MAREK BALCERZAK

PODSTAWY TEORII ERGODYCZNEJ

Politechnika Łódzka

Łódź 2022

Marek Balcerzak

Podstawy teorii ergodycznej

Politechnika Łódzka, 2022

Recenzent: dr hab. Andrzej Biś, prof. UŁ

ISBN 978-83-66741-62-1

DOI 10.34658/9788366741621

WYDAWNICTWO POLITECHNIKI ŁÓDZKIEJ
90-924 Łódź, ul. Wólczańska 223
tel. 42 631-29-52, 42 631-20-87
fax 42 631-25-38
e-mail: zamowienia@info.p.lodz.pl
www.wydawnictwo.p.lodz.pl

Spis treści

Wstęp

Teoria ergodyczna ma swoje korzenie w zagadnieniach fizycznych związanych z mechaniką statystyczną (por. [5]). Jednakże z czasem ustanowiła ona samodzielny kierunek badań, który przeżywa duży rozwój w ostatnich latach z uwagi na istotne zastosowania w innych działach matematyki. Odnotujmy, że w 2020 roku dwaj wybitni matematycy, Hillel Furstenberg i Gregory Margulis, otrzymali prestiżową nagrodę Abela w uznaniu za nowatorskie wykorzystanie metod teorii ergodycznej w teorii grup, teorii liczb i kombinatoryce.

Skrypt ogranicza się do podstawowej wiedzy z teorii ergodycznej. Tematem rozdziału 1. są odwzorowania zachowujące miarę, będące podstawowymi obiektami badań. Przedstawiono przykłady i proste własności takich odwzorowań. W rozdziale 2. omówiono twierdzenie Poincarégo o powracaniu opisujące pierwszy znaczący rezultat teorii ergodycznej. Ważnym typem odwzorowań zachowujących miarę są przekształcenia ergodyczne, którym poświęcono rozdział 3. Klasyczne twierdzenie ergodyczne Birkhoffa jest jednym z najważniejszych w tej teorii. Jego nietrywialny dowód, oparty na maksymalnym twierdzeniu ergodycznym, został przedstawiony w rozdziale 4. Jednym z wniosków jest twierdzenie ergodyczne von Neumanna, dotyczące przestrzeni L^p, gdy $1 \leqslant p < \infty$. Naturalnymi podklasami odwzorowań ergodycznych są odwzorowania mieszające i słabo mieszające, które zostały opisane w rozdziale 5., gdzie przedstawiono ciekawe charakteryzacje. Inne charakteryzacje tych odwzorowań, odwołujące się do zespolonej przestrzeni Hilberta L^2, omówiono w rozdziale 6. Odwzorowania indukowane przez czas pierwszego powrotu są tematem rozdziału 7. Ostatni rozdział 8. dotyczy izomorfizmów odwzorowań zachowujących miarę.

W skrypcie zabrakło miejsca na istotne pojęcie teorii ergodycznej, jakim jest entropia. Jednakże warto zachęcić Czytelnika do zapoznania się z materiałem na ten temat zawartym w źródłach pochodzących z bibliografii.

Teoria ergodyczna bywa traktowana jako część teorii układów dynamicznych dyskretnych, dotycząca odwzorowań zachowujących miarę. Topologiczne układy dynamiczne dyskretne stanowią zasadniczą część tej teorii, której metody są w pewnym zakresie analogiczne (por. [11], [13]). Skrypt może więc zainspirować Czytelnika do dalszego studiowania tego ważnego działu matematyki.

Autor serdecznie dziękuje Recenzentowi skryptu, Prof. Andrzejowi Bisiowi, za uważne przeczytanie tekstu, cenne uwagi i sugestie, a także za wskazanie dodatkowych pozycji bibliograficznych. Wśród nich jest ciekawa publikacja [8] pokazująca związki teorii ergodycznej z hipotezą Sarnaka.

Rozdział 1

Odwzorowania zachowujące miarę

Teoria ergodyczna powstała jako konsekwencja badań w zakresie statystycznego podejścia do zjawisk termodynamicznych zachodzących w gazach. Badania te zawdzięczamy austriackiemu fizykowi Boltzmannowi, który zakładał, że cząsteczki ciała znajdują się w nieustannym ruchu, a ich prędkości zależą od temperatury. Podstawowym pojęciem w teorii ergodycznej są odwzorowania zachowujące miarę.

Definicja 1.1. Niech $(X_1, \mathcal{A}_1, \mu_1)$, $(X_2, \mathcal{A}_2, \mu_2)$ będą przestrzeniami z miarą. Niech $T\colon X_1 \to X_2$. Mówimy, że odwzorowanie T:

(a) jest odwzorowaniem mierzalnym, gdy $T^{-1}[B] \in \mathcal{A}_1$ dla każdego $B \in \mathcal{A}_2$;

(b) zachowuje miarę, gdy T jest mierzalne oraz $\mu_1(T^{-1}[B]) = \mu_2(B)$ dla każdego $B \in \mathcal{A}_2$;

(c) zachowujące miarę jest odwracalne, gdy T jest bijekcją oraz T^{-1} zachowuje miarę.

Uwaga 1.1.

- Jeśli odwzorowania $T_1\colon X_1 \to X_2$ oraz $T_2\colon X_2 \to X_3$ zachowują miarę, to $T_2 \circ T_1\colon X_1 \to X_3$ też zachowuje miarę (wykazać).
- Przeważnie zakłada się, że $(X_1, \mathcal{A}_1, \mu_1) = (X_2, \mathcal{A}_2, \mu_2)$ i wtedy rozważa się iteracje
$$T^n := T \circ T \circ \cdots \circ T \quad (n \text{ razy}) \quad \text{dla } n \in \mathbb{N},$$
odwzorowania T zachowującego miarę. Wtedy każde odwzorowanie T^n zachowuje miarę. Często zakłada się, że rozważana miara jest skończona lub probabilistyczna.

Definicja 1.2. Rodzina $\mathcal{C} \subseteq \mathcal{P}(X)$ nazywa się klasą monotoniczną, gdy jest zamknięta względem operacji przecięcia malejącego ciągu zbiorów oraz operacji sumy rosnącego ciągu zbiorów.

Ćwiczenie 1.1. Wykazać, że najmniejsza klasa monotoniczna $\mathcal{M}(\mathcal{A})$ zawierająca ustalone ciało zbiorów $\mathcal{A} \subset \mathcal{P}(X)$ jest równa $\sigma(\mathcal{A})$, σ-ciału generowanemu przez $\mathcal{A}$; zob. [12, s. 85].

Twierdzenie 1.1. *Niech $(X_i, \mathcal{A}_i, \mu_i)$ (dla $i = 1, 2$) będą przestrzeniami z miarą skończoną oraz $T\colon X_1 \to X_2$. Niech $\mathcal{A} \subset \mathcal{A}_2$ będzie ciałem takim, że $\mathcal{A}_2 = \sigma(\mathcal{A})$ oraz dla każdego $B \in \mathcal{A}$ zachodzi*

$$T^{-1}[B] \in \mathcal{A}_1 \quad \textit{oraz} \quad \mu_1(T^{-1}[B]) = \mu_2(B).$$

Wtedy T zachowuje miarę.

Dowód. Niech

$$\mathcal{C} := \{B \in \mathcal{A}_2\colon T^{-1}[B] \in \mathcal{A}_1 \;\; \wedge \;\; \mu_1(T^{-1}[B]) = \mu_2(B)\}.$$

Z założenia mamy $\mathcal{A} \subset \mathcal{C}$. Ponadto $\mathcal{C}$ jest klasą monotoniczną. Faktycznie rozważmy ciąg zbiorów $C_1 \supset C_2 \supset \dots$, gdzie $C_n \in \mathcal{C}$ dla $n \in \mathbb{N}$. Wtedy

$$T^{-1}[C_1] \supset T^{-1}[C_2] \supset \dots \text{ oraz } T^{-1}[C_n] \in \mathcal{A}_1,\; \mu_1(T^{-1}[C_n]) = \mu_2(C_n) \text{ dla } \; n \in \mathbb{N}.$$

Zatem dla $C := \bigcap_{n\in\mathbb{N}} C_n$ otrzymujemy $T^{-1}[C] = \bigcap_{n\in\mathbb{N}} T^{-1}[C_n] \in \mathcal{A}_1$, oraz

$$\mu_1(T^{-1}[C]) = \lim_n \mu_1(T^{-1}[C_n]) = \lim_n \mu_2(C_n) = \mu_2(\bigcap_{n\in\mathbb{N}} C_n) = \mu_2(C).$$

Stąd $C \in \mathcal{C}$. Analogicznie pokazujemy, że jeśli $C_1 \subset C_2 \subset \dots$, gdzie $C_n \in \mathcal{C}$ dla $n \in \mathbb{N}$, to $D := \bigcup_{n\in\mathbb{N}} C_n \in \mathcal{C}$. Zatem na mocy Ćwiczenia 1.1 mamy $\mathcal{A}_2 = \sigma(\mathcal{A}) = \mathcal{M}(\mathcal{A}) \subset \mathcal{C}$, co daje tezę. □

Przykłady 1.1.

1. Identyczność $\mathrm{id}\colon X \to X$ zachowuje miarę dowolnej przestrzeni z miarą $(X, \mathcal{A}, \mu)$.

2. Funkcje $T_c\colon \mathbb{R} \to \mathbb{R}$, $T(x) = x + c$ ($c \in \mathbb{R}$) oraz $T_*\colon \mathbb{R} \to \mathbb{R}$, $T(x) = -x$, zachowują miarę przestrzeni $(\mathbb{R}, \mathcal{L}, \lambda)$, gdzie $\mathcal{L}$ oznacza σ-ciało zbiorów mierzalnych w sensie Lebesgue'a, zaś λ – miarę Lebesgue'a.

3. Niech $T\colon X_1 \to X_2$.
   (a) Jeśli $(X_2, \mathcal{A}_2, \mu_2)$ jest przestrzenią z miarą, to istnieje σ-ciało $\mathcal{A}_1$ podzbiorów X_1 i miara μ_1 na $\mathcal{A}_1$, dla których T zachowuje miarę. Istotnie niech $\mathcal{A}_1 := \{T^{-1}[A]\colon A \in \mathcal{A}_2\}$ oraz $\mu_1(T^{-1}[A]) := \mu_2(A)$ dla $A \in \mathcal{A}_2$.
   (b) Jeśli $(X_1, \mathcal{A}_1, \mu_1)$ jest przestrzenią z miarą, to istnieje σ-ciało $\mathcal{A}_2$ podzbiorów X_2 i miara μ_2 na $\mathcal{A}_2$, dla których T zachowuje miarę. Rzeczywiście niech $\mathcal{A}_2 := \{A \subset X_2\colon T^{-1}[A] \in \mathcal{A}_1\}$ oraz $\mu_2(A) := \mu_1(T^{-1}[A])$ dla $A \in \mathcal{A}_2$.

4. Dla $x \in \mathbb{R}$ niech
$$x \ (\mathrm{mod}\, 1) := x - \lfloor x \rfloor;$$
jest to część ułamkowa liczby x. Ustalmy liczbę $\beta \in [0,1)$ i rozważmy odwzorowanie $R_\beta \colon [0,1) \to [0,1)$ dane wzorem
$$R_\beta(x) := x + \beta \ (\mathrm{mod}\, 1) = \begin{cases} x+\beta & \text{dla} \quad x+\beta < 1; \\ x+\beta-1 & \text{dla} \quad x+\beta \geqslant 1. \end{cases}$$
Na $[0,1)$ rozważmy (odpowiednio obcięte) σ-ciało $\mathcal{L}$ zbiorów mierzalnych w sensie Lebesgue'a oraz miarę Lebesgue'a λ. Zauważmy, że R_β jest bijekcją oraz $(R_\beta)^{-1}(x) = x + 1 - \beta \ (\mathrm{mod}\, 1)$.

Pokażemy, że R_β zachowuje miarę. Niech $A \in \mathcal{L}$. Trzeba pokazać, że $R_\beta^{-1}[A] \in \mathcal{L}$ oraz $\lambda(R_\beta^{-1}[A]) = \lambda(A)$, czyli $R_{1-\beta}[A] \in \mathcal{L}$ oraz $\lambda(R_{1-\beta}[A]) = \lambda[A]$. Oznaczmy $\gamma := 1 - \beta$. Niech
$$A_0 := A \cap [0, 1-\gamma), \quad A_1 := A \cap [1-\gamma, 1).$$
Wtedy $A = A_0 \sqcup A_1$ (symbol $\sqcup$ oznacza sumę rozłączną), $R_\gamma[A] = (A_0 + \gamma) \sqcup (A_1 + \gamma - 1) \in \mathcal{L}$ oraz $A_0 + \gamma \subset [\gamma, 1)$, $A_1 + \gamma - 1 \subset [0, \gamma)$, więc
$$\lambda(R_\gamma[A]) = \lambda(A_0 + \gamma) + \lambda(A_1 + \gamma - 1) = \lambda(A_0) + \lambda(A_1) = \lambda(A).$$
Odwzorowanie R_β jest izomorficzne z obrotem $R_\beta^* \colon K \to K$ okręgu jednostkowego $K := \{z \in \mathbb{C} \colon |z| = 1\}$, gdzie $R_\beta^*(z) := \exp(i(\operatorname{Arg} z + 2\pi\beta))$ dla $z \in K$.

5. Rozważmy odwzorowanie Boole'a $T \colon \mathbb{R} \to \mathbb{R}$ dane wzorem
$$T(x) := \begin{cases} x - \frac{1}{x} & \text{dla} \quad x \neq 0; \\ 0 & \text{dla} \quad x = 0. \end{cases}$$
Zauważmy, że T nie jest odwracalne.

Pokażemy, że T zachowuje miarę Lebesgue'a na σ-ciele zbiorów mierzalnych. W tym celu zbadajmy najpierw przeciwobraz odwzorowania T dla przedziału ograniczonego $(a,b]$. Rozważmy trzy przypadki:
1^0 – gdy $a > 0$; 2^0 – gdy $b \leqslant 0$ (w obu przypadkach otrzymujemy sumę rozłącznych przedziałów o łącznej długości $b - a$) oraz 3^0 – gdy $a < 0 < b$ (wtedy $(a,b] = (a,0] \cup (0,b]$ i korzystamy z poprzednich przypadków).

Dla $n \in \mathbb{N}$ niech $Y_n := (-n, n]$ oraz $X_n := T^{-1}[Y_n]$. Można sprawdzić, że
$$X_n = \left(\frac{-n - \sqrt{n^2+4}}{2}, \frac{n - \sqrt{n^2+4}}{2} \right] \cup \left(\frac{-n + \sqrt{n^2+4}}{2}, \frac{n + \sqrt{n^2+4}}{2} \right] \cup \{0\}.$$

Ponadto dla $n \in \mathbb{N}$ oznaczmy $T_n := T|_{X_n}$. Zauważmy, że $T_n[X_n] = Y_n$ oraz T_n zachowuje miarę Lebesgue'a na zbiorach borelowskich jako odwzorowanie $T_n \colon X_n \to Y_n$. Istotnie, σ-ciało zbiorów borelowskich w Y_n jest generowane przez ciało złożone ze skończonych sum przedziałów postaci $(a, b]$ zawartych Y_n. Wystarczy więc zastosować poprzednią obserwację i Twierdzenie 1.1.

Każdy zbiór borelowski B w $\mathbb{R}$ jest sumą rosnącego ciągu zbiorów $B_n := B \cap Y_n$, $n \in \mathbb{N}$, borelowskich w Y_n. Stąd z własności miary na ciągu rosnącym zbiorów mierzalnych wnosimy, że T zachowuje miarę Lebesgue'a na zbiorach borelowskich w $\mathbb{R}$.

Ponieważ każdy zbiór mierzalny w sensie Lebesgue'a jest sumą zbioru borelowskiego (typu F_σ) i zbioru miary zero, więc dla zakończenia dowodu wystarczy zauważyć, że przeciwobraz $T^{-1}[E]$ zbioru E miary zero jest miary zero. Wynika to z faktu, że jeśli E jest pokryty przez sumę $\bigcup_n I_n$ przedziałów I_n takich że $\sum_n \lambda(I_n) < \varepsilon$, gdzie $\varepsilon > 0$ jest dowolną ustaloną liczbą, to $T^{-1}[E]$ jest pokryty przez $\bigcup_n T^{-1}[I_n]$, gdzie $\lambda(T^{-1}[I_n]) = \lambda(I_n)$ dla $n \in \mathbb{N}$, co dowodzi, że $T^{-1}[E]$ jest miary zero.

6. Odwzorowanie podwajające $T \colon [0, 1) \to [0, 1)$ dane wzorem

$$T(x) := 2x \ (\mathrm{mod}\, 1) = \begin{cases} 2x & \text{dla} \quad x \in [0, 1/2)\,; \\ 2x - 1 & \text{dla} \quad x \in [1/2, 1] \end{cases}$$

zachowuje miarę Lebesgue'a (ale nie jest odwracalne). Istotnie niech $A \subset [0, 1)$, $A \in \mathcal{L}$. Niech $S_1 \colon [0, 1) \to [0, 1/2)$, $S_1(y) := y/2$ oraz $S_2 \colon [0, 1) \to [1/2, 1)$, $S_2(y) := y/2 + 1/2$. Mamy $T^{-1}[A] = S_1[A] \sqcup S_2[A]$ oraz $S_1[A] = (1/2)A \in \mathcal{L}$, $S_2[A] = (1/2)A + 1/2 \in \mathcal{L}$ (wykonać rysunek), więc

$$\lambda(T^{-1}[A]) = \lambda((1/2)A) + \lambda((1/2)A + 1/2) = (1/2)\lambda(A) + (1/2)\lambda(A) = \lambda(A).$$

Odwzorowanie to jest izomorficzne z odwzorowaniem okręgu jednostkowego K na siebie
$T_2 \colon K \to K$ danego wzorem $T_2(z) := z^2$ dla $z \in K$. Podobnie pokazuje się, że $T_n(z) := z^n$ zachowuje miarę, gdy $n \in \mathbb{N}$ (lub ogólniej gdy $n \in \mathbb{Z}$, $n \neq 0$).

7. Odwzorowanie piekarza $T \colon [0, 1]^2 \to [0, 1]^2$ dane jest wzorem

$$T(x, y) := \begin{cases} (2x, y/2) & \text{dla} \quad x \in [0, 1/2)\,; \\ (2x - 1, (y + 1)/2) & \text{dla} \quad x \in [1/2, 1]\,. \end{cases}$$

Jest ono bijekcją. Odwzorowanie T najpierw „ściska" kwadrat $[0,1]^2$, przekształcając go na prostokąt $P := [0,2] \times [0,1/2]$, a potem „odcina" prawą połowę prostokąta P i „nakłada" ją na jego lewą połowę. Odwzorowanie T^{-1} przedstawia odwrotny proces i dane jest wzorem

$$T^{-1}(x,y) := \begin{cases} (x/2, 2y) & \text{dla} \quad y \in [0,1/2)\,; \\ ((x+1)/2, 2y-1) & \text{dla} \quad y \in [1/2,1]\,. \end{cases}$$

W podobny sposób zauważamy, że T oraz T^{-1} przekształcają dowolny prostokąt $P = [a,b] \times [c,d] \subset [0,1]^2$ na sumę rozłącznych prostokątów o łącznej mierze (dwuwymiarowej Lebesgue'a) równej polu prostokąta P (szczegóły rachunkowe pomijamy). Stąd można wywnioskować, że T oraz T^{-1} zachowują miarę Lebesgue'a. Faktycznie, z poprzedniej własności wynika, że dla dowolnego zbioru $A \subset [0,1]^2$ miara zewnętrzna Lebesgue'a każdego ze zbiorów $T[A]$ oraz $T^{-1}[A]$ jest taka sama jak miara zewnętrzna zbioru A.

8. Niech $Y := \{0,1,\dots,k-1\}$ dla $k \in \mathbb{N}$. Definiujemy miarę prawdopodobieństwa na $\mathcal{P}(Y)$ w ten sposób, że każdemu singletonowi $\{y\}$ (dla $y \in Y$) przypisujemy wartość miary $p_i > 0$, przy czym $\sum_{i=0}^{k-1} p_i = 1$. Niech $X := \prod_{n\in\mathbb{Z}} Y_n$, gdzie $Y_n = Y$ dla $n \in \mathbb{Z}$. Wyposażamy Y w σ-ciało produktowe $\mathcal{A}$ i miarę produktową μ. Niech $T\colon X \to X$ będzie dwustronnym przesunięciem (shiftem), które definiujemy wzorem $T((x_i)_{i\in\mathbb{Z}}) := (x_{i+1})_{i\in\mathbb{Z}}$. Wówczas T zachowuje miarę, bo (jak pokażemy poniżej) T zachowuje miarę zbiorów cylindrycznych postaci

$$E := \{(x_i)\colon x_{i_1} = a_1, \dots, x_{i_n} = a_n\},$$

gdzie $i_1 < \dots < i_n$ oraz $i_j \in \mathbb{Z}$, $a_i \in Y$ dla $j = 1,\dots,n$. Ponadto skończone (rozłączne) sumy takich zbiorów tworzą ciało, które generuje σ-ciało $\mathcal{A}$, zatem wystarczy zastosować Twierdzenie 1.1. Rozważmy więc dowolny zbiór cylindryczny E opisanej postaci. Wtedy $T^{-1}[E] = \{(x_i)\colon x_{i_1+1} = a_1, \dots x_{i_n+1} = a_n\}$, więc $\mu(T^{-1}[E]) = \mu(E)$.

9. Rozważmy przykład podobny do poprzedniego. Zbiór Y jest taki jak wyżej z odpowiednio zdefiniowaną miarą na $\mathcal{P}(Y)$. Niech $Y_n := Y$ dla całkowitych $n \geqslant 0$ oraz na przestrzeni $X := \prod_{n\geqslant 0} Y_n$ rozważmy odpowiednie σ-ciało produktowe i miarę produktową. Odwzorowanie $T\colon X \to X$ jest zdefiniowane wzorem $T((x_i)_{i\geqslant 0} := (x_{i+1})_{i\geqslant 0}$; jest to przesunięcie jednostronne. Podobnie jak poprzednio pokazujemy, że T zachowuje miarę. Gdy $Y = \{0,1\}$ oraz $p_0 = p_1 = 1/2$, dowodzi się, że T jest izomorficzne z odwzorowaniem podwajającym.

10. Odwzorowanie Gaussa $T\colon [0,1) \to [0,1)$ dane jest wzorem

$$T(x) := \begin{cases} 1/x - \lfloor 1/x \rfloor & \text{dla} \quad x \neq 0; \\ 0 & \text{dla} \quad x = 0. \end{cases}$$

Zauważmy, że $T(x) = 1/x - n$, gdy $x \in (1/(n+1), 1/n)$. Ponadto

$$T^{-1}[\{y\}] = \{1/(n+y)\colon n \in \mathbb{N}\}, \quad \text{gdy} \quad y \in (0,1) \setminus \{1/n\colon n \geqslant 2\}$$

oraz $T^{-1}[\{0\}] = \{0\} \cup \{1/n\colon n \geqslant 2\}$. Rozważmy miarę borelowską ν na $[0,1)$ daną wzorem

$$\nu(A) := \frac{1}{\ln 2} \int_A \frac{1}{x+1} d\,\lambda(x).$$

Wówczas T zachowuje miarę. Szczegóły można znaleźć w [11, s. 152–153].

Nieformalnie używaliśmy pojęcia izomorficznych odwzorowań zachowujących miarę w dwóch różnych przestrzeniach. Teraz doprecyzujemy to pojęcie, a później do niego wrócimy, wyjaśniając zasygnalizowane przykłady izomorfizmów.

Definicja 1.3. Jeśli dana jest przestrzeń z miarą $(X, \mathcal{S}, \mu)$ oraz odwzorowanie $T\colon X \to X$ zachowujące miarę, to czwórkę $(X, \mathcal{S}, \mu, T)$ nazywamy układem dynamicznym zachowującym miarę.

Definicja 1.4. Niech dane będą dane dwa układy dynamiczne $(X, \mathcal{A}, \mu, T)$ oraz $(Y, \mathcal{B}, \nu, S)$ na przestrzeniach probabilistycznych. Mówimy, że są one izomorficzne, gdy istnieją zbiory $X' \in \mathcal{A}$, $Y' \in \mathcal{B}$ takie, że $\mu(X') = 1$, $\nu(Y') = 1$, $T[X'] \subset X'$, $S[Y'] \subset Y'$ oraz odwracalne odwzorowanie $\varphi\colon X' \to Y'$ zachowujące miarę takie że $(\varphi \circ T)(x) = (S \circ \varphi)(x)$ dla każdego $x \in X'$.

Rozdział 2

Twierdzenie Poincarégo o powracaniu

Ustalmy przestrzeń z miarą $(X, \mathcal{S}, \mu)$.

Definicja 2.1. Niech $T\colon X \to X$ będzie odwzorowaniem mierzalnym. Mówimy, że T jest *rekurencyjne*, gdy dla każdego zbioru $A \in \mathcal{S}$ miary $\mu(A) > 0$ istnieje zbiór $M \subset A$ miary zero taki, że

$$(\forall\, x \in A \setminus M)\ (\exists n = n(x) \in \mathbb{N}) \quad T^n(x) \in A. \tag{2.1}$$

Kolejne liczby naturalne $n \in \mathbb{N}$ zliczają upływ czasu (w ustalonej jednostce), zaś T^n opisuje stan danego procesu w chwili n. Powyższy warunek mówi o tym, że dla dowolnego fragmentu A miary dodatniej danej przestrzeni, po pewnym czasie prawie każdy punkt zbioru A ponownie znajdzie się w tym zbiorze, czyli wraca do stanu pierwotnego.

Następujące twierdzenie uważane jest jako historycznie pierwszy rezultat teorii ergodycznej. Został on opublikowany w 1899 roku w pracy Poincarégo dotyczącej mechaniki nieba. Nazywany jest on twierdzeniem Poincarégo o powracaniu (lub o rekurencji).

Twierdzenie 2.1. *Niech $(X, \mathcal{S}, \mu)$ będzie przestrzenią z miarą skończoną. Wówczas każde odwzorowanie $T\colon X \to X$ zachowujące miarę jest rekurencyjne. Co więcej, dla każdego zbioru $A \in \mathcal{S}$ miary $\mu(A) > 0$ istnieje taki zbiór $M \subset A$ miary zero, że każdy punkt $x \in A \setminus M$ wraca do zbioru A nieskończenie wiele razy, tzn. istnieje taki ciąg liczb naturalnych $m_1 < m_2 < \dots$, że $T^{m_j}(x) \in A$ dla każdego $j \in \mathbb{N}$.*

Dla $T\colon X \to X$ oraz zbioru $A \subset X$ wprowadzamy oznaczenia $T^{-n}[A] := (T^n)^{-1}[A]$, gdy $n \in \mathbb{N}$, oraz $T^0 := \mathrm{id}_X$. Wzmocnienie rekurencyjności odwzorowania T w tezie powyższego twierdzenia jest pozorne, gdyż jest ono równoważne rekurencyjności. Pokazuje to następujący lemat.

Lemat 2.1. *Jeśli odwzorowanie $T\colon X \to X$ zachowujące miarę jest rekurencyjne, to dla każdego zbioru $A \in \mathcal{S}$ miary $\mu(A) > 0$ i dla prawie każdego punktu x zbioru A istnieje taki ciąg liczb naturalnych $m_1 < m_2 < \dots$, że $T^{m_j}(x) \in A$ dla wszystkich $j \in \mathbb{N}$.*

Dowód. Niech $A \in \mathcal{S}$ oraz $\mu(A) > 0$. Na mocy założenia dobierzmy taki zbiór $M_1 \subset A$ miary zero, że dla każdego $x \in A \setminus M_1$ da się dobrać liczbę naturalną $n = n(x) \in \mathbb{N}$, dla którego $T^n(x) \in A$. Niech $M := \bigcup_{j=0}^{\infty} T^{-j}[M_1]$. Mamy $\mu(T^{-j}[M_1]) = \mu(M_1) = 0$ dla $j \geqslant 0$, zatem $\mu(M) = 0$. Dla każdego $k \geqslant 0$ mamy

$$T^{-k}[M] = \bigcup_{j=0}^{\infty} T^{-(j+k)}[M_1] \subset M.$$

W konsekwencji jeśli $x \notin M$, to $T^k(x) \notin M$.

Ustalmy punkt $x \in A \setminus M$. Z powyższej obserwacji mamy $T^{n(x)}(x) \in A \setminus M$, gdzie liczba naturalna $n(x)$ jest wybrana jak we wzorze (2.1). Miech $m_1 := n(x)$. Zastosujmy teraz powtórnie powyższą obserwację do punktu $z := T^{m_1}(x) \in A \setminus M$. Wtedy znajdziemy taką liczbę naturalną $n_1 = n(z) \in \mathbb{N}$, że $T^{n_1}(z) \in A \setminus M$. Połóżmy $m_2 = m_1 + n_1$. Wówczas $T^{m_2}(x) = T^{n_1}(z) \in A \setminus M$. Dalej powtarzamy tę samą procedurę, biorąc punkt $z' = T^{m_2}(x)$. Postępując indukcyjnie, otrzymamy żądany ciąg rosnący liczb naturalnych (m_i). $\square$

Uwaga 2.1. Definicję odwzorowania rekurencyjnego można sformułować krótko poprzez warunek:

dla każdego zbioru $A \in \mathcal{S}$ zachodzi $\mu(A \setminus \bigcup_{n=1}^{\infty} T^{-n}[A]) = 0$.

Równoważność powyższej definicji z oryginalnym warunkiem definicyjnym jest łatwym ćwiczeniem.

Definicja 2.2. Odwzorowanie mierzalne $T\colon X \to X$ nazywa się konserwatywne, gdy dla dowolnego zbioru $A \in \mathcal{S}$ miary $\mu(A) > 0$ istnieje taka liczba naturalna $n \in \mathbb{N}$, że $\mu(A \cap T^{-n}[A]) > 0$.

Twierdzenie 2.2. *Odwzorowanie mierzalne jest rekurencyjne wtedy i tylko wtedy, gdy jest konserwatywne.*

Dowód. Konieczność. Niech $A \in \mathcal{S}$, $\mu(A) > 0$. Skoro T jest rekurencyjne, to na mocy Uwagi 2.1 mamy $\mu(A \setminus \bigcup_{n=1}^{\infty} T^{-n}[A]) = 0$. Stąd $\mu(A \cap \bigcup_{n=1}^{\infty} T^{-n}[A]) > 0$. Zatem $\mu(A \cap T^{-n}[A]) > 0$ dla pewnego $n \in \mathbb{N}$.

Dostateczność. Niech $A \in \mathcal{S}$, $\mu(A) > 0$. Połóżmy $M := A \setminus \bigcup_{n=1}^{\infty} T^{-n}[A]$. Gdyby $\mu(M) > 0$, to z założenia mamy $\mu(M \cap T^{-n}[M]) > 0$ dla pewnego n. Zatem istnieje taki punkt $x \in M \subset A$, że $T^n(x) \in M \subset A$ wbrew definicji M. $\square$

Przeprowadzimy teraz dowód twierdzenia Poincarégo.

Dowód. Na mocy poprzedniego twierdzenia wystarczy wykazać, że odwzorowanie T jest konserwatywne. Niech $A \in \mathcal{S}$, $\mu(A) > 0$. Przypuśćmy, że $\mu(A \cap T^{-n}[A]) = 0$ dla wszystkich $n \in \mathbb{N}$. Weźmy dowolne $k, l \in \mathbb{N}$, $k \neq l$. Załóżmy np., że $l > k$ i oznaczmy $n := l - k$. Wtedy

$$\mu(T^{-l}[A] \cap T^{-k}[A]) = \mu(T^{-n-k}[A] \cap T^{-k}[A]) = \mu(T^{-k}(T^{-n}[A] \cap A)) = \mu(T^{-n}[A] \cap A) = 0.$$

Zatem

$$\mu\left(\bigcup_{n=1}^{\infty} T^{-n}[A]\right) = \sum_{n=1}^{\infty} \mu(T^{-n}[A]) = \sum_{n=1}^{\infty} \mu(A) = \infty,$$

co daje sprzeczność z założeniem, że $\mu(X) < \infty$. □

Założenie o skończoności miary w twierdzeniu Poincarégo jest istotne, bo przesunięcie $T \colon \mathbb{R} \to \mathbb{R}$, $T(x) := x + 1$, zachowuje miarę, ale nie jest rekurencyjne (np. dla $A = [0, 1)$).

Pokażemy jeszcze jedną charakteryzację odwzorowań rekurencyjnych.

Definicja 2.3. Zbiór $C \in \mathcal{S}$ nazywa się ściśliwy względem odwzorowania mierzalnego $T \colon X \to X$, gdy $T^{-1}[C] \subset C$ oraz $\mu(C \setminus T^{-1}[C]) > 0$. Odwzorowanie T nazywa się nieściśliwe, gdy nie ma zbiorów ściśliwych.

Twierdzenie 2.3. *Odwzorowanie mierzalne $T \colon X \to X$ jest rekurencyjne wtedy i tylko wtedy, gdy jest nieściśliwe.*

Dowód. Konieczność. Załóżmy, że T jest rekurencyjne i niech zbiór $C \in \mathcal{S}$ będzie taki, że $T^{-1}[C] \subset C$. Połóżmy $A := C \setminus T^{-1}[C]$. Pokażemy, że

$$T^{-n}[A] \cap A = \emptyset \quad \text{dla każdego} \quad n \in \mathbb{N}. \tag{2.2}$$

Niech $x \in A$. Wtedy $x \in C$ oraz $T(x) \notin C$. Zatem $T(x) \notin T^{-1}[C]$ (bo $T^{-1}[C] \subset C$), czyli $T^2(x) \notin C$. Dalej przez łatwą indukcję dostajemy $T^n(x) \notin C$ dla każdego $n \in \mathbb{N}$. Zatem także $T^n(x) \notin A$ dla każdego n. To daje żądany warunek (2.2). Stąd $\mu(A) = \mu(A \setminus \bigcup_{n=1}^{\infty} T^{-n}[A]) = 0$ na mocy rekurencyjności T. Wobec tego T jest nieściśliwe.

Dostateczność. Załóżmy, że T jest nieściśliwe i niech $A \in \mathcal{S}$, $\mu(A) > 0$. Połóżmy $C := \bigcup_{n=0}^{\infty} T^{-n}[A]$. Wtedy

$$T^{-1}[C] = \bigcup_{n=1}^{\infty} T^{-n}[A] \subset C$$

oraz

$$C \setminus T^{-1}[C] = \bigcup_{n=0}^{\infty} T^{-n}[A] \setminus \bigcup_{n=1}^{\infty} T^{-n}[A] = A \setminus \bigcup_{n=1}^{\infty} T^{-n}[A].$$

Skoro T jest nieściśliwe, to wynika stąd, że $\mu(A \setminus \bigcup_{n=1}^{\infty} T^{-n}[A]) = 0$, co oznacza, że T jest rekurencyjne (por. Uwagę 2.1). $\square$

Z tego twierdzenia wynika, że dowolne odwzorowanie rekurencyjne T zachowuje miarę każdego zbioru $C \in \mathcal{S}$ spełniającego warunek $T^{-1}[C] \subset C$.

Dowód twierdzenia Poincarégo można zredagować nieco inaczej bez wprowadzania pojęć pomocniczych. Oto taki dowód:

Dowód. [1] Niech $A \in \mathcal{S}$ oraz $\mu(A) > 0$. Dla każdego $n \in \mathbb{N}$ zdefiniujmy zbiór

$$A_n := \{x \in A \colon (\forall k \in \mathbb{N})\ x \notin T^{-kn}[A]\}.$$

Zauważmy, że dla dowolnych $n, k \in \mathbb{N}$ zbiory

$$A_n,\ T^{-n}[A_n],\ \ldots,\ T^{-kn}[A_n] \tag{2.3}$$

są parami rozłączne. Faktycznie, pierwszy zbiór jest rozłączny z pozostałymi, bo dla $n, k \in \mathbb{N}$ mamy

$$A_n \cap T^{-kn}[A_n] \subset A_n \cap T^{-kn}[A] = \emptyset$$

na mocy definicji A_n. Następnie przypuśćmy, że dla pewnych $k, m, n \in \mathbb{N}$ zachodzi

$$T^{-kn}[A_n] \cap T^{-(k+m)n}[A_n] \neq \emptyset. \tag{2.4}$$

To implikuje łatwo $A_n \cap T^{-mn}[A_n] \neq \emptyset$. (Gdyby ten zbiór był pusty, to jego przeciwobraz dla odwzorowania T^{kn} też byłby pusty, co przeczy (2.4).) Jednak w myśl poprzedniej uwagi nie zachodzi warunek $A_n \cap T^{-mn}[A_n] \neq \emptyset$ nie zachodzi, co daje sprzeczność.

Ponieważ odwzorowanie T zachowuje miarę, więc miara każdego ze zbiorów (2.3) jest taka sama. Gdyby $\mu(A_n) > 0$, to na mocy rozłączności zbiorów (2.3) i dowolności k uzyskalibyśmy miarę sumy tych zbiorów większą od $\mu(X)$. Zatem $\mu(A_n) = 0$ dla każdego $n \in \mathbb{N}$. To zaś implikuje $\mu(\bigcup_{m \geqslant n} A_m) = 0$ dla każdego $n \in \mathbb{N}$ i w konsekwencji $\mu(\bigcap_{n \in \mathbb{N}} \bigcup_{m \geqslant n} A_m) = 0$.

Aby zakończyć dowód, przyjmijmy $M := \bigcap_{n \in \mathbb{N}} \bigcup_{m \geqslant n} A_m$ i weźmy dowolny element $x \in A \setminus M$. Zatem $x \in \bigcup_{n \in \mathbb{N}} \bigcap_{m \geqslant n} (A \setminus A_m)$. Istnieje więc takie n_0, że $x \notin A_m$ dla każdego $m \geqslant n_0$. Zatem znajdziemy $k_0 \in \mathbb{N}$ takie, że $x \in T^{-k_0 n_0}[A]$. Biorąc $n_1 := k_0 n_0 + 1$ znajdziemy $k_1 \in \mathbb{N}$ takie, że $x \in T^{-k_1 n_1}[A]$. Dalej postępujemy rekurencyjnie, otrzymując ciąg rosnący liczb naturalnych $(k_i n_i)_{i \geqslant 0}$ taki że $x \in T^{-k_i n_i}[A]$ dla każdego $i \geqslant 0$. $\square$

Pojęcie rekurencyjności zostało znacząco uogólnione w 1977 roku przez Furstenberga w następujący sposób:

Definicja 2.4. Odwzorowanie mierzalne $T\colon X \to X$ nazywa się wielorekurencyjne, gdy dla każdego $A \in \mathcal{S}$ miary dodatniej oraz dowolnego $k \in \mathbb{N}$ istnieje takie $n \in \mathbb{N}$, że

$$\mu(A \cap T^{-n}[A] \cap T^{-2n}[A] \cap \cdots \cap T^{-kn}[A]) > 0.$$

Z powyższej definicji wynika, że wielorekurencyjność T implikuje jego konserwatywność, czyli rekurencyjność (por. Definicję 2.2 oraz Twierdzenie 2.2).

Furstenberg w 1977 roku udowodnił następujący ogólny rezultat:

Twierdzenie 2.4. *Jeśli $T\colon X \to X$ jest odwzorowaniem zachowującym miarę w przestrzeni z miarą skończoną, to T jest wielorekurencyjne.*

Dzięki temu twierdzeniu Furstenberg uzyskał nowy dowód twierdzenia Szeremerédiego o tym, że podzbiór $\mathbb{N}$ o dodatniej gęstości zawiera skończone ciągi arytmetyczne dowolnie dużej długości.

Rozdział 3

Ergodyczność odwzorowań

Ustalmy przestrzeń probabilistyczną $(X, \mathcal{S}, \mu)$. Załóżmy, że odwzorowanie $T\colon X \to X$ zachowuje miarę. Jeśli $B \in \mathcal{S}$ jest zbiorem takim, że $T^{-1}[B] = B$, to również $T^{-1}[X \setminus B] = X \setminus B$. Badając zjawiska opisywane przez odwzorowanie T, chcemy uniknąć sytuacji, gdy pojawiają się zbiory $B, X \setminus B \in \mathcal{S}$ miary dodatniej, o powyższej własności. Istnienie takich zbiorów zmusza nas do badania odwzorowania T oddzielnie na B oraz $X \setminus B$. Ta idea prowadzi do pojęcia odwzorowania ergodycznego.

Definicja 3.1. Zbiór $B \in \mathcal{S}$ taki, że $T^{-1}[B] = B$ nazwiemy T-niezmienniczym. Odwzorowanie $T\colon X \to X$ zachowujące miarę nazywa się ergodyczne, gdy dla dowolnego zbioru T-niezmienniczego zachodzi $\mu(B) = 0$ lub $\mu(B) = 1$.

Następujące twierdzenie opisuje prostą charakteryzację ergodyczności T.

Twierdzenie 3.1. *Załóżmy, że $(X, \mathcal{S}, \mu)$ jest przestrzenią probabilistyczną i odwzorowanie $T\colon X \to X$ zachowuje miarę. Wtedy następujące warunki są równoważne:*

(i) *T jest ergodyczne;*

(ii) $(\forall\, B \in \mathcal{S})\ \ \mu(T^{-1}[B] \triangle B) = 0 \Rightarrow (\mu(B) = 0 \vee \mu(B) = 1)$*;*

(iii) $(\forall\, B \in \mathcal{S},\ \mu(B) > 0)\ \ \mu(\bigcup_{n=1}^{\infty} T^{-n}[B]) = 1$*;*

(iv) $(\forall\, A, B \in \mathcal{S},\ \mu(A) > 0, \mu(B) > 0)\ (\exists\, n \in \mathbb{N})\ \ \mu(T^{-n}[A] \cap B) > 0.$

Dowód. (i) $\Rightarrow$ (ii) Niech $B \in \mathcal{S}$ oraz $\mu(T^{-1}[B] \triangle B) = 0$. Połóżmy

$$B_\infty := \bigcap_{n=0}^{\infty} \bigcup_{i=n}^{\infty} T^{-i}[B].$$

Wtedy $B_\infty \in \mathcal{S}$ oraz (jak łatwo sprawdzić) $T^{-1}[B_\infty] = B_\infty$. Zatem na mocy (i) mamy $\mu(B_\infty) = 0$ lub $\mu(B_\infty) = 1$. Pokażemy, że

$$\mu(B_\infty \triangle B) = 0. \tag{3.1}$$

Wtedy $\mu(B) = \mu(B_\infty)$, a więc $\mu(B) = 0$ lub $\mu(B) = 1$, co daje (ii).

Dla dowodu (3.1) wykażemy przez indukcję, że dla każdego $n \in \mathbb{N}$

$$\mu(T^{-n}[B] \triangle B) = 0. \tag{3.2}$$

Dla $n = 1$ własność zachodzi z założenia. Załóżmy, że $\mu(T^{-n}[B] \triangle B) = 0$.

Mamy

$$T^{-(n+1)}[B] \triangle B \subset (T^{-(n+1)}[B] \triangle T^{-n}[B]) \cup (T^{-n}[B] \triangle B),$$

więc wystarczy pokazać, że $\mu(T^{-(n+1)}[B] \triangle T^{-n}[B]) = 0$. Wynika to z równości

$$T^{-(n+1)}[B] \triangle T^{-n}[B] = T^{-n}[T^{-1}[B] \triangle B]$$

oraz stąd, że T^n zachowuje miarę. To kończy dowód (3.2).

Niech $B_n := \bigcup_{i=n}^\infty T^{-i}[B]$ dla $n \geqslant 0$. Wtedy $B_n \triangle B \subset \bigcup_{i=n}^\infty (T^{-i}[B] \triangle B)$, więc korzystając z (3.2), mamy

$$\mu(B_n \triangle B) \leqslant \mu(\bigcup_{i=n}^\infty (T^{-i}[B] \triangle B)) = 0.$$

Na koniec

$$B_\infty \triangle B = (\bigcap_{n=0}^\infty B_n) \triangle B \subset \bigcup_{n=0}^\infty (B_n \triangle B).$$

Zatem $\mu(B_\infty \triangle B) = 0$, co daje (3.1).

(ii) $\Rightarrow$ (iii) Niech $B \in \mathcal{S}$ oraz $\mu(B) > 0$. Połóżmy $A := \bigcup_{n=1}^\infty T^{-n}[B]$. Wtedy $T^{-1}[A] \subset A$. Skoro T zachowuje miarę, to

$$\mu(A \triangle T^{-1}[A]) = \mu(A \setminus T^{-1}[A]) = \mu(A) - \mu(T^{-1}[A]) = 0.$$

Ponadto $\mu(B) > 0$ implikuje $\mu(A) > 0$ z uwagi na postać A i zachowywanie miary przez T. Zatem z założenia (ii) wynika, że $\mu(A) = 1$.

(iii) $\Rightarrow$ (iv) Niech $A, B \in \mathcal{S}$ będą takie, że $\mu(A) > 0$, $\mu(B) > 0$. Wtedy na mocy (iii) mamy $\mu(\bigcup_{n=1}^\infty T^{-n}[A]) = 1$, więc

$$\mu\left(\bigcup_{n=1}^\infty (T^{-n}[A] \cap B)\right) = \mu\left(\left(\bigcup_{n=1}^\infty T^{-n}[A]\right) \cap B\right) = \mu(B) > 0,$$

a stąd już wynika (iv).

(iv) $\Rightarrow$ (i) Przypuśćmy, że istnieje zbiór $B \in \mathcal{S}$ o własnościach $T^{-1}[B] = B$ oraz $0 < \mu(B) < 1$. Wtedy $0 < \mu(X \setminus B) < 1$ oraz $\mu(T^{-n}[B] \cap (X \setminus B)) = 0$ dla każdego $n \in \mathbb{N}$, co przeczy (iii). $\square$

Uwaga 3.1. Gdy $(X, \mathcal{S}, \mu)$ jest dowolną przestrzenią z miarą, to ergodyczność odwzorowania $T\colon X \to X$ zachowującego miarę definiuje się poprzez warunek: dla każdego zbioru $B \in \mathcal{S}$, jeśli $T^{-1}[B] = B$, to $\mu(B) = 0$ lub $\mu(X \setminus B) = 0$. Rozważmy przykład, gdy na σ-ciele $\mathcal{P}(\mathbb{Z})$ wszystkich podzbiorów zbioru $\mathbb{Z}$ dana jest miara licząca μ, tzn. $\mu(A)$ oznacza liczność zbioru A, gdy jest on skończony, zaś $\mu(A) = \infty$ w przeciwnym przypadku. Zauważmy, że odwzorowanie $T\colon \mathbb{Z} \to \mathbb{Z}$ dane wzorem $T(n) := n+1$ zachowuje miarę i nie jest rekurencyjne. Natomiast jest ono ergodyczne, bo jeśli $B \subset \mathbb{Z}$ oraz $T^{-1}[B] = B$, to albo $B = \emptyset$, albo $B = \mathbb{Z}$.

Funkcję mierzalną $f\colon X \to \mathbb{R}$ nazwiemy T-niezmienniczą (prawie T-niezmienniczą), gdy $f \circ T = f$ (odpowiednio $f \circ T = f$ prawie wszędzie).

Kolejne twierdzenie opisuje ergodyczność T w języku funkcji. Przez L^2 oznaczamy przestrzeń Banacha funkcji mierzalnych $f\colon X \to \mathbb{R}$, które są μ-całkowalne z kwadratem na X. Analogiczne definiujemy przestrzenie L^p dla $p \geqslant 1$.

Twierdzenie 3.2. *Załóżmy, że $(X, \mathcal{S}, \mu)$ jest przestrzeń probabilistyczną i odwzorowanie $T\colon X \to X$ zachowuje miarę. Wtedy następujące warunki są równoważne:*

(i) *T jest ergodyczne;*

(ii) *każda funkcja mierzalna T-niezmiennicza jest prawie wszędzie stała;*

(iii) *każda funkcja mierzalna prawie T-niezmiennicza jest prawie wszędzie stała;*

(iv) *każda funkcja $f \in L^2$, T-niezmiennicza, jest prawie wszędzie stała;*

(v) *każda funkcja $f \in L^2$, prawie T-niezmiennicza, jest prawie wszędzie stała.*

Dowód. Implikacje (ii) $\Rightarrow$ (iv), (iii) $\Rightarrow$ (v), (iii) $\Rightarrow$ (ii), (v) $\Rightarrow$ (iv) są trywialne. Zatem wystarczy wykazać (i) $\Rightarrow$ (iii) oraz (iv) $\Rightarrow$ (i).

(i) $\Rightarrow$ (iii) Rozważmy odwzorowanie ergodyczne $T\colon X \to X$ i załóżmy, że $f\colon X \to \mathbb{R}$ jest funkcją mierzalną prawie T-niezmienniczą. Dla $n \in \mathbb{N}$, $k \in \mathbb{Z}$ oznaczmy

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

Wtedy

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

a więc

$$T^{-1}[X_{k,n}] \triangle X_{k,n} \subset \{x \in X\colon (f \circ T)(x) \neq f(x)\}.$$

Skoro f jest prawie T-niezmiennicza, to ten ostatni zbiór jest miary zero. Zatem

$$\mu(T^{-1}[X_{k,n}] \triangle X_{k,n}) = 0,$$

więc na podstawie Twierdzenia 3.1 wnioskujemy, że miara $\mu(X_{k,n})$ jest równa albo 0, albo 1.

Ustalmy $n \in \mathbb{N}$. Wtedy suma $\bigcup_{k \in \mathbb{Z}} X_{k,n} = X$ składa się z parami rozłącznych składników. Zatem istnieje dokładnie jeden wskaźnik $k_n \in \mathbb{Z}$ taki, że $\mu(X_{k_n,n}) = 1$. Zauważmy, że ciąg zbiorów $(X_{k_n,n})_{n \in \mathbb{N}}$ jest malejący. Ponadto ciąg liczbowy $(k_n/2^n)_{n \in \mathbb{N}}$ jest niemalejący i ograniczony. Zatem jest on zbieżny do skończonej granicy a. Niech $Y := \bigcap_{n \in \mathbb{N}} X_{k_n,n}$. Wtedy $\mu(Y) = 1$ oraz jeśli $x \in Y$, to $|f(x) - k_n/2^n| < 1/2^n$ dla każdego $n \in \mathbb{N}$. Zatem biorąc $n \to \infty$, mamy $f(x) = a$, co oznacza, że f jest stała na Y.

(iv) $\Rightarrow$ (i) Załóżmy, że $B \in \mathcal{S}$ oraz $T^{-1}[B] = B$. Wtedy $\chi_B \in L^2$ oraz $(\chi_B \circ T)(x) = \chi_B(x)$ dla każdego $x \in X$, czyli funkcja χ_B jest T-niezmiennicza. Zatem z (iv) wynika, że χ_B jest stała prawie wszędzie na X. Jest więc ona równa prawie wszędzie 0 lub prawie wszędzie równa 1. Stąd miara $\mu(B) = \int_X \chi_B \, d\mu$ jest równa 0 lub 1. $\square$

Uwaga 3.2. W warunkach (iv) i (v) Twierdzenia 3.2 przestrzeń L^2 można zastąpić przez L^p dla $1 \leqslant p < \infty$. Teza twierdzenia pozostaje prawdziwa, gdy rozważamy funkcje f o wartościach zespolonych.

Przykład 3.1. Przesunięcie $T\colon [0,1) \to [0,1)$, $T(x) := x + 1/2 \pmod 1$ nie jest ergodyczne, bo dla $A := [0, 1/4] \cup [1/2, 3/4]$ mamy $T^{-1}[A] = A$ oraz $\mu(A) = 1/2 \notin \{0,1\}$. Podobnie jest dla $T(x) := x + 1/4 \pmod 1$. Należy przyjąć

$$A := [0, 1/8] \cup [1/4, 1/8] \cup [1/2, 5/8] \cup [3/4, 7/8].$$

W ogólnym przypadku przesunięcie $T(x) := x + p/q \pmod 1$, gdzie $0 < p < q$ oraz $p, q \in \mathbb{N}$, nie jest ergodyczne. Można to pokazać przez wskazanie stosownego przykładu lub posiłkując się Twierdzeniem 3.2 – wtedy zauważamy, że dla $f(x) := qx \pmod 1$ mamy $(f \circ T)(x) = f(x)$ dla $x \in [0,1)$, lecz funkcja f nie jest prawie wszędzie stała.

Dla odwzorowania $T\colon X \to X$ oraz punktu $x \in X$ zbiór $\{T^n(x)\colon n \geqslant 0\}$ nazywamy orbitą punktu x.

Twierdzenie 3.3. *Niech X będzie przestrzenią metryczną ośrodkową oraz μ -miarą prawdopodobieństwa określoną na zbiorach borelowskich przestrzeni X przyjmującą wartości dodatnie na niepustych otwartych podzbiorach przestrzeni X. Jeśli $T\colon X \to X$ jest ciągłym odwzorowaniem ergodycznym, to orbita prawie każdego (w sensie miary) punktu zbioru X jest zbiorem gęstym.*

Dowód. Niech zbiory U_n, $n \in \mathbb{N}$, stanowią bazę topologii przestrzeni X. Zauważmy, że orbita $\{T^n(x)\colon n \geqslant 0\}$ punktu $x \in X$ jest zbiorem gęstym wtedy i tylko wtedy, gdy $x \in \bigcap_{n=1}^{\infty} \bigcup_{k=0}^{\infty} T^{-k}[U_n]$.

Ustalmy $n \in \mathbb{N}$ i niech $A_n := \bigcup_{k=0}^{\infty} T^{-k}[U_n]$. Mamy $T^{-1}[A_n] \subset A_n$, a ponieważ T zachowuje miarę, więc jest rekurencyjne, a zatem nieściśliwe (patrz poprzedni wykład), czyli $\mu(A_n \setminus T^{-1}[A_n]) = 0$. Stąd z ergodyczności odwzorowania T i Twierdzenia 3.1 wnioskujemy, że $\mu(A_n) = 0$ lub $\mu(A_n) = 1$. Skoro $\mu(U_n) > 0$, to musi być $\mu(A_n) = 1$. W konsekwencji $\mu(\bigcap_{n=1}^{\infty} A_n) = 1$, a to zgodnie z pierwszą obserwacją daje tezę. $\square$

Twierdzenie 3.3 daje jeszcze jeden dowód braku ergodyczności przesunięcia wymiernego w Przykładzie 3.1: wystarczy zauważyć, że dla tego odwzorowania orbita dowolnego punktu jest skończona (nie jest więc zbiorem gęstym).

Natomiast przesunięcie niewymierne $T(x) := x + s \pmod 1$, gdzie $s \in [0,1) \setminus \mathbb{Q}$, jest odwzorowaniem ergodycznym. Aby to wykazać, zastosujemy Twierdzenie 3.2. Niech $f \in L^2$ będzie funkcją T-niezmienniczą. Rozważmy rozwinięcie f w (zespolony) szereg Fouriera. Dla $x \in [0,1)$ mamy:

$$f(x) = \sum_{n\in\mathbb{Z}} a_n\, e^{2\pi inx} \quad \text{oraz} \quad f(x) = (f \circ T)(x) = \sum_{n\in\mathbb{Z}} a_n\, e^{2\pi in(x+s)} = \sum_{n\in\mathbb{Z}} a_n\, e^{2\pi ins} e^{2\pi inx}.$$

Z jednoznaczności współczynników Fouriera dla funkcji f mamy $a_n(1 - e^{2\pi ins}) = 0$ dla wszystkich $n \in \mathbb{Z}$. Jeśli $n \neq 0$, to $s \notin \mathbb{Q}$ pociąga za sobą

$$e^{2\pi ins} = \cos(2\pi ns) + i \sin(2\pi ns) \neq 1.$$

Zatem $a_n = 0$ dla każdego $n \neq 0$. Stąd funkcja f jest stała równa a_0. Z Twierdzenia 3.2 wynika więc, że odwzorowanie T jest ergodyczne. Elementarny (lecz nie krótszy) dowód można znaleźć w książce Silvy [11, s. 101].

Ćwiczenie 3.1. Wykazać, że następujące odwzorowania są ergodyczne:

(a) przesunięcie dwustronne (jednostronne) w odpowiednich przestrzeniach produktowych; zob. [2, s. 25–26].

(b) odwzorowanie podwajające $T\colon [0,1) \to [0,1)$; zob. [2, s. 27].

Rozdział 4

Twierdzenia ergodyczne

Przypomnijmy, że dla danej przestrzeni z miarą $(X, \mathcal{S}, \mu)$ oraz $p \in [1, \infty)$ symbol L^p oznacza przestrzeń funkcji $f \colon X \to \mathbb{R}$ takich że $\int_X |f|^p \, d\mu < \infty$ (lub bardziej formalnie, zbiór klas abstrakcji w relacji $f = g$, μ-prawie wszędzie). Zbiór ten jest przestrzenią liniową, unormowaną przez normę

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

oraz zupełną w metryce generowanej przez tę normę. Jest więc przestrzenią Banacha. Rozważa się też analogiczną przestrzeń funkcji o wartościach zespolonych.

Udowodnimy tu podstawowe dla omawianej teorii twierdzenie Birkhoffa z 1931 roku. Przypomnijmy, że miara μ na σ-ciele podzbiorów przestrzeni X nazywa się σ-skończona, gdy istnieje ciąg zbiorów $X_n \in \mathcal{S}$, $n \in \mathbb{N}$, o mierze skończonej i takich, że $X = \bigcup_{n \in \mathbb{N}} X_n$.

Twierdzenie 4.1 (twierdzenie ergodyczne Birkhoffa)**.** *Niech $(X, \mathcal{S}, \mu)$ będzie przestrzenią z miarą σ-skończoną oraz załóżmy, że $T \colon X \to X$ jest odwzorowaniem zachowującym miarę. Jeśli $f \in L^1$, to istnieje funkcja $f^* \in L^1$ spełniająca warunek*

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} f \circ T^i = f^* \text{ prawie wszędzie na } X \tag{4.1}$$

oraz $f^* \circ T = f^*$ prawie wszędzie na X. Ponadto jeśli $\mu(X) < \infty$, to

$$\int_X f^* d\mu = \int_X f d\mu.$$

Uwaga 4.1. Jeśli $\mu(X) < \infty$ oraz odwzorowanie T jest ergodyczne, to funkcja f^* jest stała prawie wszędzie i równa $f^* = \frac{1}{\mu(X)} \int_X f d\mu$.

Opiszemy motywację, która doprowadziła do powyższego rezultatu. Niech odwzorowanie $T\colon X \to X$ zachowuje miarę przy założeniu $\mu(X) < \infty$. Dla ustalonego zbioru $E \in \mathcal{S}$ miary dodatniej oraz punktu $x \in X$ można pytać o częstotliwość należenia elementów orbity $\{x, T(x), T^2(x), \dots\}$ do zbioru E. Średnia liczność elementów orbity w zbiorze E jest równa $\frac{1}{n}\sum_{i=0}^{n-1}(\chi_E \circ T^i)(x)$. Jeśli $\mu(X) = 1$ oraz T jest ergodyczne, to z powyższej uwagi wynika, że dla prawie wszystkich $x \in X$ zachodzi

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1}(\chi_E \circ T^i)(x) = \mu(E).$$

Równość ta oznacza, że średnia czasowa odwiedzania zbioru E przez orbitę punktu x jest równa średniej przestrzennej zbioru E (czyli mierze tego zbioru).

Ogólnie granica $\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1} f \circ T^i$ nazywa się średnią czasową funkcji $f \in L^1$, zaś $\frac{1}{\mu(X)}\int_X f d\mu$ nazywa się średnią przestrzenną (fazową) funkcji f. Z powyższej uwagi wynika, że średnie te są sobie równe dla prawie wszystkich $x \in X$.

W dowodzie Twierdzenia 4.1 istotną rolę pełni operator liniowy związany z danym odwzorowaniem zachowującym miarę. Dla przestrzeni z miarą σ-skończoną $(X, \mathcal{S}, \mu)$ oznaczamy przez L^0 przestrzeń liniową funkcji rzeczywistych mierzalnych, określonych na X. Definiujemy operator $U_T\colon L^0 \to L^0$ wzorem $U_T(f) := f \circ T$. Dowód następującego lematu jest nietrudnym ćwiczeniem.

Lemat 4.1. *Operator U_T ma następujące własności:*

(i) *U_T jest operatorem liniowym;*

(ii) *$U_T(fg) = U_T(f)U_T(g)$ dla dowolnych $f, g \in L^0$;*

(iii) *$U_T(c) = c$ dla dowolnej stałej $c \in \mathbb{R}$;*

(iv) *$U_T(f) \geqslant 0$ dla dowolnej nieujemnej funkcji $f \in L^0$;*

(v) *$U_T(\chi_B) = \chi_B \circ T = \chi_{T^{-1}[B]}$ dla każdego $B \in \mathcal{S}$;*

(vi) *$\int_X U_T(f)\, d\mu = \int_X f\, d\mu$ dla wszystkich $f \in L^0$ (jeśli całka po jednej stronie nie istnieje lub jest nieskończona, to tę samą własność ma całka po drugiej stronie).*

(vii) *dla $p \geqslant 1$ zachodzi $U_T[L^p] \subset L^p$ oraz $||U_T(f)||_p = ||f||_p$ dla wszystkich $f \in L^p$.*

Powyższy operator U_T można rozważać także dla funkcji o wartościach zespolonych. Zachodzą wtedy analogiczne własności; [2], proposition 1.7.1.

Rezultatem pomocniczym w dowodzie twierdzenia Birkhoffa jest poniższe twierdzenie. Stosujemy w nim pewien abstrakcyjny operator liniowy $U\colon L^1 \to L^1$, który jest dodatni, tzn. $U(f) \geqslant 0$ dla wszystkich $f \in L^1$, $f \geqslant 0$.

Przypomnijmy, że operator liniowy $U\colon X \to Y$ między przestrzeniami unormowanymi X, Y nazywa się ograniczony, gdy istnieje liczba $C > 0$ taka, że $||U(x)||_Y \leqslant C||x||_X$ dla każdego $x \in X$. Liczbę $||U|| := \inf_{x\in X} ||U(x)||_Y$ nazywamy normą operatora U.

Twierdzenie 4.2 (maksymalne twierdzenie ergodyczne)**.** *Niech $U\colon L^1 \to L^1$ będzie dodatnim ograniczonym operatorem liniowym o normie $||U|| \leqslant 1$. Ustalmy liczbę $N \in \mathbb{N}$ oraz $f \in L^1$. Określmy*

$$f_0 := 0, \quad f_n := f + U(f) + U^2(f) + \cdots + U^{n-1}(f) \quad \textit{dla } n \geqslant 1, \; F_N := \max_{0\leqslant n\leqslant N} f_n. \tag{4.2}$$

Wtedy $\int_{\{x\in X\colon F_N(x)>0\}} f\,d\mu \geqslant 0$.

Dowód. Zauważmy, że $F_N \geqslant 0$. Korzystając z własności funkcji całkowalnych, stwierdzamy, że $F_N \in L^1$. Dla $0 \leqslant n \leqslant N$ mamy $F_N \geqslant f_n$, więc $U(F_N) \geqslant U(f_n)$, bo operator U jest liniowy i dodatni. Zatem $U(F_N) \geqslant \max_{0\leqslant n\leqslant N} U(f_n)$ oraz na mocy liniowości U dostajemy

$$\begin{aligned} f + U(F_N) &\geqslant \max_{0\leqslant n\leqslant N}(f + U(f_n)) = \max_{0\leqslant n\leqslant N}(f + U(f) + U^2(f) + \cdots + U^n(f)) \\ &= \max_{0\leqslant n\leqslant N} f_{n+1} \geqslant f_{k+1} \quad \text{dla } 0 \leqslant k \leqslant N. \end{aligned}$$

Stąd

$$f + U(F_N) \geqslant \max_{1\leqslant n\leqslant N} f_n. \tag{4.3}$$

Oznaczmy $A := \{x \in X\colon F_N(x) > 0\}$. Ustalmy dowolny $x \in A$. Wtedy $F_N(x) > 0$. Ponadto $f_0(x) = 0$, więc $F_N(x) = \max_{1\leqslant n\leqslant N} f_n(x)$. Zatem korzystając z (4.3), otrzymujemy

$$f(x) + U(F_N)(x) \geqslant \max_{1\leqslant n\leqslant N} f_n(x) = F_N(x).$$

Stąd

$$\int_A f\,d\mu \geqslant \int_A F_N\,d\mu - \int_A U(F_N)\,d\mu \geqslant \int_X F_N\,d\mu - \int_X U(F_N)\,d\mu,$$

bo $F_N(x) = 0$ dla $x \in X \setminus A$ oraz $U(F_N) \geqslant 0$. W konsekwencji $\int_A f\,d\mu \geqslant 0$, gdyż

$$\int_X U(F_N)\,d\mu = \int_X |U(F_N)|\,d\mu = ||U(F_N)||_1 \leqslant ||U||\,||F_N||_1 \leqslant ||F_N||_1 = \int_X F_N\,d\mu.$$

□

Jeśli odwzorowanie $T\colon X \to X$ zachowuje miarę, to operator U_T rozpatrywany w Lemacie 4.1 i obcięty do L^1 spełnia założenia Twierdzenia 4.2, bo jest liniowy, dodatni oraz ograniczony. Co więcej, z Lematu 4.1 (vii) wynika, że U_T jest izometrią.

Wniosek 4.1. *Niech odwzorowanie* $T\colon X \to X$ *zachowuje miarę w przestrzeni* $(X, \mathcal{S}, \mu)$ *z miarą* σ*-skończoną. Dla funkcji* $g \in L^1$ *oraz liczby* $\alpha \in \mathbb{R}$ *rozważmy zbiór*

$$B_\alpha := \left\{ x \in X\colon \sup_{n \geqslant 1} \frac{1}{n} \sum_{m=0}^{n-1} g(T^m x) > \alpha \right\}.$$

Wtedy

$$\int_{B_\alpha \cap A} g \, d\mu \geqslant \alpha \mu(B_\alpha \cap A)$$

dla dowolnego zbioru $A \in \mathcal{S}$ *takiego, że* $T^{-1}[A] = A$ *oraz* $\mu(A) < \infty$.

Dowód. Załóżmy najpierw, że $\mu(X) < \infty$ oraz $A = X$. Niech $f := g - \alpha$. Wtedy $f \in L^1$ oraz niech operator $U\colon L^1 \to L^1$ będzie dany wzorem $U := U_T$, tzn. $U(h) := h \circ T$. Zdefiniujmy funkcje f_n dla $n \geqslant 0$ oraz F_N dla $N \in \mathbb{N}$ zgodnie ze wzorami (4.2). Zauważmy, że $f_n = \sum_{m=0}^{n-1} (g(T^m x) - \alpha)$.

Pokażemy, że $B_\alpha = \bigcup_{N=1}^{\infty} \{x\colon F_N(x) > 0\}$. Niech $x \in X$ oraz $F_N(x) > 0$ dla pewnego $N \in \mathbb{N}$. Wtedy istnieje takie $n \in \{1, \ldots, N\}$, że $f_n(x) > 0$, czyli $\sum_{m=0}^{n-1} (g(T^m x) - \alpha) > 0$, skąd mamy $\frac{1}{n} \sum_{m=0}^{n-1} g(T^m x) > \alpha$, czyli $x \in B_\alpha$. Odwrotnie, jeśli $x \in B_\alpha$, to istnieje takie $N \geqslant 1$, że $\frac{1}{N} \sum_{m=0}^{N-1} g(T^m x) > \alpha$ i dlatego $F_N(x) \geqslant f_N(x) > 0$.

Na mocy Twierdzenia 4.2 otrzymujemy $\int_{\{x\colon F_N(x) > 0\}} f \, d\mu \geqslant 0$ dla każdego $N \in \mathbb{N}$. Ponadto $F_{N+1} \geqslant F_N$, więc dla $A_N := \{x\colon F_N(x) > 0\}$ mamy $A_N \subset A_{N+1}$. Połóżmy $\widetilde{f_N} := \chi_{A_N} f$ dla $N \in \mathbb{N}$. Wtedy $\widetilde{f_N} \to \chi_{B_\alpha} f$, gdy $N \to \infty$ (bo $\bigcup_{N=1}^{\infty} A_N = B_\alpha$) oraz $|\widetilde{f_N}| \leqslant |f|$ dla każdego N. Zatem z twierdzenia Lebesgue'a o zbieżności ograniczonej dostajemy

$$\int_{B_\alpha} f \, d\mu = \int_X \chi_{B_\alpha} f \, d\mu = \lim_{N \to \infty} \int_X \widetilde{f_N} \, d\mu = \lim_{N \to \infty} \int_{A_N} f \, d\mu \geqslant 0.$$

Z definicji f mamy więc

$$\int_{B_\alpha} g \, d\mu \geqslant \alpha \mu(B_\alpha)$$

(korzystamy z tego, że $\mu(B_\alpha) < \infty$). W ogólnym przypadku zastępujemy X przez A oraz T przez $T|_A$ i otrzymujemy

$$\int_{A \cap B_\alpha} g \, d\mu \geqslant \alpha \mu(A \cap B_\alpha).$$

□

Do dowodu Twierdzenia 1 będzie nam potrzebny lemat.

Lemat 4.2. *Przy założeniach Twierdzenia 4.1 dla danej funkcji $f \in L^1$ określmy funkcje f^*, f_* wzorami:*

$$f^* := \limsup_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f \circ T^i, \quad f_* := \liminf_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f \circ T^i.$$

Wtedy $f^* \circ T = f^*$ oraz $f_* \circ T = f_*$.

Dowód. Dla dowodu pierwszej równości ustalmy $x \in X$ i niech $s_n(x) := \frac{1}{n} \sum_{n=0}^{n-1} f(T^i x)$. Wtedy

$$s_n(Tx) = \frac{1}{n} \sum_{i=0}^{n-1} f(T^{i+1}x) = \frac{1}{n} \sum_{i=1}^{n} f(T^i x) = \frac{1}{n} \sum_{i=0}^{n} f(T^i x) - \frac{1}{n} f(x)$$

$$= \left(\frac{n+1}{n}\right) \frac{1}{n+1} \sum_{i=0}^{n} f(T^i x) - \frac{1}{n} f(x) = \left(\frac{n+1}{n}\right) s_{n+1}(x) - \frac{1}{n} f(x).$$

Stąd

$$(f^* \circ T)(x) = \limsup_{n\to\infty} s_n(Tx) = \limsup_{n\to\infty} s_{n+1}(x) - 0 = f^*(x).$$

Podobnie pokazujemy, że $f_* \circ T = f_*$. □

Przejdźmy do dowodu Twierdzenia 4.1.

Dowód. Ustalmy liczby rzeczywiste α, β takie, że $\beta < \alpha$. Przy oznaczeniach Lematu 4.2 niech

$$E_{\alpha,\beta} := \{x \in X \colon f_*(x) < \beta,\ \alpha < f^*(x)\}.$$

Pokażemy, że $\mu(E_{\alpha,\beta}) = 0$. Dlatego, korzystając z dowolności α, β będziemy mogli wnioskować, że $f_* = f^*$ prawie wszędzie na X.

Na początku pokażemy, że $\mu(E_{\alpha,\beta}) < \infty$. Z Lematu 4.2 wiemy, że $f^* \circ T = f^*$ oraz $f_* \circ T = f_*$, więc $T^{-1}[E_{\alpha,\beta}] = E_{\alpha,\beta}$. Ponadto

$$E_{\alpha,\beta} \subset \left\{x \in X \colon \sup_{n\in\mathbb{N}} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x) > \alpha \right\} =: B_\alpha, \tag{4.4}$$

bo jeśli $\limsup_n s_n > \alpha$, to $\sup_n s_n > \alpha$.

Załóżmy najpierw, że $\alpha > 0$. Weźmy dowolny zbiór $C \in \mathcal{S}$ taki, że $\mu(C) < \infty$ oraz $C \subset E_{\alpha,\beta}$. Pokażemy, że

$$\mu(C) \leqslant \frac{1}{\alpha} \int_X |f| \, d\mu. \tag{4.5}$$

Postępujemy podobnie jak w dowodzie Wniosku 4.1. Niech $h := f - \alpha\chi_C$. Wtedy $h \in L^1$ oraz z Twierdzenia 4.2 mamy $\int_{\{x \colon H_N(x) > 0\}} h \, d\mu \geqslant 0$ dla każdego $N \in \mathbb{N}$, gdzie

$$H_N := \max_{0 \leqslant n \leqslant N} h_n \ \text{ oraz } \ h_0 := 0,\ h_n := h + h \circ T + \cdots + h \circ T^{n-1}.$$

W konsekwencji $\int_{B_\alpha} h\,d\mu \geqslant 0$ z twierdzenia Lebesgue'a o zbieżności ograniczonej, ponieważ $B_\alpha := \bigcup_{N=1}^{\infty} \{x\colon H_N(x) > 0\}$ jest sumą wstępującego ciągu zbiorów (por. odpowiedni fragment dowodu Wniosku 4.1). Ponadto z (4.4) wynika, że $C \subset B_\alpha$. Zatem pamiętając, że $h = f - \alpha\chi_C$, mamy

$$\int_X |f|\,d\mu \geqslant \int_{B_\alpha} f\,d\mu \geqslant \alpha \int_{B_\alpha} \chi_C\,d\mu = \alpha\mu(C),$$

co daje (4.5).

Przypadek $\alpha \leqslant 0$ implikuje $\beta < 0$ i sprowadzamy go do poprzedniego, zastępując f przez $-f$. Wtedy liczby $-\beta, -\alpha$ pełnią rolę liczb α, β. Zauważamy przy tym, że dla $x \in X$

$$\liminf_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} (f \circ T)(x) < \beta \iff \limsup_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} (-f \circ T)(x) > -\beta.$$

Szczegóły pozostawiamy jako ćwiczenie.

Zatem w obu przypadkach otrzymaliśmy warunek typu (4.5), który z dowolności C implikuje $\mu(E_{\alpha,\beta}) < \infty$. Istotnie, skoro miara μ jest σ-skończona, to zbiór $E_{\alpha,\beta}$ jest sumą wstępującego ciągu zbiorów C_n o mierze skończonej i wystarczy we wzorze (4.5) przejść do granicy, gdy $n \to \infty$.

Wiedząc, że $\mu(E_{\alpha,\beta}) < \infty$ oraz $T^{-1}[E_{\alpha,\beta}] = E_{\alpha,\beta}$, możemy zastosować Wniosek 4.1. Wtedy

$$\int_{E_{\alpha,\beta}} f\,d\mu = \int_{E_{\alpha,\beta}\cap B_\alpha} f\,d\mu \geqslant \alpha\mu(E_{\alpha,\beta} \cap B_\alpha) = \alpha\mu(E_{\alpha,\beta}). \tag{4.6}$$

Jeśli zastąpimy f, α, β przez $-f, -\beta, -\alpha$ (odpowiednio) oraz uwzględnimy równości $(-f)^* = -f_*$, $(-f)_* = -f^*$, to analogicznie jak poprzednio otrzymamy $\int_{E_{\alpha,\beta}} (-f)\,d\mu \geqslant -\beta\mu(E_{\alpha,\beta})$, czyli

$$\int_{E_{\alpha,\beta}} f\,d\mu \leqslant \beta\mu(E_{\alpha,\beta}). \tag{4.7}$$

Na mocy (4.6) i (4.7) mamy $\alpha\mu(E_{\alpha,\beta}) \leqslant \beta\mu(E_{\alpha,\beta})$. Ale $\beta < \alpha$, więc $\mu(E_{\alpha,\beta}) = 0$.

Stąd dostajemy

$$\{x \in X\colon f_*(x) < f^*(x)\} = \bigcup_{\alpha,\beta\in\mathbb{Q};\, \beta<\alpha} E_{\alpha,\beta},$$

co implikuje, że $f_* = f^*$ prawie wszędzie. Zatem ciąg $(\frac{1}{n}\sum_{i=0}^{n-1} f \circ T^i)_{n\in\mathbb{N}}$ jest zbieżny prawie wszędzie. W dalszym rozumowaniu granicę tego ciągu oznaczamy przez f^*. Wiemy już, że $f^* \circ T = f^*$ na mocy Lematu 4.2.

Aby zakończyć dowód pierwszej części tezy twierdzenia, pokażemy, że $f^* \in L^1$. Rozważmy funkcje mierzalne nieujemne $g_n := \left|\frac{1}{n}\sum_{i=0}^{n-1} f \circ T^i\right|$, $n \in \mathbb{N}$. Stosując lemat Fatou, otrzymujemy

$$\int_X |f^*|\,d\mu = \int_X \lim_n g_n\,d\mu \leqslant \liminf_n \int_X g_n\,d\mu$$

$$\leqslant \liminf_n \left(\frac{1}{n} \sum_{i=0}^{n-1} \int_X |f(T^i x)| \, d\mu \right) = \liminf_n \left(\frac{1}{n} \sum_{i=0}^{n-1} \int_X |f| \, d\mu \right) = \int_X |f| \, d\mu.$$

Stąd $f^* \in L^1$. Zauważmy przy tym, że operator $S\colon L^1 \to L^1$ dany wzorem $S(g) := g^*$ jest liniowy. Ponadto jest on ograniczony, bo z powyższych nierówności wynika, że $||S(g)||_1 \leqslant ||g||_1$ dla każdego $g \in L^1$.

Na koniec pokażemy, że jeśli $\mu(X) < \infty$, to $\int_X f^* \, d\mu = \int_X f \, d\mu$. Załóżmy najpierw, że funkcja f jest ograniczona, tzn. $|f(x)| \leqslant C$ dla $x \in X$. Zastosujemy twierdzenie Lebesgue'a o zbieżności ograniczonej dla ciągu funkcji $s_n = \frac{1}{n} \sum_{i=0}^{n-1} f \circ T^i$, $n \in \mathbb{N}$. Dla wszystkich $x \in X$ mamy

$$|s_n(x)| \leqslant \frac{1}{n} \sum_{i=0}^{n-1} |f \circ T^i(x)| \leqslant \frac{n}{n} C = C.$$

Zatem

$$\int_X f^* d\mu = \int_X (\lim_n s_n) \, d\mu = \lim_n \int_X s_n \, d\mu = \lim_n \frac{1}{n} \sum_{i=1}^{n-1} \int_X f(T^i x) \, d\mu = \int_X f \, d\mu. \tag{4.8}$$

Niech teraz funkcja $f \in L^1$ będzie dowolna i dla ustalonego $\varepsilon > 0$ wybierzmy ograniczoną funkcję $g \in L^1$ taką, że $||f - g||_1 < \varepsilon/2$. Z własności operatora S wynika, że

$$||f^* - g^*||_1 = ||S(f - g)||_1 \leqslant ||f - g||_1 < \frac{\varepsilon}{2}.$$

Teraz, stosując (4.8) dla g, g^*, mamy

$$\left| \int_X f^* \, d\mu - \int_X f \, d\mu \right| \leqslant ||f^* - f||_1 \leqslant ||f^* - g^*||_1 + ||g^* - g||_1 + ||g - f||_1 < \frac{\varepsilon}{2} + 0 + \frac{\varepsilon}{2} = \varepsilon.$$

Zatem z dowolności ε mamy $\int_X f^* \, d\mu = \int_X f \, d\mu$. □

Odnotujmy, że prostszy dowód Twierdzeń 4.2 i 4.1 dla przestrzeni probabilistycznych można znaleźć w [9].

Kolejne twierdzenie ergodyczne jest wnioskiem z twierdzenia Birkhoffa.

Twierdzenie 4.3 (twierdzenie ergodyczne von Neumanna)**.** *Załóżmy, że $(X, \mathcal{S}, \mu)$ jest przestrzenią probabilistyczną oraz odwzorowanie $T\colon X \to X$ zachowuje miarę. Niech $1 \leqslant p < \infty$. Jeśli $f \in L^p$, to istnieje funkcja $f^* \in L^p$ taka, że $f^* \circ T = f^*$ prawie wszędzie oraz*

$$\lim_{n\to\infty} \left\| \frac{1}{n} \sum_{i=0}^{n-1} (f \circ T^i) - f^* \right\|_p = 0.$$

Dowód. Załóżmy najpierw, że funkcja $g\colon X \to \mathbb{R}$ jest mierzalna i ograniczona. Wtedy $g \in L^p \subset L^1$ oraz z twierdzenia Birkhoffa mamy $\lim_n \frac{1}{n}\sum_{i=0}^{n-1} g \circ T^i = g^* \in L^1$ prawie wszędzie, przy czym widać, że funkcja g^* musi być ograniczona prawie wszędzie. Zatem $g^* \in L^p$. Ponadto $\lim_n \left|\frac{1}{n}\sum_{i=0}^{n-1} g \circ T^i - g^*\right|^p = 0$ prawie wszędzie. Stąd wynika, że

$$\lim_{n\to\infty} \left\| \frac{1}{n}\sum_{i=0}^{n-1}(g \circ T^i) - g^* \right\|_p = 0$$

na mocy twierdzenia Lebesgue'a o zbieżności ograniczonej. To daje odpowiedni fragment tezy w tym przypadku.

Ciąg $\left(\frac{1}{n}\sum_{i=0}^{n-1}(g \circ T^i)\right)_{n\in\mathbb{N}}$ jako zbieżny w L^p spełnia warunek Cauchy'ego. Zatem dla każdego $\varepsilon > 0$ istnieje liczba naturalna $N = N(\varepsilon, g)$, taka że

$$\forall n,k \geqslant N \quad \left\| \frac{1}{n}\sum_{i=0}^{n-1}(g \circ T^i) - \frac{1}{k}\sum_{i=0}^{k-1}(g \circ T^i) \right\|_p < \frac{\varepsilon}{2}. \tag{4.9}$$

Dla $n \in \mathbb{N}$ określmy operator $M_n\colon L^p \to L^p$ wzorem $M_n f := \frac{1}{n}\sum_{i=0}^{n-1} f \circ T^i$. Jest on liniowy. Ponadto jest on ograniczony, bo dla każdego $f \in L^p$ mamy $||M_n f||_p \leqslant ||f||_p$. Istotnie

$$\left\| \frac{1}{n}\sum_{i=0}^{n-1} f \circ T^i \right\|_p \leqslant \frac{1}{n}\sum_{i=0}^{n-1} ||f \circ T^i||_p$$

$$= \frac{1}{n}\sum_{n=0}^{n-1}\left(\int_X |f \circ T^i|^p \, d\mu\right)^{\frac{1}{p}} = \frac{1}{n}\sum_{n=0}^{n-1}\left(\int_X |f|^p \, d\mu\right)^{\frac{1}{p}} = ||f||_p.$$

Niech teraz $f \in L^p$ oraz $\varepsilon > 0$. Wybierzmy funkcję mierzalną ograniczoną $g\colon X \to \mathbb{R}$ taką, że $||f - g||_p < \frac{\varepsilon}{4}$. Dobierzmy liczbę $N \in \mathbb{N}$ spełniającą (4.9). Wtedy korzystając ze wzoru (4.9), dla dowolnych $n, k \geqslant N$ mamy

$$||M_n f - M_k f||_p \leqslant ||M_n f - M_n g||_p + ||M_n g - M_k g||_p + ||M_k g - M_k f||_p$$

$$\leqslant ||f - g||_p + ||M_n g - M_k g||_p + ||g - f||_p < \frac{\varepsilon}{4} + \frac{\varepsilon}{2} + \frac{\varepsilon}{4} = \varepsilon.$$

To oznacza, że ciąg $(M_n f)$ spełnia warunek Cauchy'ego w L^p. Zatem istnieje funkcja $f^* \in L^p$ taka, że $||M_n f - f^*||_p \to 0$, gdy $n \to \infty$.

Należy jeszcze pokazać, że $f^* \circ T = f^*$ prawie wszędzie. Stosujemy podobny pomysł jak w dowodzie Lematu 4.2. Zauważmy, że dla każdego $n \in \mathbb{N}$ oraz $x \in X$ zachodzi zależność

$$\left(\frac{n+1}{n}\right)(M_{n+1} f)(x) - (M_n f)(Tx) = \frac{f(x)}{n}.$$

Wystarczy w tej równości przyłożyć normę w L^p i przejść do granicy, gdy $n \to \infty$. Istotnie,

$$||M_{n+1}f - M_n(f \circ T)||_p \leqslant \left\| M_{n+1}f - \left(\frac{n+1}{n}\right) M_{n+1}f \right\|_p$$

$$+ \left\| \left(\frac{n+1}{n}\right) M_{n+1}f - M_n(f \circ T) \right\|_p = \frac{1}{n}||M_{n+1}f||_p + \frac{1}{n}||f||_p \to 0.$$

Stąd otrzymujemy $||f^* - f^* \circ T||_p = 0$, co daje żądany warunek. $\square$

Na koniec wspomnijmy o jednym zastosowaniu twierdzenia Birkhoffa w teorii liczb. Liczba $x \in [0,1]$ nazywa się *prosto normalna* przy podstawie 2, gdy graniczna częstość

$$\lim_{n\to\infty} \frac{\mathrm{card}\{1 \leqslant i \leqslant n \colon x_i = 0\}}{n}$$

wystąpienia cyfry 0 w rozwinięciu binarnym $(0, x_1\, x_2\, \dots)$ liczby x wynosi $1/2$. Liczba $x \in [0,1]$ nazywa się *normalna* przy podstawie 2, gdy dla dowolnego skończonego ciągu $(a_1, \dots, a_k)$ zer i jedynek graniczna częstość

$$\lim_{n\to\infty} \frac{\mathrm{card}\{1 \leqslant i \leqslant n \colon x_i = a_1, \dots, x_{i+k-1} = a_k\}}{n}$$

wystąpienia tego ciągu skończonego w rozwinięciu binarnym liczby x wynosi $1/2^k$.

Stosując w odpowiedni sposób twierdzenie Birkhoffa do odwzorowania podwajającego na $[0,1)$, otrzymujemy następujące twierdzenie (szczegóły można znaleźć w książce Silvy [11, s. 189–190].

Twierdzenie 4.4 (Borela)**.** *Prawie każda (w sensie miary Lebesgue'a) liczba przedziału* $[0,1]$ *jest normalna przy podstawie 2.*

Na koniec dodajmy, że w literaturze, także tej najnowszej, dowodzone są inne twierdzenia ergodyczne: [2], [4], [6], [7].

Rozdział 5

Odwzorowania mieszające

Z twierdzenia ergodycznego Birkhoffa możemy wywnioskować następującą charakteryzację odwzorowań ergodycznych:

Wniosek 5.1. *Załóżmy, że* $(X, \mathcal{S}, \mu)$ *jest przestrzenią probabilistyczną. Odwzorowanie* $T\colon X \to X$ *zachowujące miarę jest ergodyczne wtedy i tylko wtedy, gdy dla dowolnych* $A, B \in \mathcal{S}$ *mamy*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mu(T^{-i}[A] \cap B) = \mu(A)\mu(B). \tag{5.1}$$

Dowód. Konieczność. Załóżmy, że $T\colon X \to X$ jest ergodyczne. Niech $A, B \in \mathcal{S}$. Zastosujmy twierdzenie Birkhoffa do funkcji $f := \chi_A$. Wtedy

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \chi_A \circ T^i = \int_X \chi_A \, d\mu = \mu(A)$$

prawie wszędzie na X (por. uwagę po Twierdzeniu 4.1). Zatem prawie wszędzie na X zachodzi

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} (\chi_A \circ T^i)\chi_B = \mu(A)\chi_B.$$

Odnotujmy, że $(\chi_A \circ T^i)\chi_B = \chi_{T^{-i}[A]\cap B}$. Dlatego na mocy twierdzenia Lebesgue'a o ograniczonym przechodzeniu do granicy, mamy

$$\frac{1}{n} \sum_{i=0}^{n-1} \mu(T^{-i}[A] \cap B) = \frac{1}{n} \sum_{i=0}^{n-1} \int_X (\chi_A \circ T^i)\chi_B \, d\mu \to \mu(A) \int_X \chi_B \, d\mu = \mu(A)\mu(B),$$

gdy $n \to \infty$.

Dostateczność. Niech $E \in \mathcal{S}$ oraz $T^{-1}[E] = E$. Przyjmijmy $A = B = E$. Stosując założenie oraz równość $T^{-1}[E] = E$, mamy $\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n-1} \mu(E) = (\mu(E))^2$. Stąd $\mu(E) = (\mu(E))^2$, więc $\mu(E) = 1$ lub $\mu(E) = 0$. □

Rozważając warunki silniejsze niż warunek (5.1), możemy zdefiniować nowe rodzaje odwzorowań zachowujących miarę.

Definicja 5.1. Niech $(X, \mathcal{S}, \mu)$ będzie przestrzenią probabilistyczną. Mówimy, że odwzorowanie $T\colon X \to X$ zachowujące miarę jest:

- słabo mieszające, gdy dla dowolnych $A, B \in \mathcal{S}$ mamy
$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} |\mu(T^{-i}[A] \cap B) - \mu(A)\mu(B)| = 0; \tag{5.2}$$
- mieszające, gdy dla dowolnych $A, B \in \mathcal{S}$ mamy
$$\lim_{n\to\infty} \mu(T^{-n}[A] \cap B) = \mu(A)\mu(B). \tag{5.3}$$

Zależność między warunkami (5.1), (5.2), (5.3) opisującymi trzy rodzaje odwzorowań zachowujących miarę wynika z implikacji między trzema poniższymi rodzajami zbieżności ciągów liczbowych.

Niech dany będzie ciąg liczbowy $(a_n)_{n\geqslant 0}$ oraz $a \in \mathbb{R}$. Wspomniane rodzaje zbieżności to:

(i) $\lim\limits_{n\to\infty} a_n = a$ – zbieżność zwykła;

(ii) $\lim\limits_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} |a_i - a| = 0$ – silna zbieżność w sensie Cesàro;

(iii) $\lim\limits_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} a_i = a$ – zbieżność w sensie Cesàro.

Ćwiczenie 5.1. Wykazać, że zachodzą implikacje (i)⇒(ii) oraz (ii)⇒(iii). Podać przykłady ciągów, dla których implikacje odwrotne nie zachodzą.

Z powyższego ćwiczenia zastosowanego dla ciągu $a_n := \mu(T^{-n}[A] \cap B)$ oraz $a := \mu(A)\mu(B)$ otrzymujemy natychmiast następujący wniosek.

Wniosek 5.2. *Dla dowolnego odwzorowania $T\colon X \to X$ zachowującego miarę w przestrzeni probabilistycznej zachodzą implikacje:*

- *jeśli T jest mieszające, to jest słabo mieszające,*
- *jeśli T jest słabo mieszające, to jest ergodyczne.*

Odnotujmy, że istnieją odwzorowania zachowujące miarę świadczące o tym, że implikacje odwrotne w powyższym wniosku nie muszą zachodzić.

W pewnych sytuacjach warunki określające mieszanie i słabe mieszanie (w Definicji 5.1) oraz charakteryzujące ergodyczność (we Wniosku 5.1) dla odwzorowań zachowujących miarę można osłabić, rozważając je tylko dla odpowiednio wybranych zbiorów $A, B \in \mathcal{S}$. Aby to pokazać, udowodnimy najpierw lemat.

Lemat 5.1. *Niech* $(X, \mathcal{S}, \mu)$ *będzie przestrzenią z miarą skończoną oraz niech* $\mathcal{A}$ *będzie ciałem generującym* σ*-ciało* $\mathcal{S}$*. Wtedy dla każdego zbioru* $A \in \mathcal{S}$ *i dowolnego* $\varepsilon > 0$ *istnieje zbiór* $B \in \mathcal{A}$ *taki, że* $\mu(A \triangle B) < \varepsilon$*.*

Dowód. Niech

$$\mathcal{D} := \{A \in \mathcal{S} \colon \forall \varepsilon > 0 \; \exists C \in \mathcal{A} \;\; \mu(A \triangle C) < \varepsilon\}.$$

Oczywiście $\mathcal{A} \subset \mathcal{D} \subset \mathcal{S}$. Pokażemy, że $\mathcal{D}$ jest klasą monotoniczną. Rozważmy ciąg $A_1 \subset A_2 \subset \ldots$ zbiorów należących do $\mathcal{D}$ i niech $A := \bigcup_{n \in \mathbb{N}} A_n$. Wtedy $\mu(A) = \lim_n \mu(A_n)$. Dla $\varepsilon > 0$ dobierzmy $k \in \mathbb{N}$ takie, że $\mu(A \triangle A_k) = |\mu(A) - \mu(A_k)| < \varepsilon/2$. Ponieważ $A_k \in \mathcal{D}$, więc istnieje zbiór $C \in \mathcal{A}$ taki, że $\mu(A_k \triangle C) < \varepsilon/2$. Stąd mamy

$$\mu(A \triangle C) \leqslant \mu(A \triangle A_k) + \mu(A_k \triangle C) < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

Podobnie pokazuje się, że jeśli $A_1 \supset A_2 \supset \ldots$ jest ciągiem zbiorów należących do $\mathcal{D}$, to $\bigcap_{n \in \mathbb{N}} A_n \in \mathcal{D}$. Skoro $\mathcal{D}$ jest klasa monotoniczną zawierającą $\mathcal{A}$ i wiemy, że klasa monotoniczna generowana przez ciało $\mathcal{A}$ pokrywa się z σ-ciałem $\mathcal{S}$, to $\mathcal{S} \subset \mathcal{D}$. Stąd otrzymujemy tezę. $\square$

Niech będzie dane σ-ciało $\mathcal{S}$ podzbiorów X. Ciało $\mathcal{A} \subset \mathcal{S}$, które spełnia warunek

$$\forall A \in \mathcal{S} \;\; \forall \varepsilon > 0 \;\; \exists B \in \mathcal{A} \;\; \mu(A \triangle B) < \varepsilon,$$

nazywamy ciałem aproksymującym dla $\mathcal{S}$.

Uwaga 5.1. Nie każde ciało aproksymujące musi generować $\mathcal{S}$, czyli implikacja odwrotna w Lemacie 5.1 nie musi zachodzić. Na przykład σ-ciało zbiorów borelowskich aproksymuje σ-ciało $\mathcal{S}$ zbiorów mierzalnych w sensie Lebesgue'a na prostej $\mathbb{R}$, ale oczywiście nie generuje $\mathcal{S}$. Rozważmy inny przykład. Weźmy zbiór $X := [0, 1)$ z σ-ciałem $\mathcal{S}$ podzbiorów X mierzalnych w sensie Lebesgue'a oraz miarą Lebesgue'a λ. Niech $\mathcal{A}$ oznacza ciało złożone ze skończonych sum przedziałów $[a, b) \subset [0, 1)$ o końcach dwójkowo-wymiernych (postaci $k/2^n$ dla $n \in \mathbb{N} \cup \{0\}$, $k \in \{0, \ldots, 2^n\}$). Wówczas $\mathcal{A}$ generuje σ-ciało zbiorów borelowskich różne od $\mathcal{S}$. Zauważmy, że dla dowolnego zbioru $E \subset [0, 1)$ jego miarę zewnętrzną Lebesgue'a można zapisać w postaci

$$\inf \left\{ \sum_n |I_n| \colon E \subset \bigcup_n I_n \right\},$$

gdzie I_n są przedziałami postaci $[a, b) \subset [0, 1)$ o końcach dwójkowo-wymiernych. Niech $A \in \mathcal{S}$ oraz $\varepsilon > 0$. Dobierzmy taki ciąg (I_n) przedziałów powyższej postaci, że $A \subset \bigcup_n I_n$

oraz $\lambda(A) > \sum_n |I_n| - \varepsilon/2$. Znajdźmy $k \in \mathbb{N}$, takie aby $\sum_{n>k} |I_n| < \varepsilon/2$. Wówczas $B := \bigcup_{n\leqslant k} I_n \in \mathcal{A}$ oraz

$$\lambda(A\triangle B) \leqslant \lambda\left(A \setminus \bigcup_{n\leqslant k} I_n\right) + \lambda\left(\bigcup_{n\in\mathbb{N}} I_n \setminus A\right) \leqslant \sum_{n>k} |I_n| + \sum_{n\in\mathbb{N}} |I_n| - \lambda(A) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Ćwiczenie 5.2. Pokazać następującą własność różnicy symetrycznej zbiorów.

$$(A\cap B)\triangle(C\cap D) \subset (A\triangle C) \cup (B\triangle D).$$

Ćwiczenie 5.3. Pokazać, że w przestrzeni $(X, \mathcal{S}, \mu)$ z miarą skończoną zachodzi oszacowanie

$$|\mu(A) - \mu(A_0)| \leqslant \mu(A\triangle A_0) \;\text{ dla } A, A_0 \in \mathcal{S}.$$

Twierdzenie 5.1. (O aproksymacji) *Niech $T\colon X \to X$ będzie odwzorowaniem zachowującym miarę w przestrzeni probabilistycznej $(X, \mathcal{S}, \mu)$ oraz $\mathcal{A} \subset \mathcal{S}$ jest ciałem aproksymującym. Wówczas zachodzą równoważności:*

(i) *T jest ergodyczne wtedy i tylko wtedy, gdy*

$$\forall A, B \in \mathcal{A} \quad \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mu(T^{-i}[A] \cap B) = \mu(A)\mu(B);$$

(ii) *T jest słabo mieszające wtedy i tylko wtedy, gdy*

$$\forall A, B \in \mathcal{A} \quad \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} |\mu(T^{-i}[A] \cap B) - \mu(A)\mu(B)| = 0;$$

(iii) *T jest mieszające wtedy i tylko wtedy, gdy*

$$\forall A, B \in \mathcal{A} \quad \lim_{i\to\infty} \mu(T^{-i}[A] \cap B) = \mu(A)\mu(B).$$

Dowód. We wszystkich zdaniach (i), (ii), (iii) implikacje $\Rightarrow$ są oczywiste. Należy więc wykazać implikacje odwrotne. Aby je udowodnić, odnotujmy własność pomocniczą. Niech $A, B \in \mathcal{S}$ oraz $\varepsilon > 0$. Ponieważ $\mathcal{A}$ jest ciałem aproksymującym, więc wybierzmy zbiory takie $A_0, B_0 \in \mathcal{A}$, że $\mu(A\triangle A_0) < \varepsilon$ oraz $\mu(B\triangle B_0) < \varepsilon$. Wtedy dla dowolnego $k \geqslant 0$ mamy (por. Ćwiczenie 5.2)

$$\mu(T^{-k}[A] \cap B)\triangle(T^{-k}[A_0] \cap B_0) \leqslant \mu((T^{-k}[A]\triangle T^{-k}[A_0]) \cup (B\triangle B_0)) < 2\varepsilon. \tag{5.4}$$

(Ostatnia nierówność wynika m.in. stąd, że $T^{-k}[A\triangle A_0] = T^{-k}[A]\triangle T^{-k}[A_0]$ oraz z tego, że T^k zachowuje miarę.) Dla dowodu implikacji $\Leftarrow$ w (i) zauważmy, że dla dowolnego n, stosując (5.4) oraz Ćwiczenie 5.3, mamy

$$\left|\frac{1}{n}\sum_{i=0}^{n-1}\mu(T^{-i}[A]\cap B)-\mu(A)\mu(B)\right| \leqslant \left|\frac{1}{n}\sum_{i=0}^{n-1}\Big(\mu(T^{-i}[A]\cap B)-\mu(T^{-i}[A_0]\cap B_0)\Big)\right|$$
$$+\left|\frac{1}{n}\sum_{i=0}^{n-1}\mu(T^{-i}[A_0]\cap B_0)-\mu(A_0)\mu(B_0)\right| + |\mu(A_0)\mu(B_0)-\mu(A_0)\mu(B)|$$
$$+|\mu(A_0)\mu(B)-\mu(A)\mu(B)| \leqslant \frac{1}{n}\sum_{i=0}^{n-1}\left|\mu(T^{-i}[A]\cap B)-\mu(A_0)\mu(B_0)\right|$$
$$+\left|\frac{1}{n}\sum_{i=0}^{n-1}\mu(T^{-i}[A_0]\cap B_0)-\mu(T^{-i}[A_0]\cap B_0)\right| + \mu(A_0)\,|\mu(B)-\mu(B_0)|$$
$$+\mu(B)\,|\mu(A)-\mu(A_0)| < 4\varepsilon + \left|\frac{1}{n}\sum_{i=0}^{n-1}\mu(T^{-i}[A_0]\cap B_0)-\mu(A_0)\mu(B_0)\right| < 5\varepsilon,$$

przy czym ostatnia nierówność zachodzi dla dostatecznie dużych n na mocy założenia dowodzonej implikacji.

Dla dowodu implikacji $\Leftarrow$ w (iii) zauważmy, że dla dowolnego całkowitego $i \geqslant 0$, stosując (5.4), mamy

$$\left|\mu(T^{-i}[A]\cap B)-\mu(A)\mu(B)\right| \leqslant \left|\mu(T^{-i}[A]\cap B)-\mu(T^{-i}[A_0]\cap B_0)\right|$$
$$+\left|\mu(T^{-i}[A_0]\cap B_0)-\mu(A_0)\mu(B_0)\right| + |\mu(A_0)\mu(B_0)-\mu(A_0)\mu(B)|$$
$$+|\mu(A_0)\mu(B)-\mu(A)\mu(B)| < 2\varepsilon + \left|\mu(T^{-i}[A_0]\cap B_0)-\mu(A_0)\mu(B_0)\right|$$
$$+\mu(A_0)\,|\mu(B)-\mu(B_0)| + \mu(B)\,|\mu(A)-\mu(A_0)|$$
$$< 4\varepsilon + \left|\mu(T^{-i}[A_0]\cap B_0)-\mu(A_0)\mu(B_0)\right| < 5\varepsilon,$$

przy czym ostatnia nierówność zachodzi dla dostatecznie dużych i na mocy założenia dowodzonej implikacji.

Dowód implikacji $\Leftarrow$ w (ii) przebiega według podobnej idei. Korzystamy z poprzednich oszacowań i dopisując stronami $\frac{1}{n}\sum_{i=0}^{n-1}$, otrzymujemy dla dowolnego n

$$\frac{1}{n}\sum_{i=0}^{n-1}\left|\mu(T^{-i}[A]\cap B)-\mu(A)\mu(B)\right|$$
$$< 4\varepsilon + \frac{1}{n}\sum_{i=0}^{n-1}\left|\mu(T^{-i}[A_0]\cap B_0)-\mu(A_0)\mu(B_0)\right| < 5\varepsilon,$$

przy czym ostatnia nierówność zachodzi dla dostatecznie dużych n na mocy założenia dowodzonej implikacji. $\square$

Przykład 5.1. Rozważmy przesunięcie jednostronne $T\colon \{0,1\}^{\mathbb{N}} \to \{0,1\}^{\mathbb{N}}$, przy czym na $\mathcal{P}(\{0,1\})$ działa miara ν dana wzorami $\nu(\{0\}) = \nu(\{1\}) = 1/2$. Na zbiorze $X := \{0,1\}^{\mathbb{N}}$ rozważane jest σ-ciało produktowe $\mathcal{S}$ generowane przez zbiory cylindryczne postaci

$$\{(x_i) \in X\colon x_{i_1} = a_1, \dots, x_{i_k} = a_k\},$$

gdzie $k \in \mathbb{N}$, $i_1 < \cdots < i_k$ oraz $a_j \in \{0,1\}$ dla $j = 1, \dots, k$. Nazwijmy je bazowymi zbiorami cylindrycznymi. Na $\mathcal{S}$ zdefiniowana jest odpowiednia miara produktowa μ. Rodzina $\mathcal{A}$ złożona ze skończonych sum bazowych zbiorów cylindrycznych jest ciałem generującym σ-ciało $\mathcal{S}$. Pokażemy, że T jest odwzorowaniem mieszającym. W tym celu na mocy Twierdzenia 5.1 wystarczy wykazać warunek

$$\forall A, B \in \mathcal{A} \quad \lim_{n\to\infty} \mu(T^{-n}[A] \cap B) = \mu(A)\mu(B). \tag{5.5}$$

Udowodnimy go najpierw w szczególnym przypadku, gdy A, B są bazowymi zbiorami cylindrycznymi. Niech

$$A := \{(x_i) \in X\colon x_{i_1} = a_1, \dots, x_{i_k} = a_k\}, \quad B := \{(x_i) \in X\colon x_{j_1} = b_1, \dots, x_{j_l} = b_l\},$$

gdzie $i_1 < \cdots < i_k$, $a_i \in \{0,1\}$ oraz $j_1 < \cdots < j_l$, $b_j \in \{0,1\}$. Znajdziemy tak duży indeks $m \in \mathbb{N}$, że $i_1 + m > j_1$. Wtedy dla każdego $n \geqslant m$ otrzymujemy bazowy zbiór cylindryczny

$$T^{-n}[A] \cap B = \{(x_i) \in X\colon x_{j_1} = b_1, \dots, x_{j_l} = b_l, x_{i_1+n} = a_1, \dots, x_{i_k+n} = a_k\}.$$

Zatem $\mu(T^{-n}[A] \cap B) = (1/2)^{k+l} = \mu(A)\mu(B)$ dla $n \geqslant m$, co dowodzi warunku (5.5) w tym przypadku. Stąd jako nietrudny wniosek otrzymujemy warunek (5.5) w ogólnym przypadku, gdy $A, B \in \mathcal{A}$. Wystarczy skorzystać z obserwacji, że zbiór należący do $\mathcal{A}$ jest rozłączną skończoną sumą bazowych zbiorów cylindrycznych. Szczegóły pozostawiamy jako ćwiczenie.

Podobną metodą można pokazać, że odwzorowanie podwajające $T\colon [0,1) \to [0,1)$, $T(x) = 2x (\mathrm{mod}\, 1)$ jest mieszające. Na $[0,1)$ rozważamy σ-ciało podzbiorów $[0,1)$ mierzalnych w sensie Lebesgue'a oraz miarę Lebesgue'a λ. Odpowiedni warunek (5.5) sprawdzamy dla przedziałów $[a,b)$ o końcach dwójkowo-wymiernych, a następnie dla skończonych rozłącznych sum takich przedziałów. Sumy tej postaci, jak wiemy z Uwagi 5.1, stanowią ciało aproksymujące dla σ-ciała $\mathcal{S}$. Zatem wystarczy zastosować Twierdzenie 5.1. Por. [11], tw. 6.6.1.

Zaprezentujemy teraz ciekawą charakteryzację odwzorowań słabo mieszających. W tym celu pokażemy najpierw charakteryzację zbieżności w sensie Cesàro do 0 ciągów nieujemnych ograniczonych. Potrzebny będzie nam pewien abstrakcyjny rodzaj zbieżności ciągów liczbowych.

Mówimy, że zbiór $E \subset \{0, 1, \dots\}$ ma gęstość $\alpha \in [0, 1]$, gdy

$$d(E) := \lim_{n\to\infty} \frac{\operatorname{card}(E \cap \{0, \dots, n-1\})}{n} = \alpha,$$

o ile ta granica istnieje. Zauważmy, że

$$\frac{\operatorname{card}(E \cap \{0, \dots, n-1\})}{n} = \frac{1}{n} \sum_{i=0}^{n-1} \chi_E(i).$$

Zbiory o gęstości 0 stanowią ideał zbiorów; jest to rodzina zamknięta względem operacji brania podzbioru oraz operacji skończonych sum zbiorów.

Mówimy, że ciąg liczbowy $(a_n)_{n\geqslant 0}$ jest zbieżny statystycznie do $a \in \mathbb{R}$, gdy istnieje zbiór E o gęstości 0 taki, że $\lim_{n\notin E} a_n = a$, co oznacza, że dla każdego $\varepsilon > 0$ istnieje liczba $k \in \mathbb{N}$ taka, że $|a_n - a| < \varepsilon$ dla wszystkich $n \geqslant k$, $n \notin E$. Zbieżność statystyczną ciągu (a_n) do a będziemy zapisywać jako (st)-$\lim_n a_n = a$. Odnotujmy, że zbieżność statystyczna zachowuje wiele własności zwykłej zbieżności, np. własności algebraiczne dotyczące dodawania i mnożenia ciągów. Natomiast nie jest prawdą, że dowolny podciąg ciągu statystycznie zbieżnego jest statystycznie zbieżny. Opisane własności można znaleźć w artykule [10].

Twierdzenie 5.2. *Ograniczony ciąg $(a_n)_{n\geqslant 0}$ liczb nieujemnych jest zbieżny w sensie Cesàro do* 0 *wtedy i tylko wtedy, gdy jest statystycznie zbieżny do* 0.

Dowód. Niech $a_n \leqslant M$ dla wszystkich n oraz załóżmy, że ciąg (a_n) jest zbieżny statystycznie do 0, więc istnieje zbiór $E \subset \{0, 1, \dots\}$ o gęstości 0 taki, że $\lim_{n\notin E} a_n = 0$. Przyjmijmy

$$\widetilde{a_n} := \begin{cases} 0 & \text{dla } n \in E \\ a_n & \text{dla } n \notin E. \end{cases}$$

Wtedy $\lim_{n\notin E} a_n = 0$ implikuje $\lim_n \widetilde{a_n} = 0$, co z kolei implikuje $\lim_n \frac{1}{n}\sum_{i=0}^{n-1} \widetilde{a_i} = 0$ (por. Ćwiczenie 5.1), czyli $\lim_n \frac{1}{n}\sum_{i<n,\, i\notin E} a_i = 0$. Zatem wiedząc, że $d(E) = 0$, otrzymujemy

$$0 \leqslant \frac{1}{n}\sum_{i=0}^{n-1} a_i = \frac{1}{n}\sum_{i<n,\, i\in E} a_i + \frac{1}{n}\sum_{i<n,\, i\notin E} a_i \leqslant \frac{M}{n}\sum_{i=0}^{n-1} \chi_E(i) + \frac{1}{n}\sum_{i<n,\, i\notin E} a_i \to M d(E) = 0.$$

Widzimy więc, że ciąg (a_n) jest zbieżny do 0 w sensie Cesàro.

Odwrotnie, załóżmy, że ciąg (a_n) jest zbieżny do 0 w sensie Cesàro. Skonstruujemy odpowiedni zbiór J o gęstości 0. Dla każdego $k \in \mathbb{N}$ niech $J_k := \{i \geqslant 0\colon a_i \geqslant 1/k\}$.

Wtedy dla każdego $k \in \mathbb{N}$ mamy $J_k \subset J_{k+1}$ oraz $\frac{1}{k}\chi_{J_k} \leqslant a_i$, gdy $i \geqslant 0$. Zatem

$$0 \leqslant \frac{1}{k}\left(\lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} \chi_{J_k}(i)\right) = \lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} \frac{1}{k}\chi_{J_k}(i) \leqslant \lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} a_i = 0,$$

co dowodzi, że $d(J_k) = 0$ dla $k \in \mathbb{N}$. Połóżmy $n_0 := 0$ oraz indukcyjnie dla każdego $k \in \mathbb{N}$ wybierzmy taką liczbę naturalną $n_k > n_{k-1}$, że

$$\frac{1}{n}\sum_{i=0}^{n-1} \chi_{J_{k+1}}(i) < \frac{1}{k+1} \quad \text{dla każdego } n \geqslant n_k. \tag{5.6}$$

Niech $J := \bigcup_{k=1}^{\infty} J_k \cap [n_{k-1}, n_k)$. Pokażemy, że $d(J) = 0$. Niech $\varepsilon > 0$ i dobierzmy takie $k_0 \in \mathbb{N}$, że $2/k_0 < \varepsilon$. Weźmy dowolne $n \geqslant n_{k_0}$. Wtedy znajdziemy $k \geqslant k_0$ takie, że $n_k \leqslant n < n_{k+1}$. Zatem

$$J \cap [0, n) = (J \cap [0, n_k)) \cup (J \cap [n_k, n)) \subset (J_k \cap [0, n)) \cup (J_{k+1} \cap [0, n)).$$

Stąd korzystając z (5.6), mamy

$$\frac{1}{n}\sum_{i=0}^{n-1} \chi_J(i) = \frac{1}{n}\sum_{i=0}^{n_k-1} \chi_J(i) + \frac{1}{n}\sum_{i=n_k}^{n-1} \chi_J(i)$$

$$\leqslant \frac{1}{n}\sum_{i=0}^{n-1} \chi_{J_k}(i) + \frac{1}{n}\sum_{i=0}^{n-1} \chi_{J_{k+1}}(i) < \frac{1}{k} + \frac{1}{k+1} < \frac{2}{k_0} < \varepsilon.$$

To dowodzi, że $d(J) = 0$. Na koniec zauważmy, że $\lim_{n\notin J} a_n = 0$. Weźmy $\varepsilon > 0$ oraz dobierzmy k_0 jak wyżej. Niech $n \geqslant n_{k_0}$. Jeśli $n \notin J$, to $n \notin J_{k_0}$, a więc (z definicji J_{k_0}) mamy $a_n < 1/k_0 < \varepsilon$. □

Poniższe twierdzenie prezentuje zapowiadaną charakteryzację odwzorowań słabo mieszających.

Twierdzenie 5.3. *Niech T będzie odwzorowaniem zachowującym miarę w przestrzeni probabilistycznej $(X, \mathcal{S}, \mu)$. Następujące warunki są równoważne:*

(i) *T jest słabo mieszające;*

(ii) *dla dowolnych zbiorów $A, B \in \mathcal{S}$ zachodzi*

$$\text{(st)-}\lim_{n} \mu(T^{-n}[A] \cap B) = \mu(A)\mu(B);$$

(iii) *dla dowolnych zbiorów $A, B \in \mathcal{S}$ zachodzi*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} (\mu(T^{-i}[A] \cap B) - \mu(A)\mu(B))^2 = 0.$$

Dowód. Równoważność warunków (i) oraz (ii) jest bezpośrednią konsekwencją Twierdzenia 5.2 zastosowanego do ciągu $a_n := |\mu(T^{-i}[A] \cap B) - \mu(A)\mu(B)|$, $n \in \mathbb{N}$.

Pokażemy implikację (ii)$\Rightarrow$(iii). Z założenia (ii) nietrudno wnioskujemy, że

$$\text{(st)-}\lim_n (\mu(T^{-i}[A] \cap B) - \mu(A)\mu(B))^2 = 0,$$

gdyż zbieżność statystyczna przy mnożeniu dwóch ciągów zachowuje się podobnie jak zwykła zbieżność. Dalej stosujemy Twierdzenie 5.2 i otrzymujemy (iii). Dowód implikacji odwrotnej jest podobny. □

Uwaga 5.2. Na podstawie twierdzenia o aproksymacji (Tw. 5.1) możemy wnioskować, że w warunkach (ii), (iii) Twierdzenia 5.3 możemy rozważać jedynie dowolne zbiory A, B z ustalonego ciała aproksymującego dla σ-ciała $\mathcal{S}$.

Z powyższego dowodu wynika, że w warunku (iii) można zastąpić drugą potęgę w odpowiednim wyrażeniu przez dowolną potęgę dodatnią.

Niech $(X, \mathcal{S}, \mu)$ będzie przestrzenią z miarą. Powiemy, że odwzorowanie $T\colon X \to X$ zachowujące miarę jest podwójnie ergodyczne, gdy dla dowolnych zbiorów $A, B \in \mathcal{S}$ miary dodatniej istnieje liczba $n \in \mathbb{N}$ taka, że $\mu(T^{-n}[A] \cap A) > 0$ oraz $\mu(T^{-n}[A] \cap B) > 0$. Ten warunek definicyjny natychmiast implikuje ergodyczność odwzorowania T (por. warunek (iv) Twierdzenia 3.1). Można dowieść, że w przestrzeni probabilistycznej pojęcia słabego mieszania i podwójnej ergodyczności T są równoważne. Pokażemy tu jedną łatwiejszą implikację.

Twierdzenie 5.4. *Niech T będzie odwzorowaniem zachowującym miarę w przestrzeni probabilistycznej $(X, \mathcal{S}, \mu)$. Jeśli T jest słabo mieszające, to jest podwójnie ergodyczne.*

Dowód. Załóżmy, że T jest słabo mieszające i niech $A, B \in \mathcal{S}$ będą zbiorami miary dodatniej. Na podstawie Twierdzenia 5.2 dobierzmy takie dwa zbiory $E_1, E_2 \subset \mathbb{N}$ o gęstości zero, że

$$\lim_{n \notin E_1} \mu(T^{-n}[A] \cap B) = \mu(A)\mu(B), \qquad \lim_{n \notin E_2} \mu(T^{-n}[A] \cap A) = \mu(A)\mu(A).$$

Skoro $\mu(A) > 0$ oraz $\mu(B) > 0$, to z powyższych warunków wynika, że możemy wybrać liczbę $n \in \mathbb{N} \setminus (E_1 \cup E_2)$ na tyle dużą, że $\mu(T^{-n}[A] \cap A) > 0$ oraz $\mu(T^{-n}[A] \cap B) > 0$. □

Przykład 5.2. Pokażemy, że istnieją podzbiory mierzalne A, B przestrzeni $[0, 1)$ z σ-ciałem zbiorów mierzalnych w sensie Lebesgue'a i miarą Lebesgue'a, które świadczą o tym, że wszystkie niezerowe obroty nie są słabo mieszające. Na mocy Twierdzenia 5.4 wystarczy, gdy pokażemy, że zbiory te nie spełniają warunku podwójnej ergodyczności dla wszystkich niezerowych obrotów. Istotnie, niech $A := [0,\, 1/8)$ oraz $B := [1/2,\, 5/8)$. Rozważmy dowolny obrót $T(x) := x + c \pmod 1$ dla $c \in (0,1)$. Wtedy dla każdej liczby $n \in \mathbb{N}$ takiej, że $T^n[A] \cap B \neq \emptyset$ mamy $T^n[A] \cap A = \emptyset$ (sprawdzić, ćwiczenie). Zbiór wszystkich obrotów pokrywa się ze zbiorem obrotów odwrotnych. Z tej obserwacji wynika, że powyższa własność zbiorów A i B świadczy o tym, że każdy obrót niezerowy nie jest odwzorowaniem podwójnie ergodycznym. W szczególności obroty niewymierne (dla $c \notin Q$) są ergodyczne i nie są słabo mieszające.

Załóżmy, że odwzorowanie $T\colon X \to X$ zachowuje miarę w przestrzeni probabilistycznej $(X, \mathcal{S}, \mu)$. Rozważmy przestrzeń produktową $(X \times X,\, \mathcal{S} \otimes \mathcal{S},\, \mu \times \mu)$, gdzie $\mathcal{S} \otimes \mathcal{S}$ jest σ-ciałem produktowym, zaś $\mu \times \mu$ oznacza miarę produktową na tym σ-ciele. Odnotujmy, że σ-ciało $\mathcal{S} \otimes \mathcal{S}$ jest generowane przez rodzinę zbiorów postaci $A \times B$, gdzie $A, B \in \mathcal{S}$. Rozważmy odwzorowanie $T \times T\colon X^2 \to X^2$ dane wzorem $(T \times T)(x, y) := (T(x), T(y))$. Zauważmy, że $T \times T$ zachowuje miarę, bo zachowuje miarę zbiorów postaci $A \times B$ (dla $A, B \in \mathcal{S}$) i ich skończonych sum, a te skończone (rozłączne) sumy tworzą ciało $\mathcal{A}$ generujące σ-ciało $\mathcal{S} \otimes \mathcal{S}$.

Następujące twierdzenie daje jeszcze jedną charakteryzację odwzorowań słabo mieszających.

Twierdzenie 5.5. *Załóżmy, że odwzorowanie $T\colon X \to X$ zachowuje miarę w przestrzeni probabilistycznej $(X, \mathcal{S}, \mu)$. Następujące warunki są równoważne:*

(a) *T jest słabo mieszające;*

(b) *$T \times T$ jest słabo mieszające;*

(c) *$T \times T$ jest ergodyczne.*

Dowód. (a)$\Rightarrow$(b) Załóżmy, że T jest słabo mieszające. Wystarczy pokazać, że warunek typu (ii) w Twierdzeniu 5.3 zachodzi dla dowolnych dwóch zbiorów postaci $A \times B$ (gdzie $A, B \in \mathcal{S}$), gdyż implikuje to, że zachodzi on także dla dowolnych dwóch zbiorów z ciała $\mathcal{A}$ (o którym wspominamy wyżej) generującego σ-ciało $\mathcal{S} \otimes \mathcal{S}$. Weźmy więc dwa zbiory $A \times C$ oraz $B \times D$, gdzie $A, B, C, D \in \mathcal{S}$. Z założenia T jest słabo mieszające, więc zachodzi warunek (i) Twierdzenia 5.3 dla zbiorów A, B i dla zbiorów C, D. Istnieją zatem takie zbiory J_1, J_2 o gęstości zero, że

$$\lim_{n \notin J_1} \mu(T^{-n}[A] \cap B) = \mu(A)\mu(B), \quad \lim_{n \notin J_2} \mu(T^{-n}[C] \cap D) = \mu(C)\mu(D).$$

Wtedy zbiór $J_1 \cup J_2$ ma gęstość zero oraz

$$\lim_{n \notin J_1 \cup J_2} (\mu \times \mu)((T \times T)^{-n}[A \times C] \cap (B \times D)) = \lim_{n \notin J_1 \cup J_2} (\mu \times \mu)((T^{-n}[A] \times T^{-n}[C]) \cap (B \times D))$$

$$= \lim_{n \notin J_1 \cup J_2} (\mu\times\mu)((T^{-n}[A]\cap B)\times(T^{-n}[C])\cap D)) = \lim_{n \notin J_1 \cup J_2} \mu(T^{-n}[A]\cap B)\,\mu(T^{-n}[C])\cap D)$$

$$= \mu(A)\,\mu(B)\,\mu(C)\,\mu(D) = (\mu\times\mu)(A\times C)\,(\mu\times\mu)(B\times D).$$

To daje warunek (ii) Twierdzenia 5.3 dla dowolnych dwóch zbiorów typu $A \times B$ (gdzie $A, B \in \mathcal{S}$) i w konsekwencji dla dowolnych dwóch zbiorów z ciała $\mathcal{A}$.

Implikacja (b)⇒(c) jest oczywista.

(c)⇒(a) Niech $A, B \in \mathcal{S}$. Zauważmy, że dla dowolnej liczby całkowitej $i \geqslant 0$, liczbę $\mu(T^{-i}[A] \cap B)$, mnożąc ją przez $\mu(X) = 1$, można zapisać w postaci

$$(\mu\times\mu)((T^{-i}[A]\cap B)\times X) = (\mu\times\mu)((T\times T)^{-i}[A\times X]\cap(B\times X)).$$

Stąd stosując (c), otrzymujemy gdy $n \to \infty$,

$$\frac{1}{n}\sum_{i=0}^{n-1}\mu(T^{-i}[A]\cap B) = \frac{1}{n}\sum_{i=0}^{n-1}(\mu\times\mu)((T\times T)^{-i}[A\times X]\cap(B\times X)) \to \mu(A)\mu(B). \quad (5.7)$$

Mamy też

$$(T^{-i}[A]\cap B)\times(T^{-i}[A]\cap B) = (T^{-i}[A]\times T^{-i}[A])\cap(B\times B) = ((T\times T)^{-i}[A\times A])\cap(B\times B).$$

Zatem przykładając $\mu \times \mu$ do skrajnych wyrażeń w powyższych równościach, mamy

$$\mu^2(T^{-i}[A]\cap B) = (\mu\times\mu)(((T\times T)^{-i}[A\times A])\cap(B\times B)).$$

Stąd dopisując $\frac{1}{n}\sum_{i=0}^{n-1}$ i stosując (c), dla $n \to \infty$ otrzymujemy

$$\frac{1}{n}\sum_{i=0}^{n-1}\mu^2(T^{-i}[A]\cap B) = \frac{1}{n}\sum_{i=0}^{n-1}(\mu\times\mu)(((T\times T)^{-i}[A\times A])\cap(B\times B)) \to \mu^2(A)\mu^2(B). \quad (5.8)$$

Aby wykazać, że T jest słabo mieszające, zastosujemy warunek (iii) Twierdzenia 5.3. Badamy granicę wyrażenia $\frac{1}{n}\sum_{i=0}^{n-1}\big(\mu(T^{-i}[A]\cap B) - \mu(A)\mu(B)\big)^2$, gdy $n \to \infty$. Wyrażenie to przekształcamy do poniższej postaci. Przechodząc do granicy, gdy $n \to \infty$, na mocy (5.7) oraz (5.8), otrzymujemy

$$\frac{1}{n}\sum_{i=0}^{n-1}\mu^2(T^{-i}[A]\cap B) - 2\mu(A)\mu(B)\frac{1}{n}\sum_{i=0}^{n-1}\mu(T^{-i}[A]\cap B) + \mu^2(A)\mu^2(B)$$

$$\to \mu^2(A)\mu^2(B) - 2\mu^2(A)\mu^2(B) + \mu^2(A)\mu^2(B) = 0.$$

To daje warunek (a). □

Rozdział 6

Charakteryzacje w zespolonej przestrzeni Hilberta L^2

6.1 Charakteryzacje odwzorowań ergodycznych w języku iloczynu skalarnego

Dla przestrzeni probabilistycznej $(X, \mathcal{S}, \mu)$ będziemy rozważać zespoloną przestrzeń L^2. Jest to nie tylko przestrzeń Banacha, ale też przestrzeń Hilberta, w której dany jest iloczyn skalarny $(\cdot, \cdot)\colon L^2 \times L^2 \to \mathbb{C}$ określony wzorem

$$(f, g) := \int_X f\overline{g}\, d\mu.$$

Przypomnijmy, kiedy przestrzeń liniowa H nad ciałem $\mathbb{C}$ nazywa się (zespoloną) przestrzenią Hilberta. Zakłada się, że dany jest iloczyn skalarny $(\cdot, \cdot)\colon H \times H \to \mathbb{C}$ spełniający warunki:

- dla każdego $y \in H$ odwzorowanie $x \mapsto (x, y)$ jest liniowe;
- $\overline{(x, y)} = (y, x)$ dla dowolnych $x, y \in H$;
- $(x, x) \geqslant 0$ dla każdego $x \in H$;
- $(x, x) = 0 \Leftrightarrow x = 0$ dla dowolnego $x \in H$.

Można udowodnić, że $|(x, y)| \leqslant \sqrt{(x, x)}\sqrt{(y, y)}$ (nierówność Schwarza) dla dowolnych $x, y \in H$ oraz że funkcja $||x|| := \sqrt{(x, x)}$ jest normą na H. Jeśli przestrzeń H z tą normą jest zupełna, to nazywa się przestrzenią Hilberta.

Jeśli $(x, y) = 0$, to mówimy, że wektory x, y są ortogonalne i piszemy $x \perp y$. Dla $A \subset H$ symbol $x \perp A$ oznacza, że $x \perp y$ dla każdego $y \in A$. Jeśli $F \subset H$ jest podprzestrzenią (liniową) przestrzeni H, to zbiór

$$F^{\perp} := \{h \in H : h \perp F\}$$

nazywamy ortogonalnym dopełnieniem podprzestrzeni F w H. Można wykazać, że $F^{\perp}$ jest domkniętą podprzestrzenią przestrzeni H oraz $F \cap F^{\perp} = \{0\}$. Co więcej, jeśli F jest domkniętą podprzestrzenią przestrzeni H, to $H = F + F^{\perp}$, przy czym każdy $x \in H$ ma jedyne przedstawienie w postaci $x = y + z$, gdzie $x \in F$, $y \in F^{\perp}$ (twierdzenie Riesza o rozkładzie).

Wracając do zespolonej przestrzeni Hilberta L^2, odnotujmy, że nierówność Schwarza dla niej ma postać $|(f, g)| \leqslant ||f||_2 \, ||g||_2$, gdy $f, g \in L^2$. Normę $|| \cdot ||_2$ w L^2 będziemy dalej zapisywać krótko przez $|| \cdot ||$. Załóżmy, że dane jest odwzorowanie $T\colon X \to X$ zachowujące miarę. Przypomnijmy znany z wcześniejszych rozważań operator liniowy $U_T\colon L^2 \to L^2$ dany wzorem $U_T(f) := f \circ T$. Będziemy zapisywać krótko $U_T f$ zamiast $U_T(f)$. Przez 1 oznaczamy funkcję stałą w L^2 równą 1.

Twierdzenie 6.1. *Załóżmy, że w przestrzeni z prawdopodobieństwem $(X, \mathcal{S}, \mu)$ dane jest odwzorowanie $T\colon X \to X$ zachowujące miarę. Wówczas*

(a) *Następujące warunki są równoważne:*

- (i) *T jest ergodyczne;*
- (ii) $\lim\limits_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} (U_T^i f, g) = (f, 1)\,(1, g)$ *dla każdego* $f, g \in L^2$*;*
- (iii) $\lim\limits_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} (U_T^i f, f) = (f, 1)\,(1, f)$ *dla każdego* $f \in L^2$.

(b) *Następujące warunki są równoważne:*

- (i) *T jest słabo mieszające;*
- (ii) $\lim\limits_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} |(U_T^i f, g) - (f, 1)\,(1, g)| = 0$ *dla każdego* $f, g \in L^2$*;*
- (iii) $\lim\limits_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} |(U_T^i f, f) - (f, 1)\,(1, f)| = 0$ *dla każdego* $f \in L^2$.

(c) *Następujące warunki są równoważne:*

- (i) *T jest mieszające;*
- (ii) $\lim\limits_{n\to\infty} (U_T^n f, g) = (f, 1)\,(1, g)$ *dla każdego* $f, g \in L^2$*;*
- (iii) $\lim\limits_{n\to\infty} (U_T^n f, f) = (f, 1)\,(1, f)$ *dla każdego* $f \in L^2$.

Dowód. Dowody tez (a), (b), (c) są podobne. Pokażemy przykładowo tezę (c).

(ii)$\Rightarrow$(i) Niech $A, B \in \mathcal{S}$ oraz $f := \chi_A$, $g := \chi_B$. Wtedy wartość iloczynu skalarnego $(U_T^n f, g)$ wynosi

$$\int_X (U_T^n f)\, \overline{g}\, d\mu = \int_X (U_T^n \chi_A)\, \chi_B\, d\mu = \int_X (\chi_A \circ T^n)\, \chi_B\, d\mu$$

$$= \int_X \chi_{T^{-n}[A]}\, \chi_B\, d\mu = \int_X \chi_{T^{-n}[A]\cap B}\, d\mu = \mu(T^{-n}[A] \cap B).$$

Ponadto

$$(f, 1)\,(1, g) = \int_X \chi_A\, d\mu \int_X \chi_B\, d\mu = \mu(A)\mu(B).$$

Na mocy założenia (ii) dostajemy więc $\lim_n \mu(T^{-n}[A] \cap B) = \mu(A)\mu(B)$, co oznacza, że odwzorowanie T jest mieszające.

(i)$\Rightarrow$(iii) Niech $A, B \in \mathcal{S}$. Stosując (i) i naśladując powyższe rozumowanie, dostajemy łatwo

$$\lim_{n\to\infty} (U_T^n \chi_A, \chi_B) = (\chi_A, 1)\,(1, \chi_B).$$

Stąd ustalając zbiór B i korzystając z liniowości operatora U_T^n oraz iloczynu skalarnego, otrzymujemy dla dowolnej funkcji prostej h

$$\lim_{n\to\infty} (U_T^n h, \chi_B) = (h, 1)\,(1, \chi_B).$$

(Przez funkcję prostą rozumiemy kombinację liniową funkcji charakterystycznych.) Następnie ustalając funkcję prostą h, na mocy liniowości iloczynu skalarnego, dostajemy

$$\lim_{n\to\infty} (U_T^n h, h) = (h, 1)\,(1, h). \tag{6.1}$$

Weźmy teraz dowolną funkcję $f \in L^2$ oraz $\varepsilon > 0$. Wybierzmy taką funkcję prostą h, że $||f-h|| < \varepsilon$. Na mocy (6.1) dobierzmy $k \in \mathbb{N}$ tak, aby $|(U_T^n h, h) - (h, 1)\,(1, h)| < \varepsilon$ dla każdego $n \geqslant k$. Stąd korzystając z nierówności Schwarza oraz równości $||U_T^n g|| = ||g||$ dla $g \in L^2$, otrzymujemy dla wszystkich $n \geqslant k$ następujące oszacowania:

$$|(U_T^n f, f) - (f, 1)\,(1, f)| \leqslant |(U_T^n f, f) - (U_T^n h, f)| + |(U_T^n h, f) - (U_T^n h, h)|$$

$$+|(U_T^n h, h) - (h, 1)\,(1, h)| + |(h, 1)\,(1, h) - (f, 1)\,(1, h)| + |(f, 1)\,(1, h) - (f, 1)\,(1, f)|$$

$$\leqslant |(U_T^n (f-h), f)| + |(U_T^n h, f-h)| + \varepsilon + |(h-f, 1)|\,|(1, h)| + |(f, 1)|\,|(1, h-f)|$$

$$\leqslant ||f-h||\,||f|| + ||f-h||\,||h|| + \varepsilon + ||h||\,||f-h|| + ||f||\,||h-f||$$

$$\leqslant 2\varepsilon||f|| + 2\varepsilon||h|| + \varepsilon \leqslant 2\varepsilon||f|| + 2\varepsilon||h-f|| + 2\varepsilon||f|| + \varepsilon \leqslant 4\varepsilon||f|| + 2\varepsilon^2 + \varepsilon.$$

Stąd $\lim\limits_{n\to\infty} (U_T^n f, f) = (f, 1)\,(1, f)$.

(iii)$\Rightarrow$(ii) Niech $f \in L^2$. Oznaczmy przez H_f najmniejszą (w sensie inkluzji) domkniętą podprzestrzeń (liniową) przestrzeni L^2 zawierającą f i wszystkie funkcje stałe oraz taką, że $U_T[H_f] \subset H_f$. Taka przestrzeń jest przecięciem wszystkich domkniętych przestrzeni o powyższych własnościach (jedną z nich jest L^2). Połóżmy

$$F_f := \left\{ g \in L^2 \colon \lim_{n\to\infty} (U_T^n f, g) = (f,1)(1,g) \right\}.$$

Nietrudno sprawdzić, że F_f jest podprzestrzenią liniową przestrzeni L^2 zawierającą wszystkie funkcje stałe oraz funkcję f (na mocy założenia (iii)). Ponadto zachodzą warunki:

1^0 podprzestrzeń F_f jest domknięta;

2^0 $U_T[F_f] \subset F_f$.

Załóżmy, że pokazaliśmy warunki 1^0 i 2^0. Wtedy $H_f \subset F_f$. Dodatkowo $H_f^\perp \subset F_f$. Istotnie, niech $g \in L^2$ oraz $g \perp H_f$. Wtedy $(U_T^n f, g) = 0$ dla każdego $n \in \mathbb{N}$ (bo $U_T^n[H_f] \subset H_f$) oraz $(1,g) = 0$ (bo H_f zawiera funkcje stałe). Zatem

$$\lim_{n\to\infty} (U_T^n f, g) = 0 = (f,1)\,(1,g),$$

więc $g \in F_f$. W konsekwencji $H_f + H_f^\perp \subset F_f$, co na mocy twierdzenia o rozkładzie daje $L^2 = F_f$, a stąd otrzymujemy tezę (ii).

Pokażemy teraz warunek 1^0. Weźmy ciąg (g_k) elementów zbioru F_f zbieżny w sensie normy do $g \in L^2$. Wykażemy, że $g \in F_f$. Ustalmy $n, k \in \mathbb{N}$. Wtedy z nierówności Schwarza i własności operatora U_T, wnosimy, że $|(U_T^n f, g) - (f,1)\,(1,g)| \leqslant$

$$|(U_T^n f, g) - (U_T^n f, g_k)| + |(U_T^n f, g_k) - (f,1)\,(1,g_k)| + |(f,1)\,(1,g_k) - (f,1)\,(1,g)|$$

$$= |(U_T^n f, g - g_k)| + |(U_T^n f, g_k) - (f,1)\,(1,g_k)| + |(f,1)|\,|(1, g_k - g)|$$

$$\leqslant ||U_T^n f||\,||g - g_k)|| + |(U_T^n f, g_k) - (f,1)\,(1,g_k)| + ||f||\,||g - g_k||$$

$$= 2||f||\,||g - g_k|| + |(U_T^n f, g_k) - (f,1)\,(1,g_k)|.$$

Niech $\varepsilon > 0$ i dobierzmy $k \in \mathbb{N}$ tak, by $2||f||\,||g - g_k|| < \varepsilon/2$. Skoro $g_k \in F_f$, to następnie dobierzmy $N \in \mathbb{N}$ tak, by

$$|(U_T^n f, g_k) - (f,1)\,(1,g_k)| < \frac{\varepsilon}{2} \quad \text{dla każdego } n \geqslant N.$$

Wtedy z poprzednich oszacowań mamy

$$|(U_T^n f, g) - (f,1)\,(1,g)| < \varepsilon \quad \text{dla każdego } n \geqslant N.$$

Zatem $g \in F_f$, co kończy dowód warunku 1^0.

Na zakończenie udowodnimy warunek 2^0. Weźmy dowolne $g \in F_f$. Wtedy $U_T g \in F_f$. Istotnie, mamy dla $n \to \infty$

$$(U_T^n f, U_T g) = \int_X (f \circ T^n)(x)\,\overline{(g \circ T)(x)}\,d\mu(x) = \int_X (f \circ T^{n-1})(y)\,\overline{g(y)}\,d\mu(y) = (U_T^{n-1} f, g)$$

$$\to (f,1)\,(1,g) = (1,f)\int_X \overline{g(y)}\,d\mu(y) = (f,1)\int_X \overline{(g \circ T)(x)}\,d\mu(x) = (f,1)\,(1,U_T g).$$

□

6.2 Własności spektralne operatora U_T

Ważnym narzędziem do badania odwzorowań zachowujących miarę T w przestrzeni probabilistycznej są własności spektralne operatora U_T.

Przypomnijmy, że dla przestrzeni liniowej V nad ciałem $\mathbb{C}$ oraz operatora liniowego $A\colon V \to V$ niezerowy wektor $v \in V$ nazywamy wektorem własnym odwzorowania A, gdy $Av = \lambda v$ dla pewnego skalara $\lambda \in \mathbb{C}$; wtedy skalar λ nazywamy wartością własną odwzorowania A.

Oznaczmy przez Id_V odwzorowanie identycznościowe na V. Zauważmy, że dla ustalonej wartości własnej λ zbiór $\{v \in V\colon Av = \lambda v\}$ jest jądrem odwzorowania liniowego $A - \lambda\,\mathrm{Id}_X$, zatem jest podprzestrzenią liniową przestrzeni V. Tę podprzestrzeń będziemy oznaczać przez E_λ.

Jeśli przestrzeń V jest skończenie wymiarowa, to zbiór wartości własnych operatora A nazywamy widmem (spektrum) operatora A. Gdy V jest nieskończenie wymiarowa, widmo operatora A definiuje się inaczej – może być to istotnie szerszy zbiór od zbioru wartości własnych operatora A.

Będziemy zakładać, że $T\colon X \to X$ jest odwzorowaniem zachowującym miarę w przestrzeni probabilistycznej $(X, \mathcal{S}, \mu)$. Rozważmy operator $U_T\colon L^2 \to L^2$ dany wzorem $U_T(f) := f \circ T$. Mówimy, że funkcja $f \in L^2$, różna od funkcji zerowej prawie wszędzie (odpowiednio, liczba $\lambda \in \mathbb{C}$) jest funkcją własną (odpowiednio, wartością własną) odwzorowania T, gdy f jest wektorem własnym (λ jest wartością własną) operatora U_T. Zatem liczba λ jest wartością własną odwzorowania T odpowiadającą funkcji własnej f, gdy $f \circ T = \lambda f$ prawie wszędzie.

Uwaga 6.1. Oczywiście liczba $\lambda := 1$ jest zawsze wartością własną odwzorowania T, dla funkcji własnej $f := 1$. Ponadto każda wartość własna λ odwzorowania T leży na okręgu jednostkowym K w $\mathbb{C}$. Aby to wykazać, załóżmy, że $f \circ T = \lambda f$ prawie wszędzie dla $f \in L^2 \setminus \{0\}$. Skoro T zachowuje miarę, to

$$\int_X |f|^2\, d\mu = \int_X |f\circ T|^2\, d\mu = \int_X |\lambda f|^2\, d\mu = |\lambda|^2 \int_X |f|^2\, d\mu,$$

skąd wnioskujemy, że $|\lambda| = 1$, bo $\int_X |f|^2\, d\mu > 0$.

Uwaga 6.2. Załóżmy, że funkcja własna $f \in L^2 \setminus \{0\}$ odwzorowania T odpowiada wartości własnej λ. Niech $N(f) := \{x\colon f(x) \neq 0\}$ będzie nośnikiem funkcji f. Wtedy funkcja $h := \frac{f}{|f|}\chi_{N(f)}$ też jest funkcją własną odpowiadającą wartości własnej λ. Istotnie, na zbiorze $N(f)$ zachodzą równości

$$\left(\frac{f}{|f|}\right)\circ T = \frac{f\circ T}{|f|\circ T} = \frac{\lambda f}{|\lambda|\,|f|} = \frac{\lambda f}{|f|},$$

gdyż $|\lambda| = 1$ na mocy Uwagi 6.1. Dlatego $h \circ T = \lambda h$. Zatem dla ustalonej wartości własnej λ odwzorowania T odpowiedni zbiór E_λ zawsze zawiera taką funkcję $h \in L^2\setminus\{0\}$, że $|h(x)| = 1$ dla $x \in N(h)$.

Uwaga 6.3. Jeśli $f \in L^2$ jest funkcją własną odwzorowania T odpowiadającą wartości własnej λ_1, zaś $g \in L^2$ jest funkcją własną odwzorowania T odpowiadającą wartości własnej $\lambda_2 \neq \lambda_1$, to f i g są ortogonalne. Istotnie,

$$\lambda_1\overline{\lambda_2}\int_X f\overline{g}\, d\mu = \int_X (\lambda_1 f)(\overline{\lambda_2 g})\, d\mu = \int_X (f\circ T)(\overline{g\circ T})\, d\mu = \int_X f\overline{g}\, d\mu.$$

Stąd $\lambda_1\overline{\lambda_2} = 1$ lub $\int_X f\overline{g}\, d\mu = 0$. Ale $\lambda_1\overline{\lambda_2} = 1$ oraz $|\lambda_1| = 1 = |\lambda_2|$ implikuje $\lambda_1 = \lambda_2$ (sprawdzić). Zatem f, g są ortogonalne.

Twierdzenie 6.2. *Załóżmy, że $T\colon X \to X$ jest odwzorowaniem ergodycznym w przestrzeni probabilistycznej $(X, \mathcal{S}, \mu)$. Wtedy zachodzą następujące własności:*

(i) *Jeśli $f \in L^2$ jest funkcją własną odwzorowania T, to funkcja $|f|$ jest prawie wszędzie stała. Ponadto jeśli T jest słabo mieszające, to funkcja f jest prawie wszędzie stała.*

(ii) *Dla każdej wartości własnej λ odwzorowania T przestrzeń liniowa $E_\lambda \subset L^2$ jest 1-wymiarowa.*

(iii) *Zbiór $E(T)$ wszystkich wartości własnych odwzorowania T jest podgrupą grupy multiplikatywnej okręgu jednostkowego $K \subset \mathbb{C}$.*

Dowód. Ad. (i). Niech $f \in L^2$ będzie funkcją własną odwzorowania T. Z założenia i Uwagi 6.1 wynika, że istnieje taka liczba $\lambda \in \mathbb{C}$, że $|\lambda| = 1$ oraz $f\circ T = \lambda f$ prawie wszędzie. W związku z tym prawie wszędzie zachodzą równości

$$|f|\circ T = |f\circ T| = |\lambda f| = |\lambda|\,|f| = |f|.$$

Dlatego z ergodyczności T wynika, że funkcja $|f|$ jest prawie wszędzie stała.

Dla dowodu drugiej części tezy załóżmy, że T jest słabo mieszające. Pamiętamy, że jest to równoważne ergodyczności odwzorowania $T \times T \colon X^2 \to X^2$. Niech $g(x, y) := f(x)\overline{f(y)}$ dla $x, y \in X$. Wówczas g należy do odpowiedniej przestrzeni $L^2(X^2)$, gdzie na X^2 rozważamy σ-ciało produktowe i miarę produktową $\mu \times \mu$. Ponadto g jest funkcją $(T \times T)$-niezmienniczą, gdyż dla prawie wszystkich $(x, y) \in X^2$ mamy

$$(g \circ (T \times T))(x, y) = (f \circ T)(x)\overline{(f \circ T)(y)} = \lambda f(x)\overline{\lambda f(y)} = |\lambda|^2 g(x, y) = g(x, y).$$

Stąd i ze znanego warunku równoważnego ergodyczności odwzorowania $T \times T$ wynika, że funkcja g jest stała prawie wszędzie na X^2, równa $c \in \mathbb{C}$. Zauważmy, że $c \neq 0$, bo $|f|$ jest stała prawie wszędzie na X i niezerowa, więc $|g(x, y)| = |f(x)||\overline{f(y)}| \neq 0$ dla prawie wszystkich $(x, y) \in X^2$. Z twierdzenia Fubiniego wynika, że

$$c = \int_{X^2} g \, d(\mu \times \mu) = \int_X f \, d\mu \int_X \overline{f} \, d\mu = \left| \int_X f \, d\mu \right|^2$$

Ponadto dla prawie każdego $x \in X$ istnieje całka

$$\int_X g(x, y) \, d\mu(y) = f(x) \int_X \overline{f(y)} \, d\mu(y) = f(x)a,$$

gdzie $a := \int_X \overline{f} \, d\mu \neq 0$. Jednocześnie $\int_X g(x, y) \, d\mu(y) = \int_X c \, d\mu = c$ dla prawie wszystkich x. Stąd $f = c/a$ prawie wszędzie, co kończy dowód.

Ad. (ii). Ustalmy wartość własną λ odwzorowania T oraz $f \in E_\lambda$. Weźmy dowolną funkcję $g \in E_\lambda$. Z tezy (i) wiemy, że funkcje $|f|$ i $|g|$ są prawie wszędzie stałe. Połóżmy $h := g/f$. Wtedy $|h|$ jest prawie wszędzie stała i niezerowa. Ponadto prawie wszędzie na X zachodzą równości

$$h \circ T = \frac{g \circ T}{f \circ T} = \frac{\lambda g}{\lambda f} = h.$$

Stąd i z ergodyczności T wynika, że $h = c$ prawie wszędzie, gdzie $c \in \mathbb{C}$ jest stałą. To daje tezę.

Ad. (iii). Niech $\lambda_1, \lambda_2 \in E(T)$. Wybierzmy funkcje $f_1 \in E_{\lambda_1}$, $f_2 \in E_{\lambda_2}$, takie że $|f_1| = 1$ i $|f_2| = 1$ prawie wszędzie (por. tezę (i) oraz Uwagę 6.2). Wtedy $f_1 f_2 \in L^2$ oraz

$$(f_1 f_2) \circ T = (f_1 \circ T)(f_2 \circ T) = (\lambda_1 f_1)(\lambda_2 f_2) = (\lambda_1 \lambda_2)(f_1 f_2).$$

Stąd $\lambda_1 \lambda_2 \in E(T)$. Ponadto $\frac{1}{f_1} \in L^2$, bo $\left|\frac{1}{f_1}\right| = 1$ prawie wszędzie oraz

$$\left(\frac{1}{f_1}\right) \circ T = \frac{1}{f_1 \circ T} = \frac{1}{\lambda_1 f_1} = \left(\frac{1}{\lambda_1}\right) \frac{1}{f_1}.$$

Stąd $\frac{1}{\lambda_1} \in E(T)$. Wobec tego $E(T)$ jest podgrupą grupy multiplikatywnej K. $\square$

Wniosek 6.1. *Załóżmy, że* $(X, \mathcal{S}, \mu)$ *jest przestrzenią probabilistyczną i odpowiednia przestrzeń* L^2 *jest ośrodkowa. Jeśli odwzorowanie* $T\colon X \to X$ *jest ergodyczne, to grupa* $E(T)$ *jest przeliczalna.*

Dowód. Przypuśćmy, że grupa $E(T)$ jest nieprzeliczalna i każdej wartości własnej $\lambda \in E(T)$ przypiszmy funkcję własną $f \in L^2$, o której możemy zakładać, że $||f|| = 1$ (gdyż na mocy Uwagi 6.2 i tezy (i) Twierdzenia 6.2 można wybrać f tak, aby $|f| = 1$ prawie wszędzie). Każde takie dwie różne funkcje własne f, g są ortogonalne na mocy Uwagi 6.3. Ponadto

$$||f-g||^2 = (f-g,\, f-g) = ||f||^2 + ||g||^2 - (f,g) - (g,f) = 2. \tag{6.2}$$

Oczywiście podprzestrzeń Z przestrzeni L^2, złożona z funkcji własnych f przypisanych wartościom własnym $\lambda \in E(T)$ i takich, że $|f| = 1$ prawie wszędzie, jest ośrodkowa. Ze wzoru (6.2) wynika, że kule w podprzestrzeni Z o środkach w tych funkcjach f i promieniach $\sqrt{2}/2$ stanowią rozłączną rodzinę nieprzeliczalną. Zatem żaden zbiór gęsty w podprzestrzeni Z nie może być przeliczalny, co przeczy jej ośrodkowości. □

Przykład 6.1. Rozważmy przestrzeń $X := [0,1]$ z σ-ciałem $\mathcal{S}$ podzbiorów $[0,1]$ mierzalnych w sensie Lebesgue'a i miarą Lebesgue'a obciętą do $\mathcal{S}$. Wtedy wiadomo, że odpowiednia przestrzeń L^2 jest ośrodkowa. (Szkic dowodu: zbiór funkcji ciągłych na X jest gęsty w L^2, a następnie trzeba użyć odpowiednich wielomianów, korzystając z twierdzenia aproksymacyjnego Weierstrassa.) Zatem na mocy Wniosku 6.1 dla każdego odwzorowania ergodycznego $T\colon X \to X$ odpowiadająca mu grupa $E(T)$ jest przeliczalna.

Odwzorowanie $T\colon X \to X$ przestrzeni probabilistycznej nazywa się całkowicie ergodyczne, gdy T^n jest ergodyczne dla każdego $n \in \mathbb{N}$. Przykładem odwzorowania całkowicie ergodycznego jest obrót niewymierny T w $[0,1)$. Faktycznie wiemy że jest to odwzorowanie ergodyczne, natomiast jego całkowita ergodyczność wynika z faktu, że T^n jest obrotem niewymiernym dla każdego $n \in \mathbb{N}$.

Wartość własna odwzorowanie T zachowującego miarę przestrzeni probabilistycznej nazywa się wymierna, gdy jest pierwiastkiem z jedności, tzn. $\lambda^k = 1$ dla pewnego $k \in \mathbb{N}$. (Wtedy λ ma postać $\lambda = \exp(2\pi r i)$ dla pewnego $r \in \mathbb{Q}$.)

Twierdzenie 6.3. *Niech* $T\colon X \to X$ *będzie odwzorowaniem ergodycznym przestrzeni probabilistycznej* $(X, \mathcal{S}, \mu)$. *Wówczas* T *jest całkowicie ergodyczne wtedy i tylko wtedy, gdy* T *nie ma wymiernych wartości własnych różnych od* 1.

Dowód. Konieczność. Niech λ będzie wartością własną dla funkcji własnej $f \in L^2$. Załóżmy, że $\lambda^k = 1$ dla pewnego $k \in \mathbb{N}$. Mamy następujące równości zachodzące prawie wszędzie: $f \circ T = \lambda f$ oraz

$$f \circ T^2 = (f \circ T) \circ T = (\lambda f) \circ T = \lambda(f \circ T) = \lambda^2 f$$

i przez łatwą indukcję dostajemy $f \circ T^k = \lambda^k f = f$. Skoro z założenia T^k jest ergodyczne, to f jest stała prawie wszędzie. Ale wtedy $\lambda = 1$, bo wartością własną odpowiadającą funkcji stałej jest 1.

Dostateczność. Mamy pokazać, że T^k jest ergodyczne dla każdego $k \geqslant 2$. Przedstawimy dowód dla $k = 2$. Przypuśćmy, że odwzorowanie T^2 nie jest ergodyczne. Wtedy istnieje taka niestała prawie wszędzie funkcja $g \in L^2$, że $g \circ T^2 = g$. Niech $f := g \circ T - g$. Oczywiście $f \in L^2$. Ponadto f nie jest zerowa prawie wszędzie, bo w przeciwnym przypadku g byłaby prawie wszędzie stała na mocy ergodyczności T. Prawie wszędzie zachodzą równości

$$f \circ T = g \circ T^2 - g \circ T = g - g \circ T = (-1)f,$$

co oznacza, że liczba -1 jest różną od 1, wymierną wartością własną odwzorowania T. Sprzeczność. Dowód dla $k > 2$ jest podobny; pozostawiamy go jako ćwiczenie. (Warto zauważyć, że zbiór pierwiastków k-tego stopnia z jedności jest grupą multiplikatywną.)

□

Mówimy, że odwzorowanie $T\colon X \to X$ zachowujące miarę w przestrzeni probabilistycznej $(X, \mathcal{S}, \mu)$ ma ciągłe widmo (spektrum), gdy $\lambda = 1$ jest jego jedyną wartością własną i jedynymi funkcjami własnymi są funkcje stałe.

Na zakończenie podamy serię warunków charakteryzujących odwzorowania mieszające. Większość z nich poznaliśmy wcześniej. Warunki (iv), (v) są nowe, a dowody ich dotyczące są nieoczywiste; zob. [3].

Twierdzenie 6.4. *Niech T będzie odwzorowaniem zachowującym miarę w przestrzeni probabilistycznej $(X, \mathcal{S}, \mu)$. Następujące warunki są równoważne:*

(i) *T jest słabo mieszające;*

(ii) *$T \times T$ jest ergodyczne względem miary $\mu \times \mu$;*

(iii) *$T \times T$ jest słabo mieszające względem miary $\mu \times \mu$;*

(iv) *dla każdego odwzorowania ergodycznego T_1 dowolnej przestrzeni probabilistycznej $(X_1, \mathcal{S}_1, \mu_1)$ odwzorowanie $T \times T_1$ jest ergodyczne względem miary $\mu \times \mu_1$;*

(v) *odwzorowanie T ma ciągłe widmo;*

(vi) *odwzorowanie T jest podwójnie ergodyczne;*

(vii) *dla dowolnych $A, B \in \mathcal{S}$ istnieje taki zbiór $J_{A,B} \subset \mathbb{N}$ o gęstości zero, że*

$$\lim_{n \notin J_{A,B}} \mu(A \cap T^{-n}[B]) = \mu(A)\mu(B);$$

(viii) *dla dowolnych $A, B \in \mathcal{S}$ zachodzi*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \Big(\mu(A \cap T^{-i}[B]) - \mu(A)\mu(B)\Big)^2 = 0.$$

Rozdział 7

Odwzorowania indukowane

Niech $T\colon X \to X$ będzie odwzorowaniem zachowującym miarę w przestrzeni probabilistycznej $(X, \mathcal{S}, \mu)$. Ustalmy zbiór $A \in \mathcal{S}$, $\mu(A) > 0$. Z twierdzenia Poincarégo wiemy, że prawie każdy punkt $x \in A$ wraca do A nieskończenie wiele razy w wyniku działania funkcji T. Usuńmy ze zbioru A odpowiedni zbiór miary zero, a otrzymany podzbiór zbioru A ponownie oznaczmy przez A. Teraz dla każdego punktu $x \in A$ określmy

$$r_A(x) := \min\{n \in \mathbb{N}\colon T^n(x) \in A\}.$$

Funkcję tę nazywamy czasem pierwszego powrotu do A. Określmy σ-ciało $\mathcal{S}_A$ podzbiorów zbioru A jako $\{E \cap A\colon E \in \mathcal{S}\}$. Ponadto niech $\mu_A(E) := \frac{\mu(E)}{\mu(A)}$ dla $E \in \mathcal{S}_A$. Wtedy $(A, \mathcal{S}_A, \mu_A)$ jest przestrzenią probabilistyczną. Odwzorowanie indukowane $T_A\colon A \to A$ określamy wzorem

$$T_A(x) := T^{r_A(x)}(x), \quad \text{dla } x \in A.$$

Lemat 7.1. *Odwzorowanie indukowane T_A jest mierzalne.*

Dowód. Zauważmy najpierw, że funkcja $r_A\colon A \to \mathbb{R}$ jest mierzalna. Wystarczy pokazać, że każdy zbiór $r_A^{-1}[\{m\}]$, $m \in \mathbb{N}$, należy do $\mathcal{S}$. To zaś wynika z poniższych wzorów:

$$r_A^{-1}[\{1\}] = A \cap T^{-1}[A], \quad r_A^{-1}[\{m\}] = A \cap T^{-m}[A] \setminus \bigcup_{n<m} T^{-n}[A] \text{ dla } m \geqslant 2.$$

Dla dowodu mierzalności T_A weźmy $U \in \mathcal{S}|_A$. Zatem $U = V \cap A$, gdzie $V \in \mathcal{S}$. Zbadamy przeciwobraz $T_A^{-1}[V] = T_A^{-1}[U]$. Dla $x \in A$ mamy

$$x \in T_A^{-1}[V] \Leftrightarrow \exists m \in \mathbb{N}\, (r_A(x) = m \ \wedge\ T^m(x) \in V)\,,$$

więc $T_A^{-1}[V] = \bigcup_{m \in \mathbb{N}} \left(r_A^{-1}[\{m\}] \cap T^{-m}[V] \right) \in \mathcal{S}$. □

Twierdzenie 7.1. *Odwzorowanie indukowane T_A zachowuje miarę w przestrzeni $(A, \mathcal{S}|_A, \mu_A)$. Ponadto jeśli T jest ergodyczne, to T_A jest też ergodyczne.*

Dowód. Dla każdego $k \in \mathbb{N}$ określmy zbiory $A_k := \{x \in A \colon r_A(x) = k\}$ oraz

$$B_k := \{x \in X \setminus A \colon T(x) \notin A \wedge \ldots \wedge T^{k-1}(x) \notin A \wedge T^k(x) \in A\}.$$

Z twierdzenia Poincarégo mamy $A = \bigcup_{k=1}^{\infty} A_k$. Ponadto

$$T^{-1}[A] = A_1 \cup B_1, \quad T^{-1}[B_n] = A_{n+1} \cup B_{n+1} \text{ dla } n \in \mathbb{N} \tag{7.1}$$

oraz sumy po prawej stronie tych równości są rozłączne.

Niech $C \in \mathcal{S}_A$. Chcemy wykazać, że $\mu_A(T_A^{-1}[C]) = \mu_A(C)$. Wiemy, że $\mu(T^{-1}[C]) = \mu(C)$. Mamy

$$T_A^{-1}[C] = \bigcup_{k=1}^{\infty} A_k \cap T_A^{-1}[C] = \bigcup_{k=1}^{\infty} (A_k \cap T^{-k}[C]) \quad \text{z definicji } T_A.$$

Zatem

$$\mu(T_A^{-1}[C]) = \sum_{k=1}^{\infty} \mu(A_k \cap T^{-k}[C]), \tag{7.2}$$

bo zbiory A_k, $k \in \mathbb{N}$, są parami rozłączne. Z drugiej strony $C \subset A$, więc $T^{-1}[C] = T^{-1}[C] \cap T^{-1}[A]$. Zatem stosując (7.1), stwierdzamy, że miara $\mu(T^{-1}[C])$ jest równa

$$\mu(A_1 \cap T^{-1}[C]) + \mu(B_1 \cap T^{-1}[C]) = \mu(A_1 \cap T^{-1}[C]) + \mu(T^{-1}(B_1 \cap T^{-1}[C])$$

$$= \mu(A_1 \cap T^{-1}[C]) + \mu(A_2 \cap T^{-2}[C]) + \mu(B_2 \cap T^{-2}[C]).$$

Powtarzając to rozumowanie, otrzymujemy na podstawie (7.1) dla każdego $n \in \mathbb{N}$

$$\mu(T^{-1}[C]) = \sum_{k=1}^{n} \mu(A_k \cap T^{-k}[C]) + \mu(B_n \cap T^{-n}[C]). \tag{7.3}$$

Zbiory B_n, $n \in \mathbb{N}$, są parami rozłączne, więc

$$1 \geqslant \mu\left(\bigcup_{n=1}^{\infty} (B_n \cap T^{-n}[C])\right) = \sum_{n=1}^{\infty} \mu(B_n \cap T^{-n}[C]),$$

a zatem $\lim_n \mu(B_n \cap T^{-n}[C]) = 0$. W konsekwencji, biorąc $n \to \infty$ w (7.3), mamy

$$\mu(T^{-1}[C]) = \sum_{k=1}^{\infty} \mu(A_k \cap T^{-k}[C]). \tag{7.4}$$

Stąd, stosując (7.2) oraz (7.4), otrzymujemy

$$\mu(C) = \mu(T^{-1}[C]) = \sum_{k=1}^{\infty} \mu(A_k \cap T^{-k}[C]) = \mu(T_A^{-1}[C]).$$

Zatem $\mu_A(C) = \mu_A(T^{-1}[C])$. To daje pierwszą część tezy.

Dla dowodu drugiej części tezy załóżmy, że odwzorowanie T jest ergodyczne. Niech $C \in \mathcal{S}_A$ oraz $T_A^{-1}[C] = C$. Mamy pokazać, że $\mu_A(C) = 0$ lub $\mu_A(C) = 1$, co jest równoważne alternatywie $\mu(C) = 0$ lub $\mu(C) = \mu(A)$. Skoro $A = \bigcup_{k=1}^{\infty} A_k$ oraz $C \subset A$, to $C = T_A^{-1}[C] = \bigcup_{k=1}^{\infty}(A_k \cap T^{-k}[C])$. Niech

$$E := \bigcup_{k=1}^{\infty}(B_k \cap T^{-k}[C]) \quad \text{oraz} \quad F := E \cup C.$$

Wtedy

$$T^{-1}[E] = \bigcup_{k=1}^{\infty}\left(T^{-1}[B_k] \cap T^{-(k+1)}[C]\right), \quad T^{-1}[C] = T^{-1}[A \cap C] = T^{-1}[A] \cap T^{-1}[C].$$

Zatem stosując (7.1), otrzymujemy

$$T^{-1}[F] = T^{-1}[E] \cup T^{-1}[C] = \bigcup_{k=1}^{\infty}\left((A_{k+1} \cup B_{k+1}) \cap T^{-(k+1)}[C]\right) \cup \left((A_1 \cup B_1) \cap T^{-1}[C]\right)$$

$$= \bigcup_{k=1}^{\infty}(A_k \cap T^{-k}[C]) \cup \bigcup_{k=1}^{\infty}(B_k \cap T^{-k}[C]) = C \cup E = F.$$

Dlatego z ergodyczności T wynika, że $\mu(F) = 0$ lub $\mu(F) = 1$. Jeśli $\mu(F) = 0$, to $\mu(C) = 0$, a więc $\mu_A(C) = 0$. Załóżmy teraz, że $\mu(F) = 1$. Wtedy $\mu(X \setminus F) = 0$. Mamy

$$X \setminus F = (A \setminus C) \cup ((X \setminus A) \setminus E) \supset A \setminus C.$$

Stąd $\mu(A \setminus C) \leqslant \mu(X \setminus F) = 0$. Ponieważ $\mu(A \setminus C) = \mu(A) - \mu(C)$, więc $\mu(A) = \mu(C)$, a zatem $\mu_A(C) = 1$. □

Idea związana z czasem pierwszego powrotu ma ciekawą interpretację, którą jest tzw. wieża Kakutaniego. Niech więc odwzorowanie $T\colon X \to X$ zachowuje miarę w przestrzeni probabilistycznej $(X, \mathcal{S}, \mu)$ oraz niech $A \in \mathcal{S}$ będzie zbiorem miary dodatniej. Będziemy tu dodatkowo zakładać, że T jest odwracalne. Oznaczmy

$$D_0 := A, \quad D_k := T^k[A] \setminus (A \cup T[A] \cup \dots \cup T^{k-1}[A]), \quad \text{dla } k \in \mathbb{N}.$$

Zbiór D_0 nazywamy parterem, zaś zbiór D_k dla $k \in \mathbb{N}$ nazywamy k-tym piętrem wieży. Pamiętamy, że $A = \bigsqcup_{n=1}^{\infty} A_n$ (suma rozłączna, po modyfikacji zbioru A), gdzie $A_n := \{x \in A\colon r_A(x) = n\}$. Z definicji r_A wynika, że dla $n \in \mathbb{N}$ mamy $T^n[A_n] \subset A$. Pokażemy, że

$$T^k[A_n] \subset D_k, \quad \text{gdy } k = 0, \dots, n-1. \tag{7.5}$$

Faktycznie inkluzja $T^k[A_n] \subset T^k[A]$ jest oczywista. Aby dokończyć dowód (7.5), wystarczy pokazać, że zbiory $T^k[A_n]$ dla $k \in \mathbb{N}$, $0 \leqslant k \leqslant n-1$, są parami rozłączne. Ustalmy $n \in \mathbb{N}$ i zauważmy najpierw, że $T^k[A_n] \cap A = \emptyset$ dla $k = 1, \ldots, n-1$. Niech teraz $1 \leqslant k < j < n$. Wtedy z iniektywności T wynika, że

$$T^j[A_n] \cap T^k[A_n] = T^k[T^{j-k}[A_n] \cap A_n] = T^k[\emptyset] = \emptyset.$$

Pokazaliśmy więc, że rodzina

$$A_n, T[A_n], \ldots, T^{n-1}[A_n]$$

jest rozłączna; nazywamy ją n-tą kolumną wieży.

Pozostaje wykazać rozłączność dowolnych dwóch zbiorów w dwóch różnych kolumnach wieży, n-tej oraz m-tej, gdzie $m \neq n$. Mamy $A_n \cap T^k[A_m] = \emptyset$ dla $0 \leqslant k < m$, bo dla $k = 0$ oczywiście $A_n \cap A_m = \emptyset$, zaś dla $k \geqslant 1$ mamy $A_n \subset A$ oraz $T^k[A_m] \cap A = \emptyset$. Ponadto $T^k[A_n] \cap T^k[A_m] = \emptyset$ dla $k < \min\{n, m\}$, co wynika z rozłączności A_n i A_m oraz iniektywności T^k. Na koniec przypuśćmy, że $x \in T^k[A_n] \cap T^j[A_m]$ dla $k, j < \min\{n, m\}$. Wtedy $T^{-k}(x) \in A_n \cap T^{j-k}[A_m]$, a to wcześniej okazało się niemożliwe.

Koncepcja wieży Kakutaniego znalazła zastosowanie w następującym twierdzeniu Marka Kaca. Najpierw udowodnimy fakt pomocniczy.

Lemat 7.2. *Jeśli odwzorowanie $T \colon X \to X$ zachowujące miarę w przestrzeni probabilistycznej $(X, \mathcal{S}, \mu)$ jest odwracalne i ergodyczne oraz $A \in \mathcal{S}$, $\mu(A) > 0$, to rodzina $\mathcal{F} := \{T^k[A_n] \colon n \in \mathbb{N},\ 0 \leqslant k \leqslant n-1\}$ jest rozłączna i $\mu(\bigcup \mathcal{F}) = 1$.*

Dowód. Rozłączność rodziny $\mathcal{F}$ została wykazana wyżej. Skoro odwzorowanie T jest ergodyczne i odwracalne, to nietrudno sprawdzić, że T^{-1} też jest ergodyczne (ćwiczenie). Ze znanej charakteryzacji ergodyczności T^{-1} dostajemy $\mu(\bigcup_{k=1}^{\infty} T^k[A]) = 1$. Z równości $A = \bigcup_{n=1}^{\infty} A_n$ mamy

$$\bigcup_{k=1}^{\infty} T^k[A] = \bigcup_{n=1}^{\infty} \bigcup_{k=1}^{\infty} T^k[A_n] \subset \bigcup_{n=1}^{\infty} \bigcup_{k=0}^{n-1} T^k[A_n] = \bigcup \mathcal{F}. \tag{7.6}$$

Stąd wynika teza $\mu(\bigcup \mathcal{F}) = 1$. Należy tylko wyjaśnić inkluzję we wzorze (7.6). W tym celu wystarczy wykazać, że dla dowolnych liczb $m, p \in \mathbb{N}$ zachodzi inkluzja

$$T^m[A_p] \subset \bigcup_{n=1}^{\infty} \bigcup_{k=0}^{n-1} T^k[A_n].$$

Ustalmy więc $m, p \in \mathbb{N}$. Mamy trzy przypadki: 1^0 gdy $m < p$; 2^0 gdy $m = p$ oraz 3^0 gdy $m > p$. W przypadku 1^0 inkluzja jest oczywista. W przypadku 2^0 mamy

$$T^m[A_m] \subset A = \bigcup_{n=1}^{\infty} T^0[A_n].$$

W przypadku 3^0 weźmy $x \in T^m[A_p]$ i wtedy

$$x \in T^{m-p}[T^p[A_p]] \subset T^{m-p}[A] = \bigcup_{n=1}^{\infty} T^{m-p}[A_n].$$

Zatem $x \in T^{m-p}[A_{p_1}]$ dla pewnego $p_1 \in \mathbb{N}$. Jeśli $m - p_1 \leqslant p_1$, to stosujemy przypadek 1^0 lub 2^0. Jeśli $m - p > p_1$, to powtarzamy rozumowanie z przypadku 3^0 i stwierdzamy, że $x \in T^{m-p-p_1}[A_{p_2}]$ dla pewnego $p_2 \in \mathbb{N}$. Rozumowanie z przypadku 3^0 może być stosowane skończoną liczbę razy, bo $m - p > m - p_1 > m - p_1 - p_2 > \dots$. □

Twierdzenie 7.2 (Kaca)**.** *Niech $T \colon X \to X$ będzie odwracalnym ergodycznym odwzorowaniem zachowującym miarę w przestrzeni probabilistycznej $(X, \mathcal{S}, \mu)$ i niech $A \in \mathcal{S}$, $\mu(A) > 0$. Wtedy $\int_A r_A \, d\mu = 1$. W konsekwencji wartość oczekiwana zmiennej losowej r_A na przestrzeni $(A, \mathcal{S}_A, \mu_A)$ wynosi $\frac{1}{\mu(A)}$.*

Dowód. Z Lematu 7.2 mamy

$$\mu\left(\bigsqcup_{n=1}^{\infty} \bigsqcup_{k=0}^{n-1} T^k[A_n]\right) = 1.$$

Dla każdego $n \in \mathbb{N}$ kolumna z indeksem n w wieży Kakutaniego składa się z n parami rozłącznych zbiorów, z których każdy ma miarę $\mu(A_n)$ (bo T^{-1} zachowuje miarę). Zatem

$$1 = \sum_{n=1}^{\infty} n\mu(A_n) = \lim_{m\to\infty} \int_A \sum_{m=1}^{m} (n\chi_{A_n}) \, d\mu = \int_A r_A \, d\mu.$$

Ostatnia równość wynika stąd, że ciąg sum $\sum_{n=1}^{m} n\chi_{A_n}$, $m \in \mathbb{N}$, zbiega punktowo w sposób niemalejący do r_A, gdy $m \to \infty$. Zatem możemy stosować twierdzenie Lebesgue'a o monotonicznym przechodzeniu do granicy pod znakiem całki. □

W dowodzie twierdzenia Kaca korzystaliśmy z tego, że jeśli T jest odwracalnym odwzorowaniem ergodycznym w przestrzeni probabilistycznej $(X, \mathcal{S}, \mu)$, to dla każdego zbioru $A \in \mathcal{S}$ miary dodatniej mamy $\mu(\bigcup_{k\geqslant 0} T^k[A]) = 1$. W szczególności dla każdego $\varepsilon \in (0, 1)$ istnieje takie $m \in \mathbb{N}$, że $\mu(A \cup T[A] \cup \dots \cup T^{m-1}[A]) > 1 - \varepsilon$.

Nasuwa się pytanie, czy dla każdej liczby $m \in \mathbb{N}$ i dowolnego $\varepsilon \in (0, 1)$ istnieje taki zbiór $A \in \mathcal{S}$, że powyższe oszacowanie zachodzi. Odpowiedź jest pozytywna przy założeniu, że miara μ jest bezatomowa. Ta odpowiedź wynika bezpośrednio z poniższego twierdzenia, które jest znane jako lemat Kakutaniego-Rokhlina. W dowodzie ponownie korzystamy z wieży Kakutaniego.

Zbiór $A \in \mathcal{S}$ nazywamy atomem miary μ na σ-ciele podzbiorów X, gdy $\mu(A) > 0$ oraz każdy podzbiór $B \subset A$, $B \in \mathcal{S}$, ma miarę 0 lub równą $\mu(A)$. Miara μ nazywa się bezatomowa, gdy nie ma atomów. Dowodzi się, że jeśli miara μ jest skończona i bezatomowa na σ-ciele $\mathcal{S}$, to dla każdego $\alpha \in [0, \mu(X)]$ istnieje taki zbiór $A \in \mathcal{S}$, że $\mu(A) = \alpha$.

Twierdzenie 7.3 (Kakutaniego–Rokhlina)**.** *Niech $T\colon X \to X$ będzie odwracalnym ergodycznym odwzorowaniem zachowującym miarę w przestrzeni probabilistycznej $(X, \mathcal{S}, \mu)$, przy czym μ jest miarą bezatomową. Wtedy dla każdego $m \in \mathbb{N}$ i dowolnego $\varepsilon \in (0, 1)$ istnieje zbiór $B \in \mathcal{S}$ taki, że zbiory*

$$B,\, T[B],\, \dots, T^{m-1}[B]$$

są parami rozłączne oraz $\mu(B \sqcup T[B] \sqcup \dots \sqcup T^{m-1}[B]) > 1 - \varepsilon$.

Dowód. Niech $m \in \mathbb{N}$ oraz $\varepsilon > 0$. Ponieważ miara μ jest bezatomowa, więc można znaleźć taki zbiór $A \in \mathcal{S}$, że $0 < \mu(A) < \varepsilon/m$. Utwórzmy wieżę Kakutaniego dla zbioru A. Jak wcześniej niech $A_n := \{x \in A\colon r_A(x) = n\}$ dla $n \in \mathbb{N}$. Położmy (por. Lemat 7.2)

$$B := \bigsqcup_{n=m}^{\infty} \bigsqcup_{j=0}^{\lfloor n/m \rfloor - 1} T^{jm}[A_n].$$

Zauważmy, że zbiory $B, T[B], \dots, T^{m-1}[B]$ są parami rozłączne i ich suma jest równa sumie całej wieży Kakutaniego z wyłączeniem co najwyżej m zbiorów w każdej kolumnie (sprawdzić). Zatem dopełnienie tej sumy ma miarę $\leqslant \sum_{n=1}^{\infty} m\mu(A_n) = m\mu(A) < \varepsilon$. Wobec tego

$$\mu(B \sqcup T[B] \sqcup \dots \sqcup T^{m-1}[B]) > 1 - \varepsilon.$$

$\square$

Rozdział 8

Izomorfizmy układów dynamicznych zachowujących miarę

Omówimy tu bardziej szczegółowo pojęcie izomorfizmu dwóch odwzorowań zachowujących miarę.

Definicja 8.1. Załóżmy, że $(X_i, \mathcal{S}_i, \mu_i)$ dla $i = 1, 2$ są przestrzeniami probabilistycznymi. Niech dane będą odwzorowania $T_i \colon X_i \to X_i$ dla $i = 1, 2$ zachowujące miarę. Mówimy, że T_1 jest izomorficzne z T_2, gdy istnieją takie zbiory $M_1 \in \mathcal{S}_1$, $M_2 \in \mathcal{S}_2$, że $\mu_1(M_1) = 1$, $\mu_2(M_2) = 1$ oraz

(i) $T_1[M_1] \subset M_1$, $T_2[M_2] \subset M_2$;

(ii) istnieje odwracalne odwzorowanie $\Phi \colon M_1 \to M_2$ zachowujące miarę oraz zachodzi równość $(\Phi \circ T_1)(x) = (T_2 \circ \Phi)(x)$ dla każdego $x \in M_1$.

Odwzorowanie Φ nazywamy izomorfizmem między T_1 i T_2. Zapisujemy $T_1 \simeq T_2$.

Uwaga 8.1. (1) W warunku (ii) dla zbiorów M_i $(i = 1, 2)$ pełnej miary rozważa się odpowiednie przestrzenie probabilistyczne $(M_i, \mathcal{S}_i|_{M_i}, \mu|_{M_i})$ dla $i = 1, 2$.
(2) Czwórkę $(X, \mathcal{S}, \mu, T)$, gdzie $(X, \mathcal{S}, \mu)$ jest przestrzenią probabilistyczną, zaś odwzorowanie $T \colon X \to X$ zachowuje miarę, nazwaliśmy układem dynamicznym zachowującym miarę i formalnie relacja $\simeq$ powinna być zdefiniowana w klasie układów dynamicznych zachowujących miarę (por. Def. 1.3, 1.4). Na ustalonej rodzinie układów dynamicznych zachowujących miarę jest to relacja równoważności.
(3) Jeśli $T_1 \simeq T_2$, to $T_1^n \simeq T_2^n$ dla każdego $n \in \mathbb{N}$ (ćwiczenie).

(4) Jeśli T_1, T_2 są odwracalnymi odwzorowaniami zachowującymi miarę, to zbiory M_1, M_2 spełniające powyższą definicję można wybrać tak, żeby $T_1[M_1] = M_1$, $T_2[M_2] = M_2$. Istotnie, mając $M_1, M_2 \in \mathcal{S}$ takie, że $\mu_1(M_1) = 1$, $\mu_2(M_2) = 1$ oraz $T_1[M_1] \subset M_1$, $T_2[M_2] \subset M_2$, połóżmy $M_1^* := \bigcap_{k\in\mathbb{Z}} T_1^k[M_1]$ oraz $M_2^* := \bigcap_{k\in\mathbb{Z}} T_2^k[M_2]$.

Przykłady.

(1) Odwzorowanie $T_2 \colon K \to K$ okręgu jednostkowego $K \subset \mathbb{C}$ dane wzorem $T_2(z) := z^2$ jest izomorficzne z odwzorowaniem podwajającym $T_1 \colon [0,1) \to [0,1)$, $T_1(x) := 2x \pmod 1$. Aby to wykazać, określmy $\Phi \colon [0,1) \to K$ wzorem $\Phi(t) := \exp(2\pi ti)$ dla $t \in [0,1)$. Na K rozważamy σ-ciało $\mathcal{S}_2$ generowane przez łuki na okręgu, przy czym miara łuku jest równa jego długości podzielonej przez 2π. Na $[0,1)$ rozważamy σ-ciało $\mathcal{S}_1$ zbiorów borelowskich z miarą Lebesgue'a. Tutaj $M_1 := [0,1)$ oraz $M_2 := K$. Nietrudno sprawdzić, że Φ jest odwracalnym odwzorowaniem zachowującym miarę (wystarczy sprawdzić, że Φ "przenosi" długość przedziałów na miarę odpowiednich łuków. Sprawdzimy, że $\Phi \circ T_2 = T_1 \circ \Phi$. Dla $t \in [0,1)$ mamy

$$(\Phi \circ T_2)(t) = \exp(2\pi i(2t \pmod 1)) = \exp(4\pi ti) = (\exp(2\pi ti))^2 = (T_1 \circ \Phi)(t).$$

(2) Rozważmy ponownie odwzorowanie $T_1 \colon [0,1) \to [0,1)$, $T_1(x) := 2x \pmod 1$ oraz niech $T_2 \colon \{0,1\}^{\mathbb{N}} \to \{0,1\}^{\mathbb{N}}$ będzie jednostronnym przesunięciem $T_2((x_i)_{i\in\mathbb{N}}) := (x_{i+1})_{i\in\mathbb{N}}$. Określmy $\Phi \colon \{0,1\}^{\mathbb{N}} \to [0,1)$ wzorem

$$\Psi(a_1, a_2, a_3, \dots) = \sum_{i=1}^{\infty} \frac{a_i}{2^i}.$$

Niech D_1 będzie podzbiorem $\{0,1\}^{\mathbb{N}}$ złożonym z ciągów stałych od pewnego miejsca, zaś przez D_2 oznaczmy zbiór liczb dwójkowo-wymiernych w przedziale $[0,1)$. Zbiory D_1, D_2 są przeliczalne, więc miary zero. Oznaczmy $M_1 := \{0,1\}^{\mathbb{N}} \setminus D_1$ oraz $M_2 := [0,1) \setminus D_2$. Wówczas $\Phi := \Psi|_{M_1}$ jest odwracalnym odwzorowaniem zachowującym miarę i spełnia ono warunki definicyjne izomorfizmu $T_1 \simeq T_2$. W szczególności zachodzi równość $\Phi \circ T_2 = T_1 \circ \Phi$ na M_1.

(3) Przypomnijmy odwzorowanie piekarza $T_1 \colon [0,1)^2 \to [0,1)^2$ dane wzorem

$$T_1(x,y) := \begin{cases} (2x, y/2) & \text{dla} \quad x \in [0, 1/2)\,; \\ (2x-1, (y+1)/2) & \text{dla} \quad x \in [1/2, 1)\,. \end{cases}$$

Opiszmy pokrótce, w jaki sposób pokazuje się, że odwzorowanie T_1 jest izomorficzne z dwustronnym przesunięciem $T_2 \colon \{0,1\}^{\mathbb{Z}} \to \{0,1\}^{\mathbb{Z}}$, które jest dane wzorem $T_2((x_i)_{i\in\mathbb{Z}}) := (x_{i+1})_{i\in\mathbb{Z}}$. Ze względów technicznych dwustronne przesuniecie T_2 przedstawmy nieco inaczej jako odwzorowanie przestrzeni $\{0,1\}^{\mathbb{N}} \times \{0,1\}^{\mathbb{N}}$ w siebie dane wzorem

$$((x_1, x_2, \dots), (y_1, y_2, \dots)) \mapsto ((x_2, x_3, \dots), (x_1, y_1, y_2, \dots)).$$

Definiując izomorfizm między T_1 i T_2, usuwamy z kwadratu $[0,1) \times [0,1)$ punkty (x,y) o współrzędnych dwójkowo-wymiernych i otrzymany zbiór oznaczamy przez M_1, zaś z przestrzeni $\{0,1\}^{\mathbb{N}} \times \{0,1\}^{\mathbb{N}}$ usuwamy pary ciągów $(x_i), (y_i)$, które są od pewnego miejsca stałe i otrzymany zbiór oznaczamy przez M_2. Dla $(x,y) \in M_1$ mamy jednoznaczne rozwinięcia dwójkowe $x = (0.x_1x_2,\dots)$ oraz $y = (0.y_1y_2\dots)$ liczb x i y. Izomorfizm $\Phi\colon M_1 \to M_2$ definiujemy w ten sposób, że przypisuje on punktowi $(x,y) \in [0,1)^2$ parę ciągów zero-jedynkowych $((x_2,x_3,\dots), (x_1,y_1,y_2,\dots))$.

Rozważmy teraz modyfikację Definicji 8.1.

Definicja 8.2. Załóżmy, że $(X_i, \mathcal{S}_i, \mu_i)$ dla $i = 1,2$ są przestrzeniami probabilistycznymi. Niech dane będą odwzorowania $T_i\colon X_i \to X_i$ dla $i = 1,2$ zachowujące miarę. Mówimy, że T_2 jest faktorem odwzorowania T_1, gdy istnieją takie zbiory $M_1 \in \mathcal{S}_1$, $M_2 \in \mathcal{S}_2$, że $\mu_1(M_1) = 1$, $\mu_2(M_2) = 1$ oraz:

(i) $T_1[M_1] \subset M_1,\ T_2[M_2] \subset M_2$;

(ii) istnieje suriektywne odwzorowanie $\Phi\colon M_1 \to M_2$ zachowujące miarę oraz zachodzi równość $(\Phi \circ T_1)(x) = (T_2 \circ \Phi)(x)$ dla każdego $x \in M_1$.

Wtedy odwzorowanie Φ nazywa się faktoryzujące.

Oczywiście każde odwzorowanie T_2 izomorficzne z T_1 jest jego faktorem. Odwzorowania faktoryzujące (a więc także izomorfizmy) zachowują ważne własności odwzorowań zachowujących miarę.

Przykłady.

(1) Przesunięcie jednostronne $T_2((x_i)_{i\in\mathbb{N}}) := (x_{i+1})_{i\in\mathbb{N}}$ w przestrzeni $\{0,1\}^{\mathbb{N}}$ jest faktorem przesunięcia dwustronnego $T_1((x_i)_{i\in\mathbb{Z}}) := (x_{i+1})_{i\in\mathbb{Z}}$ w przestrzeni $\{0,1\}^{\mathbb{Z}}$. Aby to uzasadnić, wystarczy zdefiniować odwzorowanie faktoryzujące $\Phi\colon \{0,1\}^{\mathbb{Z}} \to \{0,1\}^{\mathbb{N}}$ wzorem $\Phi((x_i)_{i\in\mathbb{Z}}) := (x_i)_{i\in\mathbb{N}}$. Wtedy Φ jest suriekcją zachowującą miarę oraz $(\Phi \circ T_1)(x) = (T_2 \circ \Phi)(x)$ dla każdego $x \in \{0,1\}^{\mathbb{Z}}$ (tutaj M_1, M_2 są całymi przestrzeniami).

(2) Niech T_1, T_2 będą odwzorowaniami zachowującymi miarę w odpowiednich przestrzeniach probabilistycznych $(X_i, \mathcal{S}_i, \mu_i)$ dla $i = 1,2$. Wtedy odwzorowanie produktowe $T_1 \times T_2\colon X_1 \times X_2 \to X_1 \times X_2$ dane wzorem $(T_1 \times T_2)(x_1,x_2) := (T_1(x_1), T_2(x_2))$ zachowuje miarę w przestrzeni $(X_1 \times X_2, \mathcal{S}_1 \otimes \mathcal{S}_2, \mu_1 \times \mu_2)$. Zauważmy, że T_1 jest faktorem odwzorowania $T_1 \times T_2$, bo rzutowanie $\Phi\colon X_1 \times X_2 \to X_1$, $\Phi(x_1,x_2) := x_1$, jest suriekcją zachowującą miarę oraz $\Phi \circ (T_1 \times T_2) = T_1 \circ \Phi$ na $X_1 \times X_2$ (tutaj M_1 i M_2 są całymi przestrzeniami). Podobnie można wykazać, że T_2 jest faktorem odwzorowania $T_1 \times T_2$.

Twierdzenie 8.1. *Jeśli T_1, T_2 są odwzorowaniami zachowującymi miarę w odpowiednich przestrzeniach probabilistycznych $(X_i, \mathcal{S}_i, \mu_i)$ dla $i = 1, 2$, przy czym T_1 jest ergodyczne (słabo mieszające, mieszające) oraz T_2 jest faktorem odwzorowania T_1, to T_2 jest ergodyczne (słabo mieszające, mieszające).*

Dowód. Dla uproszczenia załóżmy, że $M_1 = X_1$ oraz $M_2 = X_2$ i niech $\Phi\colon X_1 \to X_2$ będzie odwzorowaniem faktoryzującym. Wykażemy tezę w przypadku odwzorowań mieszających. W pozostałych przypadkach dowód jest analogiczny. Załóżmy więc, że T_1 jest mieszające. Niech $A, B \in \mathcal{S}_2$. Wtedy $\Phi^{-1}[A], \Phi^{-1}[B] \in \mathcal{S}_1$ oraz

$$\begin{aligned}\mu_2(T_2^{-n}[A] \cap B) &= \mu_1(\Phi^{-1}[T_2^{-n}[A] \cap B]) = \mu_1(\Phi^{-1}[T_2^{-n}[A]] \cap \Phi^{-1}[B]) \\ &= \mu_1((T_2^n \circ \Phi)^{-1}[A] \cap \Phi^{-1}[B]) = \mu_1((\Phi \circ T_1^n)^{-1}[A] \cap \Phi^{-1}[B]) \\ &= \mu_1(T_1^{-n}[\Phi^{-1}[A]] \cap \Phi^{-1}[B]) \to \mu_1(\Phi^{-1}[A])\mu_1(\Phi^{-1}[B]) = \mu_2(A)\mu_2(B).\end{aligned}$$

Powyższe przejście graniczne gdy $n \to \infty$ zachodzi dzięki warunkowi mieszania dla T_1. To rozumowanie daje warunek definicyjny mieszania dla T_2 i kończy dowód. $\square$

Bibliografia

[1] X. Dai, *From the First Borel–Cantelli Lemma to Poincaré's Recurrence Theorem*, “The American Mathematical Monthly” 2015, vol. 122, issue 2, p. 173–174.

[2] K. Dajani, S. Dirksin, *A Simple Introduction to Ergodic Theory*, 2008, https://webspace.science.uu.nl [dostęp: 17.10.2022].

[3] M. Einsiedler, T. Ward, *Ergodic Theory: with a view towards Number Theory*, Springer, London 2011.

[4] S. Gouëzel, A. Karlsson, *Subadditive and multiplicative ergodic theorems*, “Journal of the European Mathematical Society” 2020, vol. 22, no. 6, p. 1893-1915.

[5] J. Górnicki, *Podstawy nieliniowej teorii ergodycznej*, “Roczniki Polskiego Towarzystwa Matematycznego. Seria 2: Wiadomości Matematyczne” 2001, z. 27, s. 5–16.

[6] A. Karlsson, *Ergodic theorems for noncommuting random products*, https://www.unige.ch/math/folks/karlsson/wroclawtotal.pdf [dostęp: 17.10. 2022].

[7] A. Karlsson, F. Ledrappier, *Noncommutative Ergodic Theorems*, https://www.unige.ch/math/folks/karlsson/kl4.pdf [dostęp: 17.10.2022].

[8] J. Kułaga-Przymus, M. Lemańczyk, *Sarnak's conjecture from the ergodic theory point of view*, https://arxiv.org/pdf/2009.04757.pdf [dostęp: 17.10.2022].

[9] K. Petersen, *Easy and nearly simultaneous proofs of the ergodic theorem and maximal ergodic theorem*, https://arxiv.org/pdf/math/0004070.pdf [dostęp: 17.10.2022].

[10] T. Šalát, *On statistically convergent sequences of real numbers*, “Mathematica Slovaca”, 1980, vol. 30, no. 2, p. 139–150.

[11] C.E. Silva, *Invitation to Ergodic Theory*, American Mathematical Society, Providence 2007.

[12] S.M. Srivastava, *A Course on Borel Sets*, Springer, New York 1998.

[13] P. Walters, *Ergodic Theory – Introductory Lectures*, Springer, Berlin 1975.

ISBN 978-83-66741-62-1