d\mu. \tag{2.11}$$

Dowód. Oryginalny dowód Kitovera i Lebedeva działa nawet w przypadku, gdy $\varphi : X \to X$ jest dowolnym odwzorowaniem ciągłym (niekoniecznie homeomorfizmem), patrz [1] lub [17]. My pokażemy jak otrzymać ten wzór z Twierdzenia 2.10 w przypadku, gdy $a \neq 0$. Na mocy Stwierdzenia 2.13 mamy $r(aT_\varphi) = r(|a|T_\varphi)$. Operator T_φ jest operatorem przejścia dla układu (X,φ^{-1}), patrz Przykład 2.2. Zatem na mocy wzoru (2.9) mamy

$$\ln r(aT_\varphi) = r(|a|T_\varphi) = \lambda(\ln|a|) = \max_{\mu\in\mathcal{M}(X,\varphi^{-1})}\left(\int_X \ln|a(x)|d\mu + \tau_{T_\varphi}(\mu)\right).$$

Ale $\mathcal{M}(X,\varphi^{-1}) = \mathcal{M}(X,\varphi)$ oraz $\tau_{T_\varphi}(\mu) = 0$ na mocy Lematu 2.12. Zatem

$$\ln r(aT_\varphi) = \max_{\mu\in\mathcal{M}(X,\varphi)}\int_X \ln|a(x)|d\mu.$$

Skoro odwzorowanie $\mathcal{M}(X,\varphi) \ni \mu \mapsto \int_X \ln|a(x)|d\mu$ jest afiniczne (zachowuje kombinacje wypukłe), to powyższe maksimum realizuje się w punktach ekstremalnych zbioru $\mathcal{M}(X,\varphi)$, czyli na zbiorze miar ergodycznych $\mathrm{Erg}(X,\varphi)$, patrz Twierdzenie 1.19. □

Miar ergodycznych odwzorowania należy szukać w tak zwanych punktach niewędrujących.

Definicja 2.15. *Punktem niewędrującym* ciągłego odwzorowania $\varphi : X \to X$ nazywamy punkt x_0 dla którego istnieje taki zbiór otwarty U, że $x_0 \in U$ oraz

$$\varphi^n(U)\cap U \neq \emptyset \qquad \text{dla każdego } n\in\mathbb{N}.$$

Zbiór wszystkich punktów niewędrujących oznaczmy przez $\Omega(\varphi)$.

Definicja 2.16. *Nośnikiem miary* μ nazywamy zbiór punktów należących do X, których każde otoczenie otwarte posiada dodatnią miarę, tzn.

$$\operatorname{supp}\mu := \{x \in X : \mu(U) > 0 \text{ dla każego otwartego otoczenia } U \text{ punktu } x\}.$$

Lemat 2.17. *Dla każdej miary φ-niezmienniczej* $\operatorname{supp}\mu \subseteq \Omega(\varphi)$.

Dowód. Jeśli x_0 jest punktem wędrującym to istnieje otwarte otoczenie U punktu x_0 takie, że dla każdego $n > 0$ mamy $\varphi^n(U) \cap U = \emptyset$. Wtedy

$$\varphi^n(U) \cap \varphi^m(U) = \emptyset, \qquad \text{dla } n \neq m.$$

Ponieważ φ jest homeomorfizmem oraz korzystając z φ-niezmienniczości miary μ dostajemy, że

$$\mu(U) = \mu(\varphi(U)) = ... = \mu(\varphi^m(U)) = ... = \mu(\varphi^n(U)) = ...$$

Stąd wnioskujemy, że $\mu(U) = 0$, ponieważ, gdy założymy, że miary te nie są równe zeru tylko przykładowo pewnemu $\varepsilon > 0$, to

$$\mu\left(\bigcup_{n=0}^{\infty} \varphi^n(U)\right) = \sum_{n=0}^{\infty} \mu(\varphi^n(U)) = +\infty,$$

i stąd otrzymujemy, że nośnik miary jest zawarty w zbiorze punktów niewędrujących, gdyż rozważamy miary unormowane. □

Definicja 2.18. Jeżeli $\varphi : X \to X$ jest homeomorfizmem, to *orbitą punktu* $x \in X$ nazywamy zbiór $O(x) = \{\varphi^n(x)\}_{n\in\mathbb{Z}}$.

Uwaga. Dla każdego $x \in X$ orbita $O(x)$ jest zbiorem niezmienniczym. Punkty każdej skończonej orbity $O(x) = \{x_1, ..., x_n\}$ należą do zbioru $\Omega(\varphi)$. Rozkład równomierny na takiej orbicie zadaje miarę φ-ergodyczną

$$\mu := \frac{1}{n}\sum_{k=1}^{n} \delta_{x_k}$$

Jest to jedyna miara φ-ergodyczna, której nośnik przecina niepusto orbitę $O(x)$. Jest to również jedyna miara φ-niezmiennicza, której nośnik jest zawarty w $O(x)$.

Teraz, mając już wszystkie potrzebne narzędzia omówimy kilka konkretnych przykładów obliczenia promienia spektralnego operatorów ważonego przesunięcia.

Przykład 2.19. Niech $\varphi : [0,1] \to [0,1]$ będzie dane wzorem $\varphi(x) = x^2$ oraz niech $a \in C([0,1])$. Zauważmy, że z formuły (2.10) wynika, że

$$\ln r(aT_\varphi)^n) = \lim_{n\to\infty} \max_{x\in[0,1]} \frac{1}{n}\sum_{k=0}^{n-1} \ln |a(x^{2k})|.$$

Wyrażenie $\frac{1}{n}\sum_{k=0}^{n-1}\ln|a(x^{2k})|$ jest zbieżne albo do $\ln|a(0)|$, gdy $x<1$, bo wtedy $x^{2n}\to 0$ (porównaj Lemat 3.8 poniżej), albo do $\ln|a(1)|$, gdy $x=1$, bo wtedy $x^{2n}=1$. Sugeruje to więc, że

$$r(aT_\varphi)=\max\{|a(0)|,|a(1)|\}.$$

Jest to prawdą, chociaż w powyższym rozumowaniu formalnie popełniliśmy błąd: zamiast liczyć maksimum z granicy policzyliśmy maksimum granic! Jednakże, w pewnym sensie, taką zamianę uprawomocnia Twierdzenie 2.14. Rzeczywiście, jako że $\varphi^n(x)\to 0$ dla $x<1$ oraz $\varphi(1)=1$, to odwzorowanie φ ma tylko dwa punkty niewędrujące 0 i 1. Miary ergodyczne są więc skupione w tych punktach. Innymi słowy, $\mathrm{Erg}([0,1],\varphi)=\{\delta_0,\delta_1\}$. Zatem korzystając z Zasady Wariacyjnej (2.11)

$$\ln r(aT_\varphi)=\max_{\delta_0,\delta_1}\int_X \ln|a(x)|d\mu=\max\{\ln|a(0)|,\ln|a(1)|\}.$$

Czyli po zdjęciu logarytmu otrzymujemy przewidywany wynik.

Przykład 2.20 (Obrót wymierny)**.** Niech $X=S^1=\{z\in\mathbb{C}:|z|=1\}$ będzie okręgiem jednostkowym, a $\varphi(z)=e^{i\alpha}z$ będzie obrotem o kąt $\alpha\in\mathbb{R}$. Załóżmy, że $\frac{\alpha}{2\pi}\in\mathbb{Q}$, czyli $\varphi(z)=e^{2\pi i\frac{m}{n}}z$ dla pewnych liczb $m\in\mathbb{Z}$, $n\in\mathbb{N}$. Wtedy $\varphi^n(z)=(e^{2\pi i\frac{m}{n}})^n z=e^{2\pi i}z=z$ dla każdego $z\in X$. Innymi słowy odwzorowanie φ jest okresowe o okresie $n\in\mathbb{N}$: $\varphi^n=id$. Stąd orbity tego odwzorowania są postaci $O(z)=\{z,\varphi(z),\varphi^2,...,\varphi^{n-1}(z)\}$, $z\in S^1$. Zatem z Uwagi 2.3 mamy, że dla każdej miary ergodycznej μ istnieje $z\in S^1$ takie, że

$$\mu(z)=\mu(\varphi(z))=\mu(\varphi^2(z))=...=\mu(\varphi^{n-1}(z))=\frac{1}{n}.$$

Czyli innymi słowy $\mathrm{Erg}(S^1,\varphi)=\{\mu_z:z\in S^1\}$, gdzie

$$\mu_z=\frac{1}{n}(\delta_z+\delta_{\varphi^2(z)}+...+\delta_{\varphi^{n-1}(z)}).$$

Teraz stosując tę wiedzę do wzoru na Zasadę Wariacyjną (2.11) mamy

$$\ln r(aT_\varphi)=\max_{z\in S^1}\int_{S^1}\ln|a(z)|d\mu_z=\max_{z\in S^1}\sum_{k=0}^{n-1}\frac{1}{n}\ln|a(z^k)|=\max_{z\in S^1}\ln(\prod_{k=0}^{n-1}|a(\varphi^k(z))|)^{\frac{1}{n}}.$$

Skąd

$$r(aT_\varphi)=\max_{z\in S^1}(|a(z)\cdot a(\varphi(z))\cdot a(\varphi^2(z))\cdot...\cdot a(\varphi^{n-1}(z))|)^{\frac{1}{n}}.$$

Przykład 2.21 (Obrót niewymierny)**.** Niech $X=S^1=\{z\in\mathbb{C}:|z|=1\}$, $\varphi(z)=e^{i\alpha}z$ i $\frac{\alpha}{2\pi}\notin\mathbb{Q}$. Początkowo pokażemy, że w tym wypadku istnieje tylko jedna miara, dla której to odwzorowanie jest ergodyczne, a mianowicie jest to unormowana miara Lebesgue'a. Oznaczmy $a:=e^{i\alpha}$. Z niewymierności $\frac{\alpha}{2\pi}$ wynika, że $\overline{\{a^n\}}_{n\in\mathbb{Z}}=S^1$. Zatem, z [30, Twierdzenie 5.4] mamy, że φ jest odwzorowaniem minimalnym, więc jest ono ściśle ergodyczne, patrz [30, Twierdzenie 6.20]. Stąd, istnieje tylko jedna unormowana miara φ-niezmiennicza (i jest ona ergodyczna). Jest to unormowana miara Lebesgue'a, gdyż jest ona niezmiennicza względem obrotu.

Teraz, korzystając z Zasady Wariacyjnej (2.11) i utożsamiając $S^1 \ni z = e^{it}$ z $t \in [0, 2\pi)$ dostajemy

$$\ln r(aT_\varphi) = \int_{S^1} \ln |a(z)| d\mu = \frac{1}{2\pi} \int_0^{2\pi} \ln |a(t)| dt.$$

Niech przykładowo $a(z) = a(e^{it}) = e^{\sin t + i\psi(t)}$, gdzie $\psi(t)$ dowolna ciągła funkcja. Wtedy $\ln |a(e^{it})| = \sin t$ i z poprzedniego wzoru mamy

$$\ln r(aT_\varphi) = \frac{1}{2\pi} \int_0^{2\pi} \sin t dt = 0.$$

Zatem $r(aT_\varphi) = 1$.

Rozdział 3

Entropia

W poprzednim rozdziale omówiliśmy operatory przejścia oraz ich promienie spektralne. Następnym celem, po opisaniu Zasady Wariacyjnej (2.11) dla homeomorfizmu oraz będącej jej uogólnieniem Zasady Wariacyjnej (2.9) umożliwiającej obliczenie promienia spektralnego dla operatora przejścia (wzór jest wzbogacony o t-entropię) będzie opis promienia spektralnego dla operatorów Ruelle'a-Perrona-Frobeniusa. Jednym z istotnych elementów potrzebnych do opisu tych ostatnich jest entropia Kołmogorowa-Sinai'a. Co więcej jest on ściśle związany z ciśnieniem topologicznym, które jest uogólnieniem entropii topologicznej. Dlatego też, w tym rozdziale przedstawimy dokładnie pojęcie entropii metrycznej (Kołmogorowa-Sinai'a), i topologicznej, a także zasadę wariacyjną je wiążącą.

3.1 Entropia w teorii informacji

Niech (Ω, Σ, P) będzie przestrzenią probabilistyczną. Wyobraźmy sobie schemat obrazujący przekaz informacji. W tym celu potrzebne jest nam źródło danej informacji, przekaźnik oraz odbiorca informacji. Ilość informacji możemy utożsamić przykładowo z objętością pamięci komputera, bądź też ilością pytań potrzebną do rozkodowania zaszyfrowanej wiadomości. Właśnie tę ilość pytań oznaczmy jako funkcję $\mathcal{I}$. Formalnie jest to funkcja zadana na zdarzeniach z rodziny Σ, ale powinna zależeć tylko od prawdopodobieństwa $P(A)$ danego zdarzenia $A \in \Sigma$. Powinna też mieć następujące własności

(1) $P(A) \leqslant P(B) \Longrightarrow \mathcal{I}(A) \geqslant \mathcal{I}(B)$
(bardziej prawdopodobne zdarzenie łatwiej odgadnąć - potrzebujemy mniej pytań)

(2) $P(A \cap B) = P(A) \cdot P(B) \Longrightarrow \mathcal{I}(A \cap B) = \mathcal{I}(A) + \mathcal{I}(B)$,
(jeżeli zdarzenia są niezależne, to ilość pytań by je odgadnąć jest niezależna - sumuje się)

(3) $P(A) = 1 \Longrightarrow \mathcal{I}(A) = 0$,
(jeżeli zdarzenie jest pewne, to nie potrzebujemy żadnych pytań)

(4) $P(A) = 0 \Longrightarrow \mathcal{I}(A) = +\infty$.
(ile trzeba zadać pytań, żeby odgadnąć zdarzenie niemożliwe?)

Widać więc, że funkcja informacji $\mathcal{I}$ jest nieujemna oraz w niektórych wypadkach (gdy prawdopodobieństwo wynosi zero) jest równa nieskończoności. Co więcej traktując ją jako funkcję od prawdopodobieństw zdarzeń "zamienia iloczyn na sumę". Każda funkcja spełniająca te warunki jest funkcją logarytmiczną. Zatem $\mathcal{I}$ jest postaci

$$\mathcal{I}(A) = -\log_a P(A),$$

gdzie $a > 0$ pewną stała. My zwyczajowo będziemy przyjmować $a = e$, czyli będziemy używać logarytmu naturalnego. Wartość funkcji $\mathcal{I}$ na $A \in \Sigma$ nazywa się *entropią zdarzenia* A. Aby uzyskać entropię całego systemu, czyli jedną liczbę dla wszystkich $A \in \Sigma$, trzeba funkcję informacji $\mathcal{I}$ "uśrednić". Prowadzi to w naturalny sposób do entropii opisanej w następnym podrozdziale.

3.2 Entropia rozbicia

Niech (X, Σ, μ) będzie ustaloną przestrzenią probabilistyczną. Rozważmy skończone rozbicie $\mathcal{A} = \{A_i\}_{i=1}^n$ przestrzeni X na zbiory mierzalne, tzn. $A_i \in \Sigma$ są parami rozłączne oraz $X = \bigcup_{i=0}^n A_i$. W dalszej części sumy parami rozłącznych zbiorów będziemy oznaczać symbolem sumy rozłącznej $\bigsqcup$. Zatem $X = \bigsqcup_{i=0}^n A_i$.

Definicja 3.1 (Shannon)**.** *Entropią rozbicia* $\mathcal{A} = \{A_i\}_{i=1}^n$ nazywamy liczbę

$$\mathcal{H}(\mathcal{A}) = -\sum_{i=1}^{n} \mu(A_i) \ln \mu(A_i).$$

Definiując funkcję Hartleya $\mathcal{I}_\mathcal{A} : X \to [0, +\infty]$ wzorem $\mathcal{I}_\mathcal{A}(x) = \mathcal{I}(A_i)$, gdy $x \in A_i$ otrzymujemy, że jej wartość oczekiwna wynosi

$$E(\mathcal{I}_\mathcal{A}) = \sum_{i=1}^{n} \mu(A_i)\mathcal{I}(A_i) = \mathcal{H}(\mathcal{A}).$$

Zatem entropia rozbicia jest wartością oczekiwaną funkcji informacji. Ważną własnością $\mathcal{H}(\mathcal{A})$ jest to, że nie zależy ona od rozbicia $\mathcal{A}$, lecz od rozkładu prawdopodobieństwa.

Przykład 3.2. Niech $\mathcal{A} = \{A_1, ..., A_k\}$ oraz $\mu(A_i) = \frac{1}{k}$. Wtedy

$$\mathcal{H}(\mathcal{A}) = -\sum_{i=1}^{k} \frac{1}{k} \ln \frac{1}{k} = -k\frac{1}{k} \ln \frac{1}{k} = -\ln k^{-1} = \ln k.$$

Definicja 3.3. Niech $\mathcal{A} = \{A_i\}_{i=1}^n$ oraz $\mathcal{B} = \{B_j\}_{j=1}^m$ będą rozbiciami mierzalnymi przestrzeni X. *Połączeniem ("joinem")* $\mathcal{A}$ i $\mathcal{B}$ nazywamy rozbicie postaci

$$\mathcal{A} \vee \mathcal{B} = \{A_i \cap B_j : \quad i = 1, .., k, \quad j = 1, ..., m\}.$$

Jeśli $\mathcal{A} \vee \mathcal{B} = \mathcal{B}$ to piszemy $\mathcal{A} \preceq \mathcal{B}$ i mówimy, że rozbicie $\mathcal{B}$ jest *drobniejsze* niż $\mathcal{A}$. Zachodzi to, gdy dla każdego $A \in \mathcal{A}$ jesteśmy w stanie znaleźć takie zbiory z $\mathcal{B}$, aby $A = \bigsqcup_{i=1}^n B_i$.

Uwaga. Zawsze $\mathcal{A} \preceq \mathcal{A} \vee \mathcal{B}$.

Definicja 3.4. Entropię rozbicia $\mathcal{A}$ pod warunkiem rozbicia $\mathcal{B}$ (tzn. *entropię warunkową*) definiujemy jako średnią z entropii $\mathcal{A}$ względem prawdopodobieństw zadanych przez $\mathcal{B}$:

$$\mathcal{H}(\mathcal{A}|\mathcal{B}) := \sum_j \mu(B_j) \cdot \sum_i -\frac{\mu(A_i \cap B_j)}{\mu(B_j)} \ln \frac{\mu(A_i \cap B_j)}{\mu(B_j)} = -\sum_{i,j} \mu(A_i \cap B_j) \ln \frac{\mu(A_i \cap B_j)}{\mu(B_j)},$$

gdzie przyjmujemy konwencję, że "zero razy cokolwiek to zero", tzn. w sumach rozważamy takie j, że $\mu(A_i \cap B_j) \neq 0$ (wtedy tym bardziej $\mu(B_j) > 0$).

Uwaga. Korzystając z własności funkcji logarytmu powyższy wzór przyjmuje postać

$$\mathcal{H}(\mathcal{A}|\mathcal{B}) = \mathcal{H}(\mathcal{A} \vee \mathcal{B}) - \mathcal{H}(\mathcal{B}). \tag{3.1}$$

Stwierdzenie 3.5 (Własności entropii warunkowej)**.** *Niech $\mathcal{A}$, $\mathcal{B}$ oraz $\mathcal{C}$ będą mierzalnymi rozbiciami przestrzeni X. Wtedy*

(1) $\mathcal{H}(\mathcal{A} \vee \mathcal{B}|\mathcal{C}) = \mathcal{H}(\mathcal{A}|\mathcal{C}) + \mathcal{H}(\mathcal{B}|\mathcal{A} \vee \mathcal{C})$,

(2) $\mathcal{A} \preceq \mathcal{B} \Rightarrow \mathcal{H}(\mathcal{A}|\mathcal{C}) \leqslant \mathcal{H}(\mathcal{B}|\mathcal{C})$,

(3) $\mathcal{B} \preceq \mathcal{C} \Rightarrow \mathcal{H}(\mathcal{A}|\mathcal{B}) \geqslant \mathcal{H}(\mathcal{A}|\mathcal{C})$,

(4) $\mathcal{H}(\mathcal{A} \vee \mathcal{B}|\mathcal{C}) \leqslant \mathcal{H}(\mathcal{A}|\mathcal{C}) + \mathcal{H}(\mathcal{B}|\mathcal{C})$,

(5) $\mathcal{H}(\mathcal{A}|\mathcal{C}) \leqslant \mathcal{H}(\mathcal{A}|\mathcal{B}) + \mathcal{H}(\mathcal{B}|\mathcal{C})$.

Dowód. (1). Korzystając z (3.1) otrzymujemy, że

$$\mathcal{H}(\mathcal{A} \vee \mathcal{B}|\mathcal{C}) = \mathcal{H}(\mathcal{A} \vee \mathcal{B} \vee \mathcal{C}) - \mathcal{H}(\mathcal{C}),$$

$$\mathcal{H}(\mathcal{A}|\mathcal{C}) = \mathcal{H}(\mathcal{A} \vee \mathcal{C}) - \mathcal{H}(\mathcal{C}),$$

$$\mathcal{H}(\mathcal{B}|\mathcal{A} \vee \mathcal{C}) = \mathcal{H}(\mathcal{A} \vee \mathcal{B} \vee \mathcal{C}) - \mathcal{H}(\mathcal{A} \vee \mathcal{C}).$$

Sumując wszystko otrzymujemy tezę.

(2). Zauważmy, że jeśli $\mathcal{A} \preceq \mathcal{B}$ to $\mathcal{A} \vee \mathcal{B} = \mathcal{B}$. Więc

$$\mathcal{H}(\mathcal{B}|\mathcal{C}) = \mathcal{H}(\mathcal{A} \vee \mathcal{B}|\mathcal{C}) \overset{(1)}{=} \mathcal{H}(\mathcal{A}|\mathcal{C}) + \mathcal{H}(\mathcal{B}|\mathcal{A} \vee \mathcal{C}) \geqslant \mathcal{H}(\mathcal{A}|\mathcal{C}).$$

(3). Niech $\mathcal{B} \preceq \mathcal{C}$. Funkcja

$$k(t) = \begin{cases} t \ln t & \text{gdy } t > 0 \\ 0 & \text{gdy } t = 0 \end{cases}$$

jest wypukła. Zatem

$$\forall i, j \quad k\left(\sum_l \frac{\mu(C_l \cap B_j)}{\mu(B_j)} \cdot \frac{\mu(A_i \cap C_l)}{\mu(C_l)}\right) \leqslant \sum_l \frac{\mu(C_l \cap B_j)}{\mu(B_j)} \cdot k\left(\frac{\mu(A_i \cap C_l)}{\mu(C_j)}\right).$$

Ponieważ $\mathcal{B} \preceq \mathcal{C}$, to $\mu(C_l \cap B_j) = \mu(C_l)$. Stąd

$$k\left(\frac{\mu(A_i \cap B_j)}{\mu(B_j)}\right) \overset{\mathcal{A}\preceq\mathcal{C}}{=} k\left(\sum_{l,\mathcal{C}_l\subseteq\mathcal{B}_j} \frac{\mu(C_l \cap B_j)}{\mu(B_j)} \cdot \frac{\mu(A_i \cap C_l)}{\mu(C_l)}\right) \leqslant \sum_l \frac{\mu(C_l \cap B_j)}{\mu(B_j)} \cdot k\left(\frac{\mu(A_i \cap C_l)}{\mu(C_l)}\right).$$

Mnożąc przez $\mu(B_j)$ otrzymujemy, że

$$\mu(A_i \cap B_j) \cdot \ln \frac{\mu(A_i \cap B_j)}{\mu(B_j)} \leqslant \sum_{l,C_l\subseteq B_j} \mu(C_l \cap B_j) \cdot \frac{\mu(A_i \cap C_l)}{\mu(C_l)} \cdot \ln \frac{\mu(A_i \cap C_l)}{\mu(C_l)}.$$

Stąd

$$\mathcal{H}(\mathcal{A}|\mathcal{B}) = -\mu(A_i \cap B_j) \cdot \ln \frac{\mu(A_i \cap B_j)}{\mu(B_j)} \geqslant \sum_{i,j,l} \mu(C_l \cap B_j) \cdot \frac{\mu(A_i \cap C_l)}{\mu(C_l)} \cdot \ln \frac{\mu(A_i \cap C_l)}{\mu(C_l)} \overset{\mathcal{C}_l\subseteq\mathcal{B}_j}{=}$$

$$\overset{\mathcal{C}_l\subseteq\mathcal{B}_j}{=} -\sum_{i,l} \mu(A_i \cap C_l) \cdot \ln \frac{\mu(A_i \cap C_l)}{\mu(C_l)} = \mathcal{H}(\mathcal{A}|\mathcal{C}).$$

(4). Wynika z własności 1) oraz 3)

$$\mathcal{H}(\mathcal{A} \vee \mathcal{B}|\mathcal{C}) \overset{(1)}{=} \mathcal{H}(\mathcal{A}|\mathcal{C}) + \mathcal{H}(\mathcal{B}|\mathcal{A} \vee \mathcal{C}) \overset{(3)}{\geqslant} \mathcal{H}(\mathcal{A}|\mathcal{C}) + \mathcal{H}(\mathcal{B}|\mathcal{C}).$$

(5). Korzystamy tu z własności 1), 2) oraz 3)

$$\mathcal{H}(\mathcal{A}|\mathcal{C}) \overset{(2)}{\leqslant} \mathcal{H}(\mathcal{A} \vee \mathcal{B}|\mathcal{C}) \overset{(1)}{=} \mathcal{H}(\mathcal{B}|\mathcal{C}) + \mathcal{H}(\mathcal{A}|\mathcal{B} \vee \mathcal{C}) \overset{(3)}{\leqslant} \mathcal{H}(\mathcal{B}|\mathcal{C}) + \mathcal{H}(\mathcal{A}|\mathcal{B}).$$

□

3.3 Entropia odwzorowania

Niech tak jak poprzednio (X, Σ, μ) będzie ustaloną przestrzenią probabilistyczną. Niech $\varphi : X \to X$ będzie odwzorowaniem mierzalnym zachowującym miarę μ, natomiast $\mathcal{A} = \{A_i\}_{i=1}^k$ niech będzie rozbiciem mierzalnym X. Wtedy $\varphi^{-1}(\mathcal{A}) = \{\varphi^{-1}(A_i)\}_{i=1}^k$ również jest rozbiciem X. Dla $n \in \mathbb{N}$ rozważmy rozbicie

$$\bigvee_{i=0}^{n-1} \varphi^{-i}(\mathcal{A}) := \mathcal{A} \vee \varphi^{-1}(\mathcal{A}) \vee ... \vee \varphi^{-(n-1)}(\mathcal{A}). \tag{3.2}$$

Definicja 3.6. Entropią odwzorowania φ względem rozbicia $\mathcal{A}$ nazywamy liczbę

$$h(\varphi, \mathcal{A}) = \lim_{n\to\infty} \frac{1}{n} \mathcal{H}\left(\bigvee_{i=0}^{n-1} \varphi^{-i}(\mathcal{A})\right).$$

Istnienie granicy po prawej stronie powyższego wzoru, jak i kolejna równoważna definicja dla $h(\varphi, \mathcal{A})$ będą wykazane poniżej, w Stwierdzeniu 3.9.

W Definicji 3.6 wartość $\mathcal{H}(\bigvee_{i=0}^{n-1} \varphi^{-i}(\mathcal{A}))$ rozumiemy jako średnią ilość informacji zawartą w n eksperymentach, natomiast granicę przy $n \to \infty$ z tej wielkości podzielonej przez n jako średnią ilość informacji w jednym eksperymencie, gdyby przeprowadzać go w nieskończoność.

Lemat 3.7. *Niech $\bigvee_{i=0}^{n-1} \varphi^{-i}(\mathcal{A})$ będzie rozbiciem X danym wzorem* (3.2). *Wtedy*

$$\mathcal{H}\left(\bigvee_{i=0}^{n-1} \varphi^{-i}(\mathcal{A})\right) = \mathcal{H}(\mathcal{A}) + \sum_{j=1}^{n-1} \mathcal{H}\left(\mathcal{A}|\bigvee_{i=1}^{j} \varphi^{-i}(\mathcal{A})\right).$$

Dowód. Korzystamy z metody indukcji matematycznej po n.

$$\begin{aligned}
\mathcal{H}\left(\bigvee_{i=0}^{n} \varphi^{-i}(\mathcal{A})\right) &= \mathcal{H}\left(\mathcal{A} \vee \bigvee_{i=1}^{n} \varphi^{-i}(\mathcal{A})\right) = \mathcal{H}\left(\varphi^{-1}\left(\bigvee_{i=0}^{n-1} \varphi^{-i}(\mathcal{A})\right)\right) + \mathcal{H}\left(\mathcal{A}|\bigvee_{i=1}^{n} \varphi^{-i}(\mathcal{A})\right) \\
&= \mathcal{H}\left(\bigvee_{i=0}^{n-1} \varphi^{-i}(\mathcal{A})\right) + \mathcal{H}\left(\mathcal{A}|\bigvee_{i=i}^{n} \varphi^{-i}(\mathcal{A})\right) \\
&= \mathcal{H}(\mathcal{A}) + \sum_{j=1}^{n-1} \mathcal{H}\left(\mathcal{A}|\bigvee_{i=1}^{j} \varphi^{-i}(\mathcal{A})\right) + \mathcal{H}\left(\mathcal{A}|\bigvee_{i=1}^{n} \varphi^{-i}(\mathcal{A})\right) \\
&= \mathcal{H}(\mathcal{A}) + \sum_{j=1}^{n} \mathcal{H}\left(\mathcal{A}|\bigvee_{i=1}^{j} \varphi^{-i}(\mathcal{A})\right).
\end{aligned}$$

□

Lemat 3.8. *Niech $\{a_n\}_{n=1}^{\infty}$ będzie ciągiem liczbowym. Wtedy*

$$a_n \searrow a_0 \Rightarrow \frac{1}{n}\sum_{i=1}^{n} a_i \searrow a_0.$$

Dowód. Ciąg $\{\frac{1}{n}\sum_{i=1}^{n} a_i\}$ jest nierosnący, ponieważ

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n} a_i &= \sum_{i=1}^{n} \frac{a_i(n+1)}{(n+1)n} = \sum_{i=1}^{n} \frac{a_i n}{(n+1)n} + \sum_{i=1}^{n} \frac{a_i}{(n+1)n} \\
&\geqslant \sum_{i=1}^{n} \frac{a_i}{n+1} + \sum_{i=1}^{n} \frac{a_{n+1}}{(n+1)n} = \sum_{i=1}^{n} \frac{a_i}{n+1} + \frac{a_{n+1}}{n+1} = \frac{1}{n+1}\sum_{i=1}^{n+1} a_i.
\end{aligned}$$

Dalej mamy, że

$$a_{n+k} = \frac{1}{n+k}\sum_{i=1}^{n+k} a_{n+k} \leqslant \frac{1}{n+k}\sum_{i=1}^{n-1} a_i + \frac{k}{n+k}a_n.$$

Zatem z twierdzenia o trzech ciągach przy $k \to \infty$ ($a_{n+k} \searrow a_0$, $\frac{1}{n+k}\sum_{i=1}^{n-1} a_i \to 0$)

$$\frac{1}{n+k}\sum_{i=1}^{n+k} a_i \searrow a_0.$$

□

Z dwóch powyższych lematów wynika następujące stwierdzenie.

Stwierdzenie 3.9. *Ciągi* $\frac{1}{n}\mathcal{H}(\bigvee_{i=0}^{n-1}\varphi^{-i}(\mathcal{A}))$ *i* $\mathcal{H}(\mathcal{A}|\bigvee_{i=1}^{n-1}\varphi^{-i}(\mathcal{A}))$ *są malejące oraz*

$$h(\varphi,\mathcal{A})=\lim_{n\to\infty}\frac{1}{n}\mathcal{H}\left(\bigvee_{i=0}^{n-1}\varphi^{-i}(\mathcal{A})\right)=\lim_{n\to\infty}\mathcal{H}\left(\mathcal{A}|\bigvee_{i=1}^{n-1}\varphi^{-i}(\mathcal{A})\right)=\mathcal{H}\left(\mathcal{A}|\bigvee_{i=1}^{\infty}\varphi^{-i}(\mathcal{A})\right).$$

Następujące własności entropii odwzorowania φ względem rozbicia będą kluczowe w dowodach wielu twierdzeń.

Stwierdzenie 3.10. *(Własności entropii odwzorowania φ względem rozbicia).*

(1) $h(\varphi,\mathcal{A})\leqslant\mathcal{H}(\mathcal{A})$,

(2) $h(\varphi,\mathcal{A}\vee\mathcal{B})\leqslant h(\varphi,\mathcal{A})+h(\varphi,\mathcal{B})$,

(3) $\mathcal{A}\preceq\mathcal{B}\Rightarrow h(\varphi,\mathcal{A})\leqslant h(\varphi,\mathcal{B})$,

(4) $h(\varphi,\mathcal{A})\leqslant h(\varphi,\mathcal{B})+\mathcal{H}(\mathcal{A}|\mathcal{B})$,

(5) $k\geqslant 1\Rightarrow h(\varphi,\mathcal{A})=h(\varphi,\bigvee_{i=0}^{k-1}\varphi^{-i}(\mathcal{A}))$.

Dowód. (1). Wynika z tego, że

$$\mathcal{H}(\mathcal{A})\geqslant\frac{1}{2}\mathcal{H}(\mathcal{A}\vee\varphi^{-1}(\mathcal{A}))\geqslant\ldots\geqslant h(\varphi,\mathcal{A}).$$

(2). Korzystamy tu z subaddytywności entropii rozbicia

$$\mathcal{H}\left(\bigvee_{i=0}^{n-1}\varphi^{-i}(\mathcal{A}\vee\mathcal{B})\right)=\mathcal{H}\left(\bigvee_{i=0}^{n-1}\varphi^{-i}(\mathcal{A})\vee\bigvee_{i=0}^{n-1}\varphi^{-i}(\mathcal{B})\right)\overset{subaddyt.}{\leqslant}\mathcal{H}\left(\bigvee_{i=0}^{n-1}\varphi^{-i}(\mathcal{A})\right)+\mathcal{H}\left(\bigvee_{i=0}^{n-1}\varphi^{-i}(\mathcal{B})\right).$$

Tezę otrzymujemy dzieląc przez n oraz przechodząc z $n\to\infty$.

(3). Jeśli $\mathcal{A}\preceq\mathcal{B}$ to

$$\bigvee_{i=0}^{n-1}\varphi^{-i}(\mathcal{A})\preceq\bigvee_{i=0}^{n-1}\varphi^{-i}(\mathcal{B}).$$

Stąd mamy, że

$$\mathcal{H}\left(\bigvee_{i=0}^{n-1}\varphi^{-i}(\mathcal{A})\right)\leqslant\mathcal{H}\left(\bigvee_{i=0}^{n-1}\varphi^{-i}(\mathcal{B})\right)\Rightarrow h(\varphi,\mathcal{A})\leqslant h(\varphi,\mathcal{B}).$$

(4). Z własności entropii warunkowej

$$\begin{aligned}\mathcal{H}\left(\bigvee_{i=0}^{n-1}\varphi^{-i}(\mathcal{A})\right)&\leqslant\mathcal{H}\left(\bigvee_{i=0}^{n-1}\varphi^{-i}(\mathcal{A})\vee\bigvee_{i=0}^{n-1}\varphi^{-i}(\mathcal{B})\right)\\&=\mathcal{H}\left(\bigvee_{i=0}^{n-1}\varphi^{-i}(\mathcal{B})\right)+\mathcal{H}\left(\bigvee_{i=0}^{n-1}\varphi^{-i}(\mathcal{A})|\bigvee_{i=0}^{n-1}\varphi^{-i}(\mathcal{B})\right).\end{aligned}$$

Teraz korzystamy z subaddytywności i monotoniczności entropii rozbicia

$$\mathcal{H}\left(\bigvee_{i=0}^{n-1}\varphi^{-i}(\mathcal{A})|\bigvee_{i=0}^{n-1}\varphi^{-i}(\mathcal{B})\right) \overset{subaddyt.}{\leqslant} \sum_{i=0}^{n-1}\mathcal{H}\left(\varphi^{-i}(\mathcal{A})|\bigvee_{j=0}^{n-1}\varphi^{-j}(\mathcal{B})\right) \overset{monot.}{\leqslant}$$

$$\overset{monot.}{\leqslant} \sum_{i=0}^{n-1}\mathcal{H}\left(\varphi^{-i}(\mathcal{A})|\varphi^{-i}(\mathcal{B})\right) \overset{\varphi-niezmien.}{=} n\mathcal{H}(\mathcal{A}|\mathcal{B}).$$

Stąd

$$\frac{1}{n}\mathcal{H}\left(\bigvee_{i=0}^{n-1}\varphi^{-i}(\mathcal{A})\right) \leqslant \frac{1}{n}\mathcal{H}\left(\bigvee_{i=0}^{n-1}\varphi^{-i}(\mathcal{B})\right) + \mathcal{H}(\mathcal{A}|\mathcal{B}).$$

Przechodząc z $n \to \infty$ otrzymujemy tezę.

(5). Niech $k \geqslant 1$. Wtedy

$$\begin{aligned} h\left(\varphi, \bigvee_{i=0}^{k}\varphi^{-i}(\mathcal{A})\right) &= \lim_{n\to\infty}\frac{1}{n}\mathcal{H}\left(\bigvee_{j=0}^{n-1}\varphi^{-j}\left(\bigvee_{i=0}^{k}\varphi^{-i}(\mathcal{A})\right)\right) \\ &= \lim_{n\to\infty}\frac{1}{n}\mathcal{H}\left(\bigvee_{i=0}^{n+k-1}\varphi^{-i}(\mathcal{A})\right) = \lim_{n\to\infty}\frac{n+k}{n}\frac{1}{n+k}\mathcal{H}\left(\bigvee_{i=0}^{n+k-1}\varphi^{-i}(\mathcal{A})\right) \\ &= h(\varphi, \mathcal{A}). \end{aligned}$$

□

Definicja 3.11. Entropią odwzorowania $\varphi : X \to X$ względem miary μ nazywamy wartość

$$h(\varphi) = \sup_{\mathcal{A}} h(\varphi, \mathcal{A}),$$

gdzie supremum bierzemy po wszystkich skończonych rozbiciach mierzalnych $\mathcal{A}$.

Z własności (5) Twierdzenia 3.10 wynika następujący lemat.

Lemat 3.12. *Niech k będzie liczbą naturalną. Wtedy*

$$h(\varphi^k) = kh(\varphi).$$

3.4 Twierdzenie Kołmogorowa-Sinai'a (o rozbiciu generującym)

W tym podrozdziale zostanie przedstawione jedno z najważniejszych twierdzeń pozwalające obliczyć entropię posługując się tylko jednym rozbiciem.

Niech (X, Σ, μ) będzie przestrzenią probabilistyczną oraz niech $\mathcal{R}, \mathcal{S} \subseteq \Sigma$. Piszemy $\mathcal{R} \overset{0}{\subseteq} \mathcal{S}$, jeśli dla każdego $A \in \mathcal{R}$ istnieje taki $B \in \mathcal{S}$, że $\mu(A \triangle B) = 0$. Jeśli $\mathcal{R} \overset{0}{\subseteq} \mathcal{S}$ i $\mathcal{R} \overset{0}{\supseteq} \mathcal{S}$ to piszemy $\mathcal{R} \overset{0}{=} S$. Jako $\mathbb{P}$ oznaczmy rodzinę wszystkich skończonych rozbić mierzalnych przestrzeni X.

Lemat 3.13. *Niech $\mathcal{A}, \mathcal{B} \in \mathbb{P}$. Wtedy*

$$\mathcal{H}(\mathcal{A}|\mathcal{B}) = 0 \Longleftrightarrow \mathcal{A} \overset{0}{\subseteq} \mathcal{B}.$$

Dowód powyższego lematu znajduje się w [30, Twierdzenie 4.4].

Definicja 3.14. *Metrykę Rohlina* dla rozbić $\mathcal{A}, \mathcal{B} \in \mathbb{P}$ definiujemy wzorem

$$d(\mathcal{A}, \mathcal{B}) = \mathcal{H}(\mathcal{A}|\mathcal{B}) + \mathcal{H}(\mathcal{B}|\mathcal{A}).$$

Lemat 3.15. *Dla każdego $k \in \mathbb{N}$ i $\varepsilon > 0$ jesteśmy w stanie znaleźć takie $\delta > 0$, aby*

$$\forall_{\mathcal{A}=\{A_1,\ldots,A_k\}\in\mathbb{P}} \quad \forall_{\mathcal{B}=\{B_1,\ldots,B_k\}\in\mathbb{P}} \quad \sum_{i=1}^{k} \mu(A_i \triangle B_i) < \delta \Longrightarrow d(\mathcal{A}, \mathcal{B}) < \varepsilon.$$

Dowód. Niech $k \in \mathbb{N}$ i $\varepsilon > 0$. Należy dobrać takie δ, aby

$$\delta < \frac{1}{e}\left(1 - \frac{1}{e}\right),$$

oraz

$$-k(k-1)\delta \ln \delta - (1-\delta)\ln(1-\delta) < \frac{\varepsilon}{2}.$$

Wtedy przy $\delta \to 0$ mamy, że $\delta \ln \delta \to 0$ oraz $(1-\delta)\ln(1-\delta) \to 0$. Niech $\mathcal{C}$ będzie podziałem X na $A_i \cap B_j$ dla $i \neq j$ oraz $\bigcup_{i=1}^{k} A_i \cap B_i$. Wtedy

$$\mathcal{A} \vee \mathcal{B} = \mathcal{B} \vee \mathcal{C} = \mathcal{A} \vee \mathcal{C}.$$

Ponadto dla każdego $i \neq j$

$$A_i \cap B_j \subseteq \bigcup_{l=1}^{k} A_l \triangle B_l,$$

stąd

$$A_i \cap B_j \subseteq \bigcup_{l=1}^{k} A_l \triangle B_l \Rightarrow \mu(\bigcup_{i\neq j} A_i \cap B_j) \leqslant \sum_{l=1}^{n} \mu(A_l \triangle B_l) < \delta.$$

Otrzymujemy więc, że dla każdego $i \neq j$

$$\mu(A_i \cap B_j) < \delta.$$

Z tego, że $\mu(\bigcup_{i=1}^{k} A_i \cap B_i) = \mu(X \setminus (\bigcup_{i\neq j} A_i \cap B_j))$ mamy, że

$$\mu\left(\bigcup_{i=1}^{k} A_i \cap B_i\right) = 1 - \mu\left(\bigcup_{i\neq j} A_i \cap B_j\right) > 1 - \delta.$$

Dalej

$$\mathcal{H}(\mathcal{C}) = -\sum_{i\neq j} \mu(A_i \cap B_j) \ln \mu(A_i \cap B_j) - \mu\left(\bigcup_{i=1}^{k} A_i \cap B_i\right) \ln\left(\mu\left(\bigcup_{i=1}^{k} A_i \cap B_i\right)\right)$$
$$\leqslant -\sum_{i\neq j} \delta \ln \delta - (1-\delta)\ln(1-\delta) = -k(k-1)\delta \ln\delta - (1-\delta)\ln(1-\delta) < \frac{\varepsilon}{2}.$$

Stąd

$$\mathcal{H}(\mathcal{A}|\mathcal{B}) = \mathcal{H}(\mathcal{A}\vee\mathcal{B}) - \mathcal{H}(\mathcal{B}) \overset{\mathcal{A}\vee\mathcal{B}=\mathcal{B}\vee\mathcal{C}}{=} \mathcal{H}(\mathcal{B}\vee\mathcal{C}) - \mathcal{H}(\mathcal{B}) = \mathcal{H}(\mathcal{C}|\mathcal{B}) \leqslant \mathcal{H}(\mathcal{C}) < \frac{\varepsilon}{2}.$$

Analogicznie otrzymujemy, że $\mathcal{H}(\mathcal{B}|\mathcal{A}) \leqslant \mathcal{H}(\mathcal{C}) < \frac{\varepsilon}{2}$. Zatem

$$d(\mathcal{A},\mathcal{B}) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

□

Twierdzenie 3.16. *Niech Σ_0 będzie algebrą zbiorów taką, że jej elementy generują Σ z dokładnością do zbiorów miary zero. Wtedy*

$$\forall_{\substack{\mathcal{A}\subseteq\Sigma\\ \mathcal{A}\in\mathbb{P}}} \quad \forall_{\varepsilon>0} \quad \exists_{\substack{\mathcal{B}\subseteq\Sigma_0\\ \mathcal{B}\in\mathbb{P}}} \quad d(\mathcal{A},\mathcal{B}) < \varepsilon.$$

Dowód. Niech $\mathcal{A} = \{A_1, ..., A_k\}$ oraz niech $\varepsilon > 0$. Z lematu (3.15) istnieje takie $\delta > 0$, że dla każdego rozbicia $\mathcal{B} = \{B_1, ..., B_k\} \subseteq \Sigma$

$$\sum_{i=1}^{k} \mu(A_i \triangle B_i) < \delta \Rightarrow d(\mathcal{A},\mathcal{B}) < \varepsilon.$$

Zatem wystarczy znaleźć taki podział $\mathcal{B} = \{B_1, ..., B_k\} \subseteq \Sigma_0$, aby

$$\forall_{i=1,...,k} \quad \mu(A_i \triangle B_i) < \delta.$$

Weźmy takie λ, aby

$$\lambda(k-1)[1 + 2k(k-1)] < \delta.$$

Z [9, Twierdzenie 2.13] wiemy, że

$$\forall_{i=1,...,k}\ \exists_{C_i\in\Sigma_0} \quad \mu(A_i \triangle C_i) < \lambda.$$

Dla $i \neq j$ mamy, że $C_i \cap C_j \subseteq A_i \triangle C_i \cup A_j \triangle C_j$, więc $\mu(C_i \cap C_j) < 2\lambda$. Niech

$$\Sigma_0 \ni N = \bigcup_{i\neq j} C_i \cap C_j.$$

Dla $i \neq j$ mamy $k(k-1)$ kombinacji. Stąd też $\mu(N) < k(k-1)2\lambda$. Połóżmy

$$\forall_{i=1,...,k-1} \quad B_i := C_i \backslash N \quad \text{oraz} \quad B_k := X \backslash \bigcup_{i=1}^{k-1} B_i.$$

Wtedy $\mathcal{B} = \{B_1, ..., B_k\} \subseteq \Sigma_0$ jest podziałem. Ponadto

$$\forall_{i=1,\dots,k-1} \quad A_i \triangle B_i \subseteq A_i \triangle C_i \cup N,$$

a z tego

$$\forall_{i\leqslant k-1} \quad \mu(A_i \triangle B_i) \leqslant \mu(A_i \triangle C_i) + \mu(N) < \lambda + k(k-1)2\lambda = \lambda[1 + 2k(k-1)] < \delta.$$

Dla $A_k \triangle B_k$ mamy, że

$$A_k \triangle B_k \subseteq \bigcup_{i=1}^{k-1} A_i \triangle B_i \Rightarrow \mu(A_k \triangle B_k) < \sum_{i=1}^{k-1} \mu(A_i \triangle B_i) < \lambda(k-1)[1 + 2k(k-1)] < \delta.$$

□

Z powyższego twierdzenia wynika następujący wniosek.

Wniosek 3.17. *Niech $\mathcal{A}_1 \preceq \mathcal{A}_2 \preceq \mathcal{A}_3 \preceq ...$ i $\mathcal{B}$ będą skończonymi rozbiciami w Σ. Wtedy*

$$\mathcal{B} \overset{0}{\subseteq} \sigma\left(\bigcup_{n=1}^{\infty} \mathcal{A}_n\right) \Rightarrow \mathcal{H}(\mathcal{B}|\mathcal{A}_n) \overset{n\to\infty}{\to} 0.$$

Kołmogorow i Sinai wydzielili klasy rozbić (tak zwanych podziałów generujących), na których osiągane jest supremum w Definicji 3.11. To pozwala natychmiastowo obliczyć entropię na podstawie Definicji 3.6, a nie szukając supremum po rozbiciach w Definicji 3.11.

Definicja 3.18. Skończone rozbicie $\mathcal{A} \subseteq \Sigma$ nazywamy *rozbiciem (podziałem) generującym*, jeżeli

$$\bigvee_{n=0}^{\infty} \varphi^{-n}(\mathcal{A}) = \sigma\left(\bigcup_{n=0}^{\infty} \varphi^{-n}(\mathcal{A})\right) \overset{0}{=} \Sigma,$$

gdzie $\sigma\left(\bigcup_{n=0}^{\infty} \varphi^{-n}(\mathcal{A})\right)$ jest najmniejszą σ-algebrą generowaną przez $\bigcup_{n=0}^{\infty} \varphi^{-n}(\mathcal{A})$.

Twierdzenie 3.19 (Twierdzenie Kołmogorowa Sinai'a (o rozbiciu generującym))**.** *Niech $\varphi : X \to X$ będzie odwzorowaniem zachowującym miarę. Jeśli $\mathcal{A} \subseteq \Sigma$ jest rozbiciem generującym, to*

$$h(\varphi) = \sup_{\mathcal{B}\subseteq\Sigma} h(\varphi, \mathcal{B}) = h(\varphi, \mathcal{A}).$$

Dowód. Niech $\mathcal{B} \subseteq \Sigma$ będzie dowolnym rozbiciem. Należy wykazać, że $h(\varphi, \mathcal{B}) \leqslant h(\varphi, \mathcal{A})$. Zatem, z własności entropii odwzorowania

$$h(\varphi, \mathcal{B}) \leqslant h(\varphi, \mathcal{C}) + \mathcal{H}(\mathcal{B}|\mathcal{C}).$$

Weźmy $\mathcal{C} = \bigvee_{i=0}^{n} \varphi^{-i}(\mathcal{A})$. Wtedy

$$h(\varphi, \mathcal{B}) \leqslant h\left(\varphi, \bigvee_{i=0}^{n} \varphi^{-i}(\mathcal{A})\right) + \mathcal{H}\left(\mathcal{B}|\bigvee_{i=0}^{n} \varphi^{-i}(\mathcal{A})\right) = h(\varphi, \mathcal{A}) + \mathcal{H}\left(\mathcal{B}|\bigvee_{i=0}^{n} \varphi^{-i}(\mathcal{A})\right).$$

Z wniosku 3.17

$$\mathcal{H}\left(\mathcal{B}|\bigvee_{i=0}^{n}\varphi^{-i}(\mathcal{A})\right)\overset{n\to\infty}{\longrightarrow}0$$

□

Jeśli $\varphi : X \to X$ jest odwzorowaniem odwracalnym to Twierdzenie Kołmogorowa Sinai'a przyjmuje następującą postać.

Twierdzenie 3.20 (Twierdzenie Kołmogorowa-Sinai'a dla odwzorowań odwracalnych)**.** *Jeżeli $\varphi : X \to X$ odwracalne oraz $\mathcal{A}$-rozbicie skończone, takie, że $\bigvee_{n=-\infty}^{\infty}\varphi^{-n}(\mathcal{A}) \overset{0}{=} \Sigma$, to $h(\varphi) = h(\varphi, \mathcal{A})$.*

Z powyższego twierdzenia wynika następujący wniosek.

Wniosek 3.21. *Jeśli $\varphi : X \to X$ jest odwzorowaniem odwracalnym oraz $\bigvee_{n=0}^{\infty}\varphi^{-n}(\mathcal{A}) \overset{0}{=} \Sigma$, to $h(\varphi) = 0$.*

Dowód. Z Twierdzenia Kołmogorowa-Sinai'a, własności entropii warunkowej dla rozbicia, oraz Wniosku 3.17

$$h(\varphi) = h(\varphi, \mathcal{A}) = \mathcal{H}\left(\mathcal{A}|\bigvee_{i=1}^{\infty}\varphi^{-i}(\mathcal{A})\right) \to 0.$$

□

3.5 Przykłady obliczania entropii

Przykład 3.22 (Odwzorowanie trójkątne)**.** Niech $X = [0, 1]$ oraz niech

$$\varphi(x) = 2x[mod\,1] = \begin{cases} 2x & \text{gdy } x \in [0, \frac{1}{2}] \\ 2x-1 & \text{gdy } x \in [\frac{1}{2}, 1] \end{cases}.$$

Niech $\xi = \{\xi_1, \xi_2\} = \{[0, \frac{1}{2}], (\frac{1}{2}, 1]\}$ podział X. Zauważmy, że

$$\varphi^{-1}(\xi_1) = \left[0, \frac{1}{4}\right] \cup \left[\frac{1}{2}, \frac{3}{4}\right], \quad \varphi^{-1}(\xi_2) = \left(\frac{1}{4}, \frac{1}{2}\right] \cup \left(\frac{3}{4}, 1\right]. \tag{3.3}$$

Łatwo to można zauważyć zapisując podziały ξ_1 i ξ_2 za pomocą systemu dwójkowego: $\xi_1 = \{x = (0, \eta_2, \eta_3, ...)\}$ i $\xi_2 = \{x = (1, \eta_2, \eta_3, ...)\}$, gdzie $\eta_i \in \{0, 1\}$. Wtedy $\varphi^{-1}(\xi_1) = \{x = (\eta_1, 0, \eta_3, ...)\}$ oraz $\varphi^{-1}(\xi_1) = \{x = (\eta_1, 1, \eta_3, ...)\}$, stąd mamy (3.3). Ponadto

$$\begin{aligned}\xi \vee \varphi^{-1}(\xi) &= \{\xi_1 \cap \varphi^{-1}(\xi_1), \xi_1 \cap \varphi^{-1}(\xi_2), \xi_2 \cap \varphi^{-1}(\xi_1), \xi_2 \cap \varphi^{-1}(\xi_2)\} \\ &= \left\{\left[0, \frac{1}{4}\right], \left(\frac{1}{4}, \frac{1}{2}\right], \left[\frac{1}{2}, \frac{3}{4}\right], \left(\frac{3}{4}, 1\right]\right\}.\end{aligned}$$

Postępując indukcyjnie w ten sposób otrzymujemy, że

$$\varphi^{-i}(\xi_1) = \bigcup_{k=0}^{2^i-1} \left[\frac{k}{2^i}, \frac{2k+1}{2^{i+1}}\right], \quad \varphi^{-i}(\xi_2) = \bigcup_{k=0}^{2^i-1} \left(\frac{2k+1}{2^{i+1}}, \frac{k+1}{2^i}\right].$$

Zatem

$$\mathcal{H}\left(\bigvee_{i=0}^{n-1} \varphi^{-i}(\xi)\right) = -\sum_{\mathcal{A}\in\bigvee_{i=0}^{n-1}\varphi^{-1}(\xi)} \mu(\mathcal{A}) \ln \mu(\mathcal{A}) = -\frac{2^n}{2^n} \ln \frac{1}{2^n} = n \ln 2.$$

Stąd

$$h_\mu(\varphi, \xi) = \lim_{n\to\infty} \frac{1}{n} n \ln 2 = \ln 2.$$

Z ogólnej postaci $\bigvee_{i=0}^{n-1} \varphi^{-i}(\xi)$ widać, że ξ jest podziałem generującym. Teraz z Twierdzenia Kołmogorowa-Sinai'a, otrzymujemy, że

$$h_\mu(\varphi) = \ln 2.$$

Przykład 3.23 (Odwzorowanie okresowe)**.** Entropia odwzorowania identycznościowego $id : (X, \Sigma, \mu) \to (X, \Sigma, \mu)$ jest równa zero (tzn. $h(id) = 0$). Ponadto, jeśli $\varphi^n = 1$ dla pewnego n to $h(\varphi) = 0$, ponieważ z Lematu 3.12 mamy, że

$$h = (\varphi^n) = nh(\varphi) = 0.$$

W szczególności, każde odwzorowanie na przestrzeni skończonej ma zerową entropię.

Przykład 3.24 (Obrót o kąt wymierny)**.** Niech $X = S^1 = \{z \in \mathbb{C} : |z| = 1\}$, $\varphi(z) = e^{i\alpha}z$ (gdzie $\frac{\alpha}{2\pi} \in \mathbb{Q}$, czyli $\alpha = \frac{m}{n}, m, n \in \mathbb{Z}$). Stąd $\varphi^n(z) = z$. Zatem $h(\varphi) = 0$ z Przykładu 3.23.

Przykład 3.25 (Obrót o kąt niewymierny)**.** Niech założenia będą takie jak w Przykładzie 3.24 i niech teraz $\alpha \notin \mathbb{Q}$. To znaczy, że zbiór $\{a^n : n \in \mathbb{Z}\}$, gdzie a jest pierwiastkiem z jedynki jest gęsty w S^1 (zatem $\{a^n : n < 0\}$ także jest gęsty w S^1). Weźmy podział S^1 jako $\xi = \{\xi_1, \xi_2\}$, gdzie ξ_1 jest górną połową okręgu, a ξ_2 dolną. Dla $n > 0$ $\varphi^{-n}(\xi)$ zawiera półokręgi zaczynające się od a^{-n} i $-a^{-n}$. Skoro $\{a^{-n} : n > 0\}$ jest gęsty to dowolny półokrąg należy do $\bigvee_{n=0}^{\infty} \varphi^{-n}(\xi)$. Stąd $\bigvee_{n=0}^{\infty} \varphi^{-n}(\xi) \overset{0}{=} \Sigma$. Zatem $h(\varphi) = 0$ z Wniosku 3.21.

3.6 Entropia topologiczna

W poprzednich podrozdziałach opisana została entropia metryczna (Kołmogorowa-Sinai'a), która jest związana z miarami niezmienniczymi i pochodzi bezpośrednio z teorii informacji. W tym i w następnych podrozdziałach omówimy entropię topologiczną, która odpowiada za "mierzenie złożoności" topologicznego układu dynamicznego. W Podrozdziale 3.9 przedstawimy Zasadę Wariacyjną wiążąca entropię topologiczną z entropią Kołmogorowa-Sinai'a.

Niech X będzie zwartą przestrzenią metryczną oraz niech $\mathcal{U} = \{A_i\}_{i\in I}$, $\mathcal{V} = \{B_j\}_{j\in J}$ będą otwartymi pokryciami tej przestrzeni, tzn. zbiory A_i, B_j są otwarte oraz $X = \bigcup_{i\in I} A_i = \bigcup_{j\in J} B_j$.

Definicja 3.26. *"Join" pokryć* definiujemy jako

$$\mathcal{U} \vee \mathcal{V} = \{A_i \cap B_j : \ i \in I, j \in J\}.$$

Jeśli każdy element pokrycia $\mathcal{V}$ jest zawarty w pewnym elemencie pokrycia $\mathcal{U}$ to mówimy, że $\mathcal{V}$ jest *drobniejsze* od pokrycia $\mathcal{U}$ i oznaczamy $\mathcal{U} \preceq \mathcal{V}$. Zatem, jeżeli $\mathcal{V}$ jest podpokryciem pokrycia $\mathcal{U}$ to $\mathcal{U} \preceq \mathcal{V}$ oraz zawsze $\mathcal{U}, \mathcal{V} \preceq \mathcal{U} \vee \mathcal{V}$. Przyjmujemy oznaczenie

$$\mathcal{N}(\mathcal{U}) \text{ - liczba elementów podpokrycia } \mathcal{U} \text{ o najmniejszej mocy.}$$

Definicja 3.27. *Entropię pokrycia* $\mathcal{U}$ definiujemy jako

$$\mathcal{H}(\mathcal{U}) = \ln \mathcal{N}(\mathcal{U}).$$

Entropia pokrycia przyjmuje wartości ze zbioru $[0, +\infty)$ oraz, jeśli jest równa zeru to znaczy, że jednym z elementów pokrycia jest cała przestrzeń X.

Twierdzenie 3.28. *Niech* $\mathcal{U}$ *i* $\mathcal{V}$ *będą otwartymi pokryciami* X. *Wtedy*

(1) $\mathcal{U} \preceq \mathcal{V} \Rightarrow \mathcal{H}(\mathcal{U}) \leqslant \mathcal{H}(\mathcal{V})$,

(2) $\mathcal{H}(\mathcal{U} \vee \mathcal{V}) \leqslant \mathcal{H}(\mathcal{U}) + \mathcal{H}(\mathcal{V})$.

Dowód. (1). Niech $B_1, ..., B_{\mathcal{N}(\mathcal{V})} \in \mathcal{V}$ pokrycie X. Skoro $\mathcal{V}$ jest drobniejsze od pokrycia $\mathcal{U}$ to istnieją $A_1, ..., A_{\mathcal{N}(\mathcal{V})} \in \mathcal{U}$ takie, że dla każdego $i = 1, ..., \mathcal{N}(\mathcal{V})$, $B_i \subseteq A_i$. Zatem $\{A_i\}_{i=1}^{\mathcal{N}(\mathcal{V})} \subseteq \mathcal{U}$ także jest pokryciem X i stąd mamy tezę.

(2). Niech $A_1, ..., A_{\mathcal{N}(\mathcal{U})} \in \mathcal{U}$ oraz $B_1, ..., B_{\mathcal{N}(\mathcal{V})} \in \mathcal{V}$ będą pokryciami X. Wtedy

$$\{A_i \cap B_j : \ j = 1, ..., \mathcal{N}(\mathcal{V}), i = 1, ..., \mathcal{N}(\mathcal{U})\} \subseteq \mathcal{U} \vee \mathcal{V},$$

jest pokryciem X. Zatem $\mathcal{N}(\mathcal{U} \vee \mathcal{V}) \leqslant \mathcal{N}(\mathcal{U})\mathcal{N}(\mathcal{V})$, a więc mamy, że

$$\mathcal{H}(\mathcal{U} \vee \mathcal{V}) = \ln \mathcal{N}(\mathcal{U} \vee \mathcal{V}) \leqslant \ln(\mathcal{N}(\mathcal{U})\mathcal{N}(\mathcal{V})) = \ln \mathcal{N}(\mathcal{U}) + \ln \mathcal{N}(\mathcal{V}) = \mathcal{H}(\mathcal{U}) + \mathcal{H}(\mathcal{V}).$$

□

Niech $\varphi : X \to X$ będzie odwzorowaniem ciągłym oraz niech $\mathcal{U}$ będzie otwartym pokryciem przestrzeni X. Wtedy $\varphi^{-1}(\mathcal{U}) = \{\varphi^{-1}(A) : A \in \mathcal{U}\}$ także jest pokryciem otwartym X, ponadto $\mathcal{H}(\varphi^{-1}(\mathcal{U})) \leqslant \mathcal{H}(\mathcal{U})$ oraz $\mathcal{H}(\varphi^{-1}(\mathcal{U})) = \mathcal{H}(\mathcal{U})$, gdy φ jest surjekcją. $\bigvee_{i=0}^{n-1} \varphi^{-i}(\mathcal{U})$ definiujemy w analogiczny sposób jak w przypadku rozbić (tzn. $\bigvee_{i=0}^{n-1} \varphi^{-i}(\mathcal{U}) := \mathcal{U} \vee \varphi^{-1}(\mathcal{U}) \vee \varphi^{-2}(\mathcal{U}) \vee ... \vee \varphi^{-(n-1)}(\mathcal{U}) = \{A_{i_1} \cap \varphi^{-1}(A_{i_2}) \cap \varphi^{-2}(A_{i_3}) \cap ... \cap \varphi^{-(n-1)}(A_{i_n}) : A_{i_1}, ..., A_{i_n} \in \mathcal{U}\}$).

Lemat 3.29 (Teoplitza-Feketa)**.** *Niech* $\{a_n\}_{n=1}^{\infty} \subseteq \mathbb{R}$ *będzie ciągiem subaddytywnym, tzn.* $\forall_{n,m}\ a_{n+m} \leqslant a_n + a_m$. *Wtedy* $\lim_{n\to\infty} \frac{a_n}{n}$ *istnieje i wynosi* $\inf \frac{a_n}{n}$.

Dowód znajduje się w [30, Twierdzenie 4.9]. Powyższy lemat będzie pomocny w dowodzie następującego twierdzenia.

Twierdzenie 3.30. *Dla każdego pokrycia otwartego $\mathcal{U}$ oraz odwzorowania ciągłego $\varphi : X \to X$ istnieje granica*

$$\lim_{n\to\infty} \frac{1}{n}\mathcal{H}\left(\bigvee_{i=0}^{n-1} \varphi^{-i}(\mathcal{U})\right) \tag{3.4}$$

Dowód. Oznaczmy $a_n := \mathcal{H}(\bigvee_{i=0}^{n-1} \varphi^{-i}(\mathcal{U}))$. Wtedy

$$\begin{aligned} a_{n+m} = \mathcal{H}\left(\bigvee_{i=0}^{n+m-1} \varphi^{-i}(\mathcal{U})\right) &\leqslant \mathcal{H}\left(\bigvee_{i=0}^{n-1} \varphi^{-i}(\mathcal{U})\right) + \mathcal{H}\left(\bigvee_{i=n}^{n+m-1} \varphi^{-i}(\mathcal{U})\right) \\ = a_n + \mathcal{H}\left(\varphi^{-n}\left(\bigvee_{i=0}^{m-1} \varphi^{-i}(\mathcal{U})\right)\right) &\leqslant a_n + a_m. \end{aligned}$$

□

Granicę (3.4) w powyższym twierdzeniu nazywamy entropią odwzorowania $\varphi : X \to X$ dla pewnego pokrycia $\mathcal{U}$ i oznaczamy $h(\varphi,\mathcal{U})$.

Definicja 3.31. *Entropię topologiczną* odwzorowania $\varphi : X \to X$ definiujemy jako

$$h_{top}(\varphi) = \sup_{\mathcal{U}} h(\varphi,\mathcal{U}),$$

gdzie supremum brane jest po otwartych pokryciach $\mathcal{U}$ przestrzeni X.

3.7 Entropia topologiczna Bowena

Niech X będzie zwartą przestrzenią metryczną z metryką d i niech $A \subseteq X$. *Średnicę zbioru* A definiujemy jako $\sup_{x,y\in A} d(x,y)$ i oznaczamy $\mathrm{diam}(A)$. *Średnicą pokrycia* $\mathcal{U} = \{A_i\}_{i\in I}$ zbioru X nazywamy wartość $\sup_{i\in I} \mathrm{diam}(A_i)$ i oznaczamy $\mathrm{diam}(\mathcal{U})$.

Twierdzenie 3.32 (Lebesgue'a o pokryciu)**.** *Przy wcześniejszych założeniach, istnieje $\delta > 0$ taka, że każdy podzbiór X o średnicy mniejszej niż δ jest zawarty w pewnym elemencie pokrycia $\mathcal{U}$.*

Dowód. Patrz [30, Twierdzenie 0.20] □

Liczbę δ w powyższym twierdzeniu nazywamy *liczbą Lebesgue'a* pokycia $\mathcal{U}$. Jeśli $\mathcal{U}$ oraz $\mathcal{V}$ są pokryciami X i $\mathrm{diam}(\mathcal{U})$ jest mniejsza niż liczba Lebesgue'a pokrycia $\mathcal{V}$ to $\mathcal{V} \preceq \mathcal{U}$.

Twierdzenie 3.33. *Jeśli $\mathcal{U}_n$ jest ciągiem pokryć zbiorami otwartymi o średnicy* $\mathrm{diam}(\mathcal{U}_n) \to 0$*, to istnieje granica (właściwa lub nie)*

$$h_{top}(\varphi) = \lim_{n\to\infty} h(\varphi,\mathcal{U}_n).$$

Dowód. Patrz [30, Twierdzenie 7.6]. □

Definicja 3.34. *N-tą metrykę Bowena* definiujemy za pomocą wzoru

$$d_n(x,y) = \max_{0\leqslant k<n} d(\varphi^k(x), \varphi^k(y)).$$

Definicja 3.35. Niech $\varepsilon > 0$. Zbiór $E \subseteq X$ nazywamy *(n,ε)-rozdzielonym* jeżeli dla każdych różnych $x, y \in E$, ich odległość w metryce Bowena $d_n(x,y) > \varepsilon$, tzn.

$$\forall_{x,y\in E} \quad \exists_{0\leqslant k<n} \quad d(\varphi^k(x), \varphi^k(y)) > \varepsilon.$$

Definicja 3.36. Niech $\varepsilon > 0$. Powiemy, że $E \subseteq X$ jest *(n,ε)-rozpinający*, jeżeli kule postaci $B_n(x,\varepsilon) = \{y \in X : d_n(x,y) < \varepsilon\}$ pokrywają całą przestrzeń X, tzn.

$$\forall_{y\in X} \quad \exists_{x\in E} \quad \forall_{0\leqslant k<n} \quad d(\varphi^k(x), \varphi^k(y)) < \varepsilon.$$

Oznaczmy jako $\operatorname{sep}(n,\varepsilon,\varphi)$ liczbę elementów w najliczniejszym zbiorze (n,ε)-rozdzielonym, jako $\operatorname{span}(n,\varepsilon,\varphi)$ moc najmniej licznego zbioru (n,ε)-rozpinającego oraz jako $\operatorname{cov}(n,\varepsilon,\varphi)$ moc najmniejszego pokrycia przestrzeni X zbiorami o średnicy mniejszej niż ε w metryce d_n.

Lemat 3.37. *Przy wcześniejszych założeniach zachodzą następujące własności*

$$\operatorname{cov}(n, 2\varepsilon, \varphi) \leqslant \operatorname{span}(n,\varepsilon,\varphi) \leqslant \operatorname{sep}(n,\varepsilon,\varphi) \leqslant \operatorname{cov}(n,\varepsilon,\varphi).$$

Oznaczmy jako $h_\varepsilon(\varphi) := \lim_{n\to\infty} \frac{1}{n} \ln(\operatorname{cov}(n,\varepsilon,\varphi))$. Granica ta na mocy Lematu 3.29 istnieje, jest skończona i równa się $\inf_n \frac{1}{n} \ln(\operatorname{cov}(n,\varepsilon,\varphi))$.

Definicja 3.38. *Entropię topologiczną Bowena* definiujemy jako

$$h^B_{top}(\varphi) = \lim_{\varepsilon\to 0} h_\varepsilon(\varphi) = \sup_{\varepsilon>0} \inf_n \frac{1}{n} \ln \operatorname{cov}(n,\varepsilon,\varphi).$$

Z Lematu 3.37 mamy, że

$$\frac{1}{n} \ln \operatorname{cov}(n, 2\varepsilon, \varphi) \leqslant \frac{1}{n} \ln \operatorname{span}(n,\varepsilon,\varphi) \leqslant \frac{1}{n} \ln \operatorname{sep}(n,\varepsilon,\varphi) \leqslant \frac{1}{n} \ln \operatorname{cov}(n,\varepsilon,\varphi),$$

i przechodząc z $n \to \infty$ i $\varepsilon \to 0$ otrzymujemy, że

$$h^B_{top}(\varphi) \leqslant \lim_{\varepsilon\to 0} \lim_{n\to\infty} \sup \operatorname{span}(n,\varepsilon,\varphi) \leqslant \lim_{\varepsilon\to 0} \lim_{n\to\infty} \sup \operatorname{sep}(n,\varepsilon,\varphi) \leqslant h^B_{top}(\varphi).$$

Co oznacza, że entropię topologiczną Bowena można równie dobrze zdefiniować za pomocą $\operatorname{sep}(n,\varepsilon,\varphi)$ i $\operatorname{span}(n,\varepsilon,\varphi)$.

$$h^B_{top}(\varphi) = \lim_{\varepsilon\to 0} \lim_{n\to\infty} \sup \operatorname{span}(n,\varepsilon,\varphi) = \lim_{\varepsilon\to 0} \lim_{n\to\infty} \sup \operatorname{sep}(n,\varepsilon,\varphi).$$

Niech $\mathcal{U}$ będzie otwartym pokryciem X z liczbą Lebesgue'a δ. Jeśli dla każdego $A \subseteq X$, $\operatorname{diam}(A) < \delta$ to A zawiera się w pewnym elemencie pokrycia $\mathcal{U}$. Ponadto

$$\mathcal{N}\left(\bigvee_{i=0}^{n-1} \varphi^{-i}(\mathcal{U})\right) \leqslant \operatorname{span}\left(n, \frac{\delta}{2}, \varphi\right) \leqslant \operatorname{sep}\left(n, \frac{\delta}{2}, \varphi\right) \tag{3.5}$$

oraz, jeśli $\operatorname{diam}(\mathcal{U}) < \varepsilon$, to

$$\operatorname{span}(n,\varepsilon,\varphi) \leqslant \operatorname{sep}(n,\varepsilon,\varphi) \leqslant \mathcal{N}\left(\bigvee_{i=0}^{n-1} \varphi^{-i}(\mathcal{U})\right). \tag{3.6}$$

Niech $\mathcal{U}_\varepsilon$ i $\mathcal{V}_\varepsilon$ będą pokryciami X kulami o promieniach 2ε oraz $\frac{\varepsilon}{2}$. Wtedy

$$\mathcal{N}\left(\bigvee_{i=0}^{n-1} \varphi^{-i}(\mathcal{U}_\varepsilon)\right) \leqslant \operatorname{span}(n,\varepsilon,\varphi) \leqslant \operatorname{sep}(n,\varepsilon,\varphi) \leqslant \mathcal{N}\left(\bigvee_{i=0}^{n-1} \varphi^{-i}(\mathcal{V}_\varepsilon)\right). \tag{3.7}$$

Twierdzenie 3.39. *Jeśli (X,d) jest zwartą przestrzenią metryczną, a $\varphi : X \to X$ odwzorowaniem ciągłym, to*

$$h_{top}(\varphi) = h_{top}^B(\varphi).$$

Dowód. Niech $\varepsilon > 0$ i niech $\mathcal{U}_\varepsilon$ i $\mathcal{V}_\varepsilon$ będą takie jak w (3.7). Wtedy

$$h(\varphi,\mathcal{U}_\varepsilon) = \lim_{n\to\infty} \frac{1}{n} \ln \mathcal{N}\left(\bigvee_{i=0}^{n-1} \varphi^{-i}(\mathcal{U}_\varepsilon)\right) \leqslant \lim_{n\to\infty} \sup \frac{1}{n} \ln \operatorname{sep}(n,\varepsilon,\varphi) \leqslant \lim_{n\to\infty} \frac{1}{n} \ln \mathcal{N}\left(\bigvee_{i=0}^{n-1} \varphi^{-i}(\mathcal{V}_\varepsilon)\right).$$

Przy $\varepsilon \to 0$ z twierdzenia (3.33) otrzymujemy tezę. □

3.8 Przykłady obliczania entropii topologicznej dla ekspansywnych odwzorowań

Niech X będzie zwartą przestrzenią metryczną. Przytoczymy teraz pojęcie *ekspansywności*, które pojawi się również w dalszej części pracy.

Definicja 3.40. Powiemy, że odwzorowanie $\varphi : X \to X$ jest *dodatnio ekspansywne*, jeśli istnieje takie $\delta > 0$, że dla każdych dwóch punktów $x \neq y$ dla pewnego $n \in \mathbb{N}$ mamy

$$d(\varphi^n(x), \varphi^n(y)) > \delta.$$

W powyższym δ nazywamy *stałą ekspansywności.*.

Niech $\varphi : X \to X$ będzie odwzorowaniem ciągłym. Skończone pokrycie $\mathcal{U}$ nazywamy *generatorem* dla φ, jeśli dla każdego ciągu $\{U_i\}_{i=0}^\infty$ z $\mathcal{U}$, zbiór $\bigcap_{i=0}^\infty \varphi^{-i}(\overline{U_i})$ zawiera co najwyżej jeden punkt z X. Jeśli warunek ten zastąpimy $\bigcap_{i=0}^\infty \varphi^{-i}(U_i)$ to wtedy $\mathcal{U}$ nazywamy *słabym generatorem.* Odwzorowanie φ posiada generator wtedy i tylko wtedy, gdy posiada słaby generator, patrz [30, Twierdzenie 5.20]. Generator $\mathcal{U}$ dla odwzorowania φ zadaje topologię na X, o czym mówi poniższe twierdzenie.

Twierdzenie 3.41. *Przy wcześniejszych założeniach, dla każdego $\varepsilon > 0$ istnieje $n > 0$ takie, że każdy zbiór w $\bigvee_{i=0}^n \varphi^{-i}(\mathcal{U})$ ma średnicę mniejszą niż ε. Odwrotnie, dla każdego $n > 0$ istnieje $\varepsilon > 0$ takie, że dla $d(x,y) < \varepsilon$ zachodzi*

$$x, y \in \bigcap_{i=0}^n \varphi^{-i}(U_i),$$

dla pewnych $U_0, ..., U_n \in \mathcal{U}$.

Dowód. Załóżmy, że pierwsza część twierdzenia nie zachodzi. Wtedy istnieje takie $\varepsilon > 0$, że każdego $j > 0$ istnieją x_j, y_j i $d(x_j, y_j) > \varepsilon$ oraz istnieją $U_{j,l} \in \mathcal{U}$, gdzie $0 \leqslant l \leqslant j$ i $x_j, y_j \in \bigcap_{l=0}^{j} \varphi^{-l}(U_{j,l})$. X jest zwarta, zatem możemy wziąć podciąg $\{j_k\}$ ($j_k \in \mathbb{N}$) taki, że $x_{j_k} \to x$ i $y_{j_k} \to y$. Mamy, że $x \neq y$. Rozważmy teraz zbiory $U_{j_k,0}$. Nieskończenie wiele z nich się pokrywa, skoro $\mathcal{U}$ jest skończone. Zatem, dla nieskończenie wielu k, $x_{j_k}, y_{j_k} \in U_0$ i stąd $x, y \in \overline{U_0}$. Ponadto, dla każdego i nieskończenie wiele $U_{j_k,i}$ się pokrywa i otrzymujemy, że $U_i \in \mathcal{U}$ z $x, y \in \varphi^{-i}(\overline{U_i})$. Stąd

$$x, y \in \bigcap_{i=0}^{\infty} \varphi^{-i}(\overline{U_i}).$$

Zatem mamy sprzeczność z tym, że $\mathcal{U}$ jest generatorem.

Aby udowodnić drugą część twierdzenia, niech $n > 0$ będzie dane. Załóżmy, że $\delta > 0$ jest liczbą Lebesgue'a dla $\mathcal{U}$. Dobierzmy $\varepsilon > 0$ tak, aby przy $d(x, y) < \varepsilon$ zachodziło $d(\varphi^l(x), \varphi^l(y)) < \delta$ dla $0 \leqslant l \leqslant n$. Stąd, jeśli $d(x, y) < \varepsilon$ to $\varphi^l(x), \varphi^l(y) \in U_l$ dla pewnych $U_l \in \mathcal{U}$. Zatem

$$x, y \in \bigcup_{l=0}^{n} \varphi^{-l}(U_l).$$

□

Teraz przywołamy twierdzenie analogiczne do Twierdzenie Kołmogorowa-Sinai'a 3.19, które będzie grało istotną rolę w obliczaniu entropii topologicznej.

Twierdzenie 3.42. *Niech $\varphi : X \to X$ będzie dodatnio ekspansywnym odwzorowaniem. Wtedy*

(1) Jeżeli pokrycie otwarte $\mathcal{U}$ jest generatorem, to $h_{top}(\mathcal{U}) = h(\varphi, \mathcal{U})$.

(2) Jeżeli δ jest stałą ekspansywności dla φ, to $h_{top}(\mathcal{U}) = \operatorname{span}(0, \delta_0, \varphi) = \operatorname{sep}(0, \delta_0, \varphi)$ dla wszystkich $\delta_0 < \frac{1}{4}$.

Dowód. (1). Niech $\mathcal{V}$ będzie otwartym pokryciem i niech δ będzie jego liczbą Lebesgue'a. Z Twierdzenia 3.41 wybieramy $n > 0$ takie, aby każdy zbiór z $\bigvee_{i=0}^{n} \varphi^{-i}(\mathcal{U})$ miał średnicę mniejszą niż δ. Wtedy $\mathcal{V} \preceq \bigvee_{i=0}^{n} \varphi^{-i}(\mathcal{U})$ i

$$\begin{aligned}
h(\varphi, \mathcal{V}) &\leqslant h\left(\varphi, \bigvee_{i=0}^{n} \varphi^{-i}(\mathcal{U})\right) \\
&= \lim_{k\to\infty} \frac{1}{k} \mathcal{H}\left(\bigvee_{j=0}^{k-1} \varphi^{-j}\left(\bigvee_{i=0}^{n} \varphi^{-i}(\mathcal{U})\right)\right) \\
&= \lim_{k\to\infty} \frac{1}{k} \mathcal{H}\left(\bigvee_{i=0}^{n+k-1} \varphi^{-i}(\mathcal{U})\right) \\
&= \lim_{k\to\infty} \frac{n+k-1}{k} \cdot \frac{1}{n+k-1} \mathcal{H}\left(\bigvee_{i=0}^{n+k-1} \varphi^{-i}(\mathcal{U})\right) \\
&= h(\varphi, \mathcal{U}).
\end{aligned}$$

Stąd $h(\varphi, \mathcal{V}) \leqslant h(\varphi, \mathcal{U})$ dla wszystkich pokryć otwartych $\mathcal{V}$, a zatem

$$h_{top}(\varphi) = h(\varphi, \mathcal{U}).$$

(2). Niech $\delta_0 < \frac{\delta}{4}$. Wybierzmy $x_1, ..., x_k$ takie, że $X = \bigcup_{i=1}^{k} B(x_i, \frac{\delta}{2} - 2\delta_0)$. Pokrycie $\mathcal{U} = \{B(x_i, \frac{\delta}{2}) : 1 \leqslant i \leqslant k\}$ posiada liczbę Lebesgue'a równą $2\delta_0$, więc z (3.5) i (3.6) $h(\varphi, \mathcal{U}) \leqslant \mathrm{span}(0, \delta_0, \varphi) \leqslant \mathrm{sep}(0, \delta_0, \varphi) \leqslant h_{top}(\varphi)$. Skoro $\mathcal{U}$ jest generatorem to teza wynika z (1). $\square$

Przejdźmy teraz do przykładów.

Przykład 3.43 (Jednostronne przesunięcia)**.** Rozważmy shift $\sigma_k : \Sigma_k \to \Sigma_k$ (patrz Przykład 1.16. To znaczy $\Sigma_k := \{0, ..., k-1\}^{\mathbb{N}} = \{(x_0, x_2, ...) : x_i \in \{0, ..., k-1\}\}$) oraz $\sigma_k(x_0, x_1 ...) = (x_1, x_2, ...)$. Niech $\mathcal{U} = \{U_0, ..., U_{k-1}\}$ będzie naturalnym generatorem, tzn.

$$U_j = \{(x_0, x_2, ...) : x_0 = j\}.$$

Z Twierdzenia 3.42 dostajemy

$$\begin{aligned} h_{top}(\varphi) = h(\varphi, \mathcal{U}) &= \lim_{n\to\infty} \frac{1}{n} \ln \mathcal{N}\left(\bigvee_{i=0}^{n-1} \varphi^{-i}(\mathcal{U})\right) \\ &= \lim_{n\to\infty} \frac{1}{n} \ln k^n = \ln k. \end{aligned}$$

Przykład 3.44 (Entropia topologiczna obrotu na okręgu jednostkowym)**.** Niech $\varphi : X \to X$ będzie homeomorfizmem na okręgu jednostkowym $X = S^1$. Wiemy, że odwzorowanie φ przeprowadza przedziały (odcinki) na przedziały, ponieważ są to jedyne spójne podzbiory X. Wybierzmy $\varepsilon > 0$ takie, aby

$$d(x, y) \leqslant \varepsilon \Longrightarrow d(\varphi^{-1}(x), \varphi^{-1}(y)) \leqslant \frac{1}{4}.$$

Rozważmy zbiory rozpinające dla odwzorowania φ. Widać, że $\mathrm{span}(1, \varepsilon, \varphi) \leqslant \left[\frac{1}{\varepsilon}\right] + 1$, gdzie $\left[\frac{1}{\varepsilon}\right]$ oznacza część całkowitą $\frac{1}{\varepsilon}$. Teraz weźmy $(n-1, \varepsilon)$-rozpinający zbiór F o mocy $\mathrm{span}(n-1, \varepsilon, \varphi)$ oraz punkty z $\varphi^{n-1}(F)$ i zadane przez to odcinki. Dodajmy punkty do tego zbioru tak, aby nowe odcinki miały długość mniejszą niż ε (co najwyżej $\left[\frac{1}{\varepsilon}\right] + 1$ punktów). Oznaczmy jako E zbiór nowych punktów i połóżmy

$$F' = F \cup \varphi^{-(n-1)}(E).$$

Wtedy F jest (n, ε)-rozpinającym zbiorem na X, co pokażemy. Niech $x \in X$, wtedy istnieje $y \in F$ taki, że

$$\max_{0 \leqslant i \leqslant n-2} d(\varphi^i(x), \varphi^i(y)) \leqslant \varepsilon. \tag{3.8}$$

Jeśli $d(\varphi^{n-1}(x), \varphi^{n-1}(y)) \leqslant \varepsilon$, to F jest (n, ε)-rozpinający. Jeśli natomiast nie istnieje takie $y \in F$, że obie te własności są spełnione, to wybierzmy takie $y \in F$, aby (3.8) zachodziło. Istnieją dwa domknięte odcinki z punktami końcowymi $\varphi^{n-1}(x)$ i $\varphi^{n-1}(y)$. Wybierzmy jeden z nich i oznaczmy jako I, który jest odwzorowywany poprzez φ^{-1} na odcinek I' posiadający punkty końcowe $\varphi^{n-2}(x)$ i $\varphi^{n-2}(y)$ oraz długość mniejszą bądź równą ε. Następnie wybierzmy punkt $\varphi^{n-1}(z) \in I$, gdzie $z \in F'$ z $d(\varphi^{n-1}(x), \varphi^{n-1}(z)) \leqslant \varepsilon$. Wtedy $\varphi^{n-2}(z) \in I'$, więc $d(\varphi^{n-2}(x), \varphi^{n-2}(z)) \leqslant \varepsilon$. Odcinek I' jest odwzorowywany poprzez φ^{-1} na odcinek I'' z punktami końcowymi $\varphi^{n-3}(x)$ oraz $\varphi^{n-3}(y)$ i długością mniejszą niż $\frac{1}{4}$. Zatem długość I'' jest mniejsza, bądź równa ε. Skoro $\varphi^{n-3}(z) \in I''$ to mamy, że $d(\varphi^{n-3}(x), \varphi^{n-3}(z)) \leqslant \varepsilon$. Za pomocą indukcji dostajemy

$$d(\varphi^i(x), \varphi^i(z)) \leqslant \varepsilon \quad \text{dla} \quad 0 \leqslant i \leqslant n-1.$$

Stąd F' jest (n, ε)-rozpinającym zbiorem na X. Więc

$$\operatorname{span}(n, \varepsilon, \varphi) \leqslant \operatorname{span}(n-1, \varepsilon, \varphi) + \left[\frac{1}{\varepsilon}\right] + 1.$$

Zatem

$$\operatorname{span}(n, \varepsilon, \varphi) \leqslant n\left(\left[\frac{1}{\varepsilon}\right] + 1\right)$$

Więc mamy, że $\operatorname{span}(0, \varepsilon, \varphi) = 0$. Stąd

$$h_{top}(\varphi) = 0.$$

Przykład 3.45 (Entropia topologiczna homeomorfizmu na odcinku)**.** Rozważmy homeomorfizm $\varphi : [0,1] \to [0,1]$. Wtedy $\varphi(0) = 0$ i $\varphi(1) = 1$, bądź $\varphi(0) = 1$ i $\varphi(1) = 0$. W obu tych przypadkach obie wartości 0 i 1 są punktami stałymi odwzorowania φ^2 (tzn. $\varphi^2(0) = 0$ i $\varphi^2(1) = 1$). Niech f będzie dowolnym homeomorfizmem odcinka $[0,1]$, który przeprowadza 0 na 0 i 1 na 1. Niech $\alpha : [0,1] \to X$ będzie ciągłym odwzorowaniem zadanym wzorem $\alpha(t) = e^{2\pi i t}$. Odwzorowanie te jest injektywne na $(0,1)$ i $\alpha f \alpha^{-1}$ jest homeomorfizmem na X, który przeprowadza 1 na 1 w X. Niech $\mathcal{U}_n$ będzie otwartym pokryciem odcinka $[0,1]$ odcinkami $[0, \frac{1}{n})$, $(1-\frac{1}{n}, 1]$ i $(\frac{k}{2n}, \frac{k+2}{2n})$, gdzie $1 \leqslant k \leqslant 2n-3$. Wtedy otwarte łuki $\alpha((\frac{k}{2n}, \frac{k+2}{2n}))$, gdzie $1 \leqslant k \leqslant 2n-3$ razem z $\alpha([0, \frac{1}{n}) \cup (1-\frac{1}{n}, 1])$ tworzą pokrycie otwarte $\mathcal{V}_n$ przestrzeni X. Mamy, że

$$\mathcal{N}\left(\bigvee_{j=0}^{p-1} f^j\left(\mathcal{U}_n\right)\right) \leqslant 2^p \mathcal{N}\left(\bigvee_{j=0}^{p-1}\left(\alpha f \alpha^{-1}\right)^j(\mathcal{V}_n)\right),$$

a zatem

$$h(f, \mathcal{U}_n) \leqslant \ln 2 + h(\alpha f \alpha^{-1}, \mathcal{V}_n).$$

Gdy $n \to \infty$ to z Twierdzenia 3.33 otrzymujemy, że $h_{top}(f) \leqslant \ln 2 + h_{top}(\alpha f \alpha^{-1}) = \ln 2$. Zachodzi to dla dowolnego homeomorfizmu f odcinka $[0,1]$, który przeprowadza punkty 0 na 0 i 1 na 1. Jeśli położymy $f := \varphi^{2q}$ i $2q h_{top}(\varphi) \leqslant \ln 2$, to otrzymujemy, że

$$h_{top}(\varphi) = 0.$$

3.9 Zasada Wariacyjna dla entropii topologicznej

W tym rozdziale zostanie przedstawiona Zasada Wariacyjna wiążąca entropię topologiczną z entropią Kołmogorowa-Sinai.

W całym podrodziale zakładamy, że (X, d) jest zwartą przestrzenią metryczną, a $\varphi : X \to X$ odwzorowaniem ciągłym. Kluczowymi w drugiej części dowodu zasady wariacyjnej dla entropii będą następujące lematy.

Lemat 3.46. *Jeżeli* $\mu_n \overset{*-słabo}{\to} \mu$ *to dla każdego zbioru borelowskiego* $A \subseteq X$ *dla którego miara jego brzegu wynosi zero, zachodzi* $\mu_n(A) \to \mu(A)$. *(patrz Uwaga 1.5(3))*

Lemat 3.47. *Dla każdego* $\varepsilon > 0$ *oraz* $\mu \in \mathcal{M}(X)$ *istnieje rozbicie* $\mathcal{A} = \{A_1, ..., A_k\}$ *przestrzeni* X *na zbiory borelowskie takie, że*

$$\forall_{1 \leqslant j \leqslant k} \quad \mathrm{diam}(A_j) < \varepsilon \quad \textit{oraz} \quad \mu(\partial A_j) = 0.$$

Dowód powyższych lematów znajduje się w [30, Sekcja 8.2].

Twierdzenie 3.48 (Zasada Wariacyjna)**.** *Entropia topologiczna jest równa supremum po entropiach metrycznych, tzn.*

$$h_{top}(\varphi) = \sup_{\mu \in \mathcal{M}(X,\varphi)} h_\mu(\varphi).$$

Dowód. Dowód nierówności "$\geqslant$". Niech $\mu \in \mathcal{M}(X, \varphi)$ i niech $\mathcal{A} = \{A_1, ..., A_k\}$ będzie skończonym rozbiciem X na zbiory borelowskie. Weźmy $\varepsilon > 0$ takie, że $\varepsilon < (k \ln k)^{-1}$. Ponieważ miara μ jest regularna, to

$$\forall_{1 \leqslant j \leqslant k} \quad \exists_{B_j \subseteq A_j} \quad \mu(A_j \backslash B_j) < \varepsilon,$$

gdzie B_j jest zbiorem domkniętym. Wtedy $\mathcal{B} = \{B_0, B_1, B_2, ..., B_k\}$, gdzie $B_0 = X \backslash \bigcup_{j=1}^n B_j$ jest skończonym rozbiciem X. Zauważmy, że

$$\mu(B_0) = \mu\left(\bigcup_{i=1}^k A_i \backslash \bigcup_{j=1}^k B_j\right) = \mu\left(\bigcup_{j=1}^k A_j \backslash B_j\right) \leqslant \sum_{j=1}^k \mu\left(A_j \backslash B_j\right) < k\varepsilon.$$

Stąd też

$$\mathcal{H}(\mathcal{A}|\mathcal{B}) = -\sum_{i=0}^k \sum_{j=1}^k \mu(B_i \cap A_j) \ln\left(\frac{\mu(B_i \cap A_j)}{\mu(B_i)}\right) = -\sum_{j=1}^k \mu(B_0 \cap A_j) \ln\left(\frac{\mu(B_0 \cap A_j)}{\mu(B_0)}\right),$$

ponieważ dla $i \neq 0$ mamy $\mu(B_i \cap A_j) = 0$, bądź $\frac{\mu(B_i \cap A_j)}{\mu(B_i)} = 1$. Dalej, z własności entropii rozbicia (entropia rozbicia jest ograniczona przez $\ln k$, patrz [14, dowód na stronie 9])

$$\mathcal{H}(\mathcal{A}|\mathcal{B}) = -\mu(B_0) \sum_{j=1}^k \frac{\mu(B_0 \cap A_j)}{\mu(B_0)} \ln\left(\frac{\mu(B_0 \cap A_j)}{\mu(B_0)}\right) \leqslant \mu(B_0) \ln k < \varepsilon k \ln k < 1.$$

Zauważmy teraz, że $\forall_i\ B_0 \cup B_i = X \setminus \bigcup_{j \neq i} B_j$ jest zbiorem otwartym. Zatem możemy utworzyć pokrycie otwarte $U := \{B_0 \cup B_1, B_0 \cup B_2, ..., B_0 \cup B_k\}$. Dla każdego $n \geqslant 1$

$$\mathcal{N}\left(\bigvee_{i=0}^{n-1} \varphi^{-i}(\mathcal{B})\right) \leqslant \mathcal{N}\left(\bigvee_{i=0}^{n-1} \varphi^{-i}(U)\right) 2^n,$$

zatem

$$\mathcal{H}_\mu\left(\bigvee_{i=0}^{n-1} \varphi^{-i}(\mathcal{B})\right) \leqslant \ln \mathcal{N}\left(\bigvee_{i=0}^{n-1} \varphi^{-i}(\mathcal{B})\right) \leqslant \ln \mathcal{N}\left(\bigvee_{i=0}^{n-1} \varphi^{-i}(U)\right) + n \ln 2,$$

ponieważ $\mathcal{N}(\bigvee_{i=0}^{n-1} \varphi^{-i}(\mathcal{B}))$ to ilość wszystkich niepustych elementów rozbicia $\bigvee_{i=0}^{n-1} \varphi^{-i}(\mathcal{B})$. Teraz dzieląc przez n i przechodząc z $n \to \infty$ otrzymujemy, że

$$h_\mu(\mathcal{B}, \varphi) \leqslant h(U, \varphi) + \ln 2.$$

A stąd i z własności entropii metrycznej

$$h_\mu(\mathcal{A}, \varphi) \leqslant h_\mu(\mathcal{B}, \varphi) + \mathcal{H}_\mu(\mathcal{A}|\mathcal{B}) < h(U, \varphi) + \ln 2 + 1,$$

teraz biorąc supremum (odpowiednio po rozbiciach przy entropii metrycznej i po pokryciach otwartych w przypadku topologicznej) otrzymujemy

$$h_\mu(\varphi) \leqslant h_{top}(\varphi) + \ln 2 + 1.$$

Kładąc w miejsce φ jej n-tą potęgę ($\varphi^n = \varphi \circ \varphi \circ ... \circ \varphi$) dostajemy

$$n h_\mu(\varphi) \leqslant n h_{top}(\varphi) + \ln 2 + 1.$$

Dzieląc przez n i przechodząc z $n \to \infty$ otrzymujemy tezę.

Dowód nierówności "$\leqslant$". Niech $\varepsilon > 0$. Skonstruujmy miarę $\mu \in \mathcal{M}(X, \varphi)$ taką, aby

$$h_\mu(\varphi) \geqslant \lim_{n \to \infty} \sup \frac{1}{n} \ln \operatorname{sep}(n, \varepsilon, \varphi).$$

Niech $E_n \subseteq X$ będzie zbiorem (n, ε)-rozdzielonym o mocy $\operatorname{sep}(n, \varepsilon, \varphi)$. Rozważmy miary postaci

$$\sigma_n := \frac{1}{\operatorname{sep}(n, \varepsilon, \varphi)} \sum_{x \in E_n} \delta_x,$$

mające rozkład równomierny na E_n. Połóżmy

$$\mu_n := \frac{1}{n} \sum_{i=0}^{n-1} \sigma_n \circ \varphi^{-i}.$$

Wtedy $\{\mu_n\}_{n=1}^{\infty} \subseteq \mathcal{M}(X)$, a skoro $\mathcal{M}(X)$ jest zbiorem zwartym w $*$-słabej topologii, to istnieje podciąg zbieżny μ_{n_j} do pewnego $\mu \in \mathcal{M}(X)$ oraz

$$\lim_{j \to \infty} \frac{1}{n_j} \ln \operatorname{sep}(n_j, \varepsilon, \varphi) = \lim_{n \to \infty} \sup \frac{1}{n} \ln \operatorname{sep}(n, \varepsilon, \varphi).$$

Wtedy $\mu \in \mathcal{M}(X,\varphi)$. Z Lematu 3.47 wiemy, że istnieje rozbicie $\mathcal{A} = \{A_1, A_2, ..., A_k\}$ takie, że $\operatorname{diam}(A_i) < \varepsilon$ i $\mu(\partial A_i) = 0$ dla $1 \leqslant i \leqslant k$. Niech $n \geqslant 0$. Ponieważ każdy element rozbicia $\bigvee_{i=0}^{n-1} \varphi^{-i}(\mathcal{A}) = \{A_{i_0} \cap \varphi^{-1}(A_{i_1}) \cap ... \cap \varphi^{-(n-1)}(A_{i_{k-1}}) : i_0, i_1, ..., i_{k-1} \in \{1, ..., k\}\}$ zawiera co najwyżej jeden element zbioru E_n oraz, jeżeli zawiera, to jego miara σ_n wynosi $\frac{1}{sep(n,\varepsilon,\varphi)}$. Stąd

$$\mathcal{H}_{\sigma_n}\left(\bigvee_{i=0}^{n-1} \varphi^{-i}(\mathcal{A})\right) = \ln \operatorname{sep}(n,\varepsilon,\varphi).$$

Niech j i q będą ustalone. Wtedy istnieją liczby $a(j) = [\frac{n-j}{q}]$ oraz zbiór S taki, że

$$\{0, 1, ..., n-1\} = \{j + rq + i : 0 \leqslant r < a(j), 0 \leqslant i < q\} \cup S,$$

gdzie $|S| < 2q$. Ponadto

$$\bigvee_{i=0}^{n-1} \varphi^{-i}(\mathcal{A}) = \bigvee_{r=0}^{a(j)-1} \varphi^{-(rq+j)}\left(\bigvee_{i=0}^{q-1} \varphi^{-i}(\mathcal{A}) \vee \bigvee_{l\in S} \varphi^{-l}(\mathcal{A})\right).$$

Stąd i z podaddytywności entropii

$$\ln \operatorname{sep}(n,\varepsilon,\varphi) = \mathcal{H}_{\sigma_n}\left(\bigvee_{i=0}^{n-1} \varphi^{-i}(\mathcal{A})\right) \overset{podadd.e.}{\leqslant} \sum_{r=0}^{a(j)-1} \mathcal{H}_{\sigma_n}\left(\varphi^{-(rq+j)}\left(\bigvee_{i=0}^{q-1} \varphi^{-i}(\mathcal{A})\right)\right) +$$

$$+\sum_{l\in S} \mathcal{H}_{\sigma_n}(\varphi^{-l}(\mathcal{A})) \leqslant \sum_{r=0}^{a(j)-1} \mathcal{H}_{\sigma_n\circ\varphi^{-(rq+j)}}\left(\bigvee_{i=0}^{q-1} \varphi^{-i}(\mathcal{A})\right) + 2q\ln k.$$

Teraz sumując po j od 0 do $q-1$ oraz korzystając z faktu, że $\sum_{i=1}^{n} p_i \mathcal{H}_\mu(\mathcal{A}) = \mathcal{H}_{\sum_{i=1}^{n} p_i\mu}(\mathcal{A})$, gdzie $\sum_{i=1}^{n} p_i = 1$ i $\mu_n = \frac{1}{n}\sum_{l=0}^{n-1} \sigma_n \circ \varphi^{-l}$ otrzymujemy

$$q\ln \operatorname{sep}(n,\varepsilon,\varphi) \leqslant \sum_{p=0}^{n-1} \mathcal{H}_{\sigma_n\circ\varphi^{-p}}\left(\bigvee_{i=0}^{q-1} \varphi^{-i}(\mathcal{A})\right) + 2q^2\ln k \leqslant n\mathcal{H}_{\mu_n}\left(\bigvee_{i=0}^{q-1} \varphi^{-i}(\mathcal{A})\right) + 2q^2\ln k.$$

Stąd

$$\frac{q}{n}\ln \operatorname{sep}(n,\varepsilon,\varphi) \leqslant \mathcal{H}_{\mu_n}\left(\bigvee_{i=0}^{q-1} \varphi^{-i}(\mathcal{A})\right) + \frac{2q^2}{n}\ln k. \tag{3.9}$$

Na mocy Lematu 3.46 (elementy rozbicia $\bigvee_{i=0}^{q-1} \varphi^{-i}(\mathcal{A})$ mają brzegi, których miara jest równa zeru) dla każdego B z tego rozbicia zachodzi $\lim_{j\to\infty} \mu_{n_j}(B) = \mu(B)$, a stąd

$$\lim_{j\to\infty} \mathcal{H}_{\mu_{n_j}}\left(\bigvee_{i=0}^{q-1} \varphi^{-i}(\mathcal{A})\right) = \mathcal{H}_\mu\left(\bigvee_{i=0}^{q-1} \varphi^{-i}(\mathcal{A})\right).$$

Przechodząc z $n \to \infty$ w (3.9) i dzieląc przez q dostajemy

$$\lim_{n\to\infty}\sup \frac{1}{n}\ln \operatorname{sep}(\varepsilon,n,\varphi) \leqslant \lim_{q\to\infty} \frac{1}{q}\mathcal{H}_\mu\left(\bigvee_{i=0}^{q-1} \varphi^{-i}(\mathcal{A})\right) = h_\mu(\varphi,\mathcal{A}) \leqslant h_\mu(\varphi).$$

Lewa strona przy $\varepsilon \to 0$ jest równa entropii topologicznej Bowena, więc skoro X jest zwarty, a $\varphi : X \to X$ ciągłe to z Twierdzenia 3.39 otrzymujemy tezę. $\square$

Rozdział 4

Formalizm termodynamiczny i promień spektralny operatora Ruelle'a-Perrona-Frobeniusa

Rozdział ten, który został napisany w oparciu o publikację [4], stanowi zasadniczą część niniejszej pracy. Początkowo omówimy tu pokrótce pojęcie ciśnienia topologicznego, które jest uogólnieniem entropii topologicznej oraz elementy formalizmu termodynamicznego dla odwzorowań rozszerzających. Następnie wykażemy Zasadę Wariacyjną na promień spektralny operatorów Ruelle'a-Perrona-Frobeniusa dla dowolnych odwzorowań rozszerzających. W całym obecnym rozdziale zakładamy, że (X, d) jest zwartą przestrzenią metryczną, a $\varphi : X \to X$ odwzorowaniem ciągłym.

4.1 Ciśnienie topologiczne

Ciśnienie topologiczne jest charakterystyką liczbową, która oprócz układu dynamicznego (X, φ) dodatkowo zależy od wyboru rzeczywistej funkcji ciągłej $b \in C(X, \mathbb{R})$, nazywanej w tym kontekście *potencjałem*. W szczególności ciśnienie topologiczne, oznaczane tu jako $P(\varphi, b)$ można traktować jako odwzorowanie z $C(X, \mathbb{R})$ w $\mathbb{R} \cup \{\infty\}$. Jeśli funkcja $b \in C(X, \mathbb{R})$ jest tożsamościowo równa zeru, to ciśnienie topologiczne jest równe entropii topologicznej. Wartość $P(\varphi, b)$ można zdefiniować za pomocą zbiorów rozpinających, rozdzielających czy też pokryć otwartych (tak samo jak w przypadku entropii topologicznej). W niniejszej pracy skupimy się na dwóch pierwszych. Skoro X jest zwarty, to entropię topologiczną możemy zdefiniować, za pomocą entropii topologicznej Bowena (Twierdzenie 3.39) (stąd też nasz wybór sposobu zdefiniowania ciśnienia, jak widać w podrozdziale 3.6 entropia topologiczna pierwotnie została zdefiniowana za pomocą pokryć otwartych). Oznaczmy

$$Q_n(\varphi, b, \varepsilon) := \inf_{F \subseteq X} \sum_{x \in F} \exp\left(\sum_{i=0}^{n-1} b(\varphi^i(x))\right), \tag{4.1}$$

gdzie F jest zbiorem (n, ε)-rozpinającym, a $\varepsilon > 0$. Jest to uogólnienie naszego $\operatorname{span}(n, \varepsilon, \varphi)$ z definicji entropii topologicznej Bowena (odpowiednio entropii topologicznej, gdyż X jest zwarty) po wprowadzeniu funkcji $b \in C(X, \mathbb{R})$.

Uwaga. $Q_n(\varphi, b, \varepsilon)$ posiada szereg własności:

(1) Jest ograniczone, $0 < Q_n(\varphi, b, \varepsilon) \leqslant \| \exp(\sum_{i=0}^{n-1} b(\varphi^i)) \| \operatorname{span}(n, \varepsilon, \varphi) < \infty$, patrz [30, Uwaga(1) z sekcji 7.2].

(2) Jeśli $\varepsilon_1 < \varepsilon_2$ to $Q_n(\varphi, b, \varepsilon_1) \geqslant Q_n(\varphi, b, \varepsilon_2)$.

(3) $Q_n(\varphi, 0, \varepsilon) = \operatorname{span}(n, \varepsilon, \varphi)$.

(4) W (4.1) wystarczy wziąć infinimum po zbiorach (n, ε)-rozpinających, które nie mają podzbiorów (n, φ)-rozpinających X, ponieważ $\exp(\sum_{i=0}^{n-1} b(\varphi^i(x))) > 0$.

Dla $b \in C(X, \mathbb{R})$ i $\varepsilon > 0$ wprowadźmy nowe oznaczenie

$$Q(\varphi, b, \varepsilon) := \lim_{n \to \infty} \sup \frac{1}{n} \ln Q_n(\varphi, b, \varepsilon). \tag{4.2}$$

Odpowiednikiem powyższego dla entropii topologicznej jest $\lim_{n\to\infty} \sup \frac{1}{n} \ln \operatorname{sep}(n, \varepsilon, \varphi)$.

Uwaga. Wielkość (4.2) posiada następujące własności

(1) Jest ograniczona, $Q(\varphi, b, \varepsilon) \leqslant \|f\| + \lim_{n\to\infty} \sup \frac{1}{n} \ln \operatorname{span}(n, \varepsilon, \varphi)$, patrz Uwaga 4.1 i [30, Twierdzenie 7.7(ii)].

(2) Jeśli $\varepsilon_1 < \varepsilon_2$ to $Q(\varphi, b, \varepsilon_1) \geqslant Q(\varphi, b, \varepsilon_2)$ (Uwaga pod (4.1)).

Przejdźmy teraz do definicji ciśnienia topologicznego (zauważmy analogię z definicją dla entropii topologicznej).

Definicja 4.1. *Ciśnieniem topologicznym* odwzorowania φ nazywamy odwzorowanie $P(\varphi, b) : C(X, \mathbb{R}) \to \mathbb{R} \cup \{\infty\}$, spełniające

$$P(\varphi, b) = \lim_{\varepsilon \to 0} Q(\varphi, b, \varepsilon),$$

gdzie $b \in C(X, \mathbb{R})$.

Widzimy teraz, że jeśli weźmiemy $b \equiv 0$, to

$$P(\varphi, 0) = h_{top}(\varphi) = h_{top}^B(\varphi).$$

Ciśnienie topologiczne można również zdefiniować w inny sposób. Niech $b \in C(X, \mathbb{R})$ i $\varepsilon > 0$ jak poprzednio oraz niech $n \geqslant 1$. Oznaczmy

$$P_n(\varphi, b, \varepsilon) := \sup_{E \subseteq X} \sum_{x \in E} \exp\left(\sum_{i=0}^{n-1} b(\varphi^i(x)) \right), \tag{4.3}$$

gdzie E jest (n, ε)-rozdzielonym podzbiorem X (czyli (4.3) to odpowiednik $\operatorname{sep}(n, \varepsilon, \varphi)$ dla entropii topologicznej).

Uwaga. Wielkość (4.3) posiada następujące własności

(1) Jeśli $\varepsilon_1 < \varepsilon_2$ to $P_n(\varphi, b, \varepsilon_1) \geqslant P_n(\varphi, b, \varepsilon_2)$.

(2) $P(\varphi, 0, \varepsilon) = \text{span}(n, \varepsilon, \varphi)$.

(3) W (4.3) wystarczy wziąć supremum po wszystkich (n, ε)-rozdzielonych zbiorach, które przestają być (n, ε)-rozdzielone, gdy doda się do nich jakikolwiek punkt z X. Jest tak, ponieważ $\exp(\sum_{i=0}^{n-1} b(\varphi^i(x))) > 0$. Z tego też i z faktu, że (n, ε) jest zbiorem rozdzielonym z którego nie da się utworzyć innego (n, ε)-rozdzielonego zbioru poprzez powiększenie musi być zbiorem (n, ε)-rozpinającym X wynika, że $Q_n(\varphi, b, \varepsilon) \leqslant P_n(\varphi, b, \varepsilon)$.

(4) Jeśli istnieje $\delta > 0$ takie, że dla $x, y \in X$, $d(x, y) < \frac{\varepsilon}{2}$ implikuje $|b(x) - b(y)| < \delta$ to $P_n(\varphi, b, \varepsilon) \leqslant \exp(n\delta) Q_n(\varphi, b, \frac{\varepsilon}{2})$, patrz [30, Uwaga(12) na str. 209].

Jeśli $b \in C(X, \mathbb{R})$ i $\varepsilon > 0$, to wprowadźmy kolejne oznaczenie

$$P(\varphi, b, \varepsilon) := \lim_{n\to\infty} \sup \frac{1}{n} \ln P_n(\varphi, b, \varepsilon). \tag{4.4}$$

Uwaga. Wyrażenie (4.4) posiada następujące własności

(1) $Q(\varphi, b, \varepsilon) \leqslant P(\varphi, b, \varepsilon)$ (wynika to z Uwagi pod (4.3)).

(2) Jeśli $\varepsilon_1 < \varepsilon_2$ to $P(\varphi, b, \varepsilon_1) \geqslant P(\varphi, b, \varepsilon_2)$.

(3) Jeżeli istnieje taka $\delta > 0$, że dla $x, y \in X$, $d(x, y) < \frac{\varepsilon}{2}$ implikuje $|b(x) - b(y)|$, wtedy $P(\varphi, b, \varepsilon) \leqslant \delta + Q(\varphi, b, \varepsilon)$.

Twierdzenie 4.2. *Ciśnienie topologiczne z potencjałem b można zdefiniować w następujący sposób*

$$P(\varphi, b) := \lim_{\varepsilon\to 0} P(\varphi, b, \varepsilon) = \lim_{\varepsilon\to 0} \lim_{n\to\infty} \sup \sup_{E\subseteq X} \frac{1}{n} \ln \sum_{y\in E} \exp\left(\sum_{i=0}^{n-1} b(\varphi^i(y))\right), \tag{4.5}$$

gdzie zbiór E jest (n, ε)-rozdzielony w d_n, a d_n jest n-tą metryką Bowena.

Dowód powyższego Twierdzenia znajduje się w [30, Dowód Twierdzenia 9.1, str. 209].

Ważne jest, że ciśnienie topologiczne można również zdefiniować za pomocą miar φ-niezmienniczych, o czym mówi poniższe twierdzenie.

Twierdzenie 4.3 (Zasada Wariacyjna dla ciśnienia topologicznego Waltersa)**.** *Przy wcześniejszych założeniach*

$$P(\varphi, b) = \sup_{\mu\in\mathcal{M}(X,\varphi)} \left(\int_X b d\mu + h_\mu(\varphi)\right). \tag{4.6}$$

Dowód powyższej Zasady Wariacyjnej jest analogiczny do dowodu Zasady Wariacyjnej dla entropii topologicznej i można go znaleźć w [30, Twierdzenie 9.10]. Ze wzoru (4.6) wynika, że ciśnienie topologiczne $P(\varphi, b)$ nie zależy od wyboru metryki.

Definicja 4.4. Jeśli supremum we wzorze (4.6) realizuje się dla pewnej miary $\mu \in \mathcal{M}(X, \varphi)$, tzn. $P(\varphi, b) = \int_X b d\mu + h_\mu(\varphi)$, to miarę tę nazywamy *stanem równowagi* dla układu dynamicznego (X, φ) z potencjałem b.

Jeśli odwzorowanie entropii $\mathcal{M}(X, \varphi) \ni \mu \mapsto h_\mu(\varphi)$ jest półciągłe z góry (tzn. $h_{\mu_0}(\varphi) \geqslant \limsup_{\mu \to \mu_0} h_\mu(\varphi)$ jeśli $\mu_n \overset{*-\text{słabo}}{\longrightarrow} \mu_0$), to $P(\varphi, b) = \max_{\mu \in \mathcal{M}(X,\varphi)}(\int_X b d\mu + h_\mu(\varphi))$. Ponadto, w tym przypadku zbiór stanów równowagi jest niepusty. Co więcej, jeśli tylko $P(\varphi, b) < \infty$, to zbiór ten zawsze zawiera miarę ergodyczną i dla gęstego podzbioru funkcji w $b \in C(X, \mathbb{R})$ stan równowagi jest jednoznaczny, patrz [30, 9.13 i 9.15.1]. Badanie jednoznacznie określonych stanów równowagi zostało zapoczątkowane przez formalizm termodynamiczny.

4.2 Odwzorowania rozszerzające

Niech (X, φ) będzie topologicznym układem dynamicznym. To znaczy (X, d) jest zwartą przestrzenią metryczną, a $\varphi : X \to X$ odwzorowaniem ciągłym.

Definicja 4.5. Odwzorowanie ciągłe $\varphi : X \to X$ jest *(lokalnie) rozszerzające*, jeżeli

$$\exists_{\varepsilon>0, \theta>1} \quad \forall_{x,y \in X} \quad d(x, y) < \varepsilon \Longrightarrow d(\varphi(x), \varphi(y)) \geqslant \theta d(x, y).$$

Parę (θ, ε) będziemy nazywać *podstawowym parametrem rozszerzalności.*

W tym rozdziale będziemy rozważać odwzorowanie rozszerzające $\varphi : X \to X$, które jest dodatkowo otwarte. Z tych własności (tego, że odwzorowanie jest otwarte i rozszerzające) otrzymujemy, że $\varphi : X \to X$ jest lokalnym homeomorfizmem.

Teraz, przybliżymy nieco własności lokalnie rozszerzających układów dynamicznych. Następujące dwa stwierdzenia zostaną wykorzystane w dowodzie twierdzenia w ostatniej części pracy.

Stwierdzenie 4.6. *Niech (X, φ) będzie lokalnie rozszerzającym układem dynamicznym na zwartej przestrzeni metrycznej (X, d) z podstawowym parametrem rozszerzalności (θ, ε). Wtedy*

(1) φ jest lokalnym homeomorfizmem. Dokładniej, dla każdego $x \in X$ i $0 < \varepsilon' \leqslant \varepsilon$, $\varphi : B(x, \varepsilon') \to \varphi(B(x, \varepsilon'))$ ($B(x, \varepsilon')$ to kula w metryce d) jest homeomorfizmem.

(2) Dla każdego $y \in X$, $\varphi^{-1}(y)$ jest skończone. Ponadto, istnieje $0 < a \leqslant \varepsilon$ takie, że dla każdego $y \in X$ dla którego zachodzi $\varphi^{-1}(y) = \{x_1, ..., x_n\}$ istnieją lokalne odwrotności $g_1, ..., g_n$ odwzorowania φ zdefiniowane na kuli $B(x, a)$ takie, że $g_j(y) = x_j$ i $g_j(B(y, a))$ $(1 \leqslant j \leqslant n)$ są parami rozłączne.

(3) Niech a będzie stałą w (1). Jeśli $0 < r \leqslant a$, to odwzorowanie

$$\varphi^n : B_n(x,r) \to B(\varphi^n(x), r),$$

gdzie $B_n(x,\varepsilon)$-n-ta kula w metryce Bowena, jest homeomorfizmem.

Dowód. (1). Jasnym jest, że odwzorowanie $\varphi|_{\overline{B(x,\varepsilon')}}$ jest injektywne. Odwzorowanie φ jest ciągłe na domkniętej kuli $\overline{B(x,\varepsilon')}$ i skoro domknięta kula jest zwarta, odwrotność $\varphi|_{\overline{B(x,\varepsilon')}}$ także jest ciągłe. Skoro $\varphi : B(x,\varepsilon') \to \varphi(B(x,\varepsilon'))$ jest bijekcją to jest także homeomorfizmem.

(2). Gdyby $\varphi^{-1}(y)$ nie było skończone, to φ nie byłoby homeomorficzne w punktach granicznych $\varphi^{-1}(y)$, co przeczy (1). Niech

$$d(y) = \inf_{k \neq j} d(x_k, x_j) > 0$$

będzie najmniejszą odległością pomiędzy przeciwobrazami y. Z (1), jeśli $0 < r \leqslant \min(\varepsilon, \frac{d(y)}{2})$, to $\varphi : B(x_j, r) \to \varphi(B(x_j, r))$ jest homeomorfizmem dla każdego j. Skoro otwarty zbiór $\bigcap_{j=1}^n \varphi(B(x_j, r))$ zawiera y, to musi także zawierać wystarczająco małą kulę $B(y, r_y)$ z $r_y > 0$, taką, aby odwzorowanie odwrotne g_j odwzorowujące y w x_j spełniało

$$g_j : B(y, r_y) \to g_j(B(y, r_y)) \subseteq B(x_j, r).$$

Skoro kule $B(x_j, r)$ są rozłączne, to $g_j(B(y, r_y))$ także są rozłączne. Teraz weźmy skończoną liczbę kul $B(y_i, r_{y_i})$ takich, że $\{B(y_i, \frac{r_{y_i}}{2})\}$ są pokryciem X. Pokażemy, że

$$a = \frac{1}{2} \min_i \{r_{y_i}\}$$

jest żądaną stałą. W istocie, dla dowolnego $y \in X$ mamy, że $y \in B(y_i, \frac{r_{y_i}}{2})$ dla pewnych i. Wtedy $B(y,a) \subseteq B(y_i, r_{y_i})$. Zatem $\varphi^{-1}(B(y,a))$ posiada swoje elementy odwrotne zawarte w elementach odwrotnych $\varphi^{-1}(B(y_i, r_{y_i}))$, które są parami rozłączne.

(3). Niech $g_1, ..., g_n$ będą lokalnymi odwrotnościami φ takimi, że

$$x \xleftarrow{g_1} \varphi(x) \xleftarrow{g_2} \varphi^2(x) \xleftarrow{g_3} ... \xleftarrow{g_{n-1}} \varphi^{n-1}(x) \xleftarrow{g_n} \varphi^n(x).$$

Twierdzimy, że dla każdych $1 \leqslant k \leqslant n$ mamy

$$g_{n-k+1} \circ ... \circ g_n(B(\varphi^n(x), r)) = B_k(\varphi^{n-k}(x), r).$$

Wykażemy to za pomocą metody indukcji po k. Dla $k = 1$ należy pokazać, że $g_n(B(\varphi^n(x), r)) = B_1(\varphi^{n-1}(x), r)$. Skoro (X, φ) jest układem dynamicznym lokalnie rozszerzającym, g_n jest kontrakcją, więc

$$g_n(B(\varphi^n(x), r)) \subseteq B\left(\varphi^{n-1}(x), \frac{r}{\theta}\right) \subseteq B(\varphi^{n-1}(x), r).$$

Z drugiej strony, skoro g_n jest homeomorfizmem do którego odwzorowaniem odwrotnym jest φ, więc $y \in g_n(B(\varphi^n(x), r))$ wtedy i tylko wtedy, gdy $\varphi(y) \in B(\varphi^n(x), r)$. Z tych dwóch faktów wnioskujemy, że dla $k = 1$ równość zachodzi. Teraz załóżmy, że zachodzi dla k. Wtedy

$$g_{n-k} \circ ... \circ g_n(B(\varphi^n(x), r)) = g_{n-k}(B_k(\varphi^{n-k}(x), r)).$$

Zatem musimy pokazać, że $g_{n-k}(B_k(\varphi^{n-k}(x), r)) = B_{k+1}(\varphi^{n-k-1}(x), r)$. W istocie, $z \in g_{n-k}(B_k(\varphi^{n-k}(x), r))$ wtedy i tylko wtedy, gdy $d(z, \varphi^{n-k-1}(x)) < r$ i $\varphi(z) \in B_k(\varphi^{n-k}(x), r)$. (Korzystamy to z faktu, że kule $B_k(y, r)$ się zmniejszają, gdy k rośnie i, że φ jest lokalnym homeomorfizmem, który odwzorowuje $\varphi^{n-k-1}(x)$ w $\varphi^{n-k}(x)$ z odwzorowaniem odwrotnym g_{n-k}.). Równoważnie,

$$d(z, \varphi^{n-k-1}(x)) < r, \ \ d(\varphi(z), \varphi^{n-k}(x)) < r, \ \ ..., \ \ d(\varphi^{k+1}(z), \varphi^n(x)) < r.$$

Zatem $z \in B_{k+1}(\varphi^{n-k-1}(x), r)$. Kładąc φ^n po obu stronach

$$g_1 \circ g_2 ... \circ g_n(B(\varphi^n(x), r)) = B_n(x, r),$$

dostajemy tezę. □

Definicja 4.7. Otwarte odwzorowanie rozszerzające $\varphi : X \to X$ nazywamy *topologicznie mieszającym*, jeśli dla każdego niepustego, otwartego zbioru U możemy znaleźć takie $n \geqslant 1$, aby $\varphi^n(U) = X$.

Stwierdzenie 4.8. *Załóżmy, że $\varphi : X \to X$ jest odwzorowaniem mieszającym zdefiniowanym na zwartej przestrzeni metrycznej (X, d). Dla dowolnego $r > 0$ istnieje liczba całkowita $p = p(r) \geqslant 1$ taka, że $\varphi^p(B(x, r)) = X$ dla każdego $x \in X$.*

Dowód. Ze zwartości zbioru X, istnieje dla niego skończone pokrycie kulami $B(x_i, \frac{r}{2})$. Dla każdego i istnieje liczba całkowita $p(x_i) > 0$ taka, że

$$\varphi^{p(x_i)}\left(B\left(x_i, \frac{r}{2}\right)\right) = X.$$

Niech $p = \max_i\{p(x_i)\}$. Taki wybór spełnia nasze wymagania, ponieważ dla każdego $x \in X$ mamy, że $x \in B(x_i, \frac{r}{2})$ dla pewnych i i stąd $B(x, r) \supset B(x_i, \frac{r}{2})$. Zatem

$$\varphi^p(B(x, r)) \supseteq \varphi^{p-p(x_i)}\left(\varphi^{p(x_i)}\left(B\left(x_i, \frac{r}{2}\right)\right)\right) \supseteq \varphi^{p-p(x_i)}(X) = X.$$

□

Z pojęciem rozszerzalności związane jest inne pojęcie, a mianowicie ekspansywność (zdefiniowana w 3 rozdziale). W większości publikacji używa się częściej pojęcia ekspansywności, niż rozszerzalności. Okazuje się jednak, jak wykazał Reddy, że oba pojęcia są równoważne (i nie zależą od wyboru metryki):

Twierdzenie 4.9 (Reddy [25])**.** *Ani rozszerzalność, ani dodatnia ekspansywność odwzorowania nie zależy od wyboru metryki (zgodnej z topologią). Ponadto, $\varphi : X \to X$ jest rozszerzające wtedy i tylko wtedy, gdy $\varphi : X \to X$ jest dodatnio ekspansywne.*

Z Twierdzenia Schwartzman'a, patrz [21], [12], homeomorfizm $\varphi : X \to X$ jest odwzorowaniem rozszerzającym (dodatnio ekspansywnym) wtedy i tylko wtedy, gdy X jest zbiorem skończonym. Stąd rozszerzalność jest pojęciem przeznaczonym dla układów dynamicznych nieodwracalnych (tj. takich, gdzie odwzorowanie $\varphi : X \to X$ jest nieodwracalne). Poprawnym odpowiednikiem pojęcia rozszerzalności dla homeomorfizmów jest *obustronna ekspansywność*. W literaturze można znaleźć wiele faktów związanych z ekspansywnymi homeomorfizmami, jednakże powszechnie wiadomo, że (jednostronna) wersja takich twierdzeń zachodzi również dla odwzorowań rozszerzających (dodatnio ekspansywnych), patrz [30, Uwaga na stronie 145], [12, strona 58]. Odpowiednik [30, Twierdzenie 5.24], [12, Wniosek 3.3] mówi o tym, że każde odwzorowanie rozszerzające $\varphi : X \to X$ jest *półsprzężone z jednostronnym podshiftem* (por. Przykład 1.16). To znaczy istnieje domknięty podzbiór Z z $\Sigma_n = \{1, ..., n\}^{\mathbb{N}}$ taki, że $\sigma(Z) = Z$ oraz ciągłe oraz surjektywne odwzorowanie $\phi : Z \to X$ takie, że $\phi \circ \sigma = \sigma \circ \phi$. Co więcej, takie jednostronne podshifty (Z, σ) mogą być sklasyfikowane jako odwzorowania rozszerzające na przestrzeniach zerowo-wymiarowych.

Twierdzenie 4.10. *Dla dowolnego ekspansywnego odwzorowania $\varphi : X \to X$ odwzorowanie entropii $\mathcal{M}(X, \varphi) \ni \mu \mapsto h_\mu(\varphi)$ jest półciągłe z góry i stąd stany równowagi dla odwzorowań rozszerzających zawsze istnieją.*

Dowód. Patrz [24, Twierdzenie 3.5.6]. □

4.3 Miary Gibbsa

Modelowymi stanami równowagi (patrz Definicja 4.4) są miary Gibbsa badane przez formalizm termodynamiczny. Takie stany równowagi są również istotne w analizie spektralnej operatorów Ruelle'a-Perrona-Frobeniusa. W tym podrozdziale opiszemy te obiekty.

W literaturze zazwyczaj miary Gibsa są rozważane dla odwzorowań *topologicznie mieszających* (tzn. takich, że dla każdego niepustego zbioru U istnieje $n \geqslant 0$ takie, że $\varphi^n(U) = X$), patrz [6], [27], [31], [7]. Jednakże, jak pokazano w [24] można równie dobrze zamienić warunek na odwzorowanie $\varphi : X \to X$ na nieco słabszy. Mianowicie, będziemy tu zakładać, że odwzorowanie φ jest *topologicznie tranzytywne*, tzn. dla każdych dwóch niepustych zbiorów otwartych $U, V \subseteq X$ istnieje takie $n \in \mathbb{N}$, że $\varphi^n(U) \cap V \neq \emptyset$. Jeśli φ jest odwzorowaniem rozszerzającym i otwartym, jest to równoważne warunkowi, że istnieje $N \in \mathbb{N}$ takie, że φ^N jest topologicznie mieszające. Mówimy, że funkcja $f : X \to \mathbb{C}$ jest *hölderowsko ciągła*, jeśli istnieją stałe $L \geqslant 0$ i $\alpha \in (0, 1]$, że dla każdych $x, y \in X$ zachodzi

$$|f(x) - f(y)| \leqslant L|x - y|^\alpha.$$

Definicja 4.11. Mówimy, że miara $\mu \in \mathcal{M}(X, \varphi)$ jest *rozkładem Gibbsa* dla ciągłego odwzorowania $\varphi : X \to X$ z potencjałem $b \in C(X, \mathbb{R})$, jeśli dla dostatecznie małego $\varepsilon > 0$ możemy znaleźć takie $P \in \mathbb{R}$ i $s \geqslant 1$, aby dla każdego $x \in X$ i $n \geqslant 1$

$$s^{-1} \leqslant \frac{\mu(B_n(x,\varepsilon))}{\exp(\sum_{i=0}^{n-1} b(\varphi^i(x)) - nP)} \leqslant s,$$

gdzie $B_n(x, \varepsilon)$ jest kulą otwartą w n-tej metryce Bowena.

Twierdzenie 4.12 (Istnienie i jednoznaczność miar Gibbsa)**.** *Niech $\varphi : X \to X$ będzie odwzorowaniem rozszerzającym i niech $b \in C(X, \mathbb{R})$ jest hölderowsko ciągłe. Wtedy istnieje jednoznaczny rozkład Gibbsa $\mu_{\varphi,b} \in \mathcal{M}(X, \varphi)$ dla φ, b, oraz stałej P, gdzie P w Definicji 4.11 jest ciśnieniem topologicznym $P(\varphi, b)$. Ponadto, $\mu_{\varphi,b}$ jest jedynym stanem równowagi dla φ i b, więc dla każdego $y \in X$ mamy*

$$P = P(\varphi, b) = \int_X b d\mu_{\varphi,b} + h_{\mu_{\varphi,b}}(\varphi) = \lim_{n\to\infty} \frac{1}{n} \ln \sum_{x \in \varphi^{-1}(y)} \exp\left(\sum_{i=0}^{n-1} b(\varphi^i(x))\right).$$

Dowód. Teza wynika z połączenia [24, Stwierdzeń 3.4.3, 4.4.1, 4.1.5 i Twierdzeń 4.3.2, 4.6.2], patrz też [4]. □

Rozkład Gibbsa $\mu_{\varphi,b}$ jest skonstruowany w oparciu o Twierdzenie Ruelle'a Perrona Frobeniusa. Przypomnijmy, patrz Definicja 2.3, że operator Ruelle'a Perrona Frobeniusa jest dany wzorem

$$\mathcal{L}_c(f)(x) = \sum_{y \in \varphi^{-1}(x)} c(y) f(y), \qquad x \in X, f \in C(X),$$

gdzie $c : X \to [0, \infty)$ jest pewną ustaloną i nieujemną funkcją ciągłą. Z dodatniości, operator $\mathcal{L}_c{}^* = C(X)^* \to C(X)^*$ (dualny do operatora $\mathcal{L}_c : C(X) \to C(X)$) ogranicza się do afinicznego odwzorowania na stożku z miar bolerowskich. Następujące twierdzenie, pierwotnie było dowiedzione dla jednostronnych shiftów przez Ruelle'a, następnie zostało rozszerzone przez innych autorów na mieszające odwzorowania rozszerzające, patrz [6], [31], [27], [7]. W tej pracy zostanie ono uogólnione na topologicznie tranzytywne odwzorowania rozszerzające:

Twierdzenie 4.13 (Twierdzenie Ruelle'a-Perrona-Frobeniusa)**.** *Niech $\varphi : X \to X$ będzie otwartym, rozszerzającym oraz topologicznie tranzytywnym odwzorowaniem na zwartej przestrzeni metrycznej X. Załóżmy, że $b \in C(X, \mathbb{R})$ jest hölderowsko ciągłe. Niech $\rho := r(\mathcal{L}_{e^b})$ będzie promieniem spektralnym operatora przejścia Ruelle'a-Perrona-Frobeniusa $\mathcal{L}_{e^b} : C(X) \to C(X)$ $(c(x) = e^{b(x)})$.*

(1) Istnieje jednoznacznie wyznaczona miara probabilistyczna ν na X taka, że $\mathcal{L}_{e^b}^* \nu = \rho\nu$.

(2) ρ jest ściśle dodatnią wartością własną operatora $\mathcal{L}_{e^b}$ z jednoznacznie wyznaczoną ściśle dodatnią funkcją własną $h \in C(X)^+$ z $\nu(h) = 1$.

Z tak zadanymi h oraz ν, $a \in C(X)$, $h\nu(a) := \nu(ah)$ jest miarą Gibbsa $\mu_{\varphi,b}$ dla φ i b. Ponadto $P(\varphi, b) = \ln r(\mathcal{L}_{e^b})$.

Dowód. Dla uproszczenia położmy $\mathcal{L} = \mathcal{L}_{e^b}$ i niech $\mathcal{M}(X)$ będzie zbiorem borelowskich miar probabilistycznych.

Dowód 1). Istnienie miary $\nu \in \mathcal{M}(X)$ takiej, że $\mathcal{L}^*(\nu) = c\nu$ dla pewnych $c > 0$ wynika z Twierdzenia Tichonowa o punkcie stałym, patrz [24, Twierdzenie 5.2.8]. Załóżmy, że $\mu, \eta \in \mathcal{M}(X)$ spełniają $\mathcal{L}^*(\mu) = c_1\mu$ i $\mathcal{L}^*(\eta) = c_2\eta$ dla pewnych $c_1, c_2 > 0$. Z [24, Stwierdzenia 5.2.11, 5.1.14], $c_1 = c_2 = e^{P(\varphi,b)}$ oraz miary μ i η są równoważne z miarą Gibbsa $\mu_{\varphi,b}$. Ponadto mamy, że $P(\varphi, b) = \ln r(\mathcal{L}) = \ln \rho$, co przy wcześniejszych założeniach można udowodnić używając wzoru z Twierdzenia 4.12. Stąd $\mathcal{L}^*(\mu) = \rho\mu$, $\mathcal{L}^*(\eta) = \rho\eta$, oraz istnieje η-całkowalna funkcja h taka, że $\mu = h\eta$. Dla każdego $a \in C(X)$ mamy, że $\mathcal{L}(a)h = \mathcal{L}(a(h \circ \varphi))$ i stąd $\mathcal{L}^*(h\eta) = (h \circ \varphi)\mathcal{L}^*(\eta) = \rho(h \circ \varphi)\eta$. Z tego mamy, że

$$h\eta = \mu = \rho^{-1}\mathcal{L}^*(\mu) = \rho^{-1}\mathcal{L}^*(h\eta) = (h \circ \varphi)\eta.$$

Więc $h = (h \circ \varphi)$ η-prawie wszędzie. Jednak, ponieważ miara η jest równoważna mierze Gibbsa $\mu_{\varphi,b}$, która jest φ-ergodyczna, patrz [24, Wniosek 5.2.13], to h jest stała η-prawie wszędzie. Stąd $\eta = \mu$.

Dowód 2). Istnienie funkcji $h \in C(X)^+$ z $\nu(h) = 1$ i $\mathcal{L}(h) = \rho h$ wynika z [24, Stwierdzenie 5.3.1]. Wtedy $h\nu = \mu_{\varphi,b}$ jest miarą Gibbsa z [24, Twierdzenie 5.3.2]. Stąd, jeśli $g \in C(X)^+$ jest inną funkcją, która posiada własności h, to zachodzi $h\nu = g\nu$, więc $g = h$, ν-prawie wszędzie z czego wynika, że $g = h$, ponieważ g i h są funkcjami ciągłymi i $\operatorname{supp} \nu = X$. $\square$

Powyższe zależności mogą być użyte w bardziej ogólnych sytuacjach przy zastosowaniu rozkładu spektralnego. Przypomnijmy, że $\Omega(\varphi)$ oznacza zbiór wszystkich punktów niewędrujących. Jeśli φ jest rozszerzające to $\Omega(\varphi)$ jest równy domkniętemu zbiorowi punktów okresowych, patrz [24, Stwierdzenie 4.3.6]. Z [24, Wniosek 4.2.4] następujące twierdzenie jest specjalnym przypadkiem [24, Twierdzenie 4.3.8]:

Twierdzenie 4.14 (Rozkład spektralny). *Załóżmy, że $\varphi : X \to X$ jest otwartym i rozszerzającym odwzorowaniem takim, że zbiór punktów okresowych jest gęsty w X (to znaczy $\Omega(\varphi) = X$). Wtedy $X = \bigsqcup_{j=1}^{N} X_j$ jest zbiorem złożonym z sumy skończonej liczby φ-niezmienniczych, rozłącznych zbiorów otwartych X_j, $j = 1, ..., N$, takich, że dla poszczególnych j, odwzorowanie $\varphi|_{X_j} : X_j \to X_j$ jest topologicznie tranzytywne. Ponadto, każde X_j jest zbiorem złożonych z $k(j)$ rozłącznych zbiorów otwarto-domkniętych X_j^k, które są cyklicznie permutowane przez φ i takie, że $\varphi^{k(j)}|_{X_j^k} : X_j^k \to X_j^k$ jest topologicznie mieszające dla każdych j i k.*

Przykład 4.15. Niech $\mathbb{A} = [t_{i,j}]_{i,j=1}^n$ będzie macierzą zero-jedynkową bez wierszy złożonych z samych zer. *Topologicznym shiftem Markova* ze zbiorem stanów $\{1, ..., n\}$ i macierzą przejścia $\mathbb{A}$ nazywamy odwzorowanie $\sigma_{\mathbb{A}} : \Sigma_{\mathbb{A}} \to \Sigma_{\mathbb{A}}$, gdzie

$$\Sigma_{\mathbb{A}} := \{(\xi_1, \xi_2, ...) \in \{1, ..., n\}^{\mathbb{N}} : t_{\xi_i,\xi_{i+1}} = 1\}, \quad \sigma_{\mathbb{A}}(\xi_1, \xi_2, \xi_3, ...) := (\xi_2, \xi_3, ...).$$

Tutaj $\Sigma_{\mathbb{A}}$ jest metryzowalną i zwartą przestrzenią wyposażoną w topologię produktową otrzymaną z $\{1,...,n\}^{\mathbb{N}}$, natomiast $\sigma_{\mathbb{A}}$ jest rozszerzającym, lokalnym homeomorfizmem. Ten układ może być również rozpatrywany jako zbiór danych z których można skonstruować graf skierowany ze zbiorem wierzchołków $\{1,...,n\}$ i zbiorem krawędzi $\{(i,j) : t_{i,j} = 1\}$. Wtedy $\Sigma_{\mathbb{A}}$ jest przestrzenią złożoną z nieskończonej liczby dróg w tym grafie. Liczba dróg z i do j długości n wynosi $t_{i,j}^{(n)}$, gdzie $\mathbb{A}^n = [t_{i,j}^{(n)}]_{i,j=1}^n$ jest n-tą potęgą macierzy $\mathbb{A}$. Piszemy $i \to j$, jeśli istnieje ścieżka z i do j, to znaczy $t_{i,j}^{(n)} > 0$ dla pewnych n. Stan i jest nazywany *powtarzalnym*, jeśli $i \to j$ implikuje $j \to i$ dla każdego stanu j. Mamy, że $\Omega(\sigma_{\mathbb{A}}) = \{(\xi_1, \xi_2, ...) \in \Sigma_{\mathbb{A}} : \xi_i$ jest powtarzalny dla każdego $i = 1, 2, ...\}$. Stąd $\Omega(\sigma_{\mathbb{A}}) = \Sigma_{\mathbb{A}}$ wtedy i tylko wtedy, jeśli wszystkie stany są powtarzalne. Ponadto $\sigma_{\mathbb{A}}$ jest topologicznie tranzytywne wtedy i tylko wtedy, gdy $\mathbb{A}$ jest *nieredukowalna* (tzn. dla każdego i oraz j istnieje n takie, że $t_{i,j}^{(n)} > 0$), jeśli tylko $i \leftrightarrow j$ dla wszystkich stanów i i j. Ponadto $\sigma_{\mathbb{A}}$ jest topologicznie mieszające wtedy i tylko wtedy, gdy $\mathbb{A}$ jest *nieokresowa* (tzn. istnieje $n \in \mathbb{N}$ takie, że wszystkie kolumny i macierze w $\mathbb{A}^n$ są dodatnie). Jeśli $\Omega(\sigma_{\mathbb{A}}) = \Sigma_{\mathbb{A}}$, wtedy $i \leftrightarrow j$ jest relacją równoważności i stąd $\{1,...,n\}$ rozkłada się na rozłączne klasy równoważności. Po ewentualnym ponownym oznakowaniu, te klasy są postaci $\{k_j+1,...,k_{j+1}\}$, $j = 1,...,N$, gdzie $k_1 = 0 < k_2 < ... < k_{N+1} = n$. Wtedy $\mathbb{A} = \mathrm{diag}(\mathbb{A}_1, ..., \mathbb{A}_N)$, gdzie $\mathbb{A}_j = [t_{i,l}]_{i,l=k_j+1}^{k_j+1}$ dla $j = 1,...,n$ i $\Sigma_{\mathbb{A}} = \bigsqcup_{j=1}^{N} \Sigma_{\mathbb{A}}$ jest rozkładem z Twierdzenia 4.14 zastosowanym do $\sigma_{\mathbb{A}}$.

4.4 Zasada Wariacyjna dla promienia spektralnego operatora Ruelle'a-Perrona-Frobeniusa dla rozszerzających lokalnych homeomorfizmów

Ten końcowy podrozdział będzie poświęcony opisowi metody obliczenia promienia spektralnego operatora Ruelle'a-Perrona-Frobeniusa i jego relacji z ciśnieniem topologicznym dla układów dynamicznych generowanych przez lokalne homeomorfizmy rozszerzające.

Zakładamy, że $\varphi : X \to X$ jest lokalnym homeomorfizmem, a $c : X \to [0, \infty)$ ciągłą funkcją. Niech $\mathcal{L}_c : C(X) \to C(X)$ będzie operatorem Ruelle'a-Perrona-Frobeniusa omówionym w Rozdziale 2. Z Twierdzenia 4.13 oraz 4.3, jeśli c jest ściśle dodatnia i hölderowsko ciągła oraz φ jest odwzorowaniem rozszerzającym i topologicznie tranzytywnym, to

$$\ln r(\mathcal{L}_c) = \sup_{\mu \in \mathcal{M}(X,\varphi)} \left(\int_X \ln c d\mu + h_\mu(\varphi) \right).$$

Pokażemy, że powyższy wzór zachodzi dla dowolnej funkcji ciągłej c oraz dowolnego, otwartego odwzorowania rozszerzającego φ. Dowód nie zależy od Twierdzenia 4.13.

Korzystając z równości $\|\mathcal{L}_c^n\| = \|\mathcal{L}_c^n(1)\|$, $n \in \mathbb{N}$, (patrz Stwierdzenie 2.4) oraz wzoru Gelfanda na promień spektralny, w ogólności mamy

$$\ln r(\mathcal{L}_c) = \lim_{n \to \infty} \max_{y \in X} \frac{1}{n} \sum_{x \in \varphi^{-n}(y)} \prod_{i=0}^{n-1} c(\varphi^i(x)). \tag{4.7}$$

Następujący lemat jest częścią [20, Twierdzenie 12].

Lemat 4.16. *Jeśli $\varphi : X \to X$ jest lokalnym homeomorfizmem oraz funkcja $c : X \to (0,\infty)$ jest ściśle dodatnia, wtedy*

$$\ln r(\mathcal{L}_c) \leqslant P(\varphi, \ln c) = \sup_{\mu \in \mathcal{M}(X,\varphi)} \left(\int_X \ln c d\mu + h_\mu(\varphi) \right). \tag{4.8}$$

Dowód. Patrząc na wzory (4.5) i (4.7) wystarczy pokazać, że istnieje $\varepsilon > 0$ takie, że dla dowolnych $n \in \mathbb{N}$ i $x \in X$ zbiór $\varphi^{-n}(x)$ jest (n, ε)-rozdzielony w n-tej metryce Bowena d_n. Drugi warunek zachodzi dla każdego lokalnego homeomorfizmu. W rzeczy samej, zbiór $\Delta := \{x : |\varphi^{-1}(x)| = 1\}$ jest otwarty-domknięty w X. Funkcja $d : X \backslash \Delta \to (0,\infty)$ dana wzorem $d(x) = \min\{d(y_1, y_2) : y_1, y_2 \in \varphi^{-1}(x), y_1 \neq y_2\} > 0$ jest ciągła na zbiorze zwartym $X \backslash \Delta$. Stąd

$$\varepsilon := \frac{1}{2} \min \left\{ \min_{x \in X\backslash\Delta} d(x), d(X\backslash\Delta, \Delta) \right\} > 0.$$

Łatwo sprawdzić, że jest to oczekiwana liczba. □

W ogólności, jeśli φ nie jest rozszerzające, nierówność (4.8) może być ostra. W istocie, jeśli φ jest homeomorfizmem i $c \equiv 1$, wtedy $\ln r(\mathcal{L}_c) = 0$ oraz $P(\varphi, 0) = h_{top}(\varphi)$ może być dowolną, nieujemną liczbą, patrz podrozdział 4.1.

Twierdzenie 4.17. *Niech $\varphi : X \to X$ będzie rozszerzającym, otwartym odwzorowaniem na metryzowalnej i zwartej przestrzeni X. Dla dowolnej ciągłej funkcji $c : X \to [0,\infty)$ zachodzi*

$$\ln r(\mathcal{L}_c) = \max_{\mu \in \mathcal{M}(X,\varphi)} \left(\int_X \ln c d\mu + h_\mu(\varphi) \right) = \max_{\mu \in \mathrm{Erg}(X,\varphi)} \left(\int_X \ln c d\mu + h_\mu(\varphi) \right),$$

gdzie $h_\mu(\varphi)$ jest entropią Kołmogorowa-Sinai'a (mamy, że $h_\mu(\varphi) < \infty$).

Dowód. Z tego, że φ jest rozszerzające, odwzorowanie entropii $\mathcal{M}(X,\varphi) \ni \mu \to h_\mu(\varphi) \in [0,\infty)$ jest półciągłe z góry (patrz [30, 7.11.1 i 8.2] i Uwagi pod Twierdzeniem 4.9). Ponadto odwzorowanie $\mathcal{M}(X,\varphi) \ni \mu \to \int_X \ln c d\mu$ jest zawsze półciągłe z góry, patrz dowód z [28, Lemat 1.4]. Zatem supremum $\sup_{\mu \in \mathcal{M}(X,\varphi)}(\int_X \ln c d\mu + h_\mu(\varphi))$ jest w istocie maksimum. Co więcej, używając rozkładu ergodycznego (który umożliwia wyrażenie nie-ergodycznych miar jako średniej po miarach ergodycznych) otrzymujemy

$$\max_{\mu \in \mathcal{M}(X,\varphi)} \left(\int_X \ln c d\mu + h_\mu(\varphi) \right) = \max_{\mu \in \mathrm{Erg}(X,\varphi)} \left(\int_X \ln c d\mu + h_\mu(\varphi) \right),$$

patrz [24, Wniosek 2.4.3], bądź [30, 9.10.1].

Dowód twierdzenia będzie przebiegał poprzez kolejne rozluźnianie założeń. Początkowo załóżmy, że funkcja $c > 0$ jest ściśle dodatnia. Więc z Lematu 4.16 musimy pokazać, że $P(\varphi, \ln c) \leqslant \ln r(\mathcal{L}_c)$.

1) Początkowo, niech odwzorowanie $\varphi : X \to X$ będzie topologicznie tranzytywne. Z

Twierdzenia 4.14 $X = \bigsqcup_{j=1}^{k} X_j$ jest zbiorem złożonym z sumy k rozłącznych zbiorów X_j otwarto-domkniętych, które są cyklicznie permutowane przez φ oraz takie, że $\varphi^k : X_j \to X_j$ jest topologicznie mieszające. Więc z 4.8 dla każdego $j = 1, ..., k$ i każdego $\varepsilon > 0$ istnieje $p_j(\varepsilon) \in \mathbb{N}$ takie, że $\varphi^{kp_j(\varepsilon)}(B(x,\varepsilon) \cap X_j)) = X_j$ dla każdego $x \in X_j$. Stąd istnieje $p(\varepsilon) \geqslant \max_{j=1,...,k} p_j(\varepsilon)$ takie, że $\varphi^{p(\varepsilon)}(B(x,\varepsilon) \cap X_j)) = X_j$ dla każdego $x \in X_j$ i $j = 1, ..., k$. Z 4.6 dla wystarczająco małego $\varepsilon > 0$ mamy, że $\varphi^n(B_n(x,\varepsilon)) = B(\varphi^n(x),\varepsilon)$ dla wszystkich $n \in \mathbb{N}$ i $x \in X$ (gdzie $B_n(x,\varepsilon)$ jest kulą w n-tej metryce Bowena). Odpowiednio, dla każdych $x \in X_i$, $i = 1, ..., k$ i $n \in \mathbb{N}$ dostajemy

$$\varphi^{n+p(\varepsilon)}(B_n(x,\varepsilon)) = \varphi^{p(\varepsilon)}(B(\varphi^n(x),\varepsilon)) \supseteq X_i \quad \text{jeśli} \quad \varphi^n(x) \in X_i.$$

Z tego mamy, że dla dowolnego elementu $y_i \in X_i$, $i = 1, ..., k$, zbiór $\varphi^{-(n+p(\varepsilon))}(\{y_1, ..., y_k\})$ jest (n,ε)-rozpinający. Teraz, jak wiadomo z (4.5) ciśnienie topologiczne jest dane wzorem

$$P(\varphi, \ln c) = \lim_{\varepsilon\to 0} \lim_{n\to\infty} \sup \inf_{E\subseteq X} \frac{1}{n} \ln \sum_{y\in E} \prod_{i=0}^{n-1} c(\varphi^i(x)),$$

gdzie E jest (n,ε)-rozpinający, patrz [30, 9.3]. Dla dowolnego skończonego zbioru $E \subseteq X$ i funkcji $b : X \to [0,\infty)$ mamy, że

$$\sum_{x\in E} b(x) \leqslant \sum_{x\in\varphi^{-n}(\varphi^n(E))} b(x) = \sum_{y\in\varphi^n(E)} \sum_{x\in\varphi^{-n}(y)} b(x) \leqslant |\varphi^n(E)| \sup_{y\in X} \sum_{x\in\varphi^{-n}(y)} b(x),$$

co, biorąc $b(x) = \prod_{i=0}^{n-1} c(\varphi^i(x))$ daje

$$\frac{1}{n} \ln \sum_{x\in E} \prod_{i=0}^{n-1} c(\varphi^i(x)) \leqslant \max_{y\in X} \frac{1}{n} \ln \sum_{x\in\varphi^{-1}(y)} \prod_{i=0}^{n-1} c(\varphi^i(x)) + \ln |\varphi^n(E)|^{\frac{1}{n}}.$$

Stąd, z (4.7) i z powyższego wzoru na $P(\varphi, \ln c)$, wystarczy pokazać, że dla dostatecznie małego $\varepsilon > 0$ mamy, że

$$\lim_{n\to\infty} \sup \inf_{E\subseteq X} \ln |\varphi^n(E)|^{\frac{1}{n}} = 0,$$

gdzie E jest (n,ε)-rozpinający. Biorąc $E := \varphi^{-(n+p(\varepsilon))}(\{y_1, ..., y_k\})$ dostajemy

$$\ln |\varphi^n(E)|^{\frac{1}{n}} = \ln |\varphi^n(\varphi^{-(n+p(\varepsilon))}(\{y_1, ..., y_k\}))|^{\frac{1}{n}} = \ln |\varphi^{-p(\varepsilon)}(\{y_1, ..., y_k\})|^{\frac{1}{n}},$$

co zbiega do zera, gdy $n \to \infty$. Dowodzi to twierdzenia w rozważanym przypadku.

2) Teraz załóżmy, że zbiór punktów okresowych jest gęsty w X (równoważnie $\Omega(\varphi) = X$). Z Twierdzenia 4.14 o rozkładzie spektralnym mamy, że $X = \bigsqcup_{j=1}^{N} X_j$, gdzie $X_j = \varphi^{-1}(X_j)$ są zwarte oraz odwzorowania $\varphi|_{X_j} : X_j \to X_j$ są topologicznie tranzytywne (oraz rozszerzające i otwarte). Widać więc, że $\mathcal{L}_c = \bigoplus_{j=1}^{N} \mathcal{L}_{\ln c|_{X_j}}$ oraz $\mathrm{Erg}(X,\varphi) = \bigsqcup_{j=1}^{N} \mathrm{Erg}(X_j, \varphi|_{X_j})$. Stąd, z kroku 1) dostajemy

$$r(\mathcal{L}_c) = \max_{j=1,...,N} r(\mathcal{L}_{\ln c|_{X_j}}) = \max_{\substack{\mu\in\mathrm{Erg}(X_j,\varphi)\\ j=1,...,N}} e^{\int_{X_j} \ln c d\mu + h_\mu(\varphi)} = \max_{\mu\in\mathrm{Erg}(X,\varphi)} e^{\int_X \ln c d\mu + h_\mu(\varphi)},$$

co dowodzi twierdzenia w tym przypadku.

3) Niech $\varphi : X \to X$ będzie dowolnym rozszerzającym i otwartym odwzorowaniem. Jeśli zastosujemy warunek na φ taki, że $\Omega(\varphi) = \overline{Per(\varphi)}$ (gdzie $Per(\varphi)$ to punkty okresowe odwzorowania φ), to pozostanie ono nadal rozszerzające i otwarte, patrz [24, Lemat 3.3.10]. Zatem możemy rozważać operator przejścia $\mathcal{L}_{\ln c|_{\Omega(\varphi)}} : C(\Omega(\varphi)) \to C(\Omega(\varphi))$ związany z odwzorowaniem $\varphi : \Omega(\varphi) \to \Omega(\varphi)$. Jeśli $\|\mathcal{L}^n_{\ln c|_{\Omega(\varphi)}}(1)\| \leqslant \|\mathcal{L}^n_c(1)\|$, to $r(\mathcal{L}_{\ln c|_{\Omega(\varphi)}}) \leqslant r(\mathcal{L}_c)$. Stąd, z kroku 2) i z tego, że nośnik każdej miary $\mu \in \mathcal{M}(X, \varphi)$ zawiera się w $\Omega(\varphi)$ dostajemy

$$\max_{\mu \in \mathcal{M}(X,\varphi)} e^{\int_X \ln c d\mu + h_\mu(\varphi)} = \max_{\mu \in \mathcal{M}(\Omega(\varphi),\varphi)} e^{\int_{\Omega(\varphi)} \ln c d\mu + h_\mu(\varphi)} = r(\mathcal{L}_{\ln c|_{\Omega(\varphi)}}) \leqslant r(\mathcal{L}_c),$$

więc otrzymujemy żądaną nierówność.

4) W tym przypadku będziemy rozważać c, które może posiadać wartości zerowe. Wybierzmy ściśle dodatnie i ciągłe funkcje $c_n > 0$ takie, że $c_n \searrow c$. Łatwo zauważyć, że $\{r(\mathcal{L}_{\ln c_n})\}_{n=1}^\infty$ jest ciągiem malejącym, który nie posiada większej wartości niż $r(\mathcal{L}_c)$. Zatem, z półciągłości z góry promienia spektralnego dostajemy, że $r(\mathcal{L}_{\ln c_n}) \searrow r(\mathcal{L}_c)$. Ponadto kładąc $P := \max_{\mu \in \mathcal{M}(X,\varphi)}(\int_X \ln c d\mu + h_\mu(\varphi))$ otrzymujemy, że $\max_{\mu \in \mathcal{M}(X,\varphi)}(\int_X \ln c d\mu + h_\mu(\varphi)) \searrow P$. W istocie, zbiory $A_n := \{\mu : \int_X \ln c_n d\mu + h_\mu(\varphi) \geqslant P\}$ są niepuste i malejące. Są one zwarte, ponieważ odwzorowania $\mathcal{M}(X, \varphi) \ni \mu \to \int_X \ln c_n d\mu + h_\mu(\varphi)$ są półciągłe z góry. Stąd $\bigcap_{n=1}^\infty A_n \neq \emptyset$ i dla dowolnej miary $\mu_0 \in \bigcap_{n=1}^\infty A_n$ mamy, że $P = \int_X \ln c d\mu_0 + h_{\mu_0}(\varphi)$. Teraz, z kroku 3) dostajemy

$$r(\mathcal{L}_c) = \lim_{n\to\infty} r(\mathcal{L}_{\ln c_n}) = \lim_{n\to\infty} \max_{\mu \in \mathcal{M}(X,\varphi)} e^{\int_X \ln c_n d\mu + h_\mu(\varphi)} = \max_{\mu \in \mathcal{M}(X,\varphi)} e^{\int_X \ln c d\mu + h_\mu(\varphi)}.$$

□

Bibliografia

[1] A. B. Antonevich, *Linear Functional Equations. Operator Approach*, Birkhauser Verlag, Operator Theory Advances and Applications, V. 83, 1996.

[2] A. B. Antonevich, V. I. Bakhtin, A. V. Lebedev, *On t-entropy and variational principle for the spectral radii of transfer and weighted shift operators* Ergodic Theory Dynam. Systems **31** (2011), no. 4, 995-1042.

[3] V. I. Bakhtin, A. V. Lebedev, *A new definition of t-entropy for transfer operators*, Entropy **19** (2017), 573

[4] K. Bardadyn, B. Kwaśniewski, *Spectrum of weighted isometries: C^*-algebras, transfer operators and topological pressure* arXiv:1911.04811.

[5] K. Bardadyn, B. K. Kwaśniewski, K. S. Kurnosenko, A. V. Lebedev *Formulas of t-entropy for specific class of transfer-operators*

[6] R. Bowen, *Equilibrium states and the ergodic theory of Anosov diffeomorphisms.* Lecture Notes in Mathematics, vol.470. Berlin, Heidelberg, New York: Springer 1975.

[7] A. H. Fan, Y. P. Jiang, *On Ruelle-Perron-Frobenius Operators. I. Ruelle's Theorem*, Comun. Math. Phys. 223 (2001), 125-141.

[8] M. Fichtenholz *Rachunek różniczkowy i całkowy*, Wydawnictwo Naukowe PWN SA 1999

[9] J. K. Hunter, *Measure Theory.*, Department of Mathematics, University of California at Davis,

[10] J. L. Kelley, *General Topology*, Van Nostrand, 1955.

[11] J. L. Kelley. *The Tychonoff product theorem implies the axiom of choice.* „Fundamenta Mathematicae", 1950.

[12] H. B. Keynes, J. B. Robertson, *Generators for topological entropy and expansiveness*, Math. Systems Theory **3** (1969), 51-59.

[13] A. K. Kitover *Spectrum of automorphisms with weight and the Karnowitz-Sheinberg theorem* Funct. Anal. Appl., 13 (1) 1979, 57-58

[14] A.I. Khinchine *Mathematical Foundations of Information Theory*, Dover Publications, New York,1957

[15] T. Komorowski, *THEORY OF MEASURE AND INTEGRATION* , http://hektor.umcs.lublin.pl/ komorow/mydoc/in21.pdf

[16] B. O. Koopman, *Hamiltonian Systems and Transformation in Hilbert Space*, Proceedings of the National Academy of Sciences. 17 (5) (1931), 315–318

[17] B. K. Kwaśniewski, A. V. Lebedev, *Variational principles for spectral radius of weighted endomorphisms of* $C(X, D)$, Trans. Amer. Math. Soc. 373 (2020), no. 4, 2659–2698.

[18] B. K. Kwaśniewski *On transfer operators for* C^**-dynamical systems*, Rocky J. Math. 42, No 3 (2012), 919–938.

[19] A. Lebedev *The invertibility of elemnts in the* C^* *algebras generated by dynamical systems*, Russian Math. Surveys **34**:4 (1979) 174-175[2]

[20] A. Lebedev, O. Maslak, *The spectral radius of a weighted shift operator, variational principles, entropy and topological pressure*, Spectral and evolutionary problems. Proceedings of the Eighth Crimean Autumn Mathematical School Symposium (Simferopol, 1998). Tavria Publishers, Moscow, pp. 26–34.

[21] J. H. Mai, W. H. Sun, *Positively expansive homeomorphisms on compact metric spaces* Acta Math. Hungar., **126** (4) (2010), 366-368.

[22] Onno van Gaans, *Probability measures on metric spaces*, https://www.math.leidenuniv.nl/ vangaans/jancol1.pdf

[23] K.R. Parthasarathy, *Probability measures on metric spaces*, Academic Press, New York, 1967.

[24] F. Przytycki, M. Urbański, *Conformal Fractals: Ergodic Theory Methods*, London Mathematical Society Lecture Note Series **371**, Cambridge University Press, 2010.

[25] W. L. Reddy, *Expanding maps on compact metric spaces*, Topology Appl. **13** (1982), no. 3, 327–334.

[26] W. Rudin *Analiza rzeczywista i zespolona*, Warszawa, 2, 2009.

[27] D. Ruelle, *Thermodynamic formalism.* Encyclopedia of Math. and its Appl., vol.5, Reading, Mass: Addison-Wesley 1978.

[28] D. Ruelle, *The thermodynamic formalism for expanding maps*, Comm. Math. Phys. **125** (1989), 239-262.

[29] P.M Sołtan, *Elementy teorii operatorów na przestrzeni Hilberta*, Wydawnictwo Uniwersytetu Warszawskiego, 2017.

[30] P. Walters, *An introduction to Ergodic Theory.* Springer-Verlag 1982

[31] P. Walters, *Invariant measures and equilibrium states for some mappings which expand distances*, Trans. Am. Math. Soc. **236** (1978), 121–153.