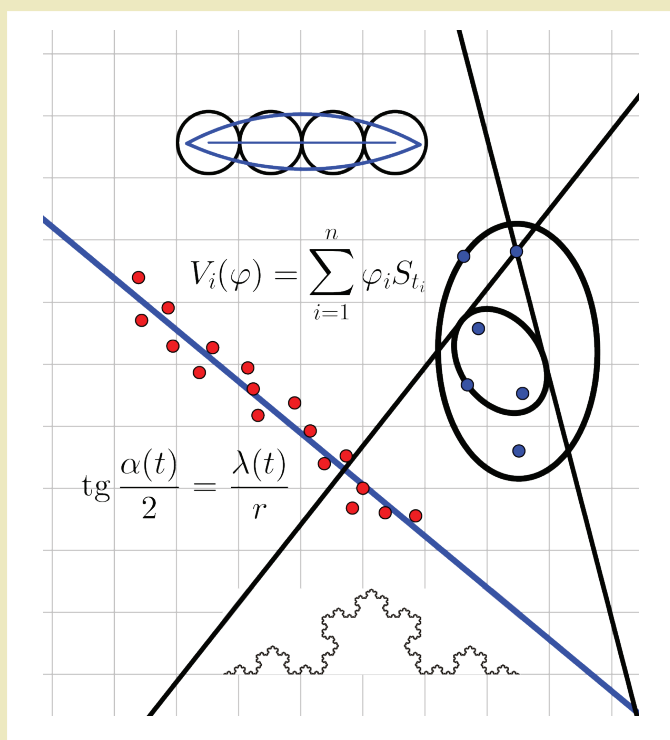




Matematyczne miscellanea

*redakcja
Izolda Gorgol*



Lublin 2016

Matematyczne miscellanea

Monografie – Politechnika Lubelska



Politechnika Lubelska
Wydział Podstaw Techniki
ul. Nadbystrzycka 38
20-618 LUBLIN

Matematyczne miscellanea

redakcja
Izolda Gorgol



Politechnika Lubelska
Lublin 2016

Recenzenci:

dr hab. Leopold Koczan, prof. Politechniki Lubelskiej

prof. dr hab. inż. Mykhaylo Pashechko

Publikacja wydana za zgodą Rektora Politechniki Lubelskiej

© Copyright by Politechnika Lubelska 2016

ISBN: 978-83-7947-231-4

Wydawca: Politechnika Lubelska

ul. Nadbystrzycka 38D, 20-618 Lublin

Realizacja: Biblioteka Politechniki Lubelskiej

Ośrodek ds. Wydawnictw i Biblioteki Cyfrowej

ul. Nadbystrzycka 36A, 20-618 Lublin

tel. (81) 538-46-59, email: wydawca@pollub.pl

www.biblioteka.pollub.pl

Druk: TOP Agencja Reklamowa Agnieszka Łuczak

www.agencjatop.pl

Elektroniczna wersja książki dostępna w Bibliotece Cyfrowej PL www.bc.pollub.pl

Nakład: 50 egz.

Spis treści

Przedmowa	7
Gry obronne	9
<i>Witold Rzymowski, Agnieszka Surowiec</i>	
1 Wstęp	9
2 Obrona koła	14
2.1 Opis gry	14
2.2 Obrona	15
2.3 Atak	19
2.4 Ucieczka z koła	24
3 Obrona odcinka przez jednego obrońcę	28
3.1 Wstępny opis gry	28
3.2 Przypadek I	29
3.3 Przypadek II	36
4 Obrona odcinka przez wielu broniących	39
Literatura	43
Poryzm Ponceleta	45
<i>Waldemar Cieślak, Witold Mozgawa</i>	
1 Wstęp	45
2 O wzorach Fussa	46
3 O indeksie obrotu dla bilardów barowych i poryzmie Ponceleta	50
4 O pewnej hipotezie związanej z poryzmem Ponceleta	56
5 Problem ekstremalny w poryzmie Ponceleta	57
6 O pewnych uogólnieniach poryzmu Ponceleta na dowolne pierścienie wypukłe	58
7 O warunkach zamkniętości łamanej Ponceleta	60
Literatura	61

Metody deflatorowe w wycenie przepływów finansowych	65
<i>Janusz J. Szuster</i>	
1 Wstęp	65
2 Niezbędne pojęcia i fakty	66
2.1 Kontrakty i opcje	67
2.2 Rynki dwustanowe – jedookresowe i wielookresowe	68
2.3 Miara martyngałowa	77
3 Stochastyczne dyskonto	81
3.1 Podstawowy model z czasem dyskretnym	81
4 Przypadek prostego modelu z czasem dyskretnym	82
5 Istota działania deflatorów	87
6 Wycena poza chwilą 0	90
7 Deflatory w szacowaniu rezerw	93
Literatura	96
Zagadnienia teorii chaosu	97
<i>Adam Stachura</i>	
1 Wstęp	97
2 Przestrzenie metryczne	98
3 Dynamiki chaotyczne	103
4 Układy dynamiczne na prostej liczbowej	108
5 Układy dynamiczne na prostej liczbowej – ciąg dalszy	113
6 Izometrie i podobieństwa w przestrzeniach euklidesowych	122
5 Samopodobieństwo. Fraktale topologiczne	129
Literatura	144

Przedmowa

Matematyka ma różne oblicza. Każdemu pokazuje tę stronę, która mu się najbardziej spodoba. Dla lubiących gimnastykę umysłu ma nieprzebrane zasoby zagadek i łamigłówek, a także poważnych hipotez oczekujących na potwierdzenie lub obalenie, dla pragmatyków ma wiele modeli opisujących rzeczywistość i metod pozwalających rozwiązać konkretne problemy praktyczne. Czasami to właśnie te praktyczne zagadnienia prowadzą do jej rozwoju, a czasami wręcz odwrotnie – rozważania teoretyczne powodują rozwój nauk doświadczalnych. Zawsze jednak matematyka znajdzie wspólny mianownik dla ciekawych świata.

Można powiedzieć, że jednym z motorów rozwoju są wojny. Potrzeba udoskonalania uzbrojenia powoduje rozwój techniki, a z drugiej strony każdy dowódca dąży do wyboru optymalnej strategii. Jak rozlokować ludzi i sprzęt, jak skutecznie atakować, a jak bronić się z powodzeniem? I na te pytania matematyka odpowiada opracowując, teorię pościgu i ucieczki. Fragmenty tej teorii czytelnik znajdzie w rozdziale *Gry obronne*.

Mamy również przykład, gdy praktyczne rozważania inżyniera z XIX wieku doprowadziły do ciągle rozwijanej teorii związanej z poryzmem Ponceleta. Można o tym przeczytać w rozdziale drugim i zapoznać się z przeglądem wiedzy na ten temat od wyników najstarszych do tych najnowszych.

Pytanie, jak uzyskać najwięcej jak najmniejszym kosztem, oczywiście kojarzy się z ekonomią. I tu matematyka ma swoją ofertę. Opracowano wiele metod pomagających w miarę bezpiecznie inwestować. Rozdział *Metody deflatorowe w wycenie przepływów finansowych* opisuje właśnie jedno z takich narzędzi. Deflatory są stochastycznymi odpowiednikami klasycznych czynników dyskontujących i pozwalają na realistyczną wycenę przyszłych niedeterministycznych przepływów finansowych.

Okazuje się, że matematyka potrafi uporządkować nawet chaos. Ostatni rozdział monografii dotyczy właśnie tego tematu. Szczególny nacisk położony jest tu na topologiczny układ dynamiczny z czasem dyskretnym oraz zbiór fraktalny (fraktal) w przestrzeni euklidesowej.

Redaktor

Gry obronne

Witold Rzymowski¹
Agnieszka Surowiec²

Streszczenie

W pracy rozważano gry, w których obrońcy i napastnik poruszają się tak zwanym ruchem prostym. Wyznaczono kresy górne promieni kół oraz długości odcinków, które mogą być obronione przez jednego lub wielu obrońców.

Abstract

In the paper games in which defenders and attacker are moving with so called simple move are considered. The suprema of radii of discs and lengths of line segments, which can be defended by one or several defenders.

Słowa kluczowe: gry obronne, gry różniczkowe

1 Wstęp

W tej pracy rozważamy problem obrony koła, ucieczki z koła i obrony odcinka przez jednego i wielu obrońców. Będziemy zakładać, że wszyscy gracze poruszają się na płaszczyźnie \mathbb{R}^2 tak zwanym „ruchem prostym”.

Pierwszą pracę z teorii gier różniczkowych [12] napisał Hugo Steinhaus w roku 1927. W drugiej połowie dwudziestego wieku modelowaniem sytuacji konfliktowych, zwłaszcza typu pościg-ucieczka, zajmowano się najbardziej intensywnie

¹Katedra Matematyki Stosowanej; Politechnika Lubelska

²Katedra Metod Ilościowych w Zarządzaniu; Politechnika Lubelska

w Stanach Zjednoczonych i Związku Radzieckim. Wydano w tym okresie wiele prac i książek z teorii gier różniczkowych, między innymi [1] (pierwsza książka z teorii gier różniczkowych) i [6].

Na początku dwudziestego pierwszego wieku zainteresowano się matematycznymi problemami związanymi z ochroną wybranych fragmentów terenu, zabudowy itp. Rozważane są między innymi problemy patrolowania danego terenu przez poruszającego się samodzielnie robota, patrz [2], [8]. Problemy ochrony militarnej rozważano już wcześniej, patrz [1], przykład 1.9.2, [7] przykłady 9.6.3-4 i [11].

Precyzyjny opis rozważanych w pracy gier wymaga wprowadzenia stosownych oznaczeń i definicji.

Działania. Dla dowolnych $x = (x_1, x_2)$, $y = (y_1, y_2) \in \mathbb{R}^2$ przyjmujemy

$$\begin{aligned} Lx &\stackrel{\text{def}}{=} (-x_2, x_1), \quad Rx \stackrel{\text{def}}{=} (x_2, -x_1), \\ \langle x, y \rangle &\stackrel{\text{def}}{=} x_1y_1 + x_2y_2, \quad \|x\| \stackrel{\text{def}}{=} \sqrt{\langle x, x \rangle} = \sqrt{x_1^2 + x_2^2}, \\ x \wedge y &\stackrel{\text{def}}{=} \det \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \end{bmatrix} = x_1y_2 - x_2y_1 = \langle Lx, y \rangle = \langle x, Ry \rangle. \end{aligned}$$

Zbiory. Dla każdego $Z \subset \mathbb{R}^2$ symbole ∂Z i $\text{int}Z$ oznaczają odpowiednio brzeg i wnętrze zbioru Z . Natomiast symbolem $\text{conv}Z$ oznaczmy otoczkę wypukłą tego zbioru. Jeżeli $\emptyset \neq Z \subset \mathbb{R}^2$, to dla każdego $x \in \mathbb{R}^2$ definiujemy

$$\text{dist}(x, Z) = \inf_{z \in Z} \|x - z\|.$$

Jeżeli $\emptyset \neq Z \subset \mathbb{R}^2$ jest zbiorem zwartym, to ([13], twierdzenie 8, s. 158) $\text{conv}Z$ jest zbiorem zwartym i wypukłym. Dla każdego $x \in \mathbb{R}^2$ istnieje wtedy ([13], twierdzenie 1, s. 189 i twierdzenie 2, s. 190) rzut punktu x na zbiór $\text{conv}Z$, czyli jedyne $\hat{x} = \Pi(x) \in \partial \text{conv}Z$ spełniające warunek

$$\|x - \hat{x}\| = \text{dist}(x, Z).$$

Dla dowolnych $x, y \in \mathbb{R}^2$ ma wówczas miejsce nierówność

$$\|\Pi(x) - \Pi(y)\| \leq \|x - y\|.$$

Dla dowolnych $z \in \mathbb{R}^2$, $r \geq 0$ przyjmujemy

$$K[z, r] = \{x \in \mathbb{R}^2 : \|x - z\| \leq r\}.$$

Oczywiście

$$\partial K[z, r] = \{x \in \mathbb{R}^2 : \|x - z\| = r\}.$$

Trajektorie. Symbolem \mathcal{M}_r , gdzie $r > 0$, oznaczymy zbiór wszystkich funkcji mierzalnych

$$v : [0, \infty) \rightarrow K[0, r].$$

Elementy rodziny \mathcal{M}_r nazwiemy sterowaniami. Funkcje stałe będziemy utożsamiać z ich wartościami. W tym sensie $K[0, r] \subset \mathcal{M}_r$. Rodzina \mathcal{M}_1 będzie zbiorem dopuszczalnych sterowań każdego z obrońców. Natomiast zbiorem dopuszczalnych sterowań napastnika będzie rodzina \mathcal{M}_θ , gdzie $\theta \geq 1$. Z wyjątkiem ostatniej części pierwszego rozdziału będziemy zakładać nierówność $\theta > 1$.

Jeżeli w chwili $t = 0$ obrońca znajduje się w punkcie $a \in \mathbb{R}^2$, to może się poruszać po dowolnej trajektorii $x : [0, \infty) \rightarrow \mathbb{R}^2$, danej wzorem

$$x(t) = x(a, u, t) = a + \int_0^t u(s) ds, \text{ gdzie } u \in \mathcal{M}_1.$$

Zbiór wszystkich takich trajektorii oznaczymy przez $X(a)$, a punkt a nazwiemy pozycją początkową obrońcy. Z twierdzenia Lebesgue'a o różniczkowaniu całki i twierdzenia Rademachera ([3], wniosek 2.9.20, s. 165 i twierdzenie 3.1.6, s. 216) wynika, że $X(a)$ jest zbiorem wszystkich funkcji $x : [0, \infty) \rightarrow \mathbb{R}^2$ spełniających warunek początkowy $x(0) = a$ i warunek Lipschitza:

$$\|x(s) - x(t)\| \leq |s - t| \text{ dla } s, t \in [0, \infty).$$

Podobnie, jeżeli w chwili $t = 0$ napastnik znajduje się w punkcie $b \in \mathbb{R}^2$, to może się poruszać po dowolnej trajektorii $y : [0, \infty) \rightarrow \mathbb{R}^2$, danej wzorem

$$y(t) = y(b, v, t) = b + \int_0^t v(s) ds, \text{ gdzie } v \in \mathcal{M}_\theta.$$

Zbiór wszystkich takich trajektorii oznaczmy przez $Y(b)$, a punkt b nazwiemy pozycją początkową napastnika. $Y(b)$ jest zbiorem wszystkich funkcji $y : [0, \infty) \rightarrow \mathbb{R}^2$ spełniających warunek początkowy $y(0) = b$ i warunek Lipschitza:

$$\|y(s) - y(t)\| \leq \theta |s - t| \text{ dla } s, t \in [0, \infty).$$

Strategie. Użyjemy uproszczonej wersji strategii z pracy [9]. Ustalmy pozycje początkowe $a, b \in \mathbb{R}^2$.

Funkcję $\varphi : Y(b) \rightarrow X(a)$ nazwiemy nieantycypującą, jeżeli dla dowolnych $y, \hat{y} \in Y(b)$ i każdego $t \geq 0$ z warunku

$$y(s) = \hat{y}(s) \text{ dla } s \in [0, t]$$

wynika warunek

$$\varphi(y)(s) = \varphi(\hat{y})(s) \text{ dla } s \in [0, t].$$

Zbiorem dopuszczalnych strategii obrońcy jest zbiór $\Phi(a, b)$ złożony ze wszystkich nieantycypujących funkcji $\varphi : Y(b) \rightarrow X(a)$.

Funkcję $\psi : X(a) \rightarrow Y(b)$ nazwiemy nieantycypującą, jeżeli dla dowolnych $x, \hat{x} \in X(a)$ i każdego $t \geq 0$ z warunku

$$x(s) = \hat{x}(s) \text{ dla } s \in [0, t]$$

wynika warunek

$$\psi(x)(s) = \psi(\hat{x})(s) \text{ dla } s \in [0, t].$$

Zbiorem dopuszczalnych strategii napastnika jest zbiór $\Psi(a, b)$ złożony ze wszystkich nieantycypujących funkcji $\psi : X(a) \rightarrow Y(b)$.

Podane wyżej definicje strategii są sformalizowaną wersją żądania, by w każdym momencie $t \geq 0$ każdy z graczy podejmował decyzję wyłącznie na podstawie informacji o dotychczasowym przebiegu gry.

Zauważmy jeszcze, że funkcje stałe

$$\varphi : Y(b) \rightarrow X(a) \quad \text{oraz} \quad \psi : X(a) \rightarrow Y(b)$$

są nieantycypujące. Utożsamiając takie funkcje z ich wartościami (w tym przypadku z trajektoriami) możemy przyjąć, że trajektorie są strategiami. W tym sensie

$$X(a) \subset \Phi(a,b) \quad \text{oraz} \quad Y(b) \subset \Psi(a,b) \quad \text{dla} \quad a,b \in \mathbb{R}^2.$$

Obrona zbioru. Ustalmy niepusty i domknięty zbiór $Z \subset \mathbb{R}^2$ oraz pozycje początkowe $a,b \in \mathbb{R}^2$, gdzie $b \notin Z$. Obrońca jest wyposażony w broń o zasięgu $\rho \geq 0$ i broni dostępu do zbioru Z . Z kolei napastnik usiłuje wtargnąć do tego zbioru. Z wyjątkiem ostatniej części pierwszego rozdziału będziemy zakładać nierówność $\rho > 0$.

Jeżeli gracze poruszają się po trajektoriach, odpowiednio $x \in X(a)$ (obrońca) i $y \in Y(b)$ (napastnik), to w każdym momencie $t \geq 0$ strefą rażenia jest koło

$$K[x(t), \rho] = x(t) + K[0, \rho].$$

Nierówność $\|y(t) - x(t)\| < \rho$ oznacza, że przez pewien czas napastnik przebywał i przez pewien czas będzie jeszcze przebywał wewnątrz strefy rażenia $K[x(t), \rho]$. Ponieważ grozi to unieszkodliwieniem (zestrzeleniem, zatopieniem, itp), to napastnik powinien utrzymywać dystans $\|y(t) - x(t)\| \geq \rho$.

Definicja 1.1. Powiemy, że $\varphi \in \Phi(a,b)$ jest strategią skutecznej obrony zbioru Z , jeżeli dla każdej trajektorii $y \in Y(b)$ z warunku $y(t) \in Z$ wynika istnienie takiego t^* , że $\|y(t^*) - \varphi(y)(t^*)\| < \rho$.

Jeżeli zbiór $\Phi(a,b)$ zawiera strategię skuteczną obrony zbioru Z , to mówimy, że obrońca broni skutecznie zbioru Z .

Definicja 1.2. Powiemy, że $\psi \in \Psi(a, b)$ jest strategią skutecznego ataku na zbiór Z , jeżeli dla każdej trajektorii $x \in X(a)$ mamy $\|\psi(x)(t) - x(t)\| \geq \rho$ i istnieje takie t^* , że $\psi(x)(t^*) \in Z$.

Jeżeli zbiór $\Psi(a, b)$ zawiera strategię skutecznego ataku na zbiór Z , to mówimy, że napastnik atakuje skutecznie zbiór Z .

Zauważmy, że zbiór Z nie może być jednocześnie skutecznie obroniony i skutecznie zaatakowany.

2 Obrona koła

2.1 Opis gry

W grze biorą udział dwaj gracze: powolny obrońca **O** i szybszy napastnik **N**. Obaj gracze poruszają się na płaszczyźnie \mathbb{R}^2 tak zwanym ruchem prostym. Obrońca jest wyposażony w broń o zasięgu rażenia $\rho > 0$. Pozycją początkową obrońcy jest punkt $a = 0$. Pozycją początkową napastnika jest punkt $b \in \mathbb{R}^2$ spełniający warunek $\|b\| > \rho$. Obrońca broni koła o środku w punkcie 0, czyli pewnego koła $K[0, r]$. Zgodnie z przyjętą definicją 1.1, jeżeli $r < \rho$, to obrona koła $K[0, r]$ nie wymaga od obrońcy żadnego działania. Z drugiej strony, z uwagi na przewagę prędkości jaką posiada napastnik, liczba r nie może być dowolnie wielka. Oczywiście nie może być $r \geq \|b\|$, bo w momencie $t = 0$ mamy

$$\rho < \|b - a\| = \|b\| \leq r,$$

więc łatwo sprawdzić, że strategia stała, czyli trajektoria dana wzorem

$$y(t) = b + \frac{\theta t}{\|b\|} b, \quad t \geq 0,$$

będzie (patrz definicja 1.2) strategią skutecznego ataku na koło $K[0, r]$. Istotny wpływ na wielkość promienia bronionego koła ma prędkość napastnika. Wykazemy dalej, że liczba

$$r^* = \frac{\theta \rho}{\theta - 1}$$

jest wielkością krytyczną.

Twierdzenie 2.1. *Jeżeli pozycją początkową obrońcy jest punkt $a = 0$, a pozycja początkowa napastnika spełnia warunek*

$$\|b\| \geq \frac{\theta\rho}{\theta-1} \left(1 + \frac{\pi}{2}\right),$$

to obrońca skutecznie obroni każde koło $K[0, r]$, gdy $r < r^$. Natomiast napastnik zaatakuje skutecznie każde koło $K[0, r]$, gdy*

$$r > r^* \text{ oraz } \|b\| > \rho.$$

Dowód twierdzenia 2.1 jest treścią dwóch następujących podrozdziałów. Dla $\alpha \in \mathbb{R}$ przyjmujemy

$$\omega(\alpha) = (\cos \alpha, \sin \alpha).$$

Zauważmy, że dla dowolnych $\alpha, \beta \in \mathbb{R}$ mają miejsce równości:

$$\begin{aligned} \omega'(\alpha) &= (-\sin \alpha, \cos \alpha) = L\omega(\alpha), \\ \langle \omega(\alpha), \omega(\beta) \rangle &= \cos(\alpha - \beta), \quad \omega(\alpha) \wedge \omega(\beta) = \sin(\beta - \alpha). \end{aligned}$$

2.2 Obrona

Zakładamy, że w chwili $t = 0$ obrońca znajduje się w punkcie 0, a napastnik w punkcie

$$b \in \mathbb{R}^2, \text{ gdzie } \|b\| \geq \frac{\theta\rho}{\theta-1} \left(1 + \frac{\pi}{2}\right).$$

Ustalamy dowolną liczbę dodatnią $r < r^*$ i dobieramy takie $\hat{\rho} > 0$, dla którego

$$\frac{\theta\hat{\rho}}{\theta-1} = r.$$

Oczywiście $0 < \hat{\rho} < \rho$. Każdej trajektorii $y \in Y(b)$ przyporządkujemy trajektorię $x = \varphi(y) \in X(0)$. Przyjmijmy

$$t_0 = \frac{\hat{\rho}}{\theta-1} \frac{\pi}{2}$$

i wybierzmy dowolne $y \in Y(b)$. Dla $t \in [0, t_0]$ mamy

$$\begin{aligned} \|y(t)\| &= \|b + y(t) - b\| \geq \|b\| - \|y(t) - b\| \geq \|b\| - \theta t \geq \|b\| - \theta t_0 \\ &\geq \frac{\theta \rho}{\theta - 1} \left(1 + \frac{\pi}{2}\right) - \frac{\theta \hat{\rho}}{\theta - 1} \frac{\pi}{2} > \frac{\theta \hat{\rho}}{\theta - 1} = r = \frac{\hat{\rho}}{\theta - 1} + \hat{\rho}, \end{aligned}$$

więc do momentu t_0 napastnik będzie się znajdował poza kołem $K[0, r]$. Dla $t \in [0, t_0]$ mamy ([4], lemat 6.3. s.82)

$$y(t) = \|y(t)\| \omega(\alpha(t)),$$

gdzie

$$\alpha(t) = \alpha^* + \int_0^t \frac{y(s) \wedge y'(s)}{\|y(s)\|^2} ds \quad \text{oraz} \quad b = \|b\| \omega(\alpha^*) = \|b\| (\cos \alpha^*, \sin \alpha^*).$$

W przedziale $[0, t_0]$ trajektorię x zdefiniujemy (wstępnie) wzorem

$$\bar{x}(t) = \lambda(t) \omega(\alpha(t)),$$

gdzie $\lambda(0) = 0$ i λ jest funkcją ściśle rosnącą i lokalnie absolutnie ciągłą. Dla prawie wszystkich $t \in [0, t_0]$ mamy

$$\begin{aligned} x'(t) &= \lambda'(t) \omega(\alpha(t)) + \alpha'(t) \lambda(t) L\omega(\alpha(t)), \\ \|x'(t)\|^2 &= (\lambda'(t))^2 + (\alpha'(t))^2 \lambda^2(t). \end{aligned}$$

Funkcję λ dobieramy tak, by dla prawie wszystkich $t \in [0, t_0]$ było $\|x'(t)\|^2 = 1$.

W takim razie dla prawie wszystkich $t \in [0, t_0]$ powinno być

$$\lambda'(t) = \sqrt{1 - (\alpha'(t))^2 \lambda^2(t)}.$$

Wynika stąd, że funkcja λ jest rozwiązaniem zadania Cauchy'ego:

$$\begin{aligned} \lambda'(t) &= \sqrt{1 - (\alpha'(t))^2 \lambda^2(t)}, \\ \lambda(0) &= 0. \end{aligned}$$

Zauważmy, że dla prawie wszystkich $t \in [0, t_0]$

$$\begin{aligned} y'(t) &= \left\langle y'(t), \frac{y(t)}{\|y(t)\|} \right\rangle \omega(\alpha(t)) + \alpha'(t) \|y(t)\| L\omega(\alpha(t)), \\ \alpha'(t) &= \left\langle y'(t), \frac{L\omega(\alpha(t))}{\|y(t)\|} \right\rangle \leq \frac{\|y'(t)\|}{\|y(t)\|} \leq \frac{\theta}{r} \leq \frac{\theta-1}{\hat{\rho}}, \end{aligned}$$

więc dla prawie wszystkich $t \in [0, t_0]$

$$\begin{aligned} \lambda'(t) &= \sqrt{1 - (\alpha'(t))^2 \lambda^2(t)} \geq \sqrt{1 - \left(\frac{\theta-1}{\hat{\rho}}\right)^2 \lambda^2(t)}, \\ \frac{d}{dt} \left(\frac{\hat{\rho}}{\theta-1} \arcsin \left(\frac{\theta-1}{\hat{\rho}} \lambda(t) \right) \right) &= \frac{\lambda'(t)}{\sqrt{1 - \left(\frac{\theta-1}{\hat{\rho}}\right)^2 \lambda^2(t)}} \geq 1. \end{aligned}$$

W rezultacie, dla $t \in [0, t_0]$,

$$\begin{aligned} \frac{\hat{\rho}}{\theta-1} \arcsin \left(\frac{\theta-1}{\hat{\rho}} \lambda(t) \right) &\geq t, \\ \lambda(t) &\geq \frac{\hat{\rho}}{\theta-1} \sin \left(\frac{\theta-1}{\hat{\rho}} t \right), \end{aligned}$$

a w szczególności

$$\lambda(t_0) \geq \frac{\hat{\rho}}{\theta-1} \sin \left(\frac{\theta-1}{\hat{\rho}} t_0 \right) = \frac{\hat{\rho}}{\theta-1} \sin \frac{\pi}{2} = \frac{\hat{\rho}}{\theta-1}.$$

Istnieje zatem takie $t_1(y) \in (0, t_0)$, że

$$\|x(t_1(y))\| = \lambda(t_1(y)) = \frac{\hat{\rho}}{\theta-1}.$$

Dla $t \in [0, t_1(y)]$ przyjmujemy

$$\varphi(y)(t) = x(t) = \lambda(t)y(t).$$

Następnie definiujemy

$$T(y) = \{t > t_1(y) : \|y(t)\| \leq r\},$$

$$t_2(y) = \begin{cases} \infty, & \text{gd}y \ T(y) = \emptyset, \\ \min T(y), & \text{gd}y \ T(y) \neq \emptyset, \end{cases}$$

$$\varphi(y)(t) = x(t) = \frac{\hat{\rho}}{\theta - 1} \frac{y(t)}{\|y(t)\|} \quad \text{dla } t \geq t_1(y), \quad \text{gd}y \ t_2(y) = \infty,$$

$$\varphi(y)(t) = x(t) = \begin{cases} \frac{\hat{\rho}}{\theta - 1} \frac{y(t)}{\|y(t)\|} & \text{dla } t \in [t_1(y), t_2(y)], \\ x(t_2(y)) & \text{dla } t \geq t_2(y), \end{cases} \quad \text{gd}y \ t_2(y) < \infty.$$

Zauważmy, że dla wszystkich $t \in (t_1(y), t_2(y))$

$$x(t) = \frac{\hat{\rho}}{\theta - 1} \frac{y(t)}{\|y(t)\|} = \frac{\hat{\rho}}{\theta - 1} \omega(\alpha(t)),$$

a dla prawie wszystkich $t \in (t_1(y), t_2(y))$

$$\begin{aligned} \|x'(t)\| &= \frac{\hat{\rho}}{\theta - 1} \|\alpha'(t) L\omega(\alpha(t))\| = \frac{\hat{\rho}}{\theta - 1} |\alpha'(t)| = \frac{\hat{\rho}}{\theta - 1} \frac{|y(t) \wedge y'(t)|}{\|y(t)\|^2} \\ &\leq \frac{\hat{\rho}}{\theta - 1} \frac{\|y'(t)\|}{\|y(t)\|} \leq \frac{\hat{\rho}}{\theta - 1} \frac{\theta}{r} = 1, \end{aligned}$$

więc $x \in X(0)$. Łatwo też sprawdzić, że $\varphi : Y(b) \rightarrow X(0)$ jest funkcją nieantycypującą, więc $\varphi \in \Phi(0, b)$. Poza tym, jeżeli $T(y) \neq \emptyset$, to

$$\|\varphi(y)(t_2(y)) - y(t_2(y))\| = \hat{\rho} < \rho.$$

Wynika stąd, że $\varphi \in \Phi(0, b)$ jest strategią skutecznej obrony koła $K[0, r]$. Ponieważ

$$r < \frac{\theta \rho}{\theta - 1} = r^*$$

było wybrane dowolnie, to obrońca broni skutecznie każdego koła o promieniu mniejszym niż r^* .

2.3 Atak

Zakładamy, że w chwili $t = 0$ obrońca znajduje się w punkcie 0 , a napastnik w punkcie b spełniającym warunek $\|b\| \geq \rho$. Wykażemy, że napastnik zaatakuje skutecznie koło $K[0, r]$, gdy $r > r^*$. Ustalmy zatem dowolne $r > r^*$. Bez zmniejszenia ogólności rozważań możemy dodatkowo założyć nierówność $\|b\| > r$. Jeżeli

$$\|b\| \leq r + \theta(r - \rho),$$

to łatwo sprawdzić, że w momencie

$$\tau = \frac{\|b\| - r}{\theta} \leq r - \rho$$

napastnik osiągnie punkt $z = \frac{r}{\|b\|}b \in \partial K[0, r]$, poruszając się po trajektorii

$$\hat{y}(t) = b - \frac{\theta t}{\|b\|}b, \quad t \in [0, \tau].$$

Dla każdego $x \in X(0)$ i każdego $t \in [0, \tau]$ będzie wtedy

$$\|\hat{y}(t) - x(t)\| \geq \|b\| - \theta t - t \geq \|b\| - \theta \tau - \tau = r - \tau \geq \rho,$$

więc wystarczy rozważyć przypadek, w którym

$$\|b\| > r + \theta(r - \rho).$$

Każdej trajektorii $x \in X(0)$ przyporządkujemy trajektorię $y = \psi(x) \in Y(b)$. Weźmy dowolne $x \in X(0)$. Napastnik postanawia poruszać się po trajektorii

$$\hat{y}(t) = b - \frac{\theta t}{\|b\|}b, \quad t \geq 0,$$

aż do momentu, w którym nastąpi równość

$$\|\hat{y}(t) - x(t)\| = \rho.$$

Jeżeli taki moment nie nastąpi, to przyjmujemy

$$y(t) = \psi(x)(t) = b - \frac{\theta b}{\|b\|}t, \quad t \geq 0.$$

W takim przypadku napastnik skutecznie zaatakuje koło $K[0, r]$, bo dla $t = \frac{\|b\|}{\theta}$ będzie

$$y(t) = 0 \in \text{int}K[0, r],$$

a dla wszystkich $t \geq 0$ będzie

$$\|y(t) - x(t)\| > \rho.$$

Załóżmy zatem, że $t_0 = t_0(x) > 0$ jest pierwszym momentem, w którym

$$\|\hat{y}(t) - x(t)\| = \rho.$$

Od tej chwili napastnik zaczyna krążyć po brzegu ruchomego koła $K[x(t), \rho]$, wypatrując okazji skutecznego zaatakowania bronionego koła $K[0, r^*]$. Dla $t \geq t_0$ niech $\hat{y} : [t_0, \infty) \rightarrow \mathbb{R}^2$ będzie rozwiązaniem równania:

$$\begin{aligned} \hat{y}'(t) &= \left\langle x'(t), \frac{\hat{y}(t) - x(t)}{\rho} \right\rangle \frac{\hat{y}(t) - x(t)}{\rho} \\ &+ \sqrt{\theta^2 - \left\langle x'(t), \frac{\hat{y}(t) - x(t)}{\rho} \right\rangle^2} \frac{L(\hat{y}(t) - x(t))}{\rho}. \end{aligned}$$

Przyjmijmy

$$z(t) = \hat{y}(t) - x(t) \quad \text{dla } t \geq 0$$

i zauważmy, że dla prawie wszystkich $t \geq t_0$

$$\hat{y}'(t) = \left\langle x'(t), \frac{z(t)}{\rho} \right\rangle \frac{z(t)}{\rho} + \sqrt{\theta^2 - \left\langle x'(t), \frac{z(t)}{\rho} \right\rangle^2} \frac{Lz(t)}{\rho},$$

$$x'(t) = \left\langle x'(t), \frac{z(t)}{\rho} \right\rangle \frac{z(t)}{\rho} + \left\langle x'(t), \frac{Lz(t)}{\rho} \right\rangle \frac{Lz(t)}{\rho},$$

$$\begin{aligned} z'(t) &= \hat{y}'(t) - x'(t) \\ &= \left(\sqrt{\theta^2 - \left\langle x'(t), \frac{z(t)}{\rho} \right\rangle^2} - \left\langle x'(t), \frac{Lz(t)}{\rho} \right\rangle \right) \frac{Lz(t)}{\rho}, \end{aligned}$$

$$\begin{aligned}
\frac{d}{dt} \|z(t)\|^2 &= 2 \langle z(t), z'(t) \rangle \\
&= 2 \left(\sqrt{\theta^2 - \left\langle x'(t), \frac{z(t)}{\rho} \right\rangle^2} - \left\langle x'(t), \frac{Lz(t)}{\rho} \right\rangle \right) \left\langle z(t), \frac{Lz(t)}{\rho} \right\rangle \\
&= 0,
\end{aligned}$$

więc

$$\|z(t)\| = \|z(t_0)\| = \rho \text{ dla } t \geq t_0.$$

Wynika stąd, że napastnik może krążyć po brzegu ruchomego koła $K[x(t), \rho]$.

Porównamy prędkość obrotu wektora $z(t)$ z prędkością obrotu wektora $x(t)$.

Dla $t \geq t_0$ mamy ([4], lemat 6.3. s.82)

$$z(t) = \|z(t)\| \omega(\gamma(t)) = \rho \omega(\gamma(t)),$$

gdzie

$$\gamma'(t) = \frac{z(t) \wedge z'(t)}{\rho^2} \text{ dla prawie wszystkich } t \geq t_0.$$

Ponieważ dla prawie wszystkich $t \geq t_0$

$$\begin{aligned}
z(t) \wedge z'(t) &= \langle Lz(t), z'(t) \rangle \\
&= \left(\sqrt{\theta^2 - \left\langle x'(t), \frac{z(t)}{\rho} \right\rangle^2} - \left\langle x'(t), \frac{Lz(t)}{\rho} \right\rangle \right) \frac{\langle Lz(t), Lz(t) \rangle}{\rho} \\
&= \rho \left(\sqrt{\theta^2 - \left\langle x'(t), \frac{z(t)}{\rho} \right\rangle^2} - \left\langle x'(t), \frac{Lz(t)}{\rho} \right\rangle \right),
\end{aligned}$$

to dla prawie wszystkich $t \geq t_0$

$$\gamma'(t) = \frac{\sqrt{\theta^2 - \left\langle x'(t), \frac{z(t)}{\rho} \right\rangle^2} - \left\langle x'(t), \frac{Lz(t)}{\rho} \right\rangle}{\rho}.$$

Dla prawie wszystkich $t \geq t_0$ mamy

$$\begin{aligned} & \sqrt{\theta^2 - \left\langle x'(t), \frac{z(t)}{\rho} \right\rangle^2} - \left\langle x'(t), \frac{Lz(t)}{\rho} \right\rangle \\ &= \frac{\theta^2 - \left\langle x'(t), \frac{z(t)}{\rho} \right\rangle^2 - \left\langle x'(t), \frac{Lz(t)}{\rho} \right\rangle^2}{\sqrt{\theta^2 - \left\langle x'(t), \frac{z(t)}{\rho} \right\rangle^2} + \left\langle x'(t), \frac{Lz(t)}{\rho} \right\rangle} \\ &= \frac{\theta^2 - \|x'(t)\|^2}{\sqrt{\theta^2 - \left\langle x'(t), \frac{z(t)}{\rho} \right\rangle^2} + \left\langle x'(t), \frac{Lz(t)}{\rho} \right\rangle} \geq \frac{\theta^2 - 1}{\theta + 1} = \theta - 1, \end{aligned}$$

więc

$$\gamma'(t) \geq \frac{\theta - 1}{\rho} \text{ dla prawie wszystkich } t \geq t_0.$$

Rozważmy teraz trajektorię x . Możliwe są dwa przypadki:

$$(a) T(x) \stackrel{\text{def}}{=} \{t \in [t_0, \infty) : \|x(t)\| \leq r - \rho\} \neq \emptyset, \quad (b) T(x) = \emptyset.$$

W przypadku (a) definiujemy

$$t_1 = t_1(x) = \min T(x),$$

otrzymując relację $y(t_1) \in K[0, r]$, bo

$$\begin{aligned} \|y(t_1)\| &= \|x(t_1) + y(t_1) - x(t_1)\| \\ &\leq \|x(t_1)\| + \|y(t_1) - x(t_1)\| = \|x(t_1)\| + \rho \leq r. \end{aligned}$$

Natomiast w przypadku (b) mamy

$$\|x(t)\| > r - \rho \text{ dla } t \geq t_0.$$

Z analogicznych powodów jak w przypadku wektora $z(t)$ mamy wówczas

$$x(t) = \|x(t)\| \omega(\alpha(t)),$$

gdzie

$$\alpha'(t) = \frac{x(t) \wedge x'(t)}{\|x(t)\|^2} \text{ dla prawie wszystkich } t \geq t_0.$$

Stąd, dla prawie wszystkich $t \geq t_0$, otrzymujemy

$$|\alpha'(t)| \leq \frac{\|x(t)\| \|x'(t)\|}{\|x(t)\|^2} = \frac{\|x'(t)\|}{\|x(t)\|} < \frac{1}{r-\rho}.$$

W rezultacie, ponieważ

$$\delta = \frac{\theta-1}{\rho} - \frac{1}{r-\rho} > \frac{\theta-1}{\rho} - \frac{1}{r^*-\rho} = \frac{\theta-1}{\rho} - \frac{\theta-1}{\rho} = 0,$$

to dla $t \geq t_0$ otrzymujemy

$$\begin{aligned} \gamma(t) - \alpha(t) &= \int_{t_0}^t (\gamma'(s) - \alpha'(s)) ds > \left(\frac{\theta-1}{\rho} - \frac{1}{r-\rho} \right) (t-t_0) \\ &= \delta(t-t_0). \end{aligned}$$

Bez zmniejszenia ogólności rozważań możemy założyć warunek

$$0 \leq \gamma(t_0) - \alpha(t_0) < \pi.$$

W pozostałych przypadkach rozumowanie jest podobne. Ponieważ różnica $\gamma - \alpha$ jest funkcją ciągłą i

$$\lim_{t \rightarrow \infty} \delta(t-t_0) = \infty,$$

to istnieje takie $t_1 = t_1(x) > t_0$, że

$$\gamma(t_1) - \alpha(t_1) = \pi \text{ oraz } 0 \leq \gamma(t) - \alpha(t) < \pi, \text{ gdy } t_0 \leq t < t_1.$$

Wobec powyższego w przypadku (b) definiujemy

$$y(t) = \psi(x)(t) = \begin{cases} \hat{y}(t), & \text{gdy } t \in [t_0, t_1], \\ \hat{y}(t_1) - \frac{\hat{y}(t_1)\theta}{\|\hat{y}(t_1)\|} (t-t_1), & \text{gdy } t \in [t_1, \infty). \end{cases}$$

Dla $t \geq t_1$ będzie

$$\|y(t) - x(t)\| \geq \rho + (\theta-1)(t-t_1) \geq \rho,$$

a dla

$$t_2 = t_2(x) = t_1 + \frac{\|\hat{y}(t_1)\|}{\theta}$$

będzie $y(t_2) = 0 \in \text{int}K[0, r]$. Ponieważ skonstruowane wyżej odwzorowanie $\psi : X(0) \rightarrow Y(b)$ jest nieantycypujące, to $\psi \in \Psi(0, b)$ jest strategią skutecznego ataku na koło $K[0, r]$. Liczba $r > r^*$ była wybrana dowolnie, więc napastnik zaatakuje skutecznie każde koło o promieniu większym niż r^* .

Uwaga 2.1. W pracy [11] dowiedziono, że pole dowolnego obszaru, który może być obroniony, nie może być większe niż pole koła o promieniu r^* .

2.4 Ucieczka z koła

W grze bierze udział jeden napastnik \mathbf{N} i n obrońców $\mathbf{O}_1, \mathbf{O}_2, \dots, \mathbf{O}_n$. W chwili $t = 0$ napastnik znajduje się w punkcie $b = 0$. Natomiast obrońcy znajdują się na brzegu koła $K[0, 1]$, czyli każdy z obrońców \mathbf{O}_k znajduje się w pewnym punkcie $a_k \in \partial K[0, 1]$. Każdy z graczy może poruszać się po trajektorii spełniającej warunek Lipschitza ze stałą $\theta = 1$, ale obrońcy nie mogą opuścić brzegu koła $K[0, 1]$. Tym razem celem napastnika jest wydostanie się z koła $K[0, 1]$, z zachowaniem pewnej dodatniej odległości $\rho > 0$ od każdego z obrońców.

Trajektorie. Zbiorem dopuszczalnych trajektorii uciekającego napastnika jest zbiór $Y(0)$ wszystkich funkcji $y : [0, \infty) \rightarrow \mathbb{R}^2$ spełniających warunek początkowy $y(0) = 0$ i warunek Lipschitza:

$$\|y(s) - y(t)\| \leq |s - t| \quad \text{dla } s, t \geq 0.$$

Zbiorem dopuszczalnych trajektorii każdego z obrońców \mathbf{O}_k jest zbiór $\check{X}(a_k)$ wszystkich funkcji $x_k : [0, \infty) \rightarrow \partial K[0, 1]$ spełniających warunek początkowy $x_k(0) = a_k$ i warunek Lipschitza:

$$\|x_k(s) - x_k(t)\| \leq |s - t| \quad \text{dla } s, t \geq 0.$$

Strategie. Dopuszczalną strategią każdego z obrońców \mathbf{O}_k jest (patrz definicja 1.1) zbiór $\Phi(a_k, 0)$. Dopuszczalną strategią napastnika jest każda nieantycypująca

funkcja

$$\psi : \prod_{k=1}^n \check{X}(a_k) \rightarrow Y(0).$$

Zbiór wszystkich takich funkcji oznaczymy przez $\Psi(a_1, a_2, \dots, a_n, 0)$. Podobnie jak w poprzednim rozdziale trajektorie ze zbioru $Y(0)$ będziemy utożsamiać ze strategiami stałymi (ze zbioru $\Psi(a_1, a_2, \dots, a_n, 0)$).

Strategia skutecznej ucieczki. Powiemy, że $\psi \in \Psi(a_1, a_2, \dots, a_n, 0)$ jest strategią skutecznej ucieczki, jeżeli

$$\inf_{(x_1, x_2, \dots, x_n) \in \prod_{k=1}^n \check{X}(a_k)} \inf_{t \geq 0} \min_{k=1, 2, \dots, n} \|\psi(x_1, x_2, \dots, x_n)(t) - x_k(t)\| > 0.$$

W przypadku $n \in \{1, 2, 3\}$ istnieje prosty sposób ucieczki. Rozważmy przypadek $n = 3$. Trzy punkty a_1, a_2, a_3 dzielą okrąg $\partial K[0, 1]$ na trzy łuki. Długość co najmniej jednego z tych łuków nie może być mniejsza niż $\frac{2\pi}{3}$. Powiedzmy, że łuk pomiędzy punktami a_1, a_2 jest takim łukiem. Niech $a \in \partial K[0, 1]$ będzie punktem dzielącym ten łuk na dwa łuki o równych długościach $\delta \geq \frac{\pi}{3}$, a x_k niech będzie trajektorią gracza \mathbf{O}_k dla $k = 1, 2, 3$. Zdefiniujmy

$$y(t) = ta \text{ dla } t \geq 0.$$

W momencie $t_1 = 1$ będzie $y(t_1) = a \in \partial K[0, 1]$. W tym samym czasie każdy z obrońców pokona drogę nie większą niż

$$1 < \frac{\pi}{3} \leq \delta.$$

Zatem

$$\min_{t \in [0, 1]} \min_{k=1, 2, 3} \|y(t) - x_k(t)\| \geq \delta - 1 \geq \frac{\pi}{3} - 1 > 0.$$

Ponieważ

$$\lim_{t \rightarrow \infty} \min_{k=1, 2, 3} \|y(t) - x_k(t)\| = \infty,$$

to

$$\inf_{t \geq 0} \min_{k=1, 2, 3} \|y(t) - x_k(t)\| > 0.$$

W przypadku $n > 3$ strategia skutecznej ucieczki z koła $K[0, 1]$ jest bardziej skomplikowana. Definiujemy najpierw pomocniczą trajektorię \hat{y} . Ustalamy $n \geq 4$ i przyjmujemy

$$\alpha = \frac{\pi}{n}, \quad \delta = \frac{\alpha - 2 \sin \frac{\alpha}{2}}{6}, \quad r = 1 - \delta, \quad \beta = 3\delta = \frac{\alpha - 2 \sin \frac{\alpha}{2}}{2}.$$

Oczywiście $\delta, r \in (0, 1)$. Zauważmy, że

$$2r \sin \frac{\alpha}{2} < 2 \sin \frac{\alpha}{2} = \alpha - 2\beta. \quad (2.1)$$

Następnie przyjmujemy

$$t_0 = r = 1 - \delta, \quad \hat{y}(t) = t\omega(0) \quad \text{dla } t \in [0, t_0],$$

$$t_j = 2rj \sin \frac{\alpha}{2}, \quad z_j = r\omega(j\alpha) \quad \text{dla } j = 0, 1, \dots, n,$$

$$\hat{y}(t) = z_{j-1} + \frac{t - t_{j-1}}{2r \sin \frac{\alpha}{2}} (z_j - z_{j-1}) \quad \text{dla } j = 1, 2, \dots, n,$$

$$\Gamma_j = \{\omega(\gamma) : \gamma \in (j\alpha - \beta, j\alpha + \beta)\} \quad \text{dla } j = 0, 1, \dots, n.$$

Łatwo sprawdzić, że $\hat{y} : [0, t_n] \rightarrow K[0, r]$ spełnia warunek Lipschitza ze stałą 1,

$$\hat{y}(t_j) = z_j \quad \text{dla } j = 0, 1, \dots, n,$$

punkty z_j są wierzchołkami $2n$ -kąta foremnego wpisanego w koło $K[0, r]$, a $\hat{y} : [0, t_n] \rightarrow K[0, r]$ jest parametryzacją górnej połowy brzegu tego wielokąta.

Dla każdego $k = 1, 2, \dots, n$ wybierzmy dowolnie trajektorię x_k obrońcy \mathbf{O}_k . Zdefiniujemy trajektorię

$$y = \Psi(x_1, x_2, \dots, x_n).$$

Dla $t \in [0, t_0]$ przyjmujemy $y(t) = \hat{y}(t)$. Przypuśćmy, że dla pewnego $k_1 \in \{1, 2, \dots, n\}$ ma miejsce relacja $x_{k_1}(t_0) \in \Gamma_0$. Z nierówności (2.1) wynika, że wtedy

$$x_{k_1}(t_j) \notin \Gamma_j \quad \text{dla } j = 1, 2, \dots, n.$$

Podobnie, jeżeli dla pewnego $k_i \in \{1, 2, \dots, n\}$ ma miejsce relacja $x_{k_i}(t_{i-1}) \in \Gamma_0$, to

$$x_{k_i}(t_j) \notin \Gamma_j \text{ dla } j = i, i+1, \dots, n.$$

Wobec powyższego dla pewnego $j \in \{0, 1, \dots, n\}$ musi być

$$x_k(t_j) \notin \Gamma_j \text{ dla } k = 1, 2, \dots, n. \quad (2.2)$$

Przyjmując, że $t_j = t_j(x_1, x_2, \dots, x_n)$ jest pierwszym momentem, w którym nastąpiły relacje (2.2) definiujemy

$$\psi(x_1, x_2, \dots, x_n)(t) = y(t) = \begin{cases} \hat{y}(t), & \text{gdy } t \in [0, t_j], \\ \hat{y}(t_j) + (t - t_j)\omega(j\alpha), & \text{gdy } t \in [t_j, \infty). \end{cases}$$

Łatwo sprawdzić, że $\psi : \prod_{k=1}^n \check{X}_k(a_k) \rightarrow Y(0)$ jest funkcją nieantycypującą oraz

$$\begin{aligned} \min_{k=1,2,\dots,n} \|y(t) - x_k(t)\| &\geq \delta \text{ dla } t \in [0, t_j(x_1, x_2, \dots, x_n)], \\ \min_{k=1,2,\dots,n} \|y(t) - x_k(t)\| &\geq 2 \sin \beta \geq \delta \text{ dla } t \in [t_j, t_j + \delta], \\ \min_{k=1,2,\dots,n} \|y(t) - x_k(t)\| &\geq 2 \sin \frac{\beta}{2} \geq \delta \text{ dla } t \in [t_j + \delta, \infty). \end{aligned}$$

Wniosek 2.1. Ponieważ

$$\begin{aligned} \frac{\alpha - 2 \sin \frac{\alpha}{2}}{6} &> \frac{1}{6} \left(\alpha - 2 \left(\frac{\alpha}{2} - \frac{(\frac{\alpha}{2})^3}{3!} + \frac{(\frac{\alpha}{2})^5}{5!} \right) \right) = \frac{1}{11520} \alpha^3 (80 - \alpha^2) \\ &> \frac{78}{11520} \alpha^3 = \frac{78}{11520} \left(\frac{\pi}{n} \right)^3 \end{aligned}$$

i mamy

$$\frac{78}{11520} \pi^3 > 0.2,$$

to uciekający z koła napastnik utrzyma dystans do każdego z obrońców większy niż $\frac{0.2}{n^3}$.

3 Obrona odcinka przez jednego obrońcę

3.1 Wstępny opis gry

Obrona odcinka może być elementem obrony większego terytorium. Zaczniemy od przypadku, w którym powolny obrońca \mathbf{O} , wyposażony w broń o zasięgu rażenia $\rho > 0$, broni odcinka przed szybszym napastnikiem \mathbf{N} . Celem napastnika jest wtargnięcie na broniony odcinek z zachowaniem bezpiecznej odległości od obrońcy. Ponieważ położenie bronionego odcinka nie ma większego znaczenia, to będziemy początkowo zakładać, że broniony będzie odcinek

$$\Delta(\ell) = [-\ell, \ell] \times \{0\} \text{ (odcinek na osi odciętych),}$$

gdzie $\ell > \rho$. Zakładamy dalej, że obrońca porusza się z prędkością 1 lub każdą mniejszą niż 1. Atak na odcinek $\Delta(\ell)$ może nastąpić z każdego punktu $b = (b_1, b_2) \in \mathbb{R}^2$ leżącego dostatecznie daleko od bronionego odcinka.

Napastnik atakujący odcinek $\Delta(\ell)$ może poruszać się po całej płaszczyźnie \mathbb{R}^2 z prędkością $\theta > 1$ lub każdą mniejszą niż θ . Inaczej mówiąc gracze poruszają się po trajektoriach opisanych na początku pierwszego rozdziału.

Przez skuteczną obronę odcinka rozumiemy takie działanie obrońcy, by przed ewentualnym wtargnięciem na broniony odcinek napastnik znalazł się wewnątrz strefy rażenia, czyli w odległości mniejszej niż ρ od obrońcy. Stacjonujący w ustalonym miejscu obrońca broni odcinków o długości mniejszej niż 2ρ , bo broni kół o promieniu mniejszym niż ρ . Natomiast obrońca ruchomy może bronić znacznie dłuższych odcinków, broniąc soczewkowatych zbiorów zawierających broniony odcinek. Rozważymy dwa przypadki.

(I) Obrońca porusza się tylko po bronionym odcinku. Oczywiście broniony odcinek nie może być zbyt długi. Liczba

$$\Lambda_{\mathbf{I}}(\theta, \rho) = 2\rho \left(1 + \frac{1}{\theta^2 - 1} \int_0^{\frac{\pi}{2}} \left(\cos \beta + \sqrt{\theta^2 - \sin^2 \beta} \right) d\beta \right)$$

jest długością krytyczną. Wykażemy, że obrońca obroni każdy odcinek o długości mniejszej niż $\Lambda_{\mathbf{I}}(\theta, \rho)$. Natomiast napastnik wtargnie bezkarnie na każdy odcinek o długości większej niż $\Lambda_{\mathbf{I}}(\theta, \rho)$.

(II) Obrońca może poruszać się po całej płaszczyźnie \mathbb{R}^2 . Oczywiście i w tym przypadku broniony odcinek nie może być zbyt długi. Liczba

$$\Lambda_{\mathbf{II}}(\theta, \rho) = \rho \frac{\theta(2\theta + \pi)}{\theta^2 - 1}$$

jest długością krytyczną. Wykażemy, że obrońca obroni każdy odcinek o długości mniejszej niż $\Lambda_{\mathbf{II}}(\theta, \rho)$. Natomiast napastnik wtargnie bezkarnie na każdy odcinek o długości większej niż $\Lambda_{\mathbf{II}}(\theta, \rho)$.

3.2 Przypadek I

Upraszczając dalsze rozważania przyjmiemy, że w chwili $t = 0$ obrońca znajduje się w punkcie $a = (0, 0)$. Schemat obrony jest następujący. Zakładamy, że $2\ell < \Lambda_{\mathbf{I}}(\theta, \rho)$ i dobieramy takie $r < \rho$, dla którego

$$\Lambda_{\mathbf{I}}(\theta, r) = 2r \left(1 + \frac{1}{\theta^2 - 1} \int_0^{\frac{\pi}{2}} \left(\cos \beta + \sqrt{\theta^2 - \sin^2 \beta} \right) d\beta \right) = 2\ell.$$

Następnie konstruujemy zbiór wypukły $\Omega(\theta, r)$ zawierający broniony odcinek $\Delta(\ell)$ i funkcję

$$F : \partial\Omega(\theta, r) \rightarrow \Delta(\ell).$$

Obrońca będzie bronił zbioru $\Omega(\theta, r)$, który w związku z tym nazwiemy obszarem bronionym. W tym celu definiujemy:

$$\begin{aligned} \omega^+(\alpha) &= (\sin \alpha, \cos \alpha), \quad \omega^-(\alpha) = (\sin \alpha, -\cos \alpha) \quad \text{dla } \alpha \in \mathbb{R}, \\ \tau(\alpha) &= \frac{1}{\theta^2 - 1} \int_{-\frac{\pi}{2}}^{\alpha} \left(\cos \beta + \sqrt{\theta^2 - \sin^2 \beta} \right) d\beta \quad \text{dla } \alpha \in \left[-\frac{\pi}{2}, \frac{\pi}{2} \right], \end{aligned}$$

$$\begin{aligned}
\Gamma^+(\alpha) &= (\Gamma_1^+(\alpha), \Gamma_2^+(\alpha)) \\
&= (-\ell, 0) + r(1 + \tau(\alpha), 0) + r\omega^+(\alpha) \text{ dla } \alpha \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right], \\
\Gamma^-(\alpha) &= (\Gamma_1^-(\alpha), \Gamma_2^-(\alpha)) \\
&= (-\ell, 0) + r(1 + \tau(\alpha), 0) + r\omega^-(\alpha) \text{ dla } \alpha \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right].
\end{aligned}$$

Łuki

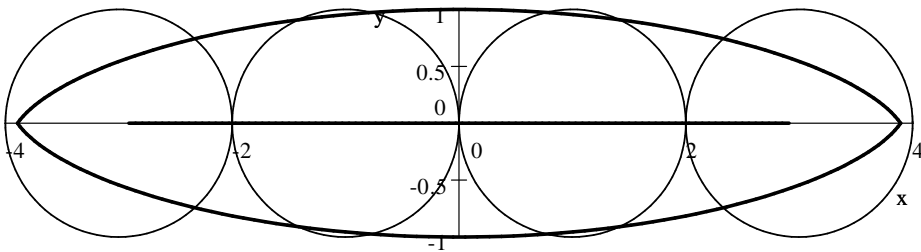
$$\partial^+\Omega(\theta, r) = \Gamma^+\left(\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]\right) \text{ oraz } \partial^-\Omega(\theta, r) = \Gamma^-\left(\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]\right)$$

są odpowiednio górnymi i dolnymi częściami brzegu bronionego obszaru $\Omega(\theta, r)$.

Przykład 3.1. Dla $\theta = \sqrt{2}$ i $r = 1$ mamy

$$\frac{1}{2}\Lambda_{\mathbf{I}}(\theta, r) = \left(1 + \frac{1}{2-1} \int_0^{\frac{\pi}{2}} \left(\cos \beta + \sqrt{2 - \sin^2 \beta}\right) d\beta\right) \approx 3.910099$$

Zauważmy, że w rozważanym przypadku poruszający się po linii prostej obrońca może obronić odcinek niewiele krótszy niż czterech obrońców stacjonujących w ustalonych punktach.



Rysunek 1: Brzeg zbioru $\Omega(\theta, r)$, przedział (obraz) $F(\partial\Omega(\theta, r))$ i strefy rażenia czterech obrońców stacjonujących w punktach $(-3, 0)$, $(-1, 0)$, $(1, 0)$, $(3, 0)$.

Lemat 3.1. *Zbiór*

$$\Omega(\theta, r) = \text{conv} \left(\Gamma^+ \left(\left[-\frac{\pi}{2}, \frac{\pi}{2} \right] \right) \cup \Gamma^- \left(\left[-\frac{\pi}{2}, \frac{\pi}{2} \right] \right) \right)$$

jest ściśle wypukły.

Dowód. Ponieważ

$$\Gamma^+ \left(-\frac{\pi}{2} \right) = \Gamma^- \left(-\frac{\pi}{2} \right) = (-\ell, 0), \quad \Gamma^+ \left(\frac{\pi}{2} \right) = \Gamma^- \left(\frac{\pi}{2} \right) = (\ell, 0),$$

a dla $\alpha \in \left(-\frac{\pi}{2}, \frac{\pi}{2} \right)$ mamy

$$\Gamma_2^-(\alpha) = -r \cos \alpha < 0 < r \cos \alpha = \Gamma_2^+(\alpha),$$

to wystarczy sprawdzić, że dla $\alpha \in \left(-\frac{\pi}{2}, \frac{\pi}{2} \right)$ mają miejsce nierówności

$$\frac{d}{d\alpha} \Gamma^+(\alpha) \wedge \frac{d^2}{d\alpha^2} \Gamma^+(\alpha) < 0 < \frac{d}{d\alpha} \Gamma^-(\alpha) \wedge \frac{d^2}{d\alpha^2} \Gamma^-(\alpha).$$

Upraszczając oznaczenia przyjmijmy

$$c(\alpha) = \cos \alpha, \quad s(\alpha) = \sin \alpha \quad \text{dla } \alpha \in \mathbb{R}.$$

Dla każdego $\alpha \in \left[-\frac{\pi}{2}, \frac{\pi}{2} \right]$ mamy

$$\begin{aligned} \Gamma^+(\alpha) &= \begin{bmatrix} -\ell \\ 0 \end{bmatrix} + r \begin{bmatrix} 1 + \tau(\alpha) \\ 0 \end{bmatrix} + r \begin{bmatrix} s(\alpha) \\ c(\alpha) \end{bmatrix} \\ &= \begin{bmatrix} -\ell \\ 0 \end{bmatrix} + r \begin{bmatrix} 1 + \tau(\alpha) + s(\alpha) \\ c(\alpha) \end{bmatrix}, \end{aligned}$$

a dla każdego $\alpha \in \left(-\frac{\pi}{2}, \frac{\pi}{2} \right)$ mamy

$$\begin{aligned} \tau'(\alpha) &= \frac{1}{\theta^2 - 1} \left(c(\alpha) + \sqrt{\theta^2 - s^2(\alpha)} \right), \\ \tau''(\alpha) &= -\frac{s(\alpha)}{\theta^2 - 1} \left(1 + \frac{c(\alpha)}{\sqrt{\theta^2 - s^2(\alpha)}} \right), \end{aligned}$$

$$\frac{d}{d\alpha}\Gamma^+(\alpha) = r \begin{bmatrix} \tau'(\alpha) + c(\alpha) \\ -s(\alpha) \end{bmatrix}, \quad \frac{d^2}{d\alpha^2}\Gamma^+(\alpha) = r \begin{bmatrix} \tau''(\alpha) - s(\alpha) \\ -c(\alpha) \end{bmatrix},$$

$$\begin{aligned} \frac{d}{d\alpha}\Gamma^+(\alpha) \wedge \frac{d^2}{d\alpha^2}\Gamma^+(\alpha) &= r^2 \det \begin{bmatrix} \tau'(\alpha) + c(\alpha) & \tau''(\alpha) - s(\alpha) \\ -s(\alpha) & -c(\alpha) \end{bmatrix} \\ &= -r^2 (1 + c(\alpha)\tau'(\alpha) - s(\alpha)\tau''(\alpha)), \end{aligned}$$

$$\begin{aligned} (\theta^2 - 1)(1 + c(\alpha)\tau'(\alpha) + s(\alpha)\tau''(\alpha)) &= \theta^2 - 1 + c^2(\alpha) \\ &+ c(\alpha)\sqrt{\theta^2 - s^2(\alpha)} + s^2(\alpha) \left(1 + \frac{c(\alpha)}{\sqrt{\theta^2 - s^2(\alpha)}}\right) > 0, \end{aligned}$$

więc

$$\frac{d}{d\alpha}\Gamma^+(\alpha) \wedge \frac{d^2}{d\alpha^2}\Gamma^+(\alpha) < 0 \quad \text{dla } \alpha \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right).$$

Podobnym sposobem dowodzimy warunku

$$\frac{d}{d\alpha}\Gamma^-(\alpha) \wedge \frac{d^2}{d\alpha^2}\Gamma^-(\alpha) > 0 \quad \text{dla } \alpha \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right).$$

□

Funkcja $F : \partial\Omega(\theta, r) \rightarrow \Delta(\ell)$ powinna mieć następującą własność. Jeżeli napastnik porusza się po trajektorii y leżącej na brzegu zbioru $\Omega(\theta, r)$, to punkt $F(y(t))$ porusza się po odcinku $\Delta(\ell)$ z prędkością co najwyżej 1, pozostając stale w odległości $r < \rho$ od punktu $y(t)$. Wobec powyższego obrońca będzie w stanie dogonić punkt $F(y(t))$, a następnie, stale za nim nadążając, utrzymywać odległość r od napastnika. W początkowej fazie ataku napastnik jest w dużej odległości od zbioru $\Omega(\theta, r)$, więc jego trajektorię y zastępujemy rzutem y na zbiór $\Omega(\theta, r)$, który oznaczymy przez \hat{y} . Prędkość punktu $\hat{y}(t)$ nie może przekroczyć prędkości punktu $y(t)$, więc obrońca będzie w stanie dogonić punkt $F(\hat{y}(t))$, a następnie, stale za nim nadążając, utrzymywać odległość r od punktu $\hat{y}(t)$. Dlatego, jeżeli napastnik postanowi wtargnąć na broniony odcinek, to wcześniej trafi na brzeg zbioru $\Omega(\theta, r)$, a wtedy znajdzie się w odległości $r < \rho$ od obrońcy.

Zdefiniujemy żadaną funkcję $F : \partial\Omega(\theta, r) \rightarrow \Delta(\ell)$. Zaczniemy od pewnej obserwacji. Przyjmijmy

$$\xi(\alpha) = r(1 + \tau(\alpha)) \quad \text{dla } \alpha \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right].$$

Dla $\alpha \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ mamy

$$\theta^2 (\xi'(\alpha))^2 = \left\| \frac{d}{d\alpha} \Gamma^+(\alpha) \right\|^2 = \left\| \frac{d}{d\alpha} \Gamma^-(\alpha) \right\|^2. \quad (3.1)$$

Dowód drugiej równości jest łatwy, więc dowiedzimy tylko pierwszej. Dla każdego $\alpha \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ mamy

$$(\xi'(\alpha))^2 = r^2 (\tau'(\alpha))^2 = \frac{r^2}{\theta^2 - 1} \left(\cos \alpha + \sqrt{\theta^2 - \sin^2 \alpha} \right)^2$$

$$\begin{aligned} \left\| \frac{d}{d\alpha} \Gamma^+(\alpha) \right\|^2 &= \left(\frac{d}{d\alpha} \Gamma_1^+(\alpha) \right)^2 + \left(\frac{d}{d\alpha} \Gamma_2^+(\alpha) \right)^2 \\ &= r^2 \left((\tau'(\alpha) + \cos \alpha)^2 \right) + r^2 \sin^2 \alpha \\ &= r^2 \left(1 + 2\tau'(\alpha) \cos \alpha + (\tau'(\alpha))^2 \right), \end{aligned}$$

$$\frac{1}{r^2} \left(\theta^2 (\xi'(\alpha))^2 - \left\| \frac{d}{d\alpha} \Gamma^+(\alpha) \right\|^2 \right) = (\theta^2 - 1) (\tau'(\alpha))^2 - 2\tau'(\alpha) \cos \alpha - 1,$$

$$(\theta^2 - 1) (\tau'(\alpha))^2 = \frac{1}{\theta^2 - 1} \left(\cos^2 \alpha + 2\sqrt{\theta^2 - \sin^2 \alpha} \cos \alpha + \theta^2 - \sin^2 \alpha \right),$$

$$2\tau'(\alpha) \cos \alpha = \frac{2}{\theta^2 - 1} \left(\cos^2 \alpha + \sqrt{\theta^2 - \sin^2 \alpha} \cos \alpha \right),$$

$$\begin{aligned} &(\theta^2 - 1) (\tau'(\alpha))^2 - 2\tau'(\alpha) \cos \alpha \\ &= \frac{1}{\theta^2 - 1} (-\cos^2 \alpha + \theta^2 - \sin^2 \alpha) = \frac{1}{\theta^2 - 1} (\theta^2 - 1) = 1, \end{aligned}$$

więc

$$\theta^2 (\xi'(\alpha))^2 - \left\| \frac{d}{d\alpha} \Gamma^+(\alpha) \right\|^2 = 0 \quad \text{dla } \alpha \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right].$$

Zauważmy dalej, że funkcja $\Gamma_1^+ : [-\frac{\pi}{2}, \frac{\pi}{2}] \xrightarrow{\text{na}} [-\ell, \ell]$ jest ściśle rosnąca. Dla $\alpha \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ mamy bowiem

$$\begin{aligned} \frac{d}{d\alpha} \Gamma_1^+(\alpha) &= \tau'(\alpha) + \cos \alpha = \frac{1}{\theta^2 - 1} \left(\cos \alpha + \sqrt{\theta^2 - \sin^2 \alpha} \right) + \cos \alpha \\ &= \frac{\theta^2}{\theta^2 - 1} \cos \alpha + \frac{1}{\theta^2 - 1} \sqrt{\theta^2 - \sin^2 \alpha} \geq \frac{\theta}{\theta^2 - 1} > 0. \end{aligned}$$

Istnieje zatem ściśle rosnąca i różniczkowalna funkcja odwrotna

$$g^+ \stackrel{\text{def}}{=} (\Gamma_1^+)^{-1} : [-\ell, \ell] \xrightarrow{\text{na}} \left[-\frac{\pi}{2}, \frac{\pi}{2}\right],$$

a poza tym

$$(g^+)'(\sigma) = \frac{1}{(\Gamma_1^+)'(g^+(\sigma))} \leq \frac{\theta^2 - 1}{\theta} \quad \text{dla } \sigma \in [-\ell, \ell].$$

Z analogicznych powodów istnieje ściśle rosnąca i różniczkowalna funkcja odwrotna

$$g^- \stackrel{\text{def}}{=} (\Gamma_1^-)^{-1} : [-\ell, \ell] \xrightarrow{\text{na}} \left[-\frac{\pi}{2}, \frac{\pi}{2}\right],$$

spełniająca warunek

$$(g^-)'(\sigma) = \frac{1}{(\Gamma_1^-)'(g^-(\sigma))} \leq \frac{\theta^2 - 1}{\theta} \quad \text{dla } \sigma \in [-\ell, \ell].$$

Dla każdego $y = (y_1, y_2) \in \partial\Omega(\theta, r)$ definiujemy

$$F(y) = \begin{cases} (\xi(g^+(y_1)), 0), & \text{gdy } y_2 \geq 0, \\ (\xi(g^-(y_1)), 0), & \text{gdy } y_2 < 0. \end{cases}$$

Wobec warunku (3.1), jeżeli $y : [0, \tau] \rightarrow \partial\Omega(\theta, r)$ spełnia warunek Lipschitza ze stałą θ , to złożenie

$$F \circ y : [0, \tau] \rightarrow [-\ell + r, \ell - r] \times \{0\}$$

spełnia spełnia warunek Lipschitza ze stałą 1.

Zdefiniujemy teraz strategię skutecznej obrony $\varphi \in \Phi(0, b)$, gdzie

$$\text{dist}(b, \Omega(\theta, r)) \geq (\ell - r)\theta.$$

Ustalmy dowolną pozycję początkową b , spełniającą powyższy warunek, dowolną trajektorię $y \in Y(b)$ i przyjmijmy

$$\hat{y}(t) = \Pi(y(t)) \quad \text{dla } t \geq 0.$$

Zdefiniujemy trajektorię $x = \varphi(y) \in X(0)$. Ponieważ rzut na domknięty zbiór wypukły jest odwzorowaniem nieoddalającym, to istnieje takie $\hat{x} \in X(0)$ i takie $t_0 = t_0(y)$, że

$$t_0 \leq \ell - r \quad \text{oraz} \quad \hat{x}(t_0) = F(\hat{y}(t_0))$$

Dla $t \geq t_0$ przyjmujemy $\hat{x}(t) = F(\hat{y}(t))$. Z warunku (1) wynika, że relacja $\hat{x} \in X(0)$ jest nadal spełniona. Przyjmujemy dalej

$$T(y) = \{t \geq t_0 : y(t) \in \partial\Omega(\theta, r)\}.$$

Możliwe są dwa przypadki:

$$(a) T(y) = \emptyset \quad \text{oraz} \quad (b) T(y) \neq \emptyset.$$

W przypadku (a) przyjmujemy $x = \varphi(y) = \hat{x}$. W przypadku (b) przyjmujemy

$$x(t) = \varphi(y)(t) = \begin{cases} \hat{x}(t), & \text{gdy } 0 \leq t \leq t_1, \\ \hat{x}(t_1), & \text{gdy } t > t_1, \end{cases} \quad \text{gdzie } t_1 = t_1(y) = \min T(y).$$

Mamy wówczas $\varphi \in \Phi(0, b)$ i φ jest strategią skuteczną obrony odcinka $[-\ell, \ell] \times \{0\}$.

Schemat ataku. Zakładamy teraz, że $2\ell > \Lambda(\theta, \rho)$ i dobieramy takie $r > \rho$, dla którego

$$\Lambda_{\text{I}}(\theta, r) = 2r \left(1 + \frac{1}{\theta^2 - 1} \int_0^{\frac{\pi}{2}} \left(\cos \beta + \sqrt{\theta^2 - \sin^2 \beta} \right) d\beta \right) = 2\ell.$$

Możemy przyjąć, że napastnik rozpoczyna atak z punktu $b = (0, r)$, a obrońca znajduje się w punkcie $a = (a_1, 0)$. Możliwe są dwa przypadki:

$$(a) a_1 \leq 0 \text{ oraz } (b) a_1 > 0.$$

W przypadku (a) napastnik rusza z prędkością θ po brzegu zbioru $\Omega(\theta, r)$, w kierunku punktu $(\ell, 0)$ (leżącego na bronionym odcinku), a po jego osiągnięciu oddala się z prędkością θ wzdłuż osi odciętych. W czasie tego manewru odległość pomiędzy napastnikiem i obrońcą nie będzie mniejsza niż $r > \rho$. Natomiast w przypadku (b) napastnik rusza z prędkością θ po brzegu zbioru $\Omega(\theta, r)$, w kierunku punktu $(-\ell, 0)$ (leżącego na bronionym odcinku), a po jego osiągnięciu oddala się z prędkością θ wzdłuż osi odciętych. Również w czasie tego manewru odległość pomiędzy napastnikiem i obrońcą nie będzie mniejsza niż $r > \rho$.

3.3 Przypadek II

W tym przypadku obrońca będzie bronił zbioru $\Omega_{\mathbf{N}}(\theta, r)$, poruszając się po brzegu zbioru $\Omega_{\mathbf{O}}(\theta, r)$, a napastnik będzie stosował strategię nękania.

Schemat obrony. Zakładamy, że $2\ell < \Lambda_{\mathbf{II}}(\theta, \rho)$ i dobieramy takie $r < \rho$, dla którego

$$\Lambda_{\mathbf{II}}(\theta, r) = r \frac{\theta(2\theta + \pi)}{\theta^2 - 1} = 2\ell.$$

Konstruujemy dwa zbiory wypukłe $\Omega_{\mathbf{N}}(\theta, r)$ i $\Omega_{\mathbf{O}}(\theta, r)$ oraz funkcję

$$F : \partial\Omega_{\mathbf{N}}(\theta, r) \rightarrow \partial\Omega_{\mathbf{O}}(\theta, r).$$

W tym celu, dla $\alpha \in [-\frac{\pi}{2}, \frac{\pi}{2}]$, definiujemy

$$\begin{aligned} \Gamma_{\mathbf{N}}^+(\alpha) &= \begin{bmatrix} -\ell \\ 0 \end{bmatrix} + \frac{r\theta^2}{\theta^2 - 1} \begin{bmatrix} 1 + \sin \alpha \\ \cos \alpha \end{bmatrix} + \frac{r\theta}{\theta^2 - 1} \begin{bmatrix} \frac{\pi}{2} + \alpha \\ 0 \end{bmatrix}, \\ \Gamma_{\mathbf{O}}^+(\alpha) &= \Gamma_{\mathbf{N}}^+(\alpha) - r \begin{bmatrix} \sin \alpha \\ \cos \alpha \end{bmatrix}, \end{aligned}$$

$$\Gamma_{\mathbf{N}}^{-}(\alpha) = \begin{bmatrix} -\ell \\ 0 \end{bmatrix} + \frac{r\theta^2}{\theta^2-1} \begin{bmatrix} 1 + \sin \alpha \\ -\cos \alpha \end{bmatrix} + \frac{r\theta}{\theta^2-1} \begin{bmatrix} \frac{\pi}{2} + \alpha \\ 0 \end{bmatrix},$$

$$\Gamma_{\mathbf{O}}^{-}(\alpha) = \Gamma_{\mathbf{N}}^{-}(\alpha) - r \begin{bmatrix} \sin \alpha \\ -\cos \alpha \end{bmatrix}.$$

Łuki

$$\partial^+ \Omega_{\mathbf{N}}(\theta, r) = \Gamma_{\mathbf{N}}^+ \left(\left[-\frac{\pi}{2}, \frac{\pi}{2} \right] \right), \quad \partial^- \Omega_{\mathbf{N}}(\theta, r) = \Gamma_{\mathbf{N}}^- \left(\left[-\frac{\pi}{2}, \frac{\pi}{2} \right] \right),$$

$$\partial^+ \Omega_{\mathbf{O}}(\theta, r) = \Gamma_{\mathbf{O}}^+ \left(\left[-\frac{\pi}{2}, \frac{\pi}{2} \right] \right), \quad \partial^- \Omega_{\mathbf{O}}(\theta, r) = \Gamma_{\mathbf{O}}^- \left(\left[-\frac{\pi}{2}, \frac{\pi}{2} \right] \right)$$

stanowią, odpowiednio górne i dolne części brzegów zbiorów $\Omega_{\mathbf{N}}(\theta, r)$ i $\Omega_{\mathbf{O}}(\theta, r)$.

Przykład 3.2. Dla $\theta = \sqrt{2}$ i $r = 1$ mamy

$$\frac{1}{2} \Lambda_{\mathbf{II}}(\theta, r) = \frac{1}{2} \frac{\sqrt{2} (2\sqrt{2} + \pi)}{2-1} \approx 4.221441,$$

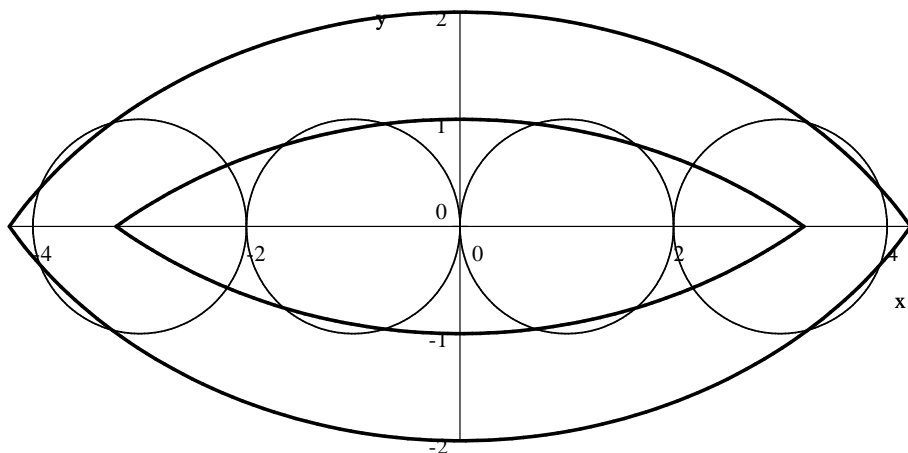
Tym razem jeden ruchomy obrońca broni odcinka dłuższego niż czterech obrońców stacjonujących w ustalonych punktach.

Zbiór $\Omega_{\mathbf{N}}(\theta, r)$ zawiera w swym wnętrzu zbiór $\Omega_{\mathbf{O}}(\theta, r)$, zawierający odcinek $\Delta(\ell)$. Natomiast funkcja F ma następującą własność. Jeżeli napastnik porusza się po trajektorii y leżącej na łuku $\partial \Omega_{\mathbf{N}}(\theta, r)$, to punkt $F(y(t))$ porusza się po łuku $\partial \Omega_{\mathbf{O}}(\theta, r)$ z prędkością co najwyżej 1, pozostając stale w odległości $r < \rho$ od punktu $y(t)$. Dalsza część schematu jest taka sama jak w poprzednim przypadku. Szczegóły zawiera praca [4].

Schemat ataku. Zakładamy, że $2\ell > \Lambda_{\mathbf{II}}(\theta, \rho)$ i dobieramy takie $r > \rho$, dla którego

$$\Lambda_{\mathbf{II}}(\theta, r) = r \frac{\theta(2\theta + \pi)}{\theta^2 - 1} = 2\ell.$$

W przypadku I napastnik mógł ujawnić całą trajektorię, po której będzie się poruszał. W tym przypadku byłoby to błędem. Wykorzystując przewagę prędkości nad



Rysunek 2: Brzegi zbiorów $\Omega_N(\theta, r)$, $\Omega_O(\theta, r)$ i strefy rażenia czterech obrońców stacjonujących w punktach $(-3, 0)$, $(-1, 0)$, $(1, 0)$, $(3, 0)$.

przeciwnikiem napastnik stosuje strategię nękania. W pierwszej fazie ataku rusza w kierunku środka bronionego odcinka, czyli w kierunku punktu $(0, 0)$. W pewnym momencie zbliży się do obrońcy na odległość $r > \rho$. Gdyby do tego nie doszło, to broniony odcinek zostałby bezkarnie zaatakowany w samym środku. Od momentu zbliżenia na odległość r napastnik zaczyna krążyć wokół obrońcy, zachowując stały dystans r i oczekując możliwości bezpiecznego ruszenia w kierunku bronionego odcinka. W tym samym czasie obrońca może wykonywać rozmaite manewry, ale przewaga prędkości pozwala napastnikowi utrzymać stałą odległość do broniącego. Ponieważ obrońca usiłuje bronić zbyt długiego odcinka, to kąt pomiędzy wektorami $y(t)$ i $x(t) - y(t)$, gdzie x, y są trajektoriami obrońcy i napastnika, stale rośnie. Dzięki temu dojdzie do takiego położenia punktów $x(t)$ i $y(t)$, z którego napastnik będzie mógł ruszyć w kierunku bronionego odcinka po linii prostej, bez obawy zmniejszenia odległości do broniącego. Szczegóły zawiera praca [5].

4 Obrona odcinka przez wielu broniących

Jeżeli każdy z obrońców $\mathbf{O}_1, \mathbf{O}_2, \dots, \mathbf{O}_n$, dysponując bronią o zasięgu ρ , może obronić odcinek o długości 2ℓ , to razem mogą obronić odcinek o długości $2n\ell$. Wystarczy, by każdy bronił „swej” części dłuższego odcinka. Pojawia się pytanie, czy tacy obrońcy mogą obronić dłuższy odcinek? Pytanie jest trudne. Jeżeli zmienimy nieco definicję obrony odcinka, to odpowiedzią na postawione pytanie będzie słowo „tak”. Zgodnie z przyjętą definicją obrony odcinka którykolwiek z obrońców powinien zbliżyć się do napastnika na odległość mniejszą niż ρ przed ewentualnym wtargnięciem napastnika na broniony odcinek. Uzupełniając zwrot „przed ewentualnym wtargnięciem” zwrotem „lub po ewentualnym wtargnięciu” otrzymujemy nową definicję obrony odcinka. W przypadku $n = 1$ obie definicje są równoważne.

Zakładamy, że obrońcy $\mathbf{O}_1, \mathbf{O}_2, \dots, \mathbf{O}_n$, gdzie $n \geq 2$, dysponujący bronią o zasięgu ρ , bronią odcinka

$$\Delta(\ell) = [0, \ell] \times \{0\}$$

i mogą się poruszać tylko po bronionym odcinku. Przyjmujemy

$$\alpha^* = \arctg \theta,$$

$$\begin{aligned} \delta_1 &= 1 + \frac{1}{\theta^2 - 1} \int_0^{\frac{\pi}{2}} \left(\cos \beta + \sqrt{\theta^2 - \sin^2 \beta} \right) d\beta, \\ \delta_2 &= \sin \alpha^* + \frac{1}{\theta^2 - 1} \int_0^{\alpha^*} \left(\cos \beta + \sqrt{\theta^2 - \sin^2 \beta} \right) d\beta, \\ \delta_3 &= \frac{2\theta}{(\theta^2 - 1)\sqrt{\theta^2 + 1}} \end{aligned}$$

i dla każdego $r > 0$ definiujemy

$$\Lambda(\theta, r) = 2r(\delta_1 + \delta_2 + \delta_3) + 2r(n-2)(\delta_2 + \delta_3).$$

Obrońcy obronią każdy odcinek o długości mniejszej niż $\Lambda(\theta, \rho)$, a napastnik wtargnie bezkarnie na każdy odcinek o długości większej niż $\Lambda(\theta, \rho)$.

Schemat obrony. Zakładamy, że $\ell < \Lambda(\theta, \rho)$ i dobieramy takie $r < \rho$, dla którego $\Lambda(\theta, r) = \ell$. Broniony odcinek dzielimy na odcinki $\Delta_1, \Delta_2, \dots, \Delta_n$. Dwa skrajne odcinki Δ_1 i Δ_n mają długość

$$\delta'(r) = r(\delta_1 + \delta_2 + \delta_3),$$

więc

$$\Delta_1 = [0, \delta'(r)] \times \{0\}, \Delta_n = [0, \ell - \delta'(r)] \times \{0\}$$

i są bronione przez graczy $\mathbf{O}_1, \mathbf{O}_n$. W przypadku $n > 2$ pozostałą część odcinka $\Delta(\ell)$ dzielimy na $n - 2$ odcinki o równych długościach

$$\delta''(r) = 2r(\delta_2 + \delta_3).$$

Każdy z odcinków $\Delta_2, \Delta_3, \dots, \Delta_{n-1}$ jest broniony przez obrońcę o tym samym numerze co odcinek. Każdy z obrońców broni swojego odcinka aż do momentu t^* , w którym napastnik wtargnie na broniony odcinek. Jeżeli to nastąpi, to napastnik znajdzie się pomiędzy dwoma sąsiednimi obrońcami, powiedzmy pomiędzy obrońcami \mathbf{O}_k i \mathbf{O}_{k+1} . W takim momencie dwaj obrońcy \mathbf{O}_k i \mathbf{O}_{k+1} ruszają z maksymalną prędkością w kierunku napastnika. Pozostali obrońcy mogą wtedy stać beczynnie. Nie później niż w momencie $t^* + \frac{r \cos \alpha^*}{\theta}$ napastnik znajdzie się w odległości r od jednego lub obu obrońców \mathbf{O}_k i \mathbf{O}_{k+1} . Dokładny, lecz dość skomplikowany, opis strategii obrony jest podany w pracy [10].

Schemat ataku. Zakładamy, że $\ell > \Lambda(\theta, \rho)$ i dobieramy takie $r > \rho$, dla którego $\Lambda(\theta, r) = \ell$. Napastnik stosuje strategię nękania opisaną w pracy [10]. Opis tej strategii wymaga wyposażenia obrońców w fikcyjne, poszerzone strefy rażenia. Ponieważ obrońcy bronią zbyt długiego odcinka, to napastnik wtargnie bezkarnie na broniony odcinek, po czym oddali się na bezpieczną odległość, pozostając stale poza strefą rażenia każdego z obrońców.

Przykład 4.1. Przyjmijmy $\theta = \sqrt{2}$, $\rho = 1$, $n = 3$. Zilustrujemy poszerzone strefy rażenia i bronione obszary każdego z trzech obrońców. Mamy

$$\delta_1 = 1 + \frac{1}{2-1} \int_0^{\frac{\pi}{2}} \left(\cos \beta + \sqrt{2 - \sin^2 \beta} \right) d\beta \approx 3.910099,$$

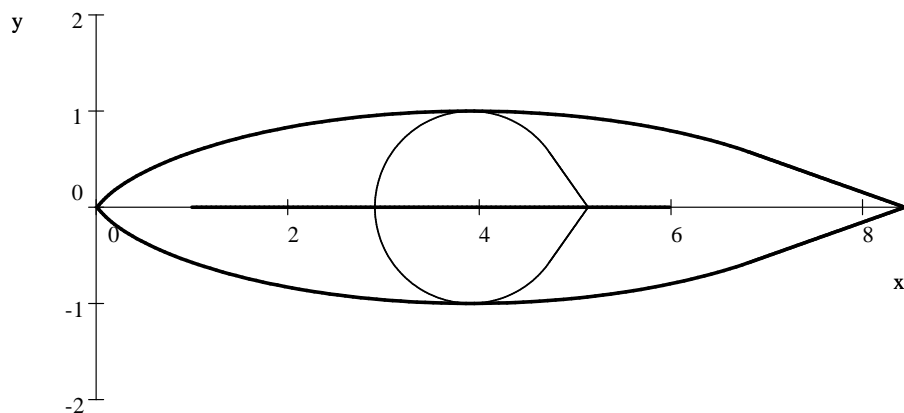
$$\delta_2 = \sin(\operatorname{arctg} \sqrt{2}) + \frac{1}{2-1} \int_0^{\operatorname{arctg} \sqrt{2}} \left(\cos \beta + \sqrt{2 - \sin^2 \beta} \right) d\beta \approx 2.893245,$$

$$\delta_3 = \frac{2\sqrt{2}}{(2-1)\sqrt{2+1}} \approx 1.632993,$$

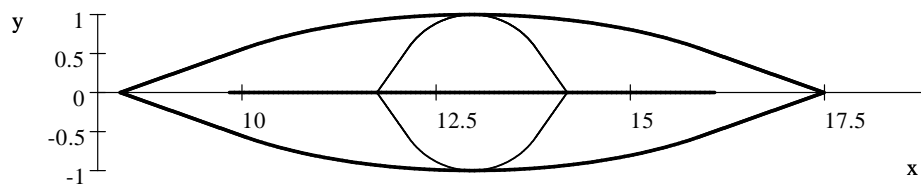
W tym przykładzie trzech obrońców broni skutecznie każdego odcinka o długości

$$2r(\delta_1 + \delta_2 + \delta_3) + 2r(\delta_2 + \delta_3) = 2r(\delta_1 + 2\delta_2 + 2\delta_3) \approx 25.92513r,$$

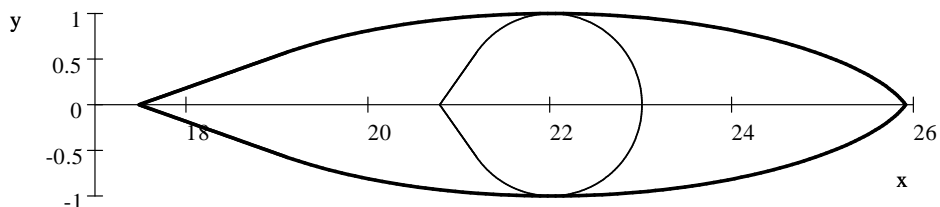
gdzie $r \in (0, 1)$.



Rysunek 3: Położenie początkowe lewego obrońcy i poszerzona strefa rażenia wraz z obszarem bronionym.



Rysunek 4: Położenie początkowe środkowego obrońcy i poszerzona strefa rażenia wraz z obszarem bronionym.



Rysunek 5: Położenie początkowe prawego obrońcy i poszerzona strefa rażenia wraz z obszarem bronionym

Gdyby każdy z nich bronił niezależnie tylko „swego” odcinka, to obroniliby każdy odcinek o długości

$$6r \left(1 + \frac{1}{2-1} \int_0^{\frac{\pi}{2}} \left(\cos \beta + \sqrt{2 - \sin^2 \beta} \right) d\beta \right) \approx 23.460593r.$$

Literatura

- [1] R. Isaacs, *Differential Games*, Wiley, New York, 1965.
- [2] V. Isler, S. Kannan, S. Khanna, Locating and Capturing an Evader in a Polygonal Environment, *Algorithmic Foundations of Robotics (Utrecht 2004)*; M. Erdmann et. all. (ed), Springer, Berlin, 2005, 251–266.
- [3] H. Federer, *Geometric Measure Theory*, Spriger-Verlag Berlin Heidelberg, 1996.
- [4] B. Kraska, *Obrona odcinka*, rozprawa doktorska, Uniwersytet Śląski w Katowicach, 2013.
- [5] B. Kraska, W. Rzymowski, Guarding a Line Segment, *Journal of Convex Analysis*, 18(3), 2011, 645–671.
- [6] N. N. Krasowskij, A. I. Subbotin, *Pozycyjne gry różniczkowe (ros)*, Nauka, Moskwa, 1974.
- [7] Ngoc-Minh Lé. On determining optimal strategies in pursuit games in the plane, *Theor. Comput. Sci.*, 197, 1998, 203–234.
- [8] R. Murieta-Cid, A. Sarmiento, S. Battacharya, S. Hutchinson, Maintaining Visibility of the Moving Target at a Fixed Distance, *Proc IEEE Int'l. Conf. on Robotics and Automation*, New Orleans, 2004, 479–484.
- [9] Cz. Ryll-Nardzewski, A theory of pursuit and evasion, *Ann. Math. Stud.*, 3, 1957, 393–405.
- [10] W. Rzymowski, A problem of guarding line segment, *SIAM J. Control Optim.*, 48, 2009, 812–838.
- [11] W. Rzymowski, A. Stachura, Solution to a problem of guarding territory, *Systems and Control Letters*, 7, 1986, 71–72.

- [12] H. Steinhaus, *Definicje potrzebne do teorii gry i pościgu*, Złota Myśl Akademicka, Lwów, 1927.
- [13] F. P. Wasiliew, *Czisljennyje metody rieszenija ekstriemalnych zadacz* (ros), Nauka, Moskwa, 1988.

Poryzm Ponceleta

Waldemar Cieślak¹

Witold Mozgawa²

Streszczenie

Problematyka związana z twierdzeniem Ponceleta skupia uwagę matematyków od prawie dwustu lat. Praca przedstawia krótkie omówienie wybranych zagadnień związanych z poryzmem Ponceleta. Przedstawiono nowe wyniki z ostatnich kilku lat. Badania pozwoliły rozwiązać niektóre problemy, a także sformułować wiele nowych hipotez.

Abstract

The problems connected with the Closed Poncelet Theorem concentrate attention of mathematicians since nearly two hundred years. The paper presents a short treatment of selected problems associated with Poncelet's problem. We present new results from last few years. The investigations allow to solve some problems and allow to formulate new hypotheses.

Słowa kluczowe: poryzm Ponceleta, miara niezmiennicza, indeks obrotu, wzory Fussa, wzór Jacobiego

1 Wstęp

Historia poryzmu Ponceleta sięga czasów napoleońskich. Jean-Victor Poncelet (1788-1867), francuski inżynier i matematyk, brał udział w kampanii Napoleona jako inżynier wojskowy. W latach 1813-1814 przebywał on w Saratowie jako więzień wojenny. W roku 1822 Poncelet, po powrocie do Francji, opublikował traktat

¹Katedra Matematyki Stosowanej, Politechnika Lubelska

²Instytut Matematyki, Uniwersytet Marii Curie-Skłodowskiej

o geometrii rzutowej [20], gdzie podał dowód bardzo ważnego twierdzenia, zwane-
go dzisiaj twierdzeniem Ponceleta o zamknięciu. Niewątpliwie, główny zarys tego
dzieła powstał podczas pobytu w Rosji.

Opiszemy krótko pojęcie poryzmu Ponceleta. Rozważmy pierścień CD na płasz-
czyźnie utworzony przez dwie elipsy C i D , gdzie C leży w obszarze ograniczonym
przez D . Na krzywych wybieramy obieg w kierunku przeciwnym do ruchu wska-
zówek zegara. Ustalmy punkt M na elipsie C . Przez punkt M prowadzimy styczną
do C . Wybieramy na stycznej półprostą o początku w M zgodną z obiegiem na
elipsie. Półprosta ta przecina D w punkcie M_1 . Z punktu M_1 prowadzimy styczną
do C . Styczna ta przecina elipsę D w punkcie M_2 . Następnie całe rozumowanie po-
wtarzamy i otrzymujemy ciąg punktów $c(M) = \{M_1, M_2, \dots\}$ na elipsie D . Może
tak się zdarzyć, że dla pewnej liczby naturalnej k mamy $M_{k+1} = M_k$. Wtedy bę-
dziemy mówić, że ciąg $c(M)$ jest zamknięty. Sformułujemy twierdzenie Ponceleta
o zamknięciu w następujący sposób:

*Dany jest pierścień CD utworzony przez dwie elipsy C i D . Jeśli dla pewnego
punktu $M \in C$ ciąg $c(M)$ jest zamknięty, to zamknięty jest każdy ciąg $c(N)$ dla
 $N \in C$.*

W literaturze istnieją dwa znaczenia słowa poryzm, które mają pewne wady.
W związku z tym proponujemy następujące znaczenie słowa poryzm: jeśli pew-
na własność zachodzi w jednym punkcie danego zbioru, to zachodzi w każdym
punkcie tego zbioru. Takie podejście było wprowadzone w pracy [7]. Dokładne
omówienie poryzmu Ponceleta można znaleźć w pracy [2]; w ostatnich latach uka-
zały się również dwie monografie o poryzmie [11] i [12].

2 O wzorach Fussa

Rozważmy pierścień kołowy złożony z dwu okręgów: wewnętrznego C , $x^2 + y^2 = r^2$
i zewnętrznego D , $(x - a)^2 + y^2 = R^2$, gdzie $a > 0$ i $a + r < R$. W roku 1792 Fuss
[13] wyprowadził wzory wiążące R , r i a dla pierścieni kołowych mających wła-
sność poryzmu Ponceleta dla $n = 4, 5, 6, 7, 8$. Przytoczymy kilka kolejnych wzorów

w postaci podanej przez Steinera w roku 1827:

$$1. \ n = 3 : a^2 = R^2 - 2rR,$$

$$2. \ n = 4 : (R^2 - a^2)^2 = 2r^2(R^2 + a^2),$$

$$3. \ n = 5 : r(R - a) = (r + a)(\sqrt{(R - r + a)(R - r - a)} + \sqrt{2R(R - r - a)}),$$

$$4. \ n = 6 : (R^2 - a^2)^4 = 4r^2(R^2 + a^2)(R^2 - a^2)^2 + 16a^2r^2R^2.$$

Dalej wzory tego typu nazywać będziemy wzorami Fussa. W roku 1923 Chaundy podał wiele wzorów Fussa dla $n \leq 20$. Istnieje wiele równoważnych postaci takich wzorów dla ustalonego n . Na przykład, dla $n = 5$ mamy wzór w formie podanej powyżej przez Steinera, jak również w formie podanej w 1830 roku przez Richelota

$$4p^2q^2(p-1)(q-1) = (p^2 + q^2 - p^2q^2)^2,$$

gdzie

$$p = \frac{R+a}{r}, \quad q = \frac{R-a}{r}.$$

Niech $z(t) = re^{it}$ dla $t \in [0, 2\pi]$ będzie parametryzacją okręgu C . Okrąg zewnętrzny D sparametryzujemy następująco $w(t) = re^{it} + \lambda(t)ie^{it}$ dla $t \in [0, 2\pi]$, gdzie funkcja λ przyjmuje wartości dodatnie i jest dana wzorem

$$\lambda(t) = \sqrt{a^2 \sin^2 t + 2ar \cos t + R^2 - r^2 - a^2} - a \sin t.$$

Punkt $w(t)$ leży na $\alpha(t)$ -izooptyce, gdzie

$$\operatorname{tg} \frac{\alpha(t)}{2} = \frac{\lambda(t)}{r}.$$

Dla funkcji $\varphi(t) = t + \alpha(t)$ przyjmujemy $\varphi^{[0]} = id$, $\varphi^{[1]} = \varphi$, $\varphi^{[2]} = \varphi \circ \varphi$, itd. Definiujemy wyrażenia zespolone

$$\Omega_m(t) = r + \lambda(\varphi^{[m]}(t))i$$

dla $m = 0, 1, 2, \dots$. Z twierdzenia 2.1 podanego w pracy [6] wynika natychmiast następujący wniosek. Załóżmy, że pierścień kołowy CD ma własność poryzmu Poncela dla ustalonej liczby naturalnej $n \geq 3$. Wtedy spełniona jest tożsamość

$$\operatorname{Im} \Omega_0 \Omega_1 \dots \Omega_{n-1} \equiv 0.$$

Twierdzenie to ma istotne znaczenie, jeśli chodzi o tzw. wzory Fussa. W pracy [6] wyprowadzono z tej tożsamości wzory Fussa dla $n = 3, 4$. Następnie przeprowadzono rachunki dla $n = 5$, jednakże nie pokazano, którą z równoważnych postaci wzoru Fussa otrzymuje się. W pracy magisterskiej pani Magdaleny Michy pt. „Wzory Fussa w poryźmie Poncela” [18] pokazano, którą dokładnie postać wzoru Fussa otrzymujemy, a ponadto wyprowadzono wzory Fussa dla $n \leq 6$. Istotne cechy tożsamości z twierdzenia 2.1 [6] to:

po pierwsze, jej prosta postać algebraiczna,

po drugie, występują tam tylko długości odcinków łączących punkty na obu okręgach i leżących na bokach wielokątów wpisano-opisanych.

W związku z tym widzimy jakie działania algebraiczne i funkcje elementarne zawierają wzory Fussa.

W przypadku $n = 5$ dla $t = 0$ podamy wyrażenia na $\lambda_m = \lambda(\varphi^{[m]}(0))$ dla $m = 0, 1, 2, 3, 4$. Wykorzystując, między innymi rysunek pierścienia kołowego złożonego z okręgów $x^2 + y^2 = r^2$, $(x - a)^2 + y^2 = R^2$, $a > 0$, oraz pięciokąta wpisano-opisanego przechodzącego przez punkt $(r, 0)$, któremu odpowiada $t = 0$, otrzymamy

$$\begin{aligned} \lambda_0 = \lambda_4 &= \sqrt{R^2 - (r - a)^2}, & \lambda_2 &= \sqrt{(R - a)^2 - r^2}, \\ \lambda_1 = \lambda_3 &= \lambda_0 \frac{R^2 - a^2 - 2ar}{R^2 - a^2 + 2ar}. \end{aligned}$$

Wykorzystując teraz tożsamość dla $t = 0$, będziemy mieli

$$\begin{aligned} 0 &= \operatorname{Im} \Omega_0 \Omega_1 \dots \Omega_4 = \operatorname{Im}(\Omega_0 \Omega_1)^2 \Omega_2 \\ &= \operatorname{Im}((r + \lambda_0 i)(r + \lambda_1 i))^2 (r + \lambda_2 i) \\ &= \operatorname{Im}[(r^2 - \lambda_0 \lambda_1)^2 + 2ri(r^2 - \lambda_0 \lambda_1)(\lambda_0 + \lambda_1) - (r^2 \lambda_0 + \lambda_1)^2](r + \lambda_2 i) \end{aligned}$$

$$= \text{Im}[(r^2 - \lambda_0 \lambda_1)^2 r - r^3(\lambda_0 + \lambda_1) - 2r\lambda_2(r^2 - \lambda_0 \lambda_1)(\lambda_0 + \lambda_1) \\ + i([(r^2 - \lambda_0 \lambda_1)^2 - r^2(\lambda_0 + \lambda_1)^2]\lambda_2 + 2r^2(r^2 - \lambda_0 \lambda_1)(\lambda_0 + \lambda_1))],$$

co daje wzór

$$[(r^2 - \lambda_0 \lambda_1)^2 - r^2(\lambda_0 + \lambda_1)^2]\lambda_2 + 2r^2(r^2 - \lambda_0 \lambda_1)(\lambda_0 + \lambda_1) = 0.$$

Podstawiając teraz wyznaczone przez nas wyrażenia na λ_m , po prostych przekształceniach, otrzymamy wzór Fussa dla $n = 5$ w postaci

$$4r^2(R - a - r)(R - r + a)(R^2 - a^2)^2 = [2r^2(R^2 + a^2)]^2.$$

Jest to dokładnie postać wzoru podana przez Richelota.

Z każdym pierścieniem kołowym zwiążemy trzy liczby, ich promienie R, r oraz odległość a między ich środkami. Jeśli $R > r$, to oczywiście $r + a < R$. W dalszej części tego rozdziału przyjmować będziemy, że $a > 0$. Warunek ten powoduje odrzucenie pierścieni utworzonych przez okręgi współśrodkowe.

Niech $\mathbb{R}_+ = (0, +\infty)$ i $\mathbf{M} = \{(R, r, a) : R, r, a \in \mathbb{R}_+, r + a < R\}$.

Funkcja $f : \mathbb{R}_+ \times \mathbf{M} \rightarrow \mathbf{M}$ dana wzorem

$$f(\xi, (R, r, a)) = ((1 - \xi)r + \xi R, r, \xi a)$$

jest lewostronnym działaniem grupy (\mathbb{R}_+, \cdot) na zbiorze \mathbf{M} .

Niestety działanie f nie jest tranzytywne. W związku z tym struktura $(\mathbb{R}_+, \mathbf{M}, f)$ nie determinuje geometrii w sensie Kleina. Pozostaje otwartym problem rozszerzenia grupy tak, by działanie było tranzytywne.

Niech $\mathbf{M}_3 = \{(R, r, a) \in \mathbf{M} : a^2 = R^2 - 2rR\}$.

Zbiór ten rozłożymy na rozłączne podzbiory. Przyjmiemy $t = \frac{R}{r}$. Wtedy równość $a^2 = R^2 - 2rR$ możemy zapisać w postaci $(\frac{a}{r})^2 = t^2 - 2t$. Dla $t > 2$ określimy zbiór

$$\mathbf{M}_3(t) = \{(rt, r, r\sqrt{t^2 - 2t}) \in \mathbf{M} : r > 0\}$$

W pracy [4] dowodzimy następującego twierdzenia:

Jeśli $(R, r, a) \in \mathbf{M}_3$ i $\xi > 0$, to istnieje $\rho > 0$ takie, że $(1 - \xi)r + \xi R, \rho r, \xi a) \in \mathbf{M}_3$.

Efektywne wyliczenie ρ pozwala zdefiniować dla każdego ustalonego ξ odwzorowanie $F_\xi : \mathbf{M}_3 \rightarrow \mathbf{M}_3$ wzorem

$$F_\xi(rt, r, r\sqrt{t^2 - 2t}) = (r_\xi t_\xi, r_\xi, r_\xi \sqrt{t_\xi^2 - 2t_\xi}),$$

gdzie

$$r_\xi = r \frac{2\xi(t-1) + 1 + \xi^2}{2\xi(t-1) + 2}$$

$$t_\xi = 2 \frac{(\xi(t-1) + 1)^2}{2\xi(t-1) + 1 + \xi^2}.$$

Wykazano, że przekształcenie F_ξ jest suriekcją. Można pokazać, że każde odwzorowanie F_ξ generuje dyfeomorfizm w zbiorze $(0, +\infty) \times (2, +\infty)$.

3 O indeksie obrotu dla bilardów barowych i poryzmie Ponceleta

W pracy [9] prezentowane są nowe wyniki o związkach pomiędzy indeksem obrotu dla bilardów barowych dla dwóch zawierających się okręgów C_R i C_r o promieniach R i r i o odległości a pomiędzy ich środkami a poryzmem Ponceleta dla tych okręgów, gdzie wymierne indeksy obrotu odpowiadają zamkniętym transwersalnym Ponceleta z lub bez samoprzecięć. Korzystając z funkcji specjalnych, wyprowadzamy interesujący podwójny szereg opisujący indeks obrotu, i którego suma, w przypadku poryzmu Ponceleta dla trójkątów, jest liczbą $\frac{\pi}{3}$. Wyprowadzamy wzór Steinera, który daje konieczny warunek na istnienie bilardu barowego, którego transwersalne Ponceleta tworzą pięciokąt z samoprzecięciami i który ma indeks obrotu równy $\frac{2}{5}$. Na koniec pokazujemy, że w pewnym otoczeniu pary zawierających się w sobie okręgów posiadających własność poryzmu Ponceleta dla indeksu $\frac{1}{3}$ istnieje zawsze para zawierających się w sobie okręgów posiadających indeksy $\frac{1}{4}$ i $\frac{1}{6}$. W przypadku $\frac{1}{4}$ ta para jest wyznaczona jednoznacznie. Niech C_R i C_r będą dwoma zawierającymi się okręgami o środkach w $(0, 0)$ i w $(a, 0)$ oraz

o promieniach, odpowiednio, R i r , takimi że $a > 0$ i $a + r < R$. Dla okręgu C_R rozważamy naturalną parametryzację daną przez $z(t) = (R \cos t, R \sin t)$, gdzie $t \in \mathbb{R}$. Z dowolnego punktu $(R \cos t, R \sin t) = Re^{it}$ na C_R prowadzimy styczną do okręgu C_r i prowadzimy ją aż do przecięcia się z okręgiem C_r i otrzymujemy punkt $Re^{i\varphi(t)}$, gdzie

$$Re^{i\varphi(t)} = f(t)e^{it} + g(t)ie^{it},$$

zaś

$$\begin{aligned} f(t) = & -\frac{R}{(a^2 + R^2 - 2aR \cos t)^2} (R^2 (3a^2 - 2r^2 + R^2) \\ & + a^2 (a^2 - 2r^2 + 3R^2) \cos 2t - 2aR (a^2 - 2r^2 + 2R^2 + a^2 \cos 2t) \cos t \\ & + 4ar(R - a \cos t) \sqrt{a^2 - r^2 + R^2 - 2aR \cos t} \sin t) \end{aligned}$$

oraz

$$\begin{aligned} g(t) = & \frac{2R}{(a^2 + R^2 - 2aR \cos t)^2} \left(rR^2 \sqrt{a^2 - r^2 + R^2 - 2aR \cos t} \right. \\ & + a \left(ar \sqrt{a^2 - r^2 + R^2 - 2aR \cos t} \cos 2t - R (2a^2 - 2r^2 + R^2 + a^2 \cos 2t) \sin t \right. \\ & \left. \left. + \left(-2rR \sqrt{a^2 - r^2 + R^2 - 2aR \cos t} + a (a^2 - 2r^2 + 3R^2) \sin t \right) \cos t \right) \right). \end{aligned}$$

Następnie bierzemy punkt $Re^{i\varphi(t)}$ jako punkt startowy następnej stycznej, która będzie się kończyć w punkcie $Re^{i\varphi(\varphi(t))}$, itd. Wtedy otrzymujemy homeomorfizm Φ okręgu C_R , który ma naturalne podniesienie $\varphi: \mathbb{R} \rightarrow \mathbb{R}$. Powyższy homeomorfizm nazywamy bilardem barowym, a ciąg łamanych łączących kolejne punkty startowe stycznych nazywamy transwersalną Ponceleta. Ponadto po pewnych rozważaniach otrzymujemy, że

$$\begin{aligned} \varphi(t) = & \arccos \left(-1 + \frac{2r^2}{(a-R)^2} \right) - 2 \arctg \left(\frac{\sqrt{a^2 - r^2 - 2aR + R^2}}{r} \right) \\ & + 2 \arctg \left(\frac{\sqrt{a^2 - r^2 + R^2 - 2aR \cos t}}{r} \right) + 2 \arctg \left(\frac{(a-R) \operatorname{ctg} \frac{t}{2}}{a+R} \right) \\ & + 2\pi \left\lfloor \frac{t}{2\pi} \right\rfloor + \pi \operatorname{sgn} t, \end{aligned}$$

gdzie

$$\varphi(0) = \arccos \left(\frac{2r^2}{(a-R)^2} - 1 \right),$$

zaś $\varphi(t + 2\pi) = \varphi(t) + 2\pi$.

Następnie pokazujemy, że 1-forma $\omega(t) = f(t)dt$ na C_R , gdzie

$$f(t) = \frac{1}{|AB|} = \frac{1}{\sqrt{a^2 - r^2 + R^2 - 2aR \cos t}}$$

jest miarą niezmienniczą względem homeomorfizmów odpowiadających poryzom z samoprzecięciami i bez samoprzecięć. Fakt ten jest już znany z pracy R. Kołodzieja [17] i J. L. Kinga [16].

Przypomnijmy, że indeks obrotu homeomorfizmu φ okręgu jest dany wzorem

$$\rho_0(\varphi, t) = \frac{1}{2\pi} \lim_{n \rightarrow \infty} \frac{\varphi^n(t) - t}{n}.$$

Korzystając z wyżej wyprowadzonej 1-formy mamy

$$\rho_0(\varphi) = \frac{\int_0^{\varphi(0)} \omega(t)}{\int_0^{\varphi(0)} \omega(t)}.$$

Po pewnych rozważaniach wykorzystujących funkcje specjalne otrzymujemy

Twierdzenie 3.1. *Dla każdego bilardu barowego φ mamy następujący wzór na indeks obrotu stowarzyszonego z nim homeomorfizmu*

$$\rho_0(\varphi) = \frac{1}{2} - \frac{\sin \varphi_0 \sum_{m=0}^{\infty} \left(\frac{(2m-1)!!}{(2m)!!} \sin^{2m} \varphi_0 \cdot \sum_{n=0}^{\infty} \frac{(2n-1)!!}{(2n)!!(2m+2n+1)} k^{2n} \sin^{2n} \varphi_0 \right)}{\pi \sum_{n=0}^{\infty} \left(\frac{(2n-1)!!}{(2n)!!} \right)^2 k^{2n}},$$

gdzie

$$\varphi_0 = \frac{\pi}{2} - \frac{\arccos \left(\frac{2r^2}{(a-R)^2} - 1 \right)}{2}.$$

Jeśli rozważymy przypadek, że transwersalna Ponceleta się zamyka po trzech krokach, czyli przypadek bicentrycznego trójkąta, to mamy

$$r = \frac{R^2 - a^2}{2R}$$

i otrzymujemy następujący wniosek.

Wniosek 3.1. Dla każdego $t \in (0, 1)$ mamy, że

$$\frac{\pi}{3} = \frac{(1+t) \sum_{m=0}^{\infty} \frac{(2m-1)!!}{(2m)!!} \cdot \left(\frac{1+t}{2}\right)^{2m} \cdot \sum_{n=0}^{\infty} \frac{(2n-1)!!}{(2n)!!(2m+2n+1)} \left(\frac{4t}{(3-t)(1+t)}\right)^n}{\sum_{n=0}^{\infty} \left(\frac{(2n-1)!!}{(2n)!!}\right)^2 \cdot \left(\frac{16t}{(3-t)(1+t)^3}\right)^n}.$$

Następnie rozważamy inną parametryzację rozważanej pary okręgów C_R i C_r . Teraz środkiem mniejszego okręgu będzie punkt $(0, 0)$, a okręgu większego punkt $(a, 0)$, gdzie $a > 0$ oraz $a + r < R$. Niech okrąg C_r będzie sparametryzowany równaniem $z(t) = re^{it}$, podczas gdy większy okrąg jest sparametryzowany przez $w(t) = re^{it} + \lambda(t)ie^{it}$, gdzie $\lambda(t) = \sqrt{R^2 - (a \cos t - r)^2} - a \sin t$ zaś $t \in \mathbb{R}$. Jeśli określimy kąt $\alpha(t)$ wzorem

$$\operatorname{tg} \frac{\alpha(t)}{2} = \frac{\lambda(t)}{r},$$

to odpowiednikiem powyższego homeomorfizmu φ jest homeomorfizm ψ określony przez

$$\psi(t) = t + \alpha(t).$$

Niech $\psi^0(t) = t$, $\psi^n(t) = \psi(\psi^{n-1}(t))$ dla $n = 1, 2, \dots$. Stąd otrzymujemy

$$e^{i\alpha(t)} = \frac{r^2 - \lambda^2(t)}{r^2 + \lambda^2(t)} + i \frac{2\lambda(t)r}{r^2 + \lambda^2(t)},$$

i

$$e^{i\psi(t)} = \left(\frac{r + i\lambda(t)}{|r + i\lambda(t)|} \right)^2 e^{it}.$$

Jeśli dla każdego $n = 0, 1, 2, \dots$ zapiszemy

$$\mu_n(t) = \frac{r + i\lambda(\psi^n(t))}{|r + i\lambda(\psi^n(t))|},$$

to otrzymujemy

$$e^{i\psi^n(t)} = (\mu_0(t)\mu_1(t)\dots\mu_{n-1}(t))^2 e^{it}.$$

Oczywiście wielokąt odpowiadający $\psi(t)$ ma własność poryzmu Ponceleta wtedy i tylko wtedy, gdy dla pewnego zredukowanego ułamka $\frac{k}{n}$, gdzie $k, n \in \mathbb{N}$, $n \geq 3$, $k < n$, mamy

$$\psi^n(t) = t + 2k\pi.$$

Z pracy [6] wynika, że jeśli wielokąt otrzymany z transversalnej Ponceleta z samoprzecięciami lub bez samoprzecięć jest zamknięty, to

$$(\mu_0(t)\mu_1(t)\dots\mu_{n-1}(t))^2 = 1.$$

Stąd mamy

$$\text{Im}(\mu_0(t)\mu_1(t)\dots\mu_{n-1}(t)) = 0.$$

Korzystając z tego wzoru wyprowadzamy zwarty wzór typu Steinera związujący promienie R , r z odległością a środków okręgów w rozważanej parze zawierających się w sobie okręgów z pięciokątnym bilardem barowym z samoprzecięciami, dla którego $\rho_0(\psi) = \frac{2}{5}$. Wzór ten ma postać

$$\begin{aligned} & 2a^4 r^2 \sqrt{R^2 - (a+r)^2} - 8a^3 r^2 R \sqrt{R^2 - (a+r)^2} + 4a^2 r^2 R^2 \sqrt{R^2 - (a+r)^2} \\ & + 8r^4 R^2 \sqrt{R^2 - (a+r)^2} + 8ar^2 R^3 \sqrt{R^2 - (a+r)^2} - 6r^2 R^4 \sqrt{R^2 - (a+r)^2} \\ & - a^6 \sqrt{(R-a)^2 - r^2} - 2a^5 r \sqrt{(R-a)^2 - r^2} - 2a^4 r^2 \sqrt{(R-a)^2 - r^2} \\ & + 3a^4 R^2 \sqrt{(R-a)^2 - r^2} + 4a^3 r R^2 \sqrt{(R-a)^2 - r^2} + 8a^2 r^2 R^2 \sqrt{(R-a)^2 - r^2} \\ & + 8ar^3 R^2 \sqrt{(R-a)^2 - r^2} + 8r^4 R^2 \sqrt{(R-a)^2 - r^2} - 3a^2 R^4 \sqrt{(R-a)^2 - r^2} \\ & - 2arR^4 \sqrt{(R-a)^2 - r^2} - 6r^2 R^4 \sqrt{(R-a)^2 - r^2} + R^6 \sqrt{(R-a)^2 - r^2} = 0. \end{aligned}$$

Na koniec prezentacji pracy *On rotation index of bar billiards and Poncelet's porism* rozważmy pierścień $C_r C_R$ utworzony przez dwa okręgi C_r i C_R , które są dane odpowiednio równaniami

$$x^2 + y^2 = r^2, \quad (x - a)^2 + y^2 = R^2,$$

gdzie

$$0 < a < R - r.$$

Przypomnimy wersję twierdzenie Ponceleta o zamknięciu, która jest odpowiednia do dalszych rozważań:

Jeśli istnieje wielokąt n o bokach jednocześnie wpisany w zewnętrzny okrąg i opisany na wewnętrznym okręgu pierścienia $C_r C_R$, to każdy punkt zewnętrznego okręgu jest wierzchołkiem takiego n -kąta.

Jeśli własność zamknięcia Ponceleta (zwana też własnością poryzmu Ponceleta) zachodzi dla $n = 3$, to znaczy, że indeks obrotu stowarzyszonego homeomorfizmu jest równy $\frac{1}{3}$ w pierścieniu $C_r C_R$, to zachodzi warunek

$$R^2 - 2Rr - a^2 = 0.$$

Dla dalszych celów wygodnie jest oznaczyć ten warunek przez $\text{Pct}(C_r C_R, 3)$. Warunki $\text{Pct}(C_r C_R, 4)$ i $\text{Pct}(C_r C_R, 6)$ to

$$(R^2 - a^2)^2 = 2r^2(R^2 + a^2),$$

oraz

$$3(R^2 - a^2)^4 = 4r^2(R^2 + a^2)(R^2 - a^2)^2 + 16r^4 a^2 R^2.$$

Łatwo widać, że dla każdej liczby $\lambda \in [0, 1]$, okrąg \mathbf{C}_λ dany równaniem

$$(x - \lambda a)^2 + y^2 = [(1 - \lambda)r + \lambda R]^2,$$

leży w pierścieniu $C_r C_R$ i $\mathbf{C}_0 = C_r$, $\mathbf{C}_1 = C_R$.

Przy tych oznaczeniach udowodniono następujące dwa twierdzenia.

Twierdzenie 3.2. *Jeśli warunek $\text{Pct}(C_r C_R, 3)$ jest spełniony, to istnieje dokładnie jedna liczba $\lambda \in (0, 1)$ taka, że warunek $\text{Pct}(C_r C_\lambda, 4)$ jest spełniony.*

Twierdzenie 3.3. *Jeśli warunek $\text{Pct}(C_r C_R, 3)$ jest spełniony, to istnieje liczba $\lambda \in (0, 1)$ taka, że warunek $\text{Pct}(C_r C_\lambda, 6)$ jest spełniony.*

Na koniec pracy postawiono hipotezę:

Jeśli warunek $\text{Pct}(C_r C_R, k)$ jest spełniony, to istnieje liczba $\lambda \in (0, 1)$ taka, że warunek $\text{Pct}(C_r C_\lambda, n)$ jest spełniony dla każdego $n > k$,

która została wykazana w pracy [10].

4 O pewnej hipotezie związanej z poryzmem Ponceleta

W tym podrozdziale pokażemy, że hipoteza przedstawiona na końcu poprzedniego rozdziału, postawiona w pracy *On rotation index of bar billiards and Poncelet's porism*, Bull. Belg. Math. Soc. Simon Stevin, 20, 2 (2013), 287-300 jest prawdziwa w następującym sensie:

dla pary okręgów posiadających własność poryzmu Ponceleta dla indeksu obrotu $\frac{1}{k}$, gdzie $k \geq 3$ jest liczbą naturalną pokazujemy, że istnieje okrąg leżący pomiędzy wyjściowymi okręgami, który wraz z mniejszym okręgiem ma własność poryzmu Ponceleta dla każdego danego indeksu $\frac{1}{n}$, gdzie n jest dowolną liczbą naturalną taką, że $n > k$.

Rozważmy pierścień $C_r C_{a,R}$ złożony z dwóch okręgów C_r i $C_{a,R}$. Okręgi C_r i $C_{a,R}$ są dane, odpowiednio, równaniami $x^2 + y^2 = r^2$ i $(x - a)^2 + y^2 = R^2$, gdzie $0 < a < R - r$. Jeśli funkcja $\lambda(t)$ jest określona, tak jak w rozdziale drugim oraz

$$B(t) = \int_0^t \frac{ds}{\sqrt{1 - \sigma^2(s)}}.$$

gdzie

$$\sigma(t) = \frac{r - a \cos t}{R},$$

to mamy następujący warunek konieczny i wystarczający:

Twierdzenie 4.1. *Twierdzenie Ponceleta o zamknięciu zachodzi w pierścieniu $C_r C_{a,R}$ dla n -kątów, gdzie $n \geq 3$ wtedy i tylko wtedy, gdy zachodzi następująca tożsamość*

$$B\left(t + 2 \operatorname{arctg} \frac{\lambda(t)}{r}\right) \equiv B(t) + \frac{1}{n} B(2\pi).$$

Korzystając z tego twierdzenia można wykazać, że

Twierdzenie 4.2. *Założmy, że twierdzenie Ponceleta o zamknięciu zachodzi w pierścieniu $A C_r C_{a,R}$ dla k -kątów, gdzie $k \geq 3$. Wtedy istnieje liczba $\gamma \in (0, 1)$ taka, że twierdzenie Ponceleta o zamknięciu zachodzi w pierścieniu $C_r C_{\gamma a, (1-\gamma)r + \gamma R}$ dla n -kątów, gdzie n jest dowolną liczbą naturalną taką, że $n > k$.*

5 Problem ekstremalny w poryzmie Ponceleta

Rozpatrzmy klasę pierścieni wypukłych, mających własność poryzmu Ponceleta dla ustalonej liczby naturalnej n . Powstaje naturalne pytanie, które to wielokąty wpisano-opisane mają ekstremalny obwód. Pytanie to postawił po raz pierwszy Radić w pracy [21] i podał rozwiązanie dla przypadku $n = 3$. W pracy [8] wyznaczone zostały czworokąty mające maksymalny i minimalny obwód. Położenie ekstremalnych czworokątów podaje następujące twierdzenie 6.1. Założmy, że pierścień kołowy $C_r C_{a,R}$ ma własność poryzmu Ponceleta. Czworokąt wpisano-opisany ma maksymalny obwód, jeśli przechodzi przez punkt $(R + a, O)$, zaś minimalny obwód, jeśli przechodzi przez punkt (r, O) . Podejście do problemu w obu pracach jest istotnie różne. Trudność w rozwiązaniu ogólnego problemu tkwi w konieczności użycia wzorów Fussa, które nie są znane w jawnej postaci. Na podstawie wyników pracy [8] i rezultatów uzyskanych przez Radića [21] możemy sformułować dwie hipotezy:

I. Wielokąt wpisano-opisany w pierścień kołowy $C_r C_{a,R}$ ma maksymalny obwód (pole), jeśli przechodzi przez punkt $(R + a, O)$.

II. Wielokąt wpisano-opisany w pierścień kołowy $C_r C_{a,R}$ ma minimalny obwód (pole), jeśli przechodzi przez punkt (r, O) .

6 O pewnych uogólnieniach poryzmu Ponceleta na dowolne pierścienie wypukłe

W pracy [19] Mozgawa udowodnił, że dla dowolnego owalu C istnieje owal C_{in} leżący w obszarze ograniczonym przez C taki, że pierścień $C_{in}C$ ma własność poryzmu Ponceleta dla prawie wszystkich liczb naturalnych n . Ponadto pokazał, że w zewnątrz C leży owal C_{out} taki, że pierścień CC_{out} ma analogiczne własności jak $C_{in}C$. Rozważmy klasę izoptyk oraz klasę krzywych równoległych do ustalonego owalu C . Jeśli p jest funkcja podparcia owalu C i $m > 0$, to funkcja podparcia $p_m(t) = p(t) + m$ wyznacza krzywą równoległą, którą oznaczymy symbolem C_m . Okazuje się, że do każdego owalu C , którego α -izoptyka C_α nie jest wypukła można dobrać krzywą równoległą C_m (będącą owalem) taką, że z kolei jej α -izoptyka $C_{m,\alpha}$ jest już wypukła. Stanowi to treść twierdzenia 2.1 [5]. Ustalmy owal C i dowolną liczbę naturalną n . Pokażemy, że C i n generują pewien pierścień, który ma własność poryzmu Ponceleta. Dla $\alpha = \frac{2\pi}{n}$ rozważmy α -izoptykę C_α owalu C . Jeśli C_α nie jest wypukła, to zgodnie z twierdzeniem 2.1 [6] istnieje krzywa równoległa C_m taka, że jej α -izoptyka $C_{m,\alpha}$ już jest wypukła. Wtedy każdy pierścień złożony z owali $C_m, C_{m,\alpha}$ ma własność poryzmu Ponceleta. Zauważmy, że obraz afiniczny pierścienia $C_m, C_{m,\alpha}$ ma również własność poryzmu Ponceleta. Stanowi to treść twierdzenia 3.1 [5]. Niech CD będzie pierścieniem, gdzie C i D są owalami. Podamy warunek konieczny i wystarczający na to, aby pierścień CD miał własność poryzmu Ponceleta. Niech owal C ma parametryzację daną za pomocą funkcji podparcia p , tj. $z(t) = p(t)e^{it} + p'(t)ie^{it}$, zaś owal D ma parametryzację $w(t) = z(t) + \lambda(t)ie^{it}$, jak to rozważaliśmy wcześniej. Skoro każdy punkt $w(t)$ należy do pewnej $\alpha(t)$ -izoptyki, to możemy napisać równanie uwikłane na funkcję α , mianowicie

$$(p'(t) + \lambda(t)) \sin \alpha - p(t + \alpha) + p(t) \cos \alpha = 0.$$

Niech $F(t, \alpha)$ oznacza lewą stronę tego równania. Można pokazać, stosując twierdzenie o funkcjach uwikłanych, że α jest funkcją różniczkowalną. Za pomocą funkcji α definiujemy funkcję φ , mianowicie $\varphi(t) = t + \alpha(t)$. Możemy teraz sformuło-

wać wspomniany warunek konieczny i wystarczający następująco (twierdzenie 4.2 [6]):

Wypukły pierścień CD ma własność poryzmu Poncela wtedy i tylko wtedy, gdy istnieje liczba naturalna $n \geq 3$ taka, że funkcja $\varphi^{[n]}(t) - t - 2\pi$ znika.

Zdefiniujemy dwie rodziny funkcji $\Phi_{L,n}$ i $\Gamma_{L,n}$. Ustalmy liczbę dodatnią L . Do rodziny $\Phi_{L,n}$ zaliczać będziemy wszystkie funkcje różniczkowalne $\varphi: \mathbb{R} \rightarrow \mathbb{R}$ spełniające warunki:

1. $0 < \varphi(0) < L$
2. $\varphi'(t) > 0$,
3. $\varphi^{[n]}(t) = t + L$,

dla dowolnego $t \in \mathbb{R}$. Do rodziny Γ_L zaliczać będziemy każdą funkcję ciągłą $b: \mathbb{R} \rightarrow \mathbb{R}$ taką, że

4. $b(t) > 0$,
5. $b(t+L) = b(t)$,

dla dowolnego $t \in \mathbb{R}$. Z każdą funkcją $b \in \Gamma_L$ stowarzyszymy dwie funkcje ciągłe, mianowicie

$$B(t) = \int_0^t b(u) du,$$

$$\beta(t) = B\left(t + \frac{L}{2}\right) - B(t).$$

Łatwo sprawdzić, że $\beta(t+L) = \beta(t)$ dla $t \in \mathbb{R}$. Oznacza to, że β osiąga minimum. W rodzinie Γ_L wyróżnimy podrodzinę $\Gamma_{L,n}$. Funkcja b z rodziny Γ_L należy do $\Gamma_{L,n}$, jeśli

$$\min_{t \in \mathbb{R}} \beta(t) > \frac{1}{n} B(L).$$

Rozpatrzmy klasę O_v krzywych ściśle wypukłych sparametryzowanych funkcją podparcia. Załóżmy, że funkcja podparcia p owalu C jest klasy C^2 i spełniona jest nierówność $p + p'' > 0$. Zachodzą następujące związki między wymienionymi rodzinami:

Twierdzenie 6.1 ([3]). *Każdy element rodziny $\Gamma_{2\pi,n}$ i krzywa $C \in Ov$ generują pierścień mający własność poryzmu Ponceleta dla liczby naturalnej n .*

Twierdzenie 6.2 ([3]). *Jeżeli $C \in Ov$, to każdy pierścień CD mający własność poryzmu Ponceleta dla liczby naturalnej n generuje element rodziny $\Gamma_{2\pi,n}$.*

Rozpatrzmy pierścień CD , gdzie C jest owalem, zaś D okręgiem. Owal C parametryzujemy za pomocą funkcji podparcia, tj. $z(t) = p(t)e^{it} + p'(t)ie^{it}$, zaś okrąg zapiszemy w postaci $w(t) = z(t) + \lambda(t)ie^{it}$. Tworzymy transversalną Ponceleta. Łamana ta definiuje dyfeomorfizm φ na owalu C . W pracy [7] wyznaczamy pochodną dyfeomorfizmu φ . Otrzymujemy następujący wzór

$$\varphi'(t) = \frac{\lambda(t)}{v(t)} \cdot \frac{v(\varphi(t))}{\eta(\varphi(t))},$$

gdzie $v(t)$ oznacza długość cięciwy okręgu stycznej do C w punkcie $z(t)$, zaś $\eta(t)$ oznacza różnicę między długością cięciwy a $\lambda(t)$. Rozpatrzmy szczególny przypadek pierścienia CD , mianowicie, gdy C jest elipsą. W tym przypadku wzór na φ' umożliwi podanie interpretacji geometrycznej miary niezmienniczej dla dyfeomorfizmu φ . Mamy następujące twierdzenie.

Twierdzenie 6.3 ([7]). *Dany jest pierścień CD , gdzie C jest elipsą, zaś D okręgiem. Miara niezmiennicza dla dyfeomorfizmu $\varphi : C \rightarrow C$ jest wyznaczona przez te krzywe i jest równa ilorazowi dodatniej stałej przez pole trójkąta o wierzchołkach na końcach cięciwy stycznej i środku elipsy.*

Ogólny przypadek, gdy C jest dowolnym owalem, pozostaje nierozwiązany.

7 O warunkach zamkniętości łamanej Ponceleta

Pierścień kołowy mający własność poryzmu Ponceleta nazwiemy pierścieniem Ponceleta. W roku 1823 Jacobi podał analityczny warunek na to, aby pierścień kołowy był pierścieniem Ponceleta.

Rozpatrzmy pierścień kołowy $C_r C_{a,R}$. Wprowadzimy następujące oznaczenia:

$$d = \frac{1}{R+a}, \quad b = \frac{1}{R-a}, \quad c = \frac{1}{r},$$

$$\omega = \cosh^{-1} \frac{d^2(b^2 - c^2) + 2c^2(d^2 - b^2)}{d^2(b^2 - c^2)},$$

$$k^2 = 1 - e^{-2\omega}.$$

Warunek Jacobiego, zgodnie z [26], ma postać:

Pierścień kołowy jest pierścieniem Ponceleta, jeśli

$$sc\left(\frac{1}{n}K(k), k\right) = \frac{c\sqrt{b^2 - d^2} + b\sqrt{c^2 - d^2}}{d(b+c)},$$

gdzie $sc(x, k)$ jest funkcją eliptyczną Jacobiego i $K(k)$ jest całką eliptyczną pierwszego rodzaju.

W pracy [3] podano następujące twierdzenie (twierdzenie 6.1):

Pierścień kołowy $C_r C_{a,R}$ jest pierścieniem Ponceleta wtedy i tylko wtedy, gdy liczba

$$n = \frac{B(2\pi)}{B(2 \operatorname{arc} \operatorname{tg} \frac{1}{r} \sqrt{R^2 - (r-a)^2})}$$

jest liczbą naturalną, gdzie

$$B(x) = \int_0^x \left[1 - \left(\frac{r - a \cos t}{R} \right)^2 \right]^{-\frac{1}{2}} dt$$

i funkcja podcałkowa należy do klasy $\Gamma_{2\pi, n}$.

Oba te twierdzenia zawierają całki, które nie są funkcjami elementarnymi. Oczywiście wzory Fussa i warunek Jacobiego muszą pokrywać się. Jednakże nieznanne jest przejście od formuły Jacobiego do wzorów Fussa.

Literatura

- [1] M. Berger, *Geometry I*, Springer-Verlag, New York, 1994.
- [2] H. J. M. Bos, C. Kers, F. Dort, D. W. Raven, Poncelet's closure theorem, *Expo. Math.*, 5, 1987, 289–364.
- [3] W. Cieślak, The Poncelet annuli, *Beitr. Algebra Geom.* 55, No. 1, 2014, 301–309.
- [4] W. Cieślak, H. Felińska, J. Szuster, W. Mozgawa, Poncelet's porism in transformation group framework, *ASTRJ*, 8, No. 24, 2014, 28–31.
- [5] W. Cieślak, E. Szczygielska, Circuminscribed polygons in a plane annulus, *Ann. Univ. Mariae Curie-Skłodowska*, 62, 2008, 49–53.
- [6] Cieślak W., Szczygielska E., On Poncelet's porism, *Ann. Univ. Mariae Curie-Skłodowska*, 64, 2, 2010, 21–28.
- [7] W. Cieślak, W. Mozgawa, In search of a measure in Poncelet's porism, *Acta Math. Hungarica*, 149 (2), 2016, 338–345.
- [8] W. Cieślak, W. Mozgawa, Extremal perimeters of quadrangles in the Poncelet porism, *Bull. Aust. Math. Soc.* 91, 2015, 487–501.
- [9] W. Cieślak, H. Martini, W. Mozgawa, On rotation index of bar billiards and Poncelet's porism, *Bull. Belg. Math. Soc. Simon Stevin*, 20(2), 2013, 287–300
- [10] W. Cieślak, H. Martini, W. Mozgawa, Rotation indices related to Poncelet's closure theorem, *Ann. Univ. Mariae Curie-Skłodowska*, Sect. A, 68, No. 2, 2014, 19–26.
- [11] V. Dragović, M. Radnović, *Poncelet Porisms and Beyond. Integrable Billiards. Hyperelliptic Jacobians and Pencils of Quadrics*, Frontiers in Mathematics, Birkhauser, Basel, 2011.

-
- [12] L. Flatto, *Poncelet's Theorem*, American Mathematical Society, Providence, RI, 2009
- [13] N. Fuss, *Nova Acta Petropol.*, 10, 1792.
- [14] C. G. J. Jacobi, Über die Anwendung der elliptischen Transcendenten auf ein bekanntes Problem der Elementargeometrie, *J. Reine Angew. Math.*, 3, 1823, 376–387.
- [15] S. M. Kerauala, Poncelet porism in two circles, *Bull. Calcutta Math. Soc.*, 39, 1947, 85–105.
- [16] J. L. King, Three Problems in Search of a Measure, *Amer. Math. Monthly*, 101, 1994, 609–628.
- [17] R. Kołodziej, The rotation number of some transformation related to the billiards on an elliptic table, *Studia Math.* 81, 1985, 293–302.
- [18] M. Michta, *Wzory Fussa w poryźmie Ponceleta*, praca magisterska, Politechnika Lubelska, Lublin 2015.
- [19] W. Mozgawa, On billiards and Poncelet's porism, *Rend. Semin. Mat. Univ. Padova*, 120, 2008, 157–166.
- [20] J. V. Poncelet, *Traité des propriétés projectives des figures: ouvrage utile a qui s'occupent des applications de la géométrie descriptive et d'opérations géométriques sur le terrain*, Vols. 1-2, 2nd ed. Paris: Gauthier-Villars, 1865-66.
- [21] M. Radić, Extreme Areas of Triangles in Poncelet's Closure Theorem, *Forum Geom.*, 4, 2004, 23–26.
- [22] M. Radić, An improved method for establishing Fuss' relations for bicentric polygons, *C. R. Math. Acad. Sci.*, 348, 2010, 415–417.
- [23] F. J. Richelot, Anwendung der elliptischen Transcendenten auf die sphärischen Polygone; welche zugleich Kreise der Kugel eingeschrieben und einem andern umgeschrieben sind, *J. Reine Angew. Math.*, 5, 1830, 250–267.

- [24] L. Santalo, *Integral geometry and geometric probability*, *Encyclopedia of Mathematics and its Applications*, Vol.1, Addison-Wesley Publishing Co., Reading, Mass., London-Amsterdam, 1976.
- [25] J. Steiner, Aufgaben und Lehrsätze, erstere aufzulösen, letztere zu beweisen, *J. Reine Angew. Math.*, 2, 1827, 289.
- [26] Weisstein E. W., *Poncelet's Porism*,
<http://mathworld.wolfram.com/PonceletsPorism.html>.

Metody deflatorowe w wycenie przepływów finansowych

Janusz J. Szuster¹

Streszczenie

Praca dotyczy koncepcji i wykorzystania deflatorów w zagadnieniach matematyki finansowej i aktuarialnej. Są one stochastycznymi odpowiednikami klasycznych czynników dyskontujących i pozwalają na realistyczną wycenę przyszłych niedeterministycznych przepływów finansowych. Zaprezentowane zostały własności tego pojęcia i przykłady ich wykorzystania.

Abstract

The paper deals with the concept and applications of deflators in financial and actuarial mathematics. Deflator is the stochastic counterpart of classical discount factor and it allows for a realistic estimate of the cost of future non-deterministic cash flows. The paper presents main properties of deflator and examples of its use.

Słowa kluczowe: wycena spójna z rynkiem, filtracja, czynnik dyskontujący, deflator, rezerwy

1 Wstęp

Dyskontowanie, stanowiące nieodłączny element modelowania matematycznego procesów finansowych, pozwala na ustalanie aktualnej wartości inwestycji, trwającej w ustalonym przedziale czasu. W trakcie tworzenia takiego modelu stajemy

¹Katedra Matematyki Stosowanej; Politechnika Lubelska

przed koniecznością wyceny efektów inwestycji na chwilę początkową, ale również na inne ważne chwile, w których podejmowane są decyzje o uruchomieniu, bądź zaniechaniu jakiegoś procesu. Umiejętność przeprowadzenia takiej oceny pozwala, co zrozumiale, na podjęcie właściwych decyzji ekonomicznych, z drugiej zaś zmusza do analizy warunków rynkowych, w których inwestycja będzie funkcjonować.

Dość często proces akumulacji, a więc proces wzrostu wartości inwestycji, opiera się na deterministycznych funkcjach akumulacji, wykorzystujących w prognozowaniu stóp informacje z rynku, reprezentowane przez struktury terminowe stóp procentowych. Jest tak, bo modelowanie deterministyczne, jakkolwiek niezbyt precyzyjne, jest nietrudne od strony technicznej i w warunkach stabilnego rynku oraz niezbyt długiego czasu trwania inwestycji, na ogół w zadowalający sposób oddaje przyszłą wartość inwestycji.

Prezentowana praca dotyczy metody kalkulacji przyszłych przepływów pieniężnych w modelach dyskretnych, bazującej na pojęciu deflatora. Tradycyjne techniki wyceny opierały się na projekcjach deterministycznych, oraz dyskontowaniu opartym na stopie dyskontowej ryzyka. Przedstawione w pracy metody, dają możliwość dokonania stochastycznej wyceny niepewnych przepływów pieniężnych, w oparciu o technikę deflatorową. W przypadku modelu właściwie skalibrowanego względem rynku, daje rezultaty pokrywające się z wycenami rynkowymi. Naturalnym użytkownikiem technik deflatorowych są specjaliści zajmujących się ekonomiczną wyceną zobowiązań lub aktywów.

2 Niezbędne pojęcia i fakty

Rozpocznijmy od uwagi, że czekające nas rozumowania wymagają przyjęcia koniecznych umów, niekiedy uwzględniając nasze doświadczenia dość idealistycznych. Uprzedzając jednak ewentualne wątpliwości, co do rzetelności proponowanych rozwiązań, warto podkreślić, że umowy te podyktowane są stroną techniczną. Zakładamy mianowicie: jednakowy dostęp do informacji dla wszystkich inwestorów, równość oprocentowania depozytów i lokat, brak kosztów transakcji, prowizji

i jakichkolwiek podatków z nią związanych, jednakową dostępność wszystkich inwestorów do kredytów oraz brak ograniczeń górnych na ich wysokość, natychmiastową wykonalność wszystkich operacji, całkowitą płynność rynku oraz niemożność wpływu na ceny rynkowe ze strony pojedynczych inwestorów.

2.1 Kontrakty i opcje

W potocznym rozumieniu nazwa instrument finansowy oznacza umowę między dwiema stronami, regulującą zależność finansową między nimi. Natura zależności między stronami jest określona przez rodzaj instrumentu finansowego. W pracy przez *instrument finansowy* rozumiemy dowolną zmienną losową, na przykład X , wyrażającą kwotę pieniędzy, którą w chwili t kupujący płaci sprzedającemu. Natomiast *instrumenty pochodne* definiujemy jako funkcje od zmiennej losowej X .

Ponieważ w pracy analizować będziemy zarówno jednookresowe jak i wielookresowe modele rynku, to warto podkreślić, że w pierwszym przypadku nie ma różnicy między instrumentem finansowym, a generowanym przez niego instrumentem pochodnym, natomiast w drugim przypadku różnica ta jest już istotna.

Wśród instrumentów finansowych interesującymi dla nas będą *kontrakty i opcje*. Zasadniczą różnicą między nimi jest kwestia symetrii. Otóż w przypadku kontraktów mamy do czynienia z instrumentami symetrycznymi, w tym rozumieniu, że strony kontraktu dokonują transakcji po ustalonym z góry czasie, podczas gdy w przypadku opcji, ich posiadacz (zgodnie z nazwą instrumentu) ma jedynie możliwość, a nie przymus (obowiązek) dokonania transakcji (kupna lub sprzedaży). Warto przypomnieć, że wyróżniamy dwie strony kontraktów i opcji: krótką i długą. *Pozycja długa* – oznacza kupującego (ang. *buyer*), natomiast *pozycja krótka* – sprzedającego (ang. *seller*).

W przypadku opcji kupna, pozycja długa daje możliwość nieograniczonego zysku, natomiast strata równa jest wtedy premii za ryzyko podczas, gdy pozycja krótka jest daje zysk równy premii za ryzyko, zaś strata może tu być nieograniczona.

Przy opcji sprzedaży, pozycja długa daje zysk ograniczony ceną wykonania,

a stratę równą premii za ryzyko, natomiast pozycja krótka, na odwrót – zysk jest równy premii za ryzyko, a strata jest ograniczona ceną wykonania. Podkreślmy, że opcje są instrumentami pochodnymi i mogą być wystawiane na różne instrumenty finansowe, rozumiane jako instrumenty bazowe. W szczególności opcje mogą być wystawiane na towary, na kontrakty, na waluty, czy na stopę procentową i nie jest to pełna lista możliwości. Jeżeli S_t oznacza cenę instrumentu bazowego, czyli aktywa, w chwili t , zaś K jest ceną wykonania, to wypłata z opcji kupna wyraża się zależnością

$$\max\{S_t - K; 0\},$$

natomiast wypłata z opcji sprzedaży wynosi

$$\max\{K - S_t; 0\}.$$

Opcje można podzielić na dwie główne kategorie: opcje waniliowe i pozostałe, czyli opcje egzotyczne. Powodem rozgraniczenia jest sposób wyznaczania wartości S_t . W przypadku opcji waniliowych S_t jest ceną realizacji opcji w chwili t . Natomiast w przypadku opcji egzotycznych S_t jest wyrażane bądź jako średnia arytmetyczna, bądź jako średnia geometryczna w ujęciu dykretnym, wartości S_{t_i} w ustalonych chwilach czasu t_i , ewentualnie jako średnie ciągłe dla każdego z tych rodzajów. Opcje waniliowe dzielimy na europejskie i amerykańskie, natomiast opcje egzotyczne mają wiele odmian, wśród których najważniejsze to opcje barierowe, azjatyckie i bermudzkie. Wymienione rodzaje opcji nie wyczerpują całego repertuaru opcji egzotycznych, a jedynie sygnalizują najczęściej spotykane.

2.2 Rynki dwustanowe – jednookresowe i wielookresowe

Ten model rynku jest najprostszym modelem rynku finansowego. Jego istotą są następujące założenia:

- (i) Istnienie dwóch jedynie momentów przeprowadzania transakcji: początku czyli chwili $t = T_0 = 0$ oraz końca czyli momentu $t = T$.

- (ii) Istnienie jedynie dwóch scenariuszy przebiegu wydarzeń, z których scenariusz korzystny tradycyjnie oznaczymy przez ω_1 natomiast scenariusz niekorzystny przez ω_2 . Zbiór scenariuszy $\{\omega_1; \omega_2\}$ będzie traktowany jako przestrzeń zdarzeń i oznaczany przez Ω . Ze względu na skończoność zbioru Ω przyjmujemy, że σ -ciałem w przestrzeni Ω jest $\mathcal{F} = 2^\Omega$. Do tego funkcję prawdopodobieństwa P na zbiorze Ω określamy równościami: $P(\omega_1) = p$ i $P(\omega_2) = 1 - p$, dla pewnego $p \in (0; 1)$.
- (iii) Istnienie dwóch papierów wartościowych, z których jeden jest ryzykowny, a więc taki, którego cena w przyszłości nie jest zdeterminowana – przykładem są akcje, czy fundusze inwestycyjne, a drugi bezryzykowny – przykładem może być lokata bankowa lub obligacje skarbowe.

Z uwagi na założenie (i) niech $t \in \{0; T\}$ oraz niech S_t oznacza cenę za jednostkę papieru ryzykownego w chwili t , zaś B_t cenę za jednostkę papieru bezryzykownego w chwili t . Dodatkowo zakładamy, że na przedziale czasu $[0; T]$ stopa procentowa r jest nieujemna. W konsekwencji uzyskujemy zależność

$$B_t = \begin{cases} 1, & \text{gdy } t = 0, \\ 1 + r, & \text{gdy } t = T, \end{cases} \quad (2.1)$$

oraz

$$S_t(\omega) = \begin{cases} s, & \text{gdy } t = 0 \text{ oraz } \omega \in \{\omega_1; \omega_2\}, \\ S^u, & \text{gdy } t = T \text{ oraz } \omega = \omega_1, \\ S^d, & \text{gdy } t = T \text{ oraz } \omega = \omega_2, \end{cases} \quad (2.2)$$

gdzie $s \in \mathbb{R}_+$ jest ustalone w konkretnym przypadku. Dodajmy dla porządku, że symbole u i d górnych indeksów S_t pochodzą od angielskich określeń „do góry” i „na dół”.

Ze względu na wcześniejszą umowę, że ω_1 jest scenariuszem korzystnym, zaś ω_2 – niekorzystnym, mamy oczywistą nierówność $S^u > S^d$.

W literaturze (np. [2, 1]) wprowadza się dodatkową funkcję $Z : \Omega \rightarrow \mathbb{R}$ określoną warunkiem

$$Z(\omega) = \begin{cases} u, & \text{gdy } \omega = \omega_1, \\ d, & \text{gdy } \omega = \omega_2. \end{cases}$$

Liczba $Z - 1$, jako wielkość niemianowana, wyraża zmianę ceny aktywa pomiędzy chwilami 0 i T , a w konsekwencji mamy $S_T(\omega) = S_0 Z = sZ$.

Osobnym problemem w przypadku rynku jednookresowego jest problem wyceny opcji. Dodajmy, że opcje traktujemy jako instrument pochodny dla pewnego instrumentu bazowego, instrumentu, którego cena musi być zgodna z wyceną rynku. Łatwo zauważyć, że opcję europejską można utożsamiać z wypłatą, co będziemy niejednokrotnie wykorzystywać. W konsekwencji tej operacji każdy instrument pochodny może być rozumiany jako wypłata generowana przez ten instrument. Ponieważ wypłata jest zjawiskiem losowym, bo zależnym od scenariusza ω – korzystnego lub niekorzystnego, to można ją traktować jako zmienną losową, którą tradycyjnie oznaczymy przez X . Zatem przez wypłatę w chwili T będziemy rozumieć każdą zmienną losową określoną na zbiorze scenariuszy $\Omega = \{\omega_1; \omega_2\}$. W omawianym modelu jednookresowym dwustanowym wypłata X jest pewną funkcją wartości S_T , czyli $X = f(S_T)$, gdzie f oznacza tę nieznaną funkcję.

Jednym z problemów jakimi zajmuje się inżynieria finansowa jest zagadnienie wyceny opcji, równoznaczne – jak zauważyliśmy, z wyceną opcji. Pomocnym w tym procesie pojęciem jest pojęcie portfela replikującego, do omówienia którego przystąpimy po sformułowaniu potrzebnych definicji.

Przez *portfel* rozumiemy każdą parę postaci $\varphi = (\alpha_0; \beta_0) \in \mathbb{R}^2$, gdzie α_0 oznacza liczbę posiadanych przez inwestora akcji, natomiast β_0 oznacza stan konta bankowego inwestora w chwili początkowej $t = 0$. Należy tu podkreślić, że zawarte w definicji informacje, że liczby α_0 oraz β_0 są dowolnymi liczbami rzeczywistymi, mają odzwierciedlenie we wcześniejszych założeniach. Dopuściliśmy mianowicie, że instrumenty finansowe są doskonale podzielne, co oznacza, że można posiadać nawet niewymierną liczbę akcji, również ujemną. Można również mieć dowolny

stan konta bankowego, w tym ujemny, bo na wstępie dopuściliśmy, że każdy inwestor ma nieograniczony dostęp do kredytów i lokat. Ujemna liczba akcji wynika z dopuszczenia tzw. *krótkiej sprzedaży*, co oznacza, że w chwili T inwestor może pożyczyć ustaloną liczbę akcji, by następnie odkupić identyczną liczbę akcji i zwrócić je w chwili końcowej T . Uczestnicy rynku finansowego mówią wtedy o zajęciu na nim krótkiej pozycji. Niedopuszczenie na rynku krótkiej sprzedaży ma formalne ograniczenie w postaci założenia, że $\alpha_0 \geq 0$. Dodajmy na zakończenie, że zbiór wszystkich portfeli oznaczamy będziemy, zgodnie z przyjętą w literaturze manierą, przez Φ .

Rozpatrzmy przypadek, gdy inwestor posiada portfel $\varphi = (\alpha_0; \beta_0)$. Przez *wartość portfela* φ w chwilach $t \in \{0; T\}$ oznaczaną przez $V_t(\varphi)$ rozumieć będziemy liczbę określoną równością

$$V_t(\varphi) = \begin{cases} \alpha_0 S_0 + \beta_0, & \text{gdy } t = 0, \\ \alpha_0 S_T + \beta_0(1+r), & \text{gdy } t = T, \end{cases}$$

gdzie r oznacza, jak poprzednio, nieujemną stopę procentową na przedziale $[0; T]$. Ponadto podkreślmy, że stan portfela nie ulega w tym czasie jakiegokolwiek zmianie.

Ponieważ inwestor sprzedaje wypłatę X , to koniecznym jest jej zabezpieczenie, wyrażające się tym, że potrafi on za uzyskane pieniądze zbudować w chwili T portfel o wartości X .

Portfel φ nazywać będziemy *portfelem replikującym wypłatę X* wtedy i tylko wtedy, gdy spełniona jest równość

$$V_T(\varphi)(\omega_i) = X(\omega_i), \quad (2.3)$$

dla każdego $\omega_i \in \{\omega_1, \omega_2\}$. Jeżeli dla instrumentu pochodnego o wypłacie X istnieje portfel $\varphi(\alpha_0, \beta_0)$, że równość (2.3) spełniona jest z prawdopodobieństwem 1, to X nazywamy *instrumentem osiągalnym*. Gdy natomiast na ustalonym rynku każdy instrument pochodny X jest osiągalny, to rynek ten nazywamy *rynkiem zupełnym*.

Portfel replikujący, poprzez wykluczenie ryzyka związanego z niepewnością, co do mogącego wystąpić scenariusza, stanowi niezawodne zabezpieczenie wypłaty X .

Dobrodziejstwo istnienia takiego portfela nie jest jednak oczywiste. Problemu tego dotyczy następujące twierdzenie.

Twierdzenie 2.1 ([1, 4]). *Dla każdego instrumentu pochodnego istnieje dokładnie jeden portfel replikujący $\varphi = (\alpha_0; \beta_0)$. Gdy X jest wypłatą z instrumentu pochodnego, to*

$$\alpha_0 = \frac{X^u - X^d}{S^u - S^d} \quad \text{oraz} \quad \beta_0 = \frac{X^d S^u - X^u S^d}{(1+r)(S^u - S^d)}$$

przy czym $X^u = X(\omega_1)$ oraz $X^d = X(\omega_2)$.

Dowód. Zauważyliśmy poprzednio, że zachodzi nierówność $S^d < S^u$, a więc $S^u - S^d \neq 0$. Z tego powodu macierz układu równań liniowych

$$\begin{aligned} \alpha_0 S^u + (1+r)\beta_0 &= X^u \\ \alpha_0 S^d + (1+r)\beta_0 &= X^d \end{aligned}$$

ma niezerowy wyznacznik, co na mocy twierdzenia Cramera oznacza, że układ taki, o nieosobliwej macierzy układu, ma dokładnie jedno rozwiązanie. Łatwe rachunki dają zależności zawarte w tezie twierdzenia. Zatem portfel replikujący istnieje dla każdego instrumentu pochodnego o wypłacie X . \square

W procesie wyznaczania portfela replikującego, fundamentalnym problemem jest wyznaczenie wartości uczciwej lub inaczej wartości godziwej wypłaty X , rozumianej jako cena inwestycji początkowej niezbędnej do budowy portfela replikującego.

W tym celu oznaczmy przez $\Pi_t(X)$ cenę instrumentu pochodnego w chwili t , przy czym $t \in \{0, T\}$. Z definicyjnych powodów, w chwili końcowej T wartość instrumentu pochodnego X musi być identyczna z wypłatą, a więc $\Pi_T(X) = X_T$.

Dla wyznaczenia wartości godziwej instrumentu pochodnego X w chwili początkowej $t = 0$, dokonajmy replikacji tego instrumentu za pomocą portfela φ , utworzonego z instrumentu bezryzykownego, na przykład obligacji zerokuponowej lub lokaty oraz aktywa S . Jako wartość instrumentu pochodnego bierzemy wartość portfela replikującego, co oznacza, że przyjmujemy $\Pi_t(X) = V_t(\varphi)$. Przy tych

oznaczeniach, przez *wartość uczciwą instrumentu pochodnego* X rozumiemy liczbę rzeczywistą $\Pi_0(X) = \varphi_0(X)$, gdzie φ_0 jest portfelem replikującym wypłatę X .

Różnica modelu rynku wielookresowego, w stosunku do poprzedniego polega na wyróżnieniu w przedziale czasu $[0; T]$ punktów dzielących go na podprzedziały równej długości. W przypadku podziału przedziału na n podprzedziałów mamy, że punktami podziału będą wtedy punkty $t_i = i \cdot \frac{T}{n}$, dla $i \in \{0, 1, \dots, n\} \stackrel{\text{df}}{=} \mathcal{I}$. Dodatkowo dopuszczamy większą niż poprzednio liczbę scenariuszy, którą oznaczymy przez m , a więc $\Omega = \{\omega_1, \omega_2, \dots, \omega_m\}$. Jeśli liczba scenariuszy jest nadal równa dwa, to w dalszym ciągu mamy do czynienia z rynkiem dwustanowym, tyle że wielookresowym. Jest on oczywiście podprzypadkiem obecnie rozpatrywanej wersji.

Skończoność zbioru Ω implikuje postać σ -ciała $\mathcal{F} = 2^\Omega$. Dla wszystkich scenariuszy ze skończonej przestrzeni Ω , ich prawdopodobieństwa są dodatnie, tj. $P(\omega_k) > 0$ dla $k \in \{1, 2, \dots, m\}$. Z σ -ciałem \mathcal{F} związany jest skończony ciąg σ -ciał $\underline{\mathcal{F}}$, którego wyrazy indeksowane są elementami zbioru \mathcal{I} , stanowiące równocześnie numery chwil $t_i \in [0; T]$. Intencją wprowadzenia ciągu σ -ciał $\underline{\mathcal{F}}$ jest rozumienie jego wyrazów jako stanów wiedzy o rynku, w poszczególnych chwilach t_i . Ponieważ wiedza ta przyrasta wraz ze wzrostem indeksu i , to ciąg ten jest wstępujący, a przy tym

$$\emptyset = \mathcal{F}_0 \subset \mathcal{F}_1 \subset \dots \subset \mathcal{F}_{m-1} \subset \mathcal{F}_m = \mathcal{F}.$$

Wiedza reprezentowana przez σ -ciała z $\underline{\mathcal{F}}$ jest zatem swego rodzaju filtacją i dlatego też jest nazywana *filtracją przestrzeni* $(\Omega, \underline{\mathcal{F}}, P)$.

Przyjmujemy również, że na analizowanym rynku znajduje się skończona liczba pierwotnych instrumentów finansowych, których liczbę oznaczymy przez $j+1$. Ceny tych instrumentów w chwili t , rozumiane jako zmienne losowe, oznaczymy zwyczajowo przez $S^0(t_i), S^1(t_i), \dots, S^j(t_i)$. Do tego podkreślmy, że zmienna losowa S^0 opisuje ceny instrumentu bezryzykowego. Zmienne $S^l(t_i)$, gdzie $l \in \{0, 1, \dots, j\}$, są mierzalne względem σ -ciała \mathcal{F}_i , ponieważ wiedza o rynku w chwili t_i nie pozwala wyrokować o zachowaniu zmiennej w chwilach z wyższymi indeksami niż i , czyli o przyszłych cenach instrumentów. Jest ona natomiast

wystarczająca, aby mówić o historycznych poziomach cen, osiągniętych w chwilach nie późniejszych niż t_i . W chwili $t_i \in [0; T]$ określony jest zatem wektor cen

$$S(t_i) = \begin{pmatrix} S^0(t_i) \\ S^1(t_i) \\ \vdots \\ S^j(t_i) \end{pmatrix} = \left(S^0(t_i), S^1(t_i), \dots, S^j(t_i) \right)^\diamond,$$

gdzie \diamond jest symbolem transpozycji macierzy. Wektor $S(t_i) \in \mathbb{R}^{j+1}$ składa się ciągu adaptowanych do filtracji \mathcal{F} zmiennych losowych.

Zakładamy ponadto, że $S^0(t_0) = 1$ oraz, że stopa procentowa r jest stała dla każdego podokresu $[t_i; t_{i+1})$ oraz, że funkcja akumulacji jest złożona, co implikuje, że $S^0(t_i) = (1+r)^i$. Przedstawione tu własności rynku, zawarte są w określeniu *rynek skończony*.

Przez *strategię inwestycyjną* rozumiemy będziemy proces prognozowalny² $\{\varphi\}_{i \in \mathcal{I}}$, gdzie

$$\varphi_i = (\varphi_i^0, \varphi_i^1, \dots, \varphi_i^j)^\diamond,$$

dla $j \in \mathcal{I}$. Przyjmujemy, że zmienne losowe φ_i^l opisują liczbę jednostek l -tego instrumentu finansowego, jakie zawiera portfel na przedziale czasu $[t_{i-1}; t_i)$. Wartości zmiennych losowych są liczbami rzeczywistymi, co jest konsekwencją wcześniejszego założenia o doskonałej podzielności walorów, a także dopuszczona jest krótka sprzedaż i nieograniczony dostęp do kredytów. Dysponując wektorami z \mathbb{R}^{j+1} , z jednej strony wektorem $S(t_i)$, a z drugiej wektorem φ_i możemy określić wielkość skalarną $V_i(\varphi)$, charakteryzującą portfel w chwili t_i i nazywaną *wartością w portfela t_i* . Mianowicie

$$V_i(\varphi) = \sum_{i=1}^n \varphi_i S_i.$$

Należy podkreślić, że na rynku doskonałym, bez prowizji i kosztów transakcyjnych, wartość portfela konstruowanego w chwili t , ale na chwilę t_{i+1} , jest równa

²Prognozowalność oznacza fakt konstruowania wektora w chwili t_i , na podstawie wiedzy dostępnej w chwili t_{i-1} i nie ulega zmianie, aż do chwili t_i . Jest to więc formalnie zależność rekurencyjna rzędu pierwszego.

$\varphi_{i+1}S(t_i)$, natomiast w chwili t_{i+1} wynosi $\varphi_{i+1}S(t_{i+1})$. Co implikuje, że zmiana wartości portfela między chwilami t_i a t_{i+1} jest równa $\varphi_{i+1}[S(t_{i+1}) - S(t_i)]$.

Rynek uznajemy za *rynek opisany* jeśli można ustalić cenę S instrumentu ryzykownego, wartość B jednostki rachunku bankowego oraz zbiór możliwych portfeli Φ . Oznacza to, że rynek \mathcal{M} rozumiemy jako trójkę

$$\mathcal{M} = (B, S, \Phi).$$

Na tak zdefiniowanym rynku można dokonać wyceny każdej wypłaty – rozumianej jako instrument pochodny. Ma on jednak jedną, ale bardzo ważną, słabą stronę. Dopuszcza mianowicie możliwość dla dodatniej wypłaty X istnienie jej ujemnej ceny. Oznacza to, że nie dokonując żadnej inwestycji kapitałowej, ale obierając odpowiednią strategię, możemy w chwili końcowej otrzymać zysk. W języku przyjętych oznaczeń wyraża się to nierównością $\Pi_0(X) < 0$. Zjawisko takie, polegające na możliwości otrzymania zysku bez nakładów nazywamy *arbitrażem*. Natomiast w sposób sformalizowany rozumiemy to zjawisko następująco.

Mówimy, że *model rynku \mathcal{M} nie dopuszcza arbitrażu* wtedy i tylko wtedy, gdy nie istnieje portfel $\varphi^* \in \Phi$, że

$$\begin{cases} V_0(\varphi^*) = 0, \\ V_T(\varphi^*) \geq 0, \\ V_t(\varphi^*)(\omega) \text{ dla pewnego } \omega \in \Omega. \end{cases} \quad (2.4)$$

Natomiast portfel $\varphi^* \in \Phi$, dla którego zachodzą warunki (2.4) nazywamy *portfelem z możliwością arbitrażu* lub krótko *możliwością arbitrażu*.

Praktyka rynkowa pokazuje, że wszelkie możliwości prowadzące do zysku bez ryzyka są natychmiast wykorzystywane przez uczestników tego rynku, zwanych *arbitrażystami*, zjawisko takie jest jednak dyskryminujące dla pozostałych inwestorów. Wskutek działania arbitrażystów, po chwilowej luce na rynku, dającej im możliwość zarobku, rynek wraca do stanu równowagi. Dlatego też eliminacja możliwości arbitrażu jest jednym z głównych wyzwań.

Twierdzenie 2.2 ([1, 4]). *Rynek $\mathcal{M} = (B, S, \Phi)$ jest wolny od arbitrażu wtedy i tylko wtedy, gdy spełniona jest na nim nierówność*

$$S^d < (1+r)S_0 < S^u. \quad (2.5)$$

Dowód. Przypomnijmy, że wprowadziliśmy poprzednio funkcję $Z : \Omega \rightarrow \mathbb{R}$, gdzie

$$Z(\omega) = \begin{cases} u, & \text{gdy } \omega = \omega_1, \\ d, & \text{gdy } \omega = \omega_2. \end{cases}$$

Jej istnienie pozwala na uproszczenie tezy twierdzenia, której można teraz nadać postać

$$d < 1+r < u. \quad (2.6)$$

Dostateczność. Przypuśćmy, że zachodzi nierówność (2.5) i chcemy dowieść nieistnienie arbitrażu. Rozważmy w tym celu portfel $\varphi = (\beta, \alpha)$, dla którego $V_0(\varphi) = \alpha S_0 + \beta = 0$. Jeśli równość ta jest spełniona i $\alpha = 0$, to również $\beta = 0$, co oznacza, że $\varphi = (0; 0)$ oraz $V_T(\varphi) = 0$. Zatem nie jest spełniona definicja portfela arbitrażowego i twierdzenie zachodzi. Jeśli zaś $\alpha \neq 0$ oraz nadal $V_0(\varphi) = \alpha S_0 + \beta = 0$, to wartość portfela $\varphi = (\alpha, \beta)$ jest równa zgodnie z wcześniejszymi ustaleniami

$$\begin{aligned} V_T(\varphi) &= \alpha S_T + (1+r)\beta = \alpha S_T - \alpha S_0(1+r) \\ &= \begin{cases} \alpha S_0(u-1-r), & \text{gdy } Z(\omega) = u, \\ \alpha S_0(d-1-r), & \text{gdy } Z(\omega) = d, \end{cases} \end{aligned}$$

ponieważ $\beta = -\alpha S_0$. Z drugiej strony odwołując się do nierówności $d < 1+r < u$ dostajemy, że $u - (1+r) > 0$ oraz $d - (1+r) < 0$

$$V_T(\varphi) = \begin{cases} \alpha S_0(u-1-r) > 0, & \text{gdy } Z(\omega) = u \text{ oraz } \alpha > 0, \\ \alpha S_0(d-1-r) < 0, & \text{gdy } Z(\omega) = d \text{ oraz } \alpha > 0, \end{cases}$$

natomiast

$$V_T(\varphi) = \begin{cases} \alpha S_0(u-1-r) < 0, & \text{gdy } Z(\omega) = u \text{ oraz } \alpha < 0, \\ \alpha S_0(d-1-r) > 0, & \text{gdy } Z(\omega) = d \text{ oraz } \alpha < 0. \end{cases}$$

Wynika stąd, że w przypadku portfela φ , dla którego $V_0(\varphi) = 0$ nie ma pewności, że $V_T(\varphi) > 0$, a to oznacza, że portfel ten nie może być portfelem arbitrażowym.

Konieczność. Załóżmy, że nie zachodzi warunek występujący w tezie, a przez to równoważny mu warunek (2.5). Oznacza to, że nie zachodzi przynajmniej jedna z nierówności składowych i niech będzie $1 + r \geq u$, a mimo to rynek jest wolny od arbitrażu.

W celu wykazania sprzeczności takiego rozumowania obierzmy portfel $\varphi = (-1, S_0)$. Taki skład portfela oznacza, że dokonujemy krótkiej sprzedaży jednego instrumentu ryzykownego, a uzyskaną w jej wyniku kwotę S_0 składamy na rachunek bankowy. Wartość portfela w chwili $t = 0$ wynosi, co oczywiste,

$$V_0(\varphi) = (-1)S_0 + S_0 = 0,$$

natomiast w chwili $t = T$ mamy, że wartość portfela jest równa

$$\begin{aligned} V_T(\varphi) &= -S_0 \cdot Z + (1+r)S_0 = S_0(1+r-Z) \\ &= \begin{cases} S_0(1+r-u) \geq 0, & \text{gdy } Z(\omega) = u, \\ S_0(1+r-d) > 0, & \text{gdy } Z(\omega) = d. \end{cases} \end{aligned}$$

Dowodzi to, że φ jest portfelem arbitrażowym i przeczy założeniu o nieistnieniu arbitrażu. \square

2.3 Miara martyngałowa

Na rynku bez możliwości arbitrażu, istnieje sposób wyceny instrumentów pochodnych oparty jedynie na wartościach instrumentów, a nie na funkcjach cen tych instrumentów, polegający na obliczaniu wartości oczekiwanej względem pewnej specjalnej miary prawdopodobieństwa. Tej właśnie miary będziemy teraz poszukiwać.

Ustaliliśmy poprzednio, że na rynku bez możliwości arbitrażu spełniona jest nierówność (2.6) czyli, że $(1+r) \in (d; u)$. Wynika stąd istnienie liczby $\lambda \in (0; 1)$, spełniającej równość

$$1 + r = \lambda d + (1 - \lambda)u,$$

gdzie, jak pamiętamy, u i d oznaczają stosunki ceny końcowej aktywa, odpowiednio do jego ceny „górnjej” i „dolnej”. Ponieważ $\lambda + (1 - \lambda) = 1$, to został określony rozkład prawdopodobieństwa dla zmiennej losowej Z postaci $\{(d, \lambda), (u; 1 - \lambda)\}$, a przez to zdefiniowana miara prawdopodobieństwa Q , w której $Q(Z = d) = \lambda$ oraz $Q(Z = u) = 1 - \lambda$. Stąd pamiętając, że $s = S_0$ jest ceną aktywa ryzykownego w chwili początkowej, czyli w $t = 0$ dostajemy, że

$$(1 + r)s = \lambda ds + (1 - \lambda)us,$$

co jest wartością oczekiwaną ceny S_T aktywa w chwili T . Otrzymaliśmy więc, że

$$E_Q S_T = \lambda ds + (1 - \lambda)us = (1 + r)s.$$

Wynika stąd, że

$$s = \frac{1}{1 + r} E_Q S_T$$

jest zdyskontowaną na moment $t = 0$ wartością oczekiwaną ceny aktywa z chwili $t = T$. Zauważamy, że wpływ na zdyskontowaną wartość oczekiwaną mają nie tyle wartości cen, ale ich stosunek. W efekcie tej obserwacji dokonuje się zabiegu, by wszystkie ceny były rozpatrywane w kontekście ustalonego aktywa, zwykle jest to jednostka waloru bezryzykownego. Oznaczając jej wartość w chwili t_i przez B_i mamy, że jej zdyskontowana wartość jej równa 1, a wtedy w miejsce procesu cen rozważamy zdyskontowany proces cen $S^*(t_i) = \frac{S(t_i)}{B(t_i)}$, co daje $S_0^* = S_0$ oraz

$$S_T^* = \frac{S_T}{1 + r}.$$

Miarę probabilistyczną P^* nazywać będziemy *miarą martyngałową* dla zdyskontowanego procesu cen S^* , jeśli miara P^* jest równoważna z P oraz S^* jest P^* -martyngałem.³

Twierdzenie 2.3 ([1, 4, 5]). *Na rynku $\mathcal{M} = (B, S, \Phi)$ istnieje miara martyngałowa P^* dla zdyskontowanego procesu cen S^* wtedy i tylko wtedy, gdy równanie*

$$S_0 \cdot (1 + r) = \lambda \cdot S^u + (1 - \lambda) \cdot S^d, \quad (2.7)$$

³Miary nazywamy *równoważnymi*, gdy mają identyczne zbiory miary zero.

ma względem λ dokładnie jedno rozwiązanie oraz $\lambda \in (0, 1)$.

Dowód. Konieczność. W przypadku, gdy P^* jest miarą martyngałową, to $S_0 = E_{P^*} S_T^*$, a ponadto $\lambda = P^*(\{\omega_1\})$ oraz $1 - \lambda = P^*(\{\omega_2\})$ są liczbami z przedziału $(0, 1)$.

Dostateczność. Jeżeli rozwiązanie λ równania (2.7) istnieje i należy do przedziału $(0, 1)$, to miarę probabilistyczną P^* można zadać układem równości

$$\begin{cases} P^*(\{\omega_1\}) = \lambda, \\ P^*(\{\omega_2\}) = 1 - \lambda. \end{cases}$$

Wtedy miara P^* jest równoważna z miarą P i spełnia równość $S_0 = E_{P^*} S_T^*$. Zatem P^* jest miarą martyngałową.

Rozwiązanie równania (2.7) ma postać:

$$\lambda = \frac{(1+r) \cdot S_0 - S^d}{S^u - S^d}, \quad (2.8)$$

i jest ono jedynym jego rozwiązaniem. Formuła opisująca rozwiązanie informuje, że miara martyngałowa P^* jest zdeterminowana poprzez stopę procentową, cenę wyjściowej i możliwe ceny końcowe. \square

Twierdzenie 2.4 ([1]). *Rynek $M = (B, S, \Phi)$ jest wolny od arbitrażu wtedy i tylko wtedy, gdy istnieje miara martyngałowa dla zdyskontowanego procesu cen S^* . Wtedy cena arbitrażowa $\Pi_0(X)$ w chwili $t = 0$ dowolnej wypłaty X dokonanej w chwili T , dana jest formułą $\Pi_0(X) = E_{P^*} \left(\frac{X}{1+r} \right)$, gdzie P^* jest miarą martyngałową.*

Dowód. Prawdziwość pierwszej części twierdzenia wynika z algebraicznej równoważności formuły (2.8) na rozwiązanie λ równania (2.7), przy warunku $\lambda \in (0; 1)$, z postacią formuły (2.5) na braku arbitrażu. Dla dowodu drugiej części twierdzenia obierzmy dowolną wypłatę X oraz odpowiadający jej portfel replikujący $\varphi = (\alpha_0; \beta_0)$. Wartość portfela φ w chwili T jest wówczas równa $V_T(\varphi) = \alpha_0 S_T + \beta_0(1+r)$. Zatem zdyskontowana na moment $t = 0$ wartość oczekiwana wypłaty X w chwili $t = T$, ma postać

$$E_{P^*} \left(\frac{X}{1+r} \right) = E_{P^*} \left(\frac{V_T(\varphi)}{1+r} \right) = E_{P^*} \left(\frac{\alpha_0 S_T + (1+r)\beta_0}{1+r} \right) = E_{P^*} \left(\alpha_0 \frac{S_T}{1+r} + \beta_0 \right).$$

Ponieważ jednak zgodnie z wcześniejszymi ustaleniami jest $\frac{S_T}{1+r} = S_T^*$, to mamy

$$E_{P^*} \left(\frac{X}{1+r} \right) = \alpha_0 \cdot E_{P^*}(S_T^*) + \beta_0,$$

a dodatkowo skoro P^* jest miarą martyngałową i skoro $E_{P^*}(S_T^*) = S_0^*$, to

$$E_{P^*} \left(\frac{X}{1+r} \right) = \alpha_0 \cdot S_0^* + \beta_0 = V_0(\varphi) = \Pi_0(X).$$

□

Jak zauważamy, cena arbitrażowa nie bierze pod uwagę ryzyka, co jest konsekwencją przyjętego modelu wyceny wolnej od ryzyka. Taka neutralność wobec ryzyka nie oznacza jego ignorowania czy nieistnienia, ale stan, w którym scenariusz prowadzący do pewnego zysku oraz scenariusz obciążony ustalonym ryzykiem są nieróżrzóżnialne pod warunkiem, że wartości oczekiwane w obu scenariuszach są identyczne.

Rozpatrzmy rynek, na którym Ω oznacza tradycyjnie zbiór możliwych scenariuszy, oraz $\omega^\circ \in \Omega$ jest wyróżnionym spośród nich. Można wtedy zdefiniować instrument finansowy X zwany instrumentem Arrow'a-Debreu, określony warunkiem

$$X(\omega) = \begin{cases} 1, & \text{gdy } \omega = \omega^\circ, \\ 0, & \text{gdy } \omega \in \Omega \setminus \{\omega^\circ\}. \end{cases}$$

Jest to w istocie „instrument charakterystyczny” scenariusza ω° , lub bardziej precyzyjnie jego indykatorem. Mianowicie, płaci on 1 jednostkę, gdy zachodzi wyróżniony scenariusz, a 0 w pozostałych przypadkach. Gdy X jest *instrumentem osiągalnym*, czyli posiadającym portfel replikujący, to jego cena w chwili $t = 0$ wynosi:

$$E_Q[X/B_1] = \sum_{\omega} Q(\omega) \cdot X(\omega) \div B_1(\omega) = Q(\omega^\circ) \div B_1(\omega^\circ). \quad (2.9)$$

Efekt działania $Q(\omega^\circ) \div B_1(\omega^\circ)$ nazywany jest *ceną stanu* i jest on zdyskontowaną wartością oczekiwaną instrumentu X . Cena dowolnego instrumentu osiągalnego w chwili $t = 0$ jest średnią wypłat z różnych scenariuszy, ważoną cenami stanów.

3 Stochastyczne dyskonto

Modelowanie procesu wartości przyszłych przepływów finansowych, opiera się na pojęciu deflatora, będącego stochastycznym odpowiednikiem deterministycznego czynnika dyskontującego. Podejście takie, zwane *metodą deflatorową*, jako cel główny stawia sobie szacowanie wartości przepływów, bazujące na metodach martyngałowych.

3.1 Podstawowy model z czasem dyskretnym

Niech ustalone $n \in \mathbb{N}_+$ oznacza horyzont czasowy, zaś (Ω, \mathcal{F}, P) niech będzie przestrzenią probabilistyczną. Poprzednio ustaliliśmy, że każdy skończony i wstępujący ciąg σ -ciał $\underline{\mathcal{F}} = (\mathcal{F}_t)_{t \in \{0, 1, \dots, n\}}$, tej przestrzeni, spełniający warunek $\emptyset = \mathcal{F}_0 \subset \mathcal{F}_1 \subset \mathcal{F}_2 \subset \dots \mathcal{F}_n = \mathcal{F}$ nazywamy *filtracją przestrzeni probabilistycznej* (Ω, \mathcal{F}, P) . Od teraz czwórkę $(\Omega, \mathcal{F}, P, \underline{\mathcal{F}})$ – nazywać będziemy *filtrowaną przestrzenią probabilistyczną z filtracją* $\underline{\mathcal{F}}$.

Przypomnijmy, że intencją wprowadzenia ciągu \mathcal{F} jest rozumienie σ_t -ciał będących jego wyrazami jako informacji posiadanej w chwili $t \in \{0, 1, \dots, n\}$. Informacje znane w chwili t są ważne tak z punktu widzenia modelowania matematycznego, jak i z punktu widzenia rynku finansowego. Składają się na nie między innymi informacje prawne, finansowe, polityczne, meteorologiczne, a także wiele innych, znanych w momencie t . Dodatkowo niech $\underline{X} = (X_0, X_1, \dots, X_n)$, oznacza ciąg $\underline{\mathcal{F}}$ -adaptowanych zmiennych losowych na przestrzeni $(\Omega, \mathcal{F}, P, \underline{\mathcal{F}})$, a więc takich, że X_t jest \mathcal{F}_t -mierzalna. Z punktu widzenia matematycznego modelu rynku, \underline{X} jest ciągiem zmiennych losowych opisujących przepływy pieniężne, przy czym X_t oznacza płatność w chwili t . Dysponując informacją \mathcal{F}_t , możemy ustalić wszystkie wartości X_s , dla $s \leq t$.

Mamy dwa równoległe cele: jeden to wyznaczenie wartości przyszłych przepływów pieniężnych X_s , dla $s > t$, w oparciu o informacje \mathcal{F}_t dostępne w chwili t , a drugi to konieczność posiadania wiedzy na temat wartości tych przepływów w dowolnej chwili $t \in \{0, 1, \dots, n\}$.

Dla ciągu \underline{X} przyjmujemy, że zmienna losowa X_t , dla dowolnego $t \in \{0, 1, \dots, n\}$, ma skończony drugi moment, czyli jest całkowalna z kwadratem, a więc

$$\underline{X} = (X_0, X_1, \dots, X_n) \in L_{n+1}^2(P),$$

gdzie $L_{n+1}^2(P)$ jest przestrzenią Hilberta określoną w następujący sposób:

- (i) $\forall \underline{X} \in L_{n+1}^2(P) \quad E \left[\sum_{t=0}^n X_t^2 \right] < \infty,$
- (ii) $\forall \underline{X} \in L_{n+1}^2(P) \forall \underline{Y} \in L_{n+1}^2(P) \quad \langle \underline{X}, \underline{Y} \rangle = E \left[\sum_{t=0}^n X_t \cdot Y_t \right],$
- (iii) $\forall \underline{X} \in L_{n+1}^2(P) \quad \|\underline{X}\| = \langle \underline{X}, \underline{X} \rangle^{1/2}.$

Fakt, że \underline{X} jest \mathcal{F} -adaptowany i całkowalny z kwadratem, wyrażamy pisząc

$$\underline{X} = (X_0, X_1, \dots, X_n) \in L_{n+1}^2(P, \mathcal{F}).$$

Jako praktyczny przykład przedstawionych faktów weźmy pojedynczą polisę, generującą w czasie jej życia przepływy pieniężne. Są wśród nich systematyczne, okresowe wpłaty w wysokości Π_t , a z drugiej ponoszone są pewne, konieczne wydatki, w dowolnych chwilach czasu na przedziale czasu $(t-1, t]$. Jeżeli wszystkie przepływy jakie mają miejsce w tym przedziale przyporządkujemy chwili t , to określimy w ten sposób dyskretną zmienną losową X_t . Zakładając, że czas trwania polisy został podzielony na n podprzedziałów $(t-1, t]$, dostajemy wraz z wpłatą początkową „uruchamiającą” polisę ciąg \underline{X} zmiennych losowych modelujących tę polisę w całym okresie jej trwania.

4 Przypadek prostego modelu z czasem dyskretnym

Zajmiemy się teraz wyceną stochastycznego przepływu pieniężnego $\underline{X} = (X_0, X_1, \dots, X_n)$, a wykorzystamy do tego celu pewną dodatnią oraz ciągłą funkcję liniową. Na wstępie przedstawimy niezbędne umowy związane przepływem pieniężnym. Otóż przepływ pieniężny \underline{X} nazwiemy

- (i) *nieujemnym* i oznaczmy: $\underline{X} \geq 0$ wtt, gdy $\forall_{k=0,1,\dots,n} X_k \geq 0$;
- (ii) *dodatnim* i piszemy: $\underline{X} > 0$ wtt, gdy $\forall_{k=0,1,\dots,n} X_k \geq 0 \wedge \exists_{k \in \{0,1,\dots,n\}} X_k > 0$ z dodatnim prawdopodobieństwem;
- (iii) *silnie dodatnim* i napiszemy: $\underline{X} \gg 0$ wtt, gdy $\forall_{k=0,1,\dots,n} X_k > 0$.

Niech teraz $Q : L_{n+1}^2(P) \rightarrow \mathbb{R}$ będzie dodatnim i ciągłym funkcjonałem liniowym na $L_{n+1}^2(P)$. Z formalnego punktu widzenia Q spełnia następujące własności

- (1) *dodatniość*, czyli $\forall_{\underline{X} \in L_{n+1}^2(P)} (\underline{X} > 0 \Rightarrow Q[\underline{X}] > 0)$
- (2) *ciągłość* — $\forall_{\underline{X}^{(j)} \in L_{n+1}^2(P)} \left\{ \left(\underline{X}^{(j)} \right)_{j \rightarrow \infty} \underline{X} \right\} \implies \lim_{k \rightarrow \infty} Q[\underline{X}^{(j)}] = Q[\underline{X}]$
- (3) *liniowość* — $\forall_{\underline{X} \in L_{n+1}^2(P)} \forall_{\underline{Y} \in L_{n+1}^2(P)} \forall_{a,b \in \mathbb{R}} \left\{ Q[a \cdot \underline{X} + b \cdot \underline{Y}] = a \cdot Q[\underline{X}] + b \cdot Q[\underline{Y}] \right\}$.

Odwzorowanie $\underline{X} \mapsto Q[\underline{X}]$, przypisuje przepływowi pieniężnemu \underline{X} , jego wartość monetarną $Q[\underline{X}] \in \mathbb{R}$ w czasie 0. Innymi słowy funkcja wyceny Q informuje o wartości dowolnego ciągu przepływów pieniężnych $\underline{X} \in L_{n+1}^2(P)$, rozumianej jako cena \underline{X} w chwili $t = 0$.

Przyjmijmy $\underline{Y}^{(j)} := \underline{X}^{(j)} - \underline{X}$, wtedy $Q[\underline{Y}^{(j)}] = Q[\underline{X}^{(j)} - \underline{X}]$. Liniowość Q gwarantuje, że $Q[\underline{Y}^{(j)}] = Q[\underline{X}^{(j)}] - Q[\underline{X}]$.

Przypuśćmy, że $\underline{Y}^{(j)}$ jest dodatni, tzn. $\underline{Y}^{(j)} \geq 0$. Wtedy ciąg $\underline{X}^{(j)}$ dąży do \underline{X} przez wartości nieujemne. Gdyby operator był ciągły, to zachodziłby warunek

$$\left(\underline{Y}^{(j)} \right)_{j \rightarrow \infty} 0 \implies Q[\underline{Y}^{(j)}] \xrightarrow{j \rightarrow \infty} 0. \quad (4.1)$$

Przypuśćmy na chwilę, że (4.1) nie jest prawdą. Dodatniość funkcjonału Q , gwarantuje istnienie $\varepsilon > 0$ i nieskończonego podciągu j^l ciągu j , że dla każdego wyrazu z j^l zachodzi:

$$Q[\underline{Y}^{(k^l)}] \geq \varepsilon. \quad (4.2)$$

Niech $\underline{Y}^{(j^m)}$, będzie podciągiem podciągu $\underline{Y}^{(j^l)}$, dla którego $\sum_{m=1}^{\infty} \|\underline{Y}^{(j^m)}\| < \infty$.

Oznaczmy przez \underline{Y} nieskończoną sumę wszystkich wyrazów ciągu $\underline{Y}^{(j^m)}$, czyli $\underline{Y} = \sum_{m=1}^{\infty} \underline{Y}^{(j^m)}$. Z zupełności przestrzeni $L_{n+1}^2(P)$ wynika, że $\underline{Y} \in L_{n+1}^2(P)$. Równocześnie zachodzą zależności $Q[\underline{Y}] = Q[\sum_{l=1}^{\infty} \underline{Y}^{(j^l)}] \geq Q[\sum_{m=1}^{n_0} \underline{Y}^{(j^m)}]$. Ponieważ dla każdego wyrazu k^l istnieje $\varepsilon > 0$, że zachodzi (4.2), to również jest prawdą, że:

$$\forall \underline{j}^m \exists \varepsilon > 0 : Q[\sum_{j^m \leq n_0} \underline{Y}^{(j^m)}] \geq n_0 \cdot \varepsilon,$$

dla każdego $n_0 \in \mathbb{N}$. Oznacza to, że $Q[\underline{Y}] = \infty$, co przeczy założonej dodatniości \underline{Y} oraz liniowości operatora Q . Zatem jeśli ciąg $\underline{Y}^{(j)}$ dąży do 0 przez wartości dodatnie, to i również ciąg $Q[\underline{Y}^{(j)}]$ spełnia ten warunek.

Niech teraz ciąg \underline{Y} będzie dowolny, niekoniecznie dodatni. Wtedy dokonując dekompozycji ciągu \underline{Y} na część dodatnią \underline{Y}_+ oraz ujemną \underline{Y}_- mamy z jednej strony naturalną zależność $\underline{Y}^{(j)} = \underline{Y}_+^{(j)} + \underline{Y}_-^{(j)}$, a z drugiej, z własności normy dostajemy następujące nierówności

$$\left(\|\underline{Y}_+^{(j)}\| \leq \|\underline{Y}^{(j)}\| \wedge \|\underline{Y}^{(j)}\| \rightarrow 0 \right) \wedge \left(\|\underline{Y}_-^{(j)}\| \leq \|\underline{Y}^{(j)}\| \wedge \|\underline{Y}^{(j)}\| \rightarrow 0 \right).$$

Zatem otrzymujemy, że $Q[\underline{Y}_+^{(j)}] \rightarrow 0$ i $Q[\underline{Y}_-^{(j)}] \rightarrow 0$. Ponownie korzystając z liniowości Q otrzymujemy:

$$\begin{aligned} \left\{ \left(Q[\underline{Y}_+^{(j)}] \xrightarrow{j \rightarrow \infty} 0 \wedge Q[\underline{Y}_-^{(j)}] \xrightarrow{j \rightarrow \infty} 0 \right) \Rightarrow \left[\left(Q[\underline{Y}_+^{(j)}] + Q[\underline{Y}_-^{(j)}] \right) \xrightarrow{j \rightarrow \infty} 0 \right] \right\} \\ \Leftrightarrow \left\{ Q[\underline{Y}^{(j)}] \xrightarrow{j \rightarrow \infty} 0 \right\}. \end{aligned}$$

Otrzymaliśmy zatem następujący wynik.

Twierdzenie 4.1 ([5]). *Założenie dodatniości i liniowości $Q[\underline{X}]$ implikuje ciągłość $Q[\underline{X}]$.*

Twierdzenie 4.2 ([3], Riesz-Frécheta o reprezentacji funkcjonału). *Jeśli \mathcal{H} jest przestrzenią Hilberta z iloczynem skalarnym $\langle \cdot | \cdot \rangle$, zaś \mathcal{H}^* – przestrzenią sprzężoną do \mathcal{H} , to dla każdego funkcjonału liniowego $y^* \in \mathbb{H}^*$, istnieje dokładnie jeden element $y \in \mathcal{H}$, spełniający dla wszystkich $x \in \mathcal{H}$ równość: $y^*(x) = \langle x | y \rangle$.*

W odniesieniu do ważnej dla nas przestrzeni $L_{n+1}^2(P)$ twierdzenie Riesz-Frécheta zapiszemy w postaci następującego wniosku.

Wniosek 4.1. ([5]) Dla dodatniego i ciągłego operatora $Q : L_{n+1}^2(P) \rightarrow \mathbb{R}$ zachodzi warunek

$$\exists \underline{D} \in L_{n+1}^2(P) \quad \forall \underline{X} \in L_{n+1}^2(P) \quad \left\{ Q[\underline{X}] = \langle \underline{X} | \underline{D} \rangle = E \left[\sum_{t=0}^n X_t \cdot D_t \right] \right\}.$$

Ponieważ praktyczne zastosowanie ciągów zmiennych losowych $\underline{X} \in L_{n+1}^2(P)$ zostało zasygnalizowane poprzednio, to w połączeniu z twierdzeniem (4.2) i następującym po nim wnioskiem, możemy teraz zdefiniować pojęcie deflatora.

Ciąg (wektor) $\underline{D} \in L_{n+1}^2(P)$, którego istnienie gwarantuje twierdzenie (4.2) nazywamy *deflatorem* [2, 5].

Uwaga 4.1. Ponieważ n występujące w indeksie przestrzeni $L_{n+1}^2(P)$ może być dowolną liczbą naturalną, to dla $n = 0$, ciąg \underline{D} składa się z jednego tylko wyrazu D_0 . Tak określony ciąg \underline{D} jest deflatorem, dlatego dla dowolnego $n \in \mathbb{N}$ każdy pojedynczy wyraz D_t ciągu (wektora) \underline{D} jest również nazywany deflatorem.

Bezpośrednio z definicji deflatora otrzymujemy wniosek.

Wniosek 4.2. Deflator $D \in L_{n+1}^2(P)$ ma następujące własności

W1. Z dodatniości funkcjonału Q wynika, że $D \gg 0$.

W2. Jeśli zmienna $\underline{X} \in L_{n+1}^2(P, \mathcal{F})$ jest \mathcal{F} -adaptowana, to istnieje \mathcal{F} -adaptowany deflator \tilde{D} .

W3. Nie istnieją dwa ortogonalne i \mathbb{F} -adaptowane wektory $\underline{D}, \underline{D}^* \in L_{n+1}^2(P)$.

Dowód. Zauważamy, że \mathcal{F} -adaptowalność deflatora oznacza w konsekwencji dla deflatora składowego \tilde{D}_t jego \mathcal{F}_t -mierzalność dla każdego $t \in \{0, 1, \dots, n\}$. Zachodzi zatem równość $\tilde{D}_t = E[D_t | \mathcal{F}_t]$. Wtedy dla każdego \mathcal{F} -adaptowanego, całkowitego z kwadratem przepływu pieniężnego \underline{X} zachodzi

$$Q[\underline{X}] = E \left[\sum_{t=0}^n X_t \cdot D_t \right] = \sum_{t=0}^n E[X_t \cdot D_t] = \sum_{t=0}^n E[E[X_t \cdot D_t | \mathcal{F}_t]] =$$

$$= \sum_{t=0}^n E [X_t \cdot E [D_t | \mathcal{F}_t]] = E \left[\sum_{t=0}^n X_t \cdot E [D_t | \mathcal{F}_t] \right] = E \left[\sum_{t=0}^n X_t \cdot \tilde{D}_t \right].$$

Stąd najczęściej będziemy zakładać, że D_t jest \mathcal{F}_t -adaptowany.

Założmy, że nie jest to prawdziwa własność W3. Zatem istnieją dwa liniowo niezależne \mathcal{F} -adaptowane ciągi zmiennych losowych \underline{D} oraz \underline{D}^* , że

$$\forall \underline{X} \in L_{n+1}^2(P) \quad Q[\underline{X}] = \langle \underline{X}, \underline{D} \rangle = \langle \underline{X}, \underline{D}^* \rangle$$

Wtedy \underline{D} oraz \underline{D}^* muszą spełniać zależności

$$\begin{cases} Q[\underline{D}] = \|\underline{D}\|^2 = \langle \underline{D}, \underline{D}^* \rangle, \\ Q[\underline{D}^*] = \langle \underline{D}^*, \underline{D} \rangle = \|\underline{D}^*\|^2. \end{cases}$$

Wynika z tego, że

$$\langle \underline{D}^*, \underline{D} \rangle = \|\underline{D}^*\| \cdot \|\underline{D}\|. \quad (4.3)$$

Z nierówności Schwarz'a $\langle u, v \rangle \leq \|u\| \cdot \|v\|$ mamy, że równość zachodzi wtedy i tylko wtedy, gdy wektory u i v są liniowo zależne, a więc istnieje dla nich $\lambda \in \mathbb{C}$, że zachodzą równości:

$$u = \lambda \cdot v \text{ lub } v = \lambda \cdot u.$$

W konsekwencji, równość (4.3) zachodzi jedynie dla \underline{D} oraz \underline{D}^* – równych z dokładnością do stałej, a więc wyklucza założoną ortogonalność wektorów \underline{D} i \underline{D}^* . \square

Poza przedstawionymi własnościami deflatora, przyjmujemy umowę, że $D_0 \equiv 1$ – tożsamościowo. Oznacza to, że dla deterministycznej zmiennej, będącej płatnością w kwocie X_0 – w chwili $t = 0$, funkcjonał liniowy Q przyporządkowuje tej płatności jej wartość nominalną, a więc $Q[(X_0, 0, 0, \dots, 0)] = X_0$.

Uwaga 4.2. W praktyce występuje zagadnienie zwane *zadaniem modelowania deflatorów*. Jego istota sprowadza się do znalezienia takiego funkcjonału Q lub równoważnie, znalezienia \mathcal{F} -adaptowanego deflatora \underline{D} , o tej własności, że generuje on wartości pokrywające się z wielkościami rzeczywistymi. \mathcal{F}_t -adaptowalność

oznacza, że deflator D_t , będący stochastycznym czynnikiem dyskontującym, jest znany w chwili t i przez to można znaleźć bezpośredni związek pomiędzy zachowaniem D_t rynku finansowego, a \mathcal{F}_t -adaptowanym przepływem pieniężnym X_t w chwili t .

5 Istota działania deflatorów

Można opisowo powiedzieć, że deflator D_t odpowiada za transfer do chwili 0 wartości przepływu pieniężnego dokonanego w chwili t . Możemy go zatem traktować jako stochastyczny transport wartości, czyli stochastyczne dyskonto. Z tego powodu przepływ pieniężny \underline{X} charakteryzowany jest przez ciąg (wektor) zmiennych losowych postaci $\underline{X}_t = (0, \dots, 0, X_t, 0, \dots, 0)$. Zmienna losowa X_t nie musi być konieczne niezależna ze zmienną losową D_t , a to oznacza, a więc może zachodzić warunek

$$Q[\underline{X}] = E[X_t \cdot D_t] \neq E[X_t] \cdot E[D_t]. \quad (5.1)$$

Dokonajmy dekompozycji deflatora D_t , na *deflatory segmentowe* odpowiadające za transport wartości z chwili t do chwili $t - 1$, które oznaczamy przez Y_t . Z faktu, że $D_t > 0$ wynika, że deflatory segmentowe można wyrazić poprzez iloraz deflatora D_t oraz jego poprzednika D_{t-1} , dla każdego $t \in \{1, \dots, n\}$. Zatem

$$Y_t = \frac{D_t}{D_{t-1}} \quad \text{dla } t \in \{1, \dots, n\}. \quad (5.2)$$

Stąd otrzymujemy naturalny i równocześnie ważny ciąg zależności

$$D_t = Y_t \cdot D_{t-1} = Y_t \cdot Y_{t-1} \cdot D_{t-2} = Y_t \cdot Y_{t-1} \cdot \dots \cdot Y_2 \cdot Y_1 = \prod_{i=1}^t Y_i.$$

Podkreślmy przy tym, że deflatory segmentowe Y_t , dla $t \geq 1$, w sposób stochastyczny dyskontują wartość przepływu pieniężnego dokonanego z chwili t do chwili $t - 1$. Naturalnym wydaje się w tej sytuacji pytanie o związek między deflatorem \underline{D} , a zerokuponowymi obligacjami i klasycznym dyskontem finansowym.

Oznaczmy przez $U^{(t)} = (0, \dots, 0, 1, 0, \dots, 0)$ pojedynczy przepływ pieniężny, dotyczący obligacji zerokuponowej płaćcej jedną jednostkę w chwili t . Wartość obligacji w chwili 0, określa formuła

$$P_{0,t} = Q[U^{(t)}] = E[D_t].$$

Oczywiście funkcja $P_{0,t}$ także transferuje wartość przepływu pieniężnego z chwili t na chwilę 0. Różnica polega jednak na tym, że $P_{0,t}$ jest \mathcal{F}_0 -mierzalna, podczas gdy D_t jest \mathcal{F}_t -mierzalną zmienną losową. Z tego powodu deterministyczny czynnik dyskontujący $P_{0,t}$ znany jest jedynie na początku okresu, podczas gdy D_t jest znane pod koniec okresu $(0, t]$. Jeżeli rozpatrujemy deterministyczne przepływy pieniężne \underline{X} , to możemy przy ich wycenie posługiwać się zarówno obligacjami zerokuponowymi jak i deflatorami, a w obu przypadkach uzyskamy ten sam wynik. Sytuacja ulega radykalnej zmianie, gdy dyskontowanie dotyczy przepływów pieniężnych mających losową naturę. W takim przypadku koniecznością jest odwołanie się do deflatorów, a wynika to z możliwości silnego skorelowania zmiennych D_t oraz X_t z tym samym czynnikiem. Konieczność tę wyraża warunek (5.1). Przykładem takiego scenariusza może być opcja X_t , zależna od faktycznej realizacji D_t .

Tradycyjne, podejście aktuarialne do dyskonta, sprowadza się do wyznaczenia stałej stopy zwrotu r , dla całego okresu $(0, T]$. Wtedy przepływ pieniężny o wartości 1 w chwili t , ma w chwili 0 wartość wynoszącą

$$v_t = (1 + r)^{-t}.$$

Modele oparte o zdefiniowany w ten sposób czynnik dyskontujący v_t , modelują procesy finansowe niezgodne z ich rzeczywistymi rezultatami. Ta właśnie ułomność jest powodem poszukiwania modelu dyskontowania, uwzględniającego czynniki niedeterministyczne.

Przykład 5.1. Oprzemy się tu na przypadku prezentowanym w [5]. Należy na wstępie zaznaczyć, że przed wykorzystaniem omawianych narzędzi, niezbędnym jest posiadanie dobrze skalibrowanego modelu rynku. Przeanalizujemy model rynku jednookresowego dwustanowego. Przestrzenią zdarzeń niech tradycyjnie będzie

$\Omega = \{\omega_1, \omega_2\}$. Znalezienie wartości deflatorów, sprowadza się wtedy do rozwiązania pewnego układu warunków. Przypuśćmy, że mamy dwa aktywa A_1 oraz A_2 , o cenach rynkowych odpowiednio Q_{A_1} i Q_{A_2} oraz wypłatach S_{A_1} S_{A_2} . Mamy przy tym następujące informacje: $Q(A_1) = 1.65$, $S_{A_1}^u = 3$, $S_{A_2}^d = 1$ oraz $Q(A_2) = 1.65$, $S_{A_2}^u = 2$, $S_{A_2}^d = 0.5$

Dla wyznaczenia ceny stanu, dla poszczególnych scenariuszy ω_1, ω_2 , rozwiążemy układ równań

$$\begin{aligned} 3 \cdot \psi(\omega_1) + 1 \cdot \psi(\omega_2) &= 1.65, \\ 2 \cdot \psi(\omega_1) + 0.5 \cdot \psi(\omega_2) &= 1. \end{aligned}$$

Stąd już bezpośrednio otrzymujemy ceny stanu: $\psi(\omega_1) = 0.35$, $\psi(\omega_2) = 0.60$.

Dołączmy do już posiadanych jeszcze jedno aktywo ryzykowne A_3 , płacące $S_{A_3}^u = 2$ – czyli w stanie ω_1 oraz $S_{A_3}^d = 1$ – czyli w stanie ω_2 . Jego cenę wyznaczmy z następującego równania:

$$Q[C] = 2 \cdot \psi(\omega_1) + 1 \cdot \psi(\omega_2) = 2 \cdot 0.35 + 1 \cdot 0.6 = 1.3.$$

Niech $U^{(1)}$ będzie obligacją zerokuponową. Instrument ten płaci 1 w każdym z obu scenariuszy, stąd jego wartość w chwili 0 jest równa

$$Q[U^{(1)}] = 1 \cdot 0.35 + 1 \cdot 0.6 = 0.95. \quad (5.3)$$

Informacja ta pozwala wyznaczyć stopę zwrotu z lokaty jako $(0.95)^{-1} - 1 \approx 0.526$. Niech $\psi(\omega_i)$ będzie ceną stanu ω_i w chwili 0. Symbol $X_1(\omega_i)$ oznaczał będzie wypłatę z nowego ryzykownego aktywa \underline{X} w czasie 1, w przypadku wystąpienia scenariusza ω_i . W chwili 0 cenę rynkową aktywa \underline{X} opisuje zależność

$$Q[\underline{X}] = \sum_{i=1}^2 \psi(\omega_i) \cdot X_1(\omega_i). \quad (5.4)$$

Przyjmijmy, że w chwili 1 wystąpienie scenariusza ω_1 następuje z prawdopodobieństwem $p(\omega_1) \in (0, 1)$, a scenariusza ω_2 z prawdopodobieństwem $p(\omega_2) = 1 - p(\omega_1)$ w stanie ω_2 . Zatem zależności (5.4) można nadać teraz postać

$$Q[\underline{X}] = \sum_{i=1}^2 \psi(\omega_i) \cdot X_1(\omega_i) = \sum_{i=1}^2 \frac{p(\omega_i)}{p(\omega_i)} \cdot \psi(\omega_i) \cdot X_i(\omega_i) =$$

$$= E \left[\sum_{i=1}^2 \left(\frac{\psi(\omega_i)}{p(\omega_i)} \right) \cdot (p(\omega_i) \cdot X_i(\omega_i)) \right] = E \left[\left(\frac{\Psi}{p} \right) \cdot \tilde{X} \right], \quad (5.5)$$

gdzie

$$\left(\frac{\Psi}{p} \right) = \left(\frac{\psi(\omega_1)}{p(\omega_1)}, \frac{\psi(\omega_2)}{p(\omega_2)} \right) \quad \text{oraz} \quad \tilde{X} = (p(\omega_2) \cdot X_2(\omega_2), p(\omega_2) \cdot X_2(\omega_2))$$

Określiśmy w ten sposób ciąg zmiennych losowych \tilde{D} , którego wyrazy są ilorazami ceny stanu i prawdopodobieństwa wystąpienia tego stanu oraz ciąg \tilde{X} jako zmodyfikowany wektor cen, uwzględniający prawdopodobieństwo wystąpienia ceny i mamy

$$\tilde{D} := \frac{\Psi}{p} \quad \text{oraz} \quad \tilde{X} := p \cdot X.$$

Wykorzystując zdefiniowane zmienne losowe \tilde{D} oraz \tilde{X} równanie (5.5) przybiera postać

$$Q[X] = E \left[\tilde{D} \cdot \tilde{X} \right]. \quad (5.6)$$

Dla szczególnego doboru prawdopodobieństw $p(\omega_i)$, deflator \tilde{D} przyjmuje w chwili 1 wartości: dla stanu ω_1 przy $\psi(\omega_1) = 0.35$ mamy $P(\omega_1) = 0.5$ oraz $\tilde{D}_1 = \tilde{D}(\omega_1) = 0.7$ natomiast dla stanu ω_2 przy $\psi(\omega_2) = 0.65$ mamy $P(\omega_2) = 0.5$ oraz $\tilde{D}_2 = \tilde{D}(\omega_2) = 1.2$.

Uzyskane wyniki pozwalają w alternatywny do (5.3) sposób wyznaczyć wartość obligacji zerokuponowej.

$$Q[U^{(1)}] = E[\tilde{D}] = \sum_{i=1}^2 \tilde{D}(\omega_i) \cdot p(\omega_i) = 0.7 \cdot 0.5 + 1.2 \cdot 0.5 = 0.95.$$

6 Wycena poza chwilą 0

Rozpatrzmy \mathcal{F} -adaptowany deflator $D \in L_{n+1}^2(P, \mathcal{F})$. Dla losowego wektora $X \in L_{n+1}^2(P, \mathcal{F})$ proces cen definiujemy układ równości

$$Q_t[X] = Q[X | \mathcal{F}_t] = \frac{1}{D_t} \cdot E \left[\sum_{k=0}^n D_k \cdot X_k \middle| \mathcal{F}_t \right], \quad \text{dla } t \in \{0, 1, \dots, n\}. \quad (6.1)$$

Silna dodatniość deflatora ($D \gg 0$) zapewnia, że $Q_t[\underline{X}]$ jest dobrze określone. Zgodnie z definicją operatora Q , funkcja $\underline{X} \mapsto Q_t[\underline{X}]$ przyporządkowuje w chwili t przepływowi pieniężnemu \underline{X} wartość $Q_t[\underline{X}]$, wyrażoną w jednostkach monetarnych (\mathcal{F}_t -mierzalną). Z tego powodu wartość ta jest wielkością losową widzianą z punktu $t = 0$, uzależnioną od \mathcal{F}_t . Przypuśćmy, że aktywo \underline{X} ma w czasie t cenę $Q_t[\underline{X}]$. W konsekwencji generowany jest ciąg płatności:

$$Q_t[\underline{X}] \cdot U^{(t)} = (0, \dots, 0, Q_t[\underline{X}], 0, \dots, 0).$$

Z dzisiejszego punktu widzenia, ciąg płatności ma wartość

$$Q_0 [Q_t[\underline{X}] \cdot Z^{(t)}].$$

Obecnie dysponujemy jedynie informacją \mathcal{F}_0 w momencie 0 o cenach $Q_t[\underline{X}]$ aktywa \underline{X} w czasie t . Zachodzi więc równość

$$Q_0[\underline{X}] = Q_0 [Q_t[\underline{X}] \cdot U^{(t)}]. \quad (6.2)$$

Mamy alternatywę: albo kupić \underline{X} dziś i zapłacić za niego obecną cenę, albo też kupić \underline{X} w chwili t , a więc po cenie $Q_t[\underline{X}]$ obowiązującej w chwili t . Ze względu na jednakową ilość informacji \mathcal{F}_0 o obu kontraktach, a także z powodu tego, że otrzymujemy to samo aktywo \underline{X} – kontrakty te powinny mieć jednakową cenę.

Przeanalizujmy alternatywne postępowanie i przypuśćmy, że decydujemy się kupić przepływ pieniężny \underline{X} wtedy i tylko wtedy, gdy wystąpi pewne zdarzenie $C_t \in \mathcal{F}$. Aktualnie nie wiemy czy zdarzenie C_t zajdzie, czy może nie, dlatego należy zmodyfikować formułę wyceny (6.2) i nadać jej postać

$$Q_0[\underline{X} \cdot \mathbb{1}_{C_t}] = Q_0 [Q_t[\underline{X}] \cdot U^{(t)} \cdot \mathbb{1}_{C_t}]. \quad (6.3)$$

Posługując się deflatorami możemy ją przekształcić do postaci

$$E \left[\sum_{i=0}^n D_i \cdot X_i \cdot \mathbb{1}_{C_t} \right] = E [D_t \cdot Q_t[\underline{X}] \cdot \mathbb{1}_{C_t}]. \quad (6.4)$$

Ponieważ $(D_t \cdot Q_t[\underline{X}])$ jest \mathcal{F}_t -mierzalna, to równanie (6.4) jest prawdziwe dla każdego $C_t \in \mathcal{F}_t$. W efekcie otrzymaliśmy warunek definicyjny warunkowej wartości oczekiwanej, przy zadanym σ -ciele \mathcal{F}_t . Posługując się teraz własnościami warunkowej wartości oczekiwanej, przekształcimy obie strony równania (6.4) wskutek czego dostajemy

$$E \left[\sum_{i=0}^n D_i \cdot X_k \middle| \mathcal{F}_t \right] \cdot P(F_t) = E \left[D_t \cdot Q_t[\underline{X}] \middle| \mathcal{F}_t \right] \cdot P(F_t),$$

i dalej

$$E \left[\sum_{i=0}^n D_i \cdot X_i \middle| \mathcal{F}_t \right] = E \left[D_t \cdot Q_t[\underline{X}] \middle| \mathcal{F}_t \right].$$

Zmienna losowa $(D_t \cdot Q_t[\underline{X}])$ jest \mathcal{F}_t -mierzalna, stąd dla każdego $\mathcal{F}_t \in \mathbb{F}$ zachodzi równość

$$E \left[\sum_{i=0}^n D_i \cdot X_i \middle| \mathcal{F}_t \right] = E \left[D_t \cdot Q_t[\underline{X}] \middle| \mathcal{F}_t \right] = D_t \cdot Q_t[\underline{X}].$$

W efekcie otrzymujemy, że

$$\frac{1}{D_t} \cdot E \left[\sum_{k=0}^n D_k \cdot X_k \middle| \mathcal{F}_t \right] = Q_t[\underline{X}]. \quad (6.5)$$

Równania (6.5) i (6.1) są równoważne, co uzasadnia ekonomiczne znaczenie zależności (6.1). Dotychczas został zdefiniowany klasyczny czynnik dyskontujący $DF(0, m)$ na przedziale czasu $[0; m]$

$$DF(0, m) = Q_0[U^{(m)}] = E[D_m],$$

jako ustalony w chwili 0, dla zerokuponowej obligacji o terminie zapadalności m . Dla $t < m$ symbol $DF(t, m)$ oznacza czynnik dyskontujący na chwilę t , ustaloną wartość wycenioną w gotówce w chwili m . Wobec tego dla każdego $t \in [0; m]$ zachodzi równość

$$DF(0, t) \cdot DF(t, m) = DF(0, m),$$

nazywana *warunkiem zgodności rynku*.

Lewa strona ostatniego równania wyraża cenę na chwilę 0 zakupu $DF(t, m)$, dokonanego w momencie t , zaś $DF(t, m)$ jest ustaloną w chwili 0 ceną zakupu jednej jednostki, dokonanego w chwili t i otrzymanej w chwili m . Natomiast prawa strona równości jest ceną, którą płacimy w chwili 0, za jedną jednostkę otrzymaną w momencie m . Wynika stąd, że $D(k, m)$ jako czynnik dyskontujący aktywo, o wartości w chwili m równej jednej jednostce monetarnej, będzie miał w momencie k wartość określoną formułą

$$DF(k, m) = \frac{DF(0, m)}{DF(0, k)}.$$

Jest to cena zerokuponowej obligacji z terminem zapadalności m , ustalona w chwili 0 i płacona w momencie k .

Jednocześnie ten sam walor można wycenić w oparciu o technikę deflatorową. Mamy mianowicie zależność

$$Q_k [U^{(m)}] = \frac{1}{D_k} \cdot E [D_m | \mathcal{F}_k] = E \left[\frac{D_m}{D_k} | \mathcal{F}_k \right].$$

Jest to w istocie 6.1 dla pojedynczej deterministycznej płatności.

7 Deflatory w szacowaniu rezerw

Rozpatrzmy kontrakt ubezpieczeniowy, reprezentowany przez ciąg stochastycznych przepływów pieniężnych \underline{X} . Wtedy dla dowolnego $i \leq n$ mamy

$$\underline{X}^{(i)} = (0, \dots, 0, X_i, X_{i+1}, \dots, X_n),$$

przy czym $\underline{X}^{(i)}$ jest rozumiany jako wektor losowy, płatności pozostałych po czasie $i - 1$. Stąd rezerwy, tzn. środki które należy zgromadzić do chwili i , aby móc wywiązać się ze zobowiązań wynikających z kontraktu ubezpieczeniowego dla $t \leq i - 1$ opisuje formuła

$$R_i^{(t)} = R [\underline{X}_{(i)} | \mathcal{F}] = Q_t [\underline{X}_{(i)}] = \frac{1}{D_t} \cdot E \left[\sum_{j=i}^n D_j \cdot X_j | \mathcal{F}_t \right]. \quad (7.1)$$

Wartość $R_i^{(t)}$ odpowiada z jednej strony warunkowej wartości oczekiwanej przepływu pieniężnego $\underline{X}^{(i)}$, wycenianej w chwili t . Z drugiej zaś, przy pomocy $R_i^{(t)}$ można przewidzieć, w rozumieniu „wycenić”, przyszłą wartość zmiennej losowej $\underline{X}^{(i)}$. Stąd $R_i^{(t)}$ jest często nazywane *rezerwą najlepszego oszacowania*.

Przeanalizujmy teraz przypadek dwóch firm ubezpieczeniowych, mających identyczny portfel zobowiązań, ale stosujących różne strategie zarządzania rezerwami, a mianowicie niech

- (i) firma U_1 utrzymuje kontrakt, aż do momentu realizacji ostatniej płatności
- (ii) firma U_2 decyduje się w chwili $i - 1$ odsprzedać przyszłe zobowiązania (następujące zgodnie z planem w chwilach i, \dots, n), pod warunkiem zajścia pewnego zdarzenia $C_{i-1} \in \mathcal{F}_{i-1}$.

W efekcie, każda ze strategii będzie generować różne dla firm U_1 i U_2 przepływy pieniężne, odpowiednio $\underline{X}^{(U_1)}$ i $\underline{X}^{(U_2)}$. Naturalnie $\underline{X}^{(U_1)}$ oraz $\underline{X}^{(U_2)}$ są wektorami.

Różnica wartości obu tych strategii jest równa

$$\begin{aligned}
 Q[\underline{X}^{(U_1)} - \underline{X}^{(U_2)}] &= Q[\underline{X}^{(U_1)}] - Q[\underline{X}^{(U_2)}] \\
 &= E\left[\sum_{j=0}^n D_j \cdot X_j\right] - \left\{ E\left[\sum_{j=0}^{i-1} D_j \cdot X_j\right] + D_{i-1} \cdot R[\underline{X}^{(i)} | \mathcal{F}_{i-1}] \cdot \mathbb{1}_{C_{i-1}} + E\left[\sum_{j=i}^n D_j \cdot X_j \cdot (1 - \mathbb{1}_{C_{j-1}})\right] \right\} \\
 &= E\left[\sum_{j=0}^n D_j \cdot X_j\right] - E\left[\sum_{j=0}^{i-1} D_j \cdot X_j\right] - E\left[\sum_{j=i}^n D_j \cdot X_j\right] - D_{j-1} \cdot R[\underline{X}^{(j)} | \mathcal{F}_{j-1}] \cdot \mathbb{1}_{C_{j-1}} \\
 &\quad + E\left[\sum_{j=i}^n D_j \cdot X_j \cdot \mathbb{1}_{C_{i-1}}\right] \\
 &= -E\left[D_{i-1} \cdot R[\underline{X}^{(i)} | \mathcal{F}_{i-1}] \cdot \mathbb{1}_{C_{i-1}}\right] + E\left[\sum_{j=i}^n D_j \cdot X_j \cdot \mathbb{1}_{C_{i-1}}\right]. \tag{7.2}
 \end{aligned}$$

Przy obu zaprezentowanych strategiach dysponujemy tą samą informacją \mathcal{F}_0 , a zatem podobnie jak w równaniu (6.2), ceny początkowe w chwili 0 przepływów pieniężnych $\underline{X}^{(U_1)}$ i $\underline{X}^{(U_2)}$ powinny być równe. Wynika stąd, że dla każdego scenariusza $C_{i-1} \in \mathcal{F}_{i-1}$ musi zachodzić następująca sekwencja implikacji

$$\begin{aligned}
& \left\{ Q[\underline{X}^{(U_1)}] = Q[\underline{X}^{(U_2)}] \right\} \implies \left\{ Q[\underline{X}^{(U_1)} - \underline{X}^{(U_2)}] = 0 \right\} \\
& \implies \left\{ -E \left[D_{i-1} \cdot R[\underline{X}^{(i)} | \mathcal{F}_{i-1}] \cdot \mathbb{1}_{C_{i-1}} \right] + E \left[\sum_{j=i}^n D_j \cdot X_j \cdot \mathbb{1}_{C_{i-1}} \right] = 0 \right\} \\
& \implies \left\{ E \left[D_{i-1} \cdot R[\underline{X}^{(i)} | \mathcal{F}_{i-1}] \cdot \mathbb{1}_{C_{i-1}} \right] = E \left[\sum_{j=i}^n D_j \cdot X_j \cdot \mathbb{1}_{C_{i-1}} \right] \right\}. \quad (7.3)
\end{aligned}$$

Odwołując się do własności warunkowej wartości oczekiwanej i rozumując analogicznie jak w przypadku przejścia od (6.4) do (6.5), zmodyfikujemy zależność (7.2). Pozwala na to równość $D_{k-1} \cdot R[\underline{X}^{(j)} | C_{i-1}] = D_{i-1} \cdot Q_{i-1}[\underline{X}_{(i)}]$ oraz fakt, że $D_{i-1} \cdot Q_{i-1}[\underline{X}_{(i)}]$ jest zmienną C_{t-1} -mierzalną. W konsekwencji otrzymujemy nową zależność w postaci

$$D_{i-1} \cdot R[\underline{X}^{(i)} | \mathcal{F}^{i-1}] = E \left[\sum_{j=i}^n D_j \cdot X_j | \mathcal{F}_{j-1} \right].$$

lub postaci równoważnej

$$R[\underline{X}^{(i)} | \mathcal{F}_{i-1}] = \frac{1}{D_{i-1}} \cdot E \left[\sum_{j=i}^n D_j \cdot X_j | \mathcal{F}_{i-1} \right] = Q_{i-1}[\underline{X}_{(i)}] \stackrel{(7.1)}{=} R_j^{(i-1)}.$$

Przedstawimy teraz w formie twierdzenia równość, zwaną własnością samofinansownia.

Twierdzenie 7.1 ([4]). *Dla przepływów charakteryzujących polisę ubezpieczeniową, rezerw oraz deflatorów zachodzi następująca zależność rekurencyjna*

$$E \left[D_t \cdot (R_{t+1}^{(t)} + X_t) | \mathcal{F}_{t-1} \right] = D_{t-1} \cdot R_t^{(t-1)}. \quad (7.4)$$

Dowód. Przy założeniu \mathcal{F}_t -mierzalności X_t zachodzi następujący układ równości

$$\begin{aligned}
& E \left[D_t \cdot (R_{t+1}^{(t)} + X_t) | \mathcal{F}_{t-1} \right] = E \left[D_t \cdot R_{t+1}^{(t)} + D_t \cdot X_t | \mathcal{F}_{t-1} \right] \\
& = E \left[D_t \cdot \frac{1}{D_t} \cdot E \left[\sum_{i=t+1}^n D_i \cdot X_i | \mathcal{F}_t \right] + D_t \cdot X_t | \mathcal{F}_{t-1} \right] =
\end{aligned}$$

$$\begin{aligned}
&= E \left[E \left[\sum_{i=t+1}^n D_i \cdot X_i \mid \mathcal{F}_t \right] + E [D_t \cdot X_t \mid \mathcal{F}_t] \mid \mathcal{F}_{t-1} \right] \\
&= E \left[\sum_{i=t}^n D_i \cdot X_i \mid \mathcal{F}_{t-1} \right] = D_{t-1} \cdot R_t^{(t-1)}.
\end{aligned}$$

Porównując formuły początkową i końcową otrzymujemy tezę twierdzenia. \square

Literatura

- [1] J. Jakubowski, *Modele matematyczne rynków instrumentów pochodnych I*, Warszawa 2011.
- [2] S. Jarvis, F. Southall, Elliot V., *Modern Valuation Techniques Presented to the Staple Inn Actuarial Society*, 2001.
- [3] W. Mlak, *Wstęp do teorii przestrzeni Hilberta*, PWN, Warszawa 1972
- [4] S. R. Pliska *Wprowadzenie do matematyki finansowej, modele z czasem dyskretnym*, Warszawa 2005.
- [5] M. V. Wüthrich, H. Bühlmann, H. Furrer, *Market-Consistent Actuarial Valuation*, Berlin 2007.

Wybrane zagadnienia teorii chaosu

Adam Stachura¹

Streszczenie

W pracy są rozważane wybrane zagadnienia teorii topologicznych systemów dynamicznych oraz zbiorów fraktalnych w przestrzeniach euklidesowych. Zamieszczone są także odpowiednie przykłady.

Abstract

The paper deals with selected problems involving topological dynamical systems and fractal subsets of euclidean space. Several examples are included.

Słowa kluczowe: topologiczny system dynamiczny, iteracje, chaos matematyczny (funkcja chaotyczna), fraktale, samopodobieństwo

1 Wstęp

Przedmiotem niniejszej pracy jest przegląd wybranej problematyki dotyczącej modnej obecnie dziedziny, jaką jest teoria chaosu. Rzecz jasna pełny przegląd wszystkich związanych z tą tematyką zagadnień wymagałby o wiele obszerniejszego opracowania. Dlatego też prezentowaną pracę można potraktować raczej jako próbę zademonstrowania (na wybranych przykładach) technik i metod dowodzenia twierdzeń pojawiających się w tej teorii. Wybór omawianych zagadnień jest oczywiście subiektywny i odzwierciedla w dużym stopniu upodobania autora.

¹Katedra Matematyki Stosowanej; Politechnika Lubelska

Jeżeli mielibyśmy wskazać terminy kluczowe tej pracy, to są nimi: topologiczny układ dynamiczny z czasem dyskretnym oraz zbiór fraktalny (fraktal) w przestrzeni euklidesowej. Obydwa te pojęcia można oczywiście uogólniać w rozmaity sposób. Układy dynamiczne z czasem dyskretnym są szczególnym przypadkiem ogólnych układów dynamicznych; zainteresowanych Czytelników odsyłamy np. do pozycji [8], [10], [16]. Fraktale z kolei, rozpatrywane w tej pracy, są również dość szczególne. Więcej informacji o tego typu zbiorach można znaleźć np. w [1], [3], [17]. Interesujące rozważania dotyczące związków omawianych zagadnień z problemami takimi jak sztuczna inteligencja, perspektywy algorytmizacji „myślenia” itp. można znaleźć np. w [18]. Związana z teorią układów dynamicznych jest również problematyka tak zwanych automatów komórkowych, którą tylko sygnalizujemy ([15], zob. też [22]).

Mówiąc o teorii chaosu nie sposób nie wspomnieć o książce B.B. Mandelbrota [12], pozycji uważanej przez wielu za zapoczątkowującą zainteresowanie zagadnieniami chaosu. My wymienimy w tym kontekście także pozycje [17] i [20].

Przejdziemy teraz „do rzeczy”.

Układem dynamicznym (systemem dynamicznym) nazywać będziemy parę (X, S) , gdzie $X \neq \emptyset$ i $S : X \rightarrow X$. Zbiór X nazywamy *przestrzenią fazową* rozważanego układu, a odwzorowanie S nazywamy *dynamiką*.

Dla $x \in X$ ciąg $(S^n(x))_{n \in \mathbb{N}_0}$ nazywamy *trajektorią* punktu x . (Tzw. n -tą iterację odwzorowania S określają wzory: $S^0(x) = x$, $S^{n+1}(x) = S(S^n(x))$, gdzie $x \in X$, $n \in \mathbb{N}_0 = \{n \in \mathbb{Z} : n \geq 0\}$.)

Pojawiają się (wśród wielu innych) następujące problemy:

1. Jak zachowuje się ciąg $(S^n(x))$ przy $n \rightarrow \infty$? (Ewolucja trajektorii).
2. Jak zachowuje się ciąg $(S^n(X))$ przy $n \rightarrow \infty$? (Ewolucja układu).

2 Przestrzenie metryczne

Przypomnimy następujące definicje (zob. np. [9], [19]):

Przestrzenią metryczną nazywamy parę (X, d) , gdzie $X \neq \emptyset$, zaś $d : X \times X \rightarrow \mathbb{R}$ jest funkcją spełniającą następujące warunki:

$$(a) d(x, y) \geq 0 \text{ i } d(x, y) = 0 \Leftrightarrow x = y,$$

$$(b) d(x, y) = d(y, x),$$

$$(c) d(x, z) \leq d(x, y) + d(y, z)$$

dla wszystkich $x, y, z \in X$. Funkcję d nazywamy metryką. Dla $x, y \in X$ liczbę $d(x, y)$ nazywamy odległością punktów x i y . Jeżeli $Y \subset X, Y \neq \emptyset$, to parę (Y, d') nazywamy podprzestrzenią przestrzeni (X, d) , gdzie $d' = d|_{Y \times Y}$ jest metryką d zawężoną do zbioru Y , tzn. $d'(x, y) = d(x, y)$ dla $x, y \in Y$.

Niech (X, d) będzie przestrzenią metryczną. Dla $x \in X$ i $r \geq 0$ zbiory:

$$K(x; r) = \{y \in X : d(x, y) < r\},$$

$$K[x; r] = \{y \in X : d(x, y) \leq r\},$$

$$S[x; r] = \{y \in X : d(x, y) = r\}$$

nazywamy, odpowiednio, kulą otwartą, kulą domkniętą i sferą o środku w punkcie x i promieniu r . Jeżeli $x \in X$, to każdą kulę otwartą $K(x; r)$ o dodatnim promieniu $r > 0$ nazywamy otoczeniem punktu x .

Niech (X, d) będzie przestrzenią metryczną. Zbiór niepusty $A \subset X$ nazywamy otwartym, jeżeli dla każdego $x \in A$ istnieje otoczenie punktu x zawarte w zbiorze A . Zbiór pusty także zaliczamy do zbiorów otwartych. Zbiór $B \subset X$ nazywamy domkniętym, jeżeli jego dopełnienie $X \setminus B$ jest otwarte.

Niech (X, d) będzie przestrzenią metryczną. Niech $A \subset X, A \neq \emptyset$. Średnicę $diam(A)$ tego zbioru określamy wzorem:

$$diam(A) = \sup \{d(x, y) : x, y \in A\}.$$

Zbiór A nazywamy ograniczonym, jeżeli $diam(A) < +\infty$. (Zbiór pusty także uważamy za ograniczony). Jeżeli $diam(A) = +\infty$, to zbiór A nazywamy nieograniczonym.

Niech (X, d) będzie przestrzenią metryczną i niech $A \subset X$. Punkt $x \in X$ nazywamy punktem wewnętrznym zbioru A , jeżeli istnieje $r > 0$ takie, że $K(x; r) \subset A$. Zbiór wszystkich punktów wewnętrznych zbioru A nazywamy wnętrzem tego zbioru i oznaczamy symbolem $Int(A)$. Punkt $y \in X$ nazywamy punktem zewnętrznym

zbioru A , jeżeli istnieje $r > 0$ takie, że $K(y; r) \subset X \setminus A$. Punkt $z \in X$ nazywamy punktem brzegowym zbioru A , jeżeli dla każdego $r > 0$ istnieją punkty $a, b \in K(z; r)$ takie, że $a \in A$ i $b \notin A$. Zbiór wszystkich punktów brzegowych zbioru A nazywamy brzegiem tego zbioru i oznaczamy symbolem $Fr(A)$ lub ∂A . Sumę mnogościową

$$\bar{A} = A \cup Fr(A)$$

nazywamy domknięciem zbioru A .

Prawdziwe są następujące stwierdzenia.

Fakt 2.1. Niech (X, d) będzie przestrzenią metryczną i niech $A \subset X$. Wtedy $Int(A) \subset A \subset \bar{A}$. Zbiór A jest otwarty wtedy i tylko wtedy, gdy $A = Int(A)$. Zbiór A jest domknięty wtedy i tylko wtedy, gdy $A = \bar{A}$.

Fakt 2.2. Niech (X, d) będzie przestrzenią metryczną i niech $A \subset X$.

(a) Wnętrze zbioru A jest sumą mnogościową wszystkich zbiorów otwartych zawartych w zbiorze A .

(b) Wnętrze zbioru A jest największym (w sensie inkluzji) zbiorem otwartym zawartym w zbiorze A .

(c) Domknięcie zbioru A jest iloczynem mnogościowym wszystkich zbiorów domkniętych zawierających zbiór A .

(d) Domknięcie zbioru A jest najmniejszym (w sensie inkluzji) zbiorem domkniętym zawierającym w sobie zbiór A .

Niech (X, d) będzie przestrzenią metryczną. Powiemy, że ciąg (x_n) punktów tej przestrzeni jest zbieżny do punktu $x \in X$ (zwanym granicą ciągu), jeżeli $\lim_{n \rightarrow \infty} d(x, x_n) = 0$. Piszemy wtedy: $\lim_{n \rightarrow \infty} x_n = x$.

W przestrzeni metrycznej każdy ciąg zbieżny jest ograniczony (lecz niekoniecznie na odwrót).

Fakt 2.3. Niech (X, d) będzie przestrzenią metryczną i niech $A \subset X$.

(a) Zbiór A jest domknięty wtedy i tylko wtedy, gdy dla każdego ciągu zbieżnego punktów ze zbioru A granica tego ciągu należy do A .

(b) Punkt $x \in X$ należy do domknięcia \bar{A} zbioru A wtedy i tylko wtedy, gdy istnieje ciąg (x_n) punktów zbioru A zbieżny do punktu x .

Niech (X, d) będzie przestrzenią metryczną i niech $A \subset X$. Zbiór A nazywamy gęstym w tej przestrzeni (krótko: gęstym), jeżeli $\bar{A} = X$. Przestrzeń X nazywamy ośrodkową, jeżeli istnieje zbiór $A \subset X$ gęsty i przeliczalny.

Niech (X, d) będzie przestrzenią metryczną. Powiemy, że ciąg (x_n) punktów tej przestrzeni spełnia warunek Cauchy'ego (jest ciągiem Cauchy'ego), jeżeli

$$\bigwedge_{\varepsilon > 0} \bigvee_{n_0 \in \mathbb{N}} \bigwedge_{n, m \geq n_0} d(x_n, x_m) < \varepsilon$$

lub, co jest równoważne, gdy $\lim_{n, m \rightarrow \infty} d(x_n, x_m) = 0$.

W przestrzeni metrycznej każdy ciąg zbieżny jest ciągiem Cauchy'ego (lecz niekoniecznie na odwrót).

Niech (X, d) będzie przestrzenią metryczną. Przestrzeń nazywamy zupełną, jeżeli każdy ciąg punktów tej przestrzeni spełniający warunek Cauchy'ego jest w tej przestrzeni zbieżny.

Niech (X, d) będzie przestrzenią metryczną i niech $A \subset X$. Zbiór A nazywamy zwartym, jeżeli z każdego ciągu (x_n) punktów tego zbioru można wybrać podciąg zbieżny, którego granica należy do zbioru A .

Fakt 2.4. Niech (X, d) będzie przestrzenią metryczną i niech $A \subset X$. Następujące warunki są równoważne:

(1) A jest zbiorem zwartym.

(2) Z każdego pokrycia otwartego zbioru A można wybrać podpokrycie skończone, tzn. jeżeli $\{G_t : t \in T\}$ jest dowolną rodziną indeksowaną podzbiorów otwartych przestrzeni X taką, że

$$A \subset \bigcup_{t \in T} G_t,$$

to istnieje skończona podrodzina $\{G_{t_1}, G_{t_2}, \dots, G_{t_n}\}$ taka, że

$$A \subset \bigcup_{k=1}^n G_{t_k}.$$

Fakt 2.5. Niech (X, d) będzie przestrzenią metryczną. Następujące warunki są równoważne:

- (1) X jest przestrzenią zwartą.
- (2) Jeżeli $\{A_t : t \in T\}$ jest dowolną rodziną indeksowaną podzbiorów domkniętych przestrzeni X taką, że przekrój każdej skończonej podrodziny $\{A_{t_1}, A_{t_2}, \dots, A_{t_n}\}$ jest niepusty, to

$$\bigcap_{t \in T} A_t \neq \emptyset.$$

Niech (X, d_X) oraz (Y, d_Y) będą przestrzeniami metrycznymi i niech $f : X \rightarrow Y$ będzie dowolnym odwzorowaniem. Mówimy, że odwzorowanie to jest ciągłe w punkcie $x \in X$, jeżeli dla każdego ciągu (x_n) punktów przestrzeni X prawdziwa jest implikacja:

$$\lim_{n \rightarrow \infty} x_n = x \Rightarrow \lim_{n \rightarrow \infty} f(x_n) = f(x).$$

Przekształcenie f nazywamy ciągłym, jeżeli jest ono ciągłe w każdym punkcie $x \in X$.

Fakt 2.6. Niech (X, d_X) oraz (Y, d_Y) będą przestrzeniami metrycznymi i niech $f : X \rightarrow Y$. Następujące warunki są równoważne:

- (1) Funkcja f jest ciągła.
- (2) Przeciwobraz każdego otwartego podzbioru przestrzeni Y poprzez funkcję f jest otwarty w przestrzeni X .
- (3) Przeciwobraz każdego domkniętego podzbioru przestrzeni Y poprzez funkcję f jest domknięty w przestrzeni X .

Fakt 2.7. Niech (X, d) będzie przestrzenią metryczną i niech $f : X \rightarrow \mathbb{R}$ będzie funkcją ciągłą (w zbiorze \mathbb{R} rozważamy zwykłą, euklidesową metrykę $\rho(a, b) = |a - b|$, $a, b \in \mathbb{R}$). Jeżeli zbiór $A \subset X$ jest zwarty, to funkcja f jest w tym zbiorze ograniczona i osiąga w nim swoje kresy: górny i dolny.

Niech (X, d_X) oraz (Y, d_Y) będą przestrzeniami metrycznymi i niech $f : X \rightarrow Y$. Mówimy, że odwzorowanie f spełnia warunek Lipschitza (jest odwzorowaniem

lipschitzowskim), jeżeli istnieje liczba $L > 0$ taka, że

$$d_Y(f(x), f(y)) \leq L d_X(x, y)$$

dla wszystkich $x, y \in X$. Jeżeli, dodatkowo, liczba L (tzw. stała Lipschitza) spełnia warunek: $0 \leq L < 1$, to f nazywamy kontrakcją.

Fakt 2.8. *Każde odwzorowanie lipschitzowskie jest ciągłe.*

3 Dynamiki chaotyczne

W tym paragrafie ponownie zajmiemy się badaniem ogólnych układów dynamicznych. Przypomnijmy (zob. par. 1), że układem dynamicznym jest para (X, S) , gdzie $S: X \rightarrow X$. O zbiorze X (tzw. przestrzeni fazowej rozważanego układu) zakładamy zawsze, że jest niepusty.

Jeżeli $x \in X$, to ciąg $(S^n(x))$, $n \in \mathbb{N}_0$, punktów przestrzeni fazowej nazywamy trajektorią punktu x , czyli punktu początkowego tej trajektorii (przypominamy, że przyjmujemy umowę: $S^0(x) = x$). Traktując wskaźnik n jako czas możemy interpretować trajektorię w ten sposób, że punkt początkowy „wędruje” po, przestrzeni fazowej, a punkty trajektorii $x_n = S^n(x)$ wyznaczają obserwowane położenie tego punktu w kolejnych chwilach $n \in \mathbb{N}_0$. Czas przy tej interpretacji zmienia się w sposób dyskretny: $t = 0, 1, 2, \dots$, dlatego mówimy, że rozważamy układ dynamiczny z czasem dyskretnym.

Będziemy w tym paragrafie zakładać, że przestrzeń fazowa rozważanego układu jest przestrzenią metryczną (X, d) . Taki układ dynamiczny będziemy nazywać *topologicznym*. O dynamice $S: X \rightarrow X$ zakłada się zwykle, że jest przekształceniem ciągłym. Nie jest to jednak wymóg konieczny i będziemy studiować także nieciągłe dynamiki.

Tytułowa „chaotyczność” dynamiki S (lub układu dynamicznego (X, S)) związana jest z pewnymi cechami trajektorii punktów układu. Nie ma niestety jednej, ogólnie przyjętej definicji pojęcia chaosu. Matematycy zajmujący się tą problematyką przyjmują rozmaite założenia i definicje, różniące się w szczegółach (co

prawda zwykle w niewielkim stopniu). Prawie zawsze jednak o chaosie mówi się wtedy, gdy trajektorie mają dwie ważne cechy. Jedną z nich jest tzw. wrażliwość na zmianę położenia początkowego. Nie wchodząc w tym miejscu w szczegóły (sprecyzujemy to pojęcie później) chodzi o to, że trajektorie startujące w chwili $t = 0$ z dowolnie bliskich punktów powinny (po dłuższym bądź krótszym czasie) „rozejść się” na ustalony dystans. Drugim wymogiem jest propagacja - trajektorie startujące z otoczenia dowolnego punktu powinny (znów po dłuższym bądź krótszym czasie) trafić w otoczenie każdego innego punktu. Przyjmuje się zwykle także dalsze założenia, o czym później.

Decydując się na wybór konkretnej definicji pojęcia chaosu, sprecyzujemy najpierw ważne pojęcie dotyczące układów dynamicznych. W dalszym ciągu zwrot „układ dynamiczny” będzie oznaczać topologiczny układ dynamiczny z czasem dyskretnym, w sensie opisanym na początku tego paragrafu.

Niech (X, S) będzie układem dynamicznym i niech $x \in X$. Przypuśćmy, że dla pewnej liczby $n \in \mathbb{N}$ (a więc $n \neq 0$) jest

$$S^n(x) = x.$$

Oczywiście wtedy także jest $S^{2n}(x) = x$, $S^{3n}(x) = x$, ogólnie $S^{kn}(x) = x$ dla wszystkich $k \in \mathbb{N}$. Co więcej, musi wtedy istnieć najmniejsza liczba naturalna n , dla której zachodzi powyższa równość. Niech $p \in \mathbb{N}$ będzie tą liczbą. Wtedy $S^p(x) = x$ oraz $S^k(x) \neq x$ dla każdego $k \in \{0, 1, \dots, p-1\}$. Nietrudno sprawdzić, że wtedy $S^j(x) \neq S^k(x)$ dla wszystkich $j, k \in \{0, 1, \dots, p-1\}$ takich, że $j \neq k$. Ponadto dla każdego $n \in \mathbb{N}_0$ istnieją liczby $q, r \in \mathbb{N}_0$ takie, że $n = pq + r$ i $r \in \{0, 1, \dots, p-1\}$. Mamy wtedy:

$$S^n(x) = S^{pq+r}(x) = S^r[S^{qp}(x)] = S^r(x).$$

Oznacza to, że cała trajektoria $(S^n(x))$, $n \in \mathbb{N}_0$, jest ciągiem, którego zbiór wyrazów składa się z dokładnie p parami różnych punktów i jest zbiorem postaci

$$\{x, S(x), S^2(x), \dots, S^{p-1}(x)\},$$

przy czym $S^p(x) = x$.

W związku z tym przyjmujemy następującą definicję.

Definicja 3.1. Niech (X, S) będzie układem dynamicznym. Punkt $x \in X$ nazywamy okresowym, jeżeli istnieje liczba $n \in \mathbb{N}$ taka, że $S^n(x) = x$. Mówimy, że punkt $x \in X$ jest punktem okresowym o okresie p (gdzie $p \in \mathbb{N}$), jeżeli $S^p(x) = x$ oraz $S^k(x) \neq x$ dla każdego $k \in \{0, 1, \dots, p-1\}$.

Punkty o okresie 1, tj. punkty $x \in X$ takie, że $S(x) = x$ nazywamy punktami stałymi dynamiki S (lub układu (X, S)).

Definicja 3.2. Niech (X, S) będzie układem dynamicznym. Jeżeli zbiór wszystkich punktów okresowych tego układu jest gęstym podzbiorem przestrzeni metrycznej (X, d) (zob. par. 3), to będziemy mówić, że rozważany układ spełnia warunek gęstości zbioru punktów okresowych, który będziemy w dalszym ciągu nazywać krótko warunkiem [G].

Definicja 3.3. Niech (X, S) będzie układem dynamicznym. Przypuśćmy, że istnieje liczba $\eta > 0$ taka, że dla każdego punktu $x \in X$ i każdego otoczenia O_x tego punktu istnieje $y \in O_x$ oraz liczba $n \in \mathbb{N}_0$, dla których

$$d(S^n(x), S^n(y)) \geq \eta.$$

Powiemy wtedy, że układ spełnia warunek wrażliwości na położenia początkowe, który będziemy w dalszym ciągu nazywać krótko warunkiem [W].

Stałą dodatnią η nazywamy stałą wrażliwości układu (X, S) .

Definicja 3.4. Niech (X, S) będzie układem dynamicznym. Powiemy, że ten układ spełnia warunek tranzytywności, który będziemy w dalszym ciągu nazywać krótko warunkiem [T], jeżeli dla dowolnych zbiorów otwartych i niepustych $A, B \subset X$ istnieje punkt $x \in A$ i liczba $m \in \mathbb{N}$ takie, że $S^m(x) \in B$.

Bez dodatkowych założeń niewiele można powiedzieć o zależnościach pomiędzy zdefiniowanymi powyżej własnościami. Nieco lepiej jest w („modelowej” zresztą) sytuacji, gdy X jest przedziałem w zbiorze \mathbb{R} liczb rzeczywistych ze zwykłą, euklidesową metryką daną wzorem: $d(x, y) = |x - y|$, $x, y \in X$, a dynamika $S: X \rightarrow X$ jest funkcją ciągłą przekształcającą przedział X w X . (Będziemy badać ten przypadek w następnych paragrafach).

Wprowadzimy teraz pewne terminy i oznaczenia. Jeżeli (X, d) jest przestrzenią metryczną, $A \subset X$, $A \neq \emptyset$ i $x \in X$, to odległość pomiędzy punktem x i zbiorem A jest liczbą daną wzorem

$$\text{dist}(x, A) = \inf_{a \in A} d(x, a).$$

Jeżeli $A, B \subset X$ są zbiorami niepustymi, to przyjmujemy, że

$$\text{dist}(A, B) = \inf \{d(a, b) : a \in A \wedge b \in B\}.$$

Zauważmy, że dla każdego $x \in X$ zachodzi nierówność:

$$\text{dist}(A, B) \leq \text{dist}(x, A) + \text{dist}(x, B). \quad (3.1)$$

Ponadto jeżeli zbiory A, B są zwarte, w szczególności gdy są skończone, to

$$A \cap B = \emptyset \Leftrightarrow \text{dist}(A, B) > 0. \quad (3.2)$$

Niech (X, S) będzie układem dynamicznym. Dla każdego $x \in X$ symbolem

$$T_x = \{S^n(x) : n \in \mathbb{N}_0\}.$$

będziemy oznaczać zbiór wszystkich punktów trajektorii $(S^n(x))$, $n \in \mathbb{N}_0$, punktu x .

Twierdzenie 3.1. . Niech (X, S) będzie układem dynamicznym z ciągłą dynamiką S , spełniającym warunki **[G]** i **[T]**. Jeżeli w przestrzeni X istnieją dwa punkty okresowe a, b takie, że $T_a \cap T_b = \emptyset$, to układ (X, S) spełnia także warunek **[W]**.

Dowód. Niech $\delta = \text{dist}(T_a, T_b)$. Z założenia i z (3.2) wynika, że $\delta > 0$. Pokażemy, że układ (X, S) spełnia warunek **[W]** ze stałą wrażliwości $\eta = \frac{1}{8}\delta$.

Niech $x \in X$ będzie dowolnym punktem. Z (3.1) wynika, że

$$\text{dist}(x, T_a) + \text{dist}(x, T_b) \geq 8\eta,$$

więc przynajmniej jedna z liczb $\text{dist}(x, T_a)$ oraz $\text{dist}(x, T_b)$ jest większa bądź równa 4η . Przypuśćmy np., że

$$\text{dist}(x, T_a) \geq 4\eta \quad (3.3)$$

(jeżeli $\text{dist}(x, T_b) \geq 4\eta$, rozumiemy podobnie). Niech O_x będzie dowolnym otoczeniem punktu x . Zbiór

$$U = O_x \cap K(x; \eta) \quad (3.4)$$

jest otwarty w przestrzeni X i niepusty ($x \in U$). Ponieważ układ spełnia warunek [G], więc w zbiorze U istnieje punkt okresowy z . Niech $p \in \mathbb{N}$ będzie okresem tego punktu. Połóżmy

$$W = \bigcap_{j=0}^p S^{-j} [K(S^j(a); \eta)].$$

Wtedy W jest zbiorem otwartym (jako iloczyn skończonej liczby zbiorów otwartych

$S^{-j} [K(S^j(a); \eta)]$, (zob. fakt 2.6) i niepustym (gdyż $a \in W$). Ponieważ układ spełnia warunek [T], więc istnieje $y \in U$ i $m \in \mathbb{N}$ takie, że $S^m(y) \in W$.

Niech $k = \text{Ent}\left(\frac{m}{p}\right) + 1$. Wtedy $k - 1 \leq \frac{m}{p} < k$, a stąd

$$1 \leq pk - m \leq p.$$

Wobec tego

$$S^{pk}(y) = S^{pk-m}(S^m(y)) \in S^{pk-m}(W) \subset K(S^{pk-m}(a); \eta). \quad (3.5)$$

Z nierówności trójkąta dla metryki d otrzymujemy:

$$d(x, S^{pk-m}(a)) \leq d(x, z) + d(z, S^{pk}(y)) + d(S^{pk}(y), S^{pk-m}(a)),$$

czyli

$$d(S^{pk}(z), S^{pk}(y)) \geq d(x, S^{pk-m}(a)) - d(S^{pk}(y), S^{pk-m}(a)) - d(x, z), \quad (3.6)$$

gdyż z jest punktem okresowym o okresie p i $S^{pk}(z) = z$.

Ponieważ $z \in K(x; \eta)$, więc $d(x, z) < \eta$. Podobnie, $S^{pk}(y) \in K(S^{pk-m}(a); \eta)$ (zob. wzór (3.5)), więc

$$d(S^{pk}(y), S^{pk-m}(a)) < \eta.$$

Na koniec zauważmy, że $S^{pk-m}(a) \in T_a$ i wobec (3.3) jest $d(x, S^{pk-m}(a)) \geq 4\eta$. Z nierówności (3.6) wynika zatem, że

$$d(S^{pk}(z), S^{pk}(y)) \geq 4\eta - \eta - \eta = 2\eta.$$

Kładąc $n = pk$ i ponownie stosując nierówność trójkąta, stwierdzamy, że

$$d(S^n(z), S^n(x)) + d(S^n(x), S^n(y)) \geq d(S^n(z), S^n(y)) \geq 2\eta.$$

Wobec tego $d(S^n(z), S^n(x)) \geq \eta$ lub $d(S^n(x), S^n(y)) \geq \eta$. Ale $z, y \in U$, więc także $z, y \in O_x$ (zob. wzór (3.4)). W obydwu przypadkach oznacza to zatem, że układ (X, S) spełnia warunek [W] ze stałą wrażliwości η . \square

Definicja 3.5. Niech (X, S) będzie układem dynamicznym. Powiemy, że układ ten jest chaotyczny (lub że dynamika układu jest chaotyczna), jeżeli (X, S) spełnia warunki [G], [W] i [T].

4 Układy dynamiczne na prostej liczbowej

W tym paragrafie zajmujemy się topologicznymi układami dynamicznymi z czasem dyskretnym typu (P, f) , gdzie $f : P \rightarrow P$, zaś $P \subset \mathbb{R}$ jest przedziałem o niepustym wnętrzu, tj. jednym ze zbiorów postaci: (a, b) , $[a, b)$, $(a, b]$, $[a, b]$, $(a, +\infty)$, $[a, +\infty)$, $(-\infty, a)$, $(-\infty, a]$, $(-\infty, +\infty)$, gdzie $a, b \in \mathbb{R}$ i $a < b$. Rodzinę wszystkich takich przedziałów oznaczymy przez \mathcal{P} . Metryka w przedziale $P \in \mathcal{P}$ będzie zawsze metryką euklidesową, daną wzorem $d(x, y) = |x - y|$ dla wszystkich $x, y \in P$.

Zacniemy od przypomnienia własności Darboux funkcji ciągłych.

Fakt 4.1. Niech $P \in \mathcal{P}$ i niech $f : P \rightarrow \mathbb{R}$ będzie funkcją ciągłą. Niech $x_1, x_2 \in P$, $x_1 < x_2$. Wtedy dla każdej liczby y zawartej pomiędzy $f(x_1)$ oraz $f(x_2)$ (tzn. takiej, że $f(x_1) < y < f(x_2)$ lub $f(x_2) > y > f(x_1)$) istnieje taki $x \in (x_1, x_2)$, że $f(x) = y$.

Wnioskami z własności Darboux są następujące fakty.

Fakt 4.2. Niech $P \in \mathcal{P}$ i niech $f : P \rightarrow \mathbb{R}$ będzie funkcją ciągłą. Jeżeli f nie jest funkcją stałą, to obraz $f(P)$ przedziału P poprzez funkcję f jest przedziałem należącym do rodziny \mathcal{P} .

Fakt 4.3. Niech $a, b \in \mathbb{R}$, $a < b$, i niech $f : [a, b] \rightarrow \mathbb{R}$ będzie funkcją ciągłą. Jeżeli funkcja f przyjmuje na końcach przedziału $[a, b]$ wartości różnych znaków (tzn. $f(a) < 0$ i $f(b) > 0$ lub $f(a) > 0$ i $f(b) < 0$), to istnieje taki $x \in (a, b)$, że $f(x) = 0$.

Zacniemy od następującego lematu, dotyczącego wszystkich topologicznych układów dynamicznych z czasem dyskretnym. Jeżeli (X, S) jest takim układem, $A \subset X$ i $m \in \mathbb{N}_0$, to będziemy stosować oznaczenie: $S^{-m}(A) = (S^m)^{-1}(A)$.

Lemat 4.1. Niech (X, S) będzie układem dynamicznym z ciągłą dynamiką S , spełniającym warunek [T]. Dla dowolnych zbiorów otwartych i niepustych $A, B \subset X$ i dowolnego $m \in \mathbb{N}$ istnieje punkt $x \in A$ i liczba $n \in \mathbb{N}$ takie, że $n > m$ i $S^n(x) \in B$.

Dowód. Niech $A, B \subset X$, $m \in \mathbb{N}$. Wobec ciągłości dynamiki S zbiór $S^{-m}(B)$ jest otwarty. Z warunku [T] wynika, że istnieje $x \in A$ i liczba $k \in \mathbb{N}$ taka, że $S^k(x) \in S^{-m}(B)$. Wtedy $S^m(S^k(x)) \in B$, czyli $S^{m+k}(x) \in B$ i wystarczy przyjąć, że $n = m + k$. \square

Lemat 4.2. Niech $a, b \in \mathbb{R}$, $a < b$, i niech $([a, b], f)$ będzie układem dynamicznym z ciągłą dynamiką f . Wtedy istnieje punkt stały dynamiki f w przedziale $[a, b]$.

Łatwy dowód pomijamy.

Lemat 4.3. Niech $P \in \mathcal{P}$ i niech (P, f) będzie układem dynamicznym z ciągłą dynamiką f . Jeżeli $Q \subset P$, $Q \in \mathcal{P}$ jest przedziałem nie zawierającym punktów okresowych i $x_0 \in Q$ oraz $p, n \in \mathbb{N}$ są takie, że $n < p$ i $f^n(x_0) \in Q$, $f^p(x_0) \in Q$, to $x_0 < f^n(x_0) < f^p(x_0)$ lub $x_0 > f^n(x_0) > f^p(x_0)$.

Dowód. Niech $Q \subset P$ będzie przedziałem nie zawierającym punktów okresowych i przypuśćmy, że $x_0 \in Q$, $n < p$, $f^n(x_0) \in Q$ i $f^p(x_0) \in Q$. Wtedy $f^n(x_0) \neq f^p(x_0)$,

gdyż w przeciwnym razie punkt $f^n(x_0)$ byłby punktem okresowym w przedziale Q o okresie nie większym od $p - n$, wbrew założeniu. Niech będzie $x_0 < f^n(x_0)$ (w przypadku, gdy $x_0 > f^n(x_0)$, rozumiemy analogicznie). Pokażemy, że wtedy $f^n(x_0) < f^p(x_0)$.

Przypuśćmy, że tak nie jest, czyli że $f^p(x_0) < f^n(x_0)$. Dalsze rozumowanie podzielimy na etapy.

Etap I. Niech $g(x) = f^n(x)$ dla wszystkich $x \in P$. Wtedy jest $g(x_0) = f^n(x_0)$, czyli $g(x_0) \in Q$. Wykażemy indukcyjnie, że $x_0 < g^k(x_0)$ dla wszystkich $k \in \mathbb{N}$. Z założenia jest $x_0 < g(x_0)$. Przypuśćmy, że dla pewnego $k \in \mathbb{N}$ jest

$$x_0 < g^k(x_0), \quad g^{k+1}(x_0) \leq g(x_0).$$

Równość $g^{k+1}(x_0) = g(x_0)$ oznaczałaby, że $f^{nk}(g(x_0)) = g(x_0)$, czyli że $g(x_0)$ jest punktem okresowym o okresie nie przekraczającym nk , wbrew założeniu o przedziale Q . Zatem musi być $g^{k+1}(x_0) < g(x_0)$.

Niech $\varphi(x) = g^k(x) - x$ dla $x \in P$. Wtedy

$$\varphi(x_0) = g^k(x_0) - x_0 > 0$$

oraz

$$\varphi(g(x_0)) = g^{k+1}(x_0) - g(x_0) < 0,$$

więc (zob. fakt 4.3) istnieje $c \in (x_0, g(x_0)) \subset Q$ taki, że $\varphi(c) = 0$, czyli $g^k(c) = f^{nk}(c) = c$. Oznacza to, że punkt c jest punktem okresowym o okresie nie przekraczającym nk , wbrew założeniu o przedziale Q . Otrzymana sprzeczność dowodzi, że musi być $g(x_0) < g^{k+1}(x_0)$, a ponieważ $x_0 < g(x_0)$, więc $x_0 < g^{k+1}(x_0)$.

Etap II. Niech $h(x) = f^{p-n}(x)$ dla wszystkich $x \in P$. Wtedy jest $h(f^n(x_0)) = f^p(x_0)$, czyli $h(f^n(x_0)) \in Q$. Wykażemy indukcyjnie, że $h^k(f^n(x_0)) < f^n(x_0)$ dla wszystkich $k \in \mathbb{N}$. Z założenia jest $h(f^n(x_0)) < f^n(x_0)$. Przypuśćmy, że dla pewnego $k \in \mathbb{N}$ jest

$$h^k(f^n(x_0)) < f^n(x_0), \quad h^{k+1}(f^n(x_0)) \geq h(f^n(x_0)).$$

Równość $h^{k+1}(f^n(x_0)) = h(f^n(x_0))$ znaczyłaby, że $f^{(p-n)k}[h(f^n(x_0))] = h(f^n(x_0))$, czyli że $h(f^n(x_0))$ jest punktem okresowym o okresie nie przekraczającym $(p-n)k$, wbrew założeniu o przedziale Q . Zatem musi być $h^{k+1}(f^n(x_0)) > h(f^n(x_0))$.

Niech $\psi(x) = h^k(x) - x$ dla $x \in P$. Wtedy

$$\psi(f^n(x_0)) = h^k(f^n(x_0)) - f^n(x_0) < 0$$

oraz

$$\psi[h(f^n(x_0))] = h^{k+1}(f^n(x_0)) - h(f^n(x_0)) > 0,$$

więc (zob. fakt 4.3) istnieje punkt

$$c \in (h(f^n(x_0)), f^n(x_0)) = (f^p(x_0), f^n(x_0)) \subset Q$$

taki, że $\psi(c) = 0$, czyli $h^k(c) = f^{(p-n)k}(c) = c$. Oznacza to, że punkt c jest punktem okresowym o okresie nie przekraczającym $(p-n)k$, wbrew założeniu o przedziale Q . Otrzymana sprzeczność dowodzi, że musi być $h^{k+1}(f^n(x_0)) < h(f^n(x_0))$, a ponieważ $h(f^n(x_0)) < f^n(x_0)$, więc $h^{k+1}(f^n(x_0)) < f^n(x_0)$.

Etap III. Z rozważań etapu I wynika, że $x_0 < g^{p-n}(x_0)$, czyli $x_0 < f^{(p-n)n}(x_0)$. Z kolei z etapu II wynika, że $h^n(f^n(x_0)) < f^n(x_0)$, czyli $f^{(p-n)n}(f^n(x_0)) < f^n(x_0)$. Rozważmy funkcję ω określoną wzorem:

$$\omega(x) = f^{(p-n)n}(x) - x$$

dla $x \in P$. Wtedy

$$\omega(x_0) = f^{(p-n)n}(x_0) - x_0 > 0$$

oraz

$$\omega(f^n(x_0)) = f^{(p-n)n}(f^n(x_0)) - f^n(x_0) < 0,$$

więc (zob. fakt 4.3) istnieje punkt $c \in (x_0, f^n(x_0)) \subset Q$ taki, że $\omega(c) = 0$, czyli $f^{(p-n)n}(c) = c$. Oznacza to, że punkt c jest punktem okresowym o okresie nie przekraczającym $(p-n)n$, wbrew założeniu o przedziale Q . Otrzymaliśmy więc sprzeczność, co dowodzi, że musi być $f^n(x_0) < f^p(x_0)$. \square

Twierdzenie 4.1. . Niech $P \in \mathcal{P}$. Każdy układ dynamiczny (P, f) z ciągłą dynamiką f spełniający warunek [T] spełnia także warunek [G].

Dowód. Należy udowodnić, że punkty okresowe dynamiki f tworzą zbiór gęsty w przedziale P . Załóżmy, że tak nie jest. Wtedy istnieje punkt w przedziale P oraz otoczenie tegoż punktu nie zawierające punktów okresowych. Innymi słowy, istnieje przedział otwarty niepusty $U \subset P$ nie zawierający punktów okresowych. Ponieważ układ (P, f) spełnia warunek [T], więc istnieje $x \in U$ oraz liczba $n \in \mathbb{N}$ taka, że $f^n(x) \in U$.

Niech $y = f^n(x)$. Ponieważ U z założenia nie zawiera punktów okresowych, więc $y \neq x$. Wobec ciągłości funkcji f^n , możemy znaleźć otoczenie W punktu y i rozłączne z nim otoczenie Q punktu x takie, że $f^n(Q) \subset W$ oraz $Q \subset U$ i $W \subset U$. Z lematu 4.1 wynika, że istnieje liczba $p \in \mathbb{N}$, $p > n$, oraz punkt $x_0 \in Q$ taki, że $f^p(x_0) \in Q$. Wtedy $x_0 \in U$, $f^n(x_0) \in U$, $f^p(x_0) \in U$, $n < p$ i przedział U nie zawiera punktów okresowych dynamiki f , możemy więc zastosować lemat 4.3. Wynika z niego, że $x_0 < f^n(x_0) < f^p(x_0)$ lub $x_0 > f^n(x_0) > f^p(x_0)$. Ponieważ $x_0 \in Q$ i $f^p(x_0) \in Q$, zaś Q jest przedziałem, w obydwu przypadkach jest $f^n(x_0) \in Q$. Z drugiej strony $f^n(x_0) \in W$ i $W \cap Q = \emptyset$, czyli $f^n(x_0) \notin Q$.

Otrzymaliśmy sprzeczność - tym samym dowód został zakończony. \square

Wniosek 4.1. Niech $P \in \mathcal{P}$. Każdy układ dynamiczny (P, f) z ciągłą dynamiką f spełniający warunek [T] jest chaotyczny.

Dowód. Na mocy twierdzenia 4.1 rozważany układ spełnia warunek [G]. Ponieważ przedział P ma niepuste wnętrze, więc łatwo można pokazać, że istnieją w tym przedziale punkty okresowe a, b takie, że $T_a \cap T_b = \emptyset$, gdzie T_a oraz T_b oznaczają zbiory wszystkich punktów trajektorii punktów a i b , odpowiednio (zob. oznaczenia z par. 3). Wystarczy teraz skorzystać z twierdzenia 3.1. \square

Zasygnalizujemy w tym miejscu dwa twierdzenia związane z własnościami dynamik ciągłych w przedziałach na osi liczbowej. Pierwsze z nich, mówiąc bardzo ogólnie i nie wchodząc w szczegóły, stwierdza chaotyczność dynamiki, dla której

istnieją punkty okresowe o okresie 3 (zob. [11]). Uogólnieniem tego rezultatu jest słynne twierdzenie Szarkowskiego z 1964 r. (Zauważmy w tym miejscu, że oryginalna praca Szarkowskiego [21] przez długi czas pozostawała „niezauważona” przez ogół matematyków).

5 Układy dynamiczne na prostej liczbowej – ciąg dalszy

Niech $P = [0, 1] \subset \mathbb{R}$. W tym paragrafie będziemy studiować układ dynamiczny (P, S) z trzema konkretnymi dynamikami S .

Definicja 5.1. Funkcję $D_2 : [0, 1] \rightarrow [0, 1]$ określoną wzorem

$$D_2(x) = \begin{cases} 2x, & \text{gdy } 0 \leq x < \frac{1}{2}, \\ 2x - 1, & \text{gdy } \frac{1}{2} \leq x \leq 1 \end{cases} \quad (5.1)$$

dla wszystkich $x \in [0, 1]$ nazywamy funkcją 2-adyczną.

Funkcja D_2 jest oczywiście nieciągła w punkcie $\frac{1}{2}$. W celu zbadania własności układu dynamicznego $([0, 1], D_2)$ zaczniemy od przypomnienia niezbędnych wiadomości dotyczących szeregów liczbowych.

Szereg

$$\sum_{j=0}^{\infty} a_0 q^j, \quad (5.2)$$

gdzie $a_0 \in \mathbb{R}$ i $q \in \mathbb{R}$, nazywamy szeregiem geometrycznym. Liczbę q nazywamy ilorazem tego szeregu.

Fakt 5.1. Jeżeli $a_0 \neq 0$, to szereg geometryczny (5.2) jest zbieżny wtedy i tylko wtedy, gdy $|q| < 1$. Zachodzi wtedy równość:

$$\sum_{j=0}^{\infty} a_0 q^j = \frac{a_0}{1 - q}.$$

Jeżeli $a_0 = 0$, to $\sum_{j=0}^{\infty} a_0 q^j = 0$.

Fakt 5.2. Niech $\sum_{j=0}^{\infty} x_j$ będzie szeregiem o wyrazach nieujemnych. Jeżeli istnieje szereg geometryczny zbieżny o wyrazach nieujemnych (5.2) taki, że $0 \leq x_j \leq a_0 q^j$ dla wszystkich $j \in \mathbb{N}_0$, to szereg $\sum_{j=0}^{\infty} x_j$ jest zbieżny i $\sum_{j=0}^{\infty} x_j \leq \sum_{j=0}^{\infty} a_0 q^j$.

Lemat 5.1. Dla każdej liczby $x \in [0, 1]$ istnieje ciąg (x_j) , $j \in \mathbb{N}_0$, liczb ze zbioru dwuelementowego $\{0, 1\}$ taki, że

$$x = \sum_{j=0}^{\infty} \frac{x_j}{2^j}. \quad (5.3)$$

Zauważmy, że szereg po prawej stronie wzoru (5.3) jest zawsze zbieżny na mocy faktu 5.2.

Definicja 5.2. Niech $x \in [0, 1]$. Ciąg (x_j) , $j \in \mathbb{N}_0$, o którym mowa w lemacie 5.1 nazywamy ciągiem kolejnych cyfr w zapisie liczby x w systemie pozycyjnym o podstawie 2 lub krótko, w zapisie dwójkowym tej liczby. (Mówi się także o rozwinięciu dwójkowym liczby x). Będziemy stosować oznaczenie:

$$x = (x_0, x_1 x_2 x_3 \dots)_2, \quad (5.4)$$

gdzie ciąg (x_j) , $j \in \mathbb{N}_0$, liczb ze zbioru dwuelementowego $\{0, 1\}$ spełnia równość (5.3).

Niekiedy mamy więcej niż jedną możliwość. Tak np. łatwo można się przekonać, że:

$$\frac{3}{4} = (0, 11000\dots)_2 = (0, 10111\dots)_2$$

(w pierwszym zapisie dwójkowym tylko dwie cyfry są równe 1, a pozostałe są równe 0, podczas gdy w drugim tylko dwie cyfry są równe 0, a pozostałe są równe 1).

Zapis dwójkowy (5.4) liczby $x \in [0, 1]$ nazywamy skończonym, gdy ciąg (x_j) , $j \in \mathbb{N}_0$, kolejnych cyfr w zapisie dwójkowym liczby x jest od pewnego miejsca stały i jego wyrazy od tego miejsca są równe 0, tj. gdy istnieje taka liczba $j_0 \in \mathbb{N}_0$, że $x_j = 0$ dla wszystkich $j \geq j_0$.

Lemat 5.2. Liczba $x \in [0, 1]$, $x \neq 0$, ma dwa różne zapisy dwójkowe wtedy i tylko wtedy, gdy jeden z nich jest skończony:

$$x = (x_0, x_1 x_2 \dots x_{m-1} 1000 \dots),$$

tj. ciąg (x_j) , $j \in \mathbb{N}_0$, kolejnych cyfr w zapisie dwójkowym liczby x spełnia warunek: istnieje $m \in \mathbb{N}_0$ takie, że $x_m = 1$ oraz $x_j = 0$ dla wszystkich $j > m$. Drugim zapisem liczby x jest

$$x = (x_0, x_1 x_2 \dots x_{m-1} 0111 \dots),$$

tj. ciąg (y_j) , $j \in \mathbb{N}_0$, kolejnych cyfr w tym zapisie spełnia warunek: $y_j = x_j$ dla wszystkich $j \in \{0, 1, \dots, m-1\}$, $y_m = 0$ oraz $y_j = 1$ dla wszystkich $j > m$.

Zapis dwójkowy (5.4) liczby $x \in [0, 1]$ nazywamy *okresowym*, gdy istnieją takie liczby $s \in \mathbb{N}_0$ i $p \in \mathbb{N}$, że dla każdego $j \in \mathbb{N}_0$, gdzie

$$j = qp + r \quad \text{i} \quad q \in \mathbb{N}_0, r \in \{0, 1, \dots, p-1\},$$

zachodzi równość: $x_{s+j} = x_{s+r}$. Inaczej mówiąc, okresowy zapis dwójkowy liczby x ma postać:

$$x = (x_0, x_1 x_2 \dots x_s x_{s+1} \dots x_{s+p-1} x_s x_{s+1} \dots x_{s+p-1} x_s x_{s+1} \dots x_{s+p-1} \dots)_2.$$

Liczbę p nazywamy *długością* okresu w tym zapisie.

Jak z tego widać, zapis dwójkowy skończony uważamy za okresowy, podobnie jak zapis dwójkowy, w którym wszystkie cyfry od pewnego miejsca równe są 1.

Następujący fakt jest dobrze znany.

Fakt 5.3. Liczba $x \in [0, 1]$ ma zapis dwójkowy okresowy wtedy i tylko wtedy, gdy jest liczbą wymierną.

Zajmiemy się teraz działaniem dynamiki 2-adycznej D_2 na liczby $x \in [0, 1]$ zapisane w postaci dwójkowej.

Definicja 5.3. Niech $x \in [0, 1]$. Standardowym zapisem dwójkowym liczby x nazywać będziemy jej zapis dwójkowy skończony w przypadku, gdy liczba x ma dwa zapisy dwójkowe (zob. lemat 5.2), lub jedyny istniejący zapis dwójkowy tej liczby w pozostałym przypadku.

Tak więc w standardowym zapisie dwójkowym nie zdarzy się nigdy tak, że wszystkie od pewnego miejsca cyfry zapisu będą równe 1.

Lemat 5.3. Niech $0 \leq x < \frac{1}{2}$ i niech $x = (0, 0x_2x_3\dots)_2$ będzie standardowym zapisem dwójkowym liczby x . Wtedy

$$D_2(x) = (0, x_2x_3x_4\dots)_2 \quad (5.5)$$

jest standardowym zapisem dwójkowym liczby $D_2(x)$.

Dowód. Niech (x_j) , $j \in \mathbb{N}_0$, będzie ciągiem kolejnych cyfr w standardowym zapisie dwójkowym liczby x . Wtedy $x_0 = x_1 = 0$ i ze wzoru (5.3) mamy

$$x = \sum_{j=0}^{\infty} \frac{x_j}{2^j} = \sum_{j=2}^{\infty} \frac{x_j}{2^j},$$

a z definicji 5.1 otrzymujemy:

$$D_2(x) = 2x = 2 \left(\sum_{j=2}^{\infty} \frac{x_j}{2^j} \right) = \sum_{j=2}^{\infty} \frac{x_j}{2^{j-1}} = \sum_{j=1}^{\infty} \frac{x_{j+1}}{2^j},$$

skąd już wynika wzór (5.5). Zapis dany tym wzorem jest oczywiście standardowy. \square

W analogiczny sposób możemy udowodnić następujący lemat.

Lemat 5.4. Niech $\frac{1}{2} \leq x < 1$ i niech $x = (0, 1x_2x_3\dots)_2$ będzie standardowym zapisem dwójkowym liczby x . Wtedy

$$D_2(x) = (0, x_2x_3x_4\dots)_2 \quad (5.6)$$

jest standardowym zapisem dwójkowym liczby $D_2(x)$.

Z lematów 5.3 i 5.4 wynika

Wniosek 5.1. Niech $0 \leq x < 1$ i niech (x_j) , $j \in \mathbb{N}_0$, będzie ciągiem kolejnych cyfr standardowego zapisu dwójkowego liczby x . Wtedy $x_0 = 0$ i (y_j) , $j \in \mathbb{N}_0$, gdzie $y_0 = 0$, $y_j = x_{j+1}$ dla każdego $j \in \mathbb{N}$ jest ciągiem kolejnych cyfr standardowego zapisu dwójkowego liczby $y = D_2(x)$.

Dynamika D_2 działa więc nadzwyczaj prosto – gdy $0 \leq x < 1$ i $x = (0, x_1 x_2 x_3 \dots)_2$ w postaci standardowej, to

$$D_2(x) = (0, x_2 x_3 x_4 \dots)_2 \quad (5.7)$$

jest standardowym zapisem dwójkowym liczby $D_2(x)$. Wyjątkiem jest $x = 1$, gdzie przyjmujemy, że $D_2(1) = 1$.

Wniosek 5.2. Układ dynamiczny $([0, 1], D_2)$ spełnia warunek [G].

Dowód. Niech A będzie dowolnym otwartym podzbiorem przedziału $[0, 1]$. Mamy pokazać, że istnieje punkt okresowy $x \in A$. Wybierzmy $r > 0$ i punkt $y \in A$ taki, że $(y - r, y + r) \subset A$. (Zatem $y \neq 1$). Niech (y_j) , $j \in \mathbb{N}_0$, będzie ciągiem kolejnych cyfr standardowego zapisu dwójkowego liczby y . Wtedy $y_0 = 0$, $y = (0, y_1 y_2 y_3 \dots)_2$. Niech p będzie liczbą naturalną tak dobraną, że $\frac{1}{2^p} < r$. Rozważmy liczbę x , której standardowy zapis dwójkowy jest okresowy o okresie długości p i ma postać: $x = (0, x_1 x_2 x_3 \dots)_2$, gdzie $x_0 = 0$ i dla każdego $j = pk - r$, $k \in \mathbb{N}$, $r \in \{0, \dots, p - 1\}$ kładziemy

$$x_j = x_{pk-r} = y_{p-r}.$$

Standardowy zapis dwójkowy liczby x ma więc postać

$$x = (0, y_1 y_2 \dots y_p y_1 y_2 \dots y_p y_1 y_2 \dots y_p \dots)_2$$

i ze wzoru (5.7) łatwo wynika, że $D_2^p(x) = x$. Punkt x jest więc okresowy. Poza tym mamy z (5.3):

$$y = \sum_{j=0}^{\infty} \frac{y_j}{2^j} = \sum_{j=1}^p \frac{y_j}{2^j} + \sum_{j=p+1}^{\infty} \frac{y_j}{2^j},$$

$$x = \sum_{j=0}^{\infty} \frac{x_j}{2^j} = \sum_{j=1}^p \frac{x_j}{2^j} + \sum_{j=p+1}^{\infty} \frac{x_j}{2^j} = \sum_{j=1}^p \frac{y_j}{2^j} + \sum_{j=p+1}^{\infty} \frac{x_j}{2^j},$$

więc

$$|y-x| = \left| \sum_{j=p+1}^{\infty} \frac{y_j}{2^j} - \sum_{j=p+1}^{\infty} \frac{x_j}{2^j} \right| \leq \sum_{j=p+1}^{\infty} \frac{|y_j - x_j|}{2^j} \leq \sum_{j=p+1}^{\infty} \frac{1}{2^j} = \frac{1}{2^p} < r$$

na mocy faktów 5.1 i f 4.2, gdyż $|y_j - x_j| \leq 1$ dla wszystkich $j \in \mathbb{N}_0$. Wobec tego $x \in (y-r, y+r) \subset A$. \square

Wniosek 5.3. Układ dynamiczny $([0, 1], D_2)$ spełnia warunek [T].

Dowód. Weźmy dwa zbiory otwarte $A, B \subset [0, 1]$ i niech $z \in B$. Wybierzmy $r > 0$ i punkt $y \in A$ taki, że $(y-r, y+r) \subset A$. (Zatem $y \neq 1$). Niech (z_j) , $j \in \mathbb{N}_0$, będzie ciągiem kolejnych cyfr standardowego zapisu dwójkowego liczby z . Podobnie, niech (y_j) , $j \in \mathbb{N}_0$, będzie ciągiem kolejnych cyfr standardowego zapisu dwójkowego liczby y . Wtedy $y_0 = 0$, $y = (0, y_1 y_2 y_3 \dots)_2$. Wybierzmy liczbę naturalną m tak dużą, że $\frac{1}{2^{m-1}} < r$. Rozważmy liczbę x , której standardowy zapis dwójkowy ma postać:

$$x = (0, y_1 y_2 \dots y_{m-1} z_1 z_2 z_3 \dots)_2$$

(czyli ciągiem kolejnych cyfr standardowego zapisu dwójkowego liczby x jest ciąg (x_j) , $j \in \mathbb{N}_0$, gdzie $x_0 = 0$, $x_j = y_j$ dla $j \in \{1, \dots, m-1\}$ i $x_j = z_{j-m+1}$ dla $j \geq m$). Rozumując podobnie jak w dowodzie wniosku 5.1 stwierdzamy łatwo, że $|y-x| < r$, czyli że $x \in (y-r, y+r) \subset A$. Zarazem z (5.7) łatwo wynika, że $D_2^m(x) = z$, czyli $D_2^m(x) \in B$. \square

Wniosek 5.4. Układ dynamiczny $([0, 1], D_2)$ spełnia warunek [W].

Dowód. Wybierzmy $x \in [0, 1]$ i otoczenie O_x punktu x . Rozpatrzmy przypadek, gdy $x \neq 1$ (dowód w przypadku gdy $x = 1$ jest łatwy). Wybierzmy liczbę $r > 0$ taką, że $(x-r, x+r) \subset O_x$ i niech (x_j) , $j \in \mathbb{N}_0$, będzie ciągiem kolejnych cyfr standardowego zapisu dwójkowego liczby x . Niech n będzie liczbą naturalną tak dużą, że $\frac{1}{2^n} < r$. Rozważmy liczbę y taką, że ciągiem kolejnych cyfr standardowego zapisu

dwójkowego tej liczby jest ciąg (y_j) , $j \in \mathbb{N}_0$, gdzie $y_j = x_j$ dla wszystkich $j \neq n$, natomiast

$$y_n = \begin{cases} 1, & \text{gdy } x_n = 0, \\ 0, & \text{gdy } x_n = 1. \end{cases}$$

Tak jak w dowodzie wniosku 5.1 stwierdzamy łatwo, że $|y - x| < r$, czyli że $y \in (x - r, x + r) \subset O_x$. Natomiast $|D_2^n(x) - D_2^n(y)| = \frac{1}{2}$. Oznacza to, że układ spełnia warunek [W] ze stałą wrażliwości $\frac{1}{2}$. \square

Wobec definicji 3.5 mamy oczywisty wniosek.

Wniosek 5.5. Układ dynamiczny $([0, 1], D_2)$ jest chaotyczny.

Przejdziemy teraz do drugiego przekształcenia przedziału $[0, 1]$.

Definicja 5.4. Funkcję $T : [0, 1] \rightarrow [0, 1]$ określoną wzorem

$$T(x) = 1 - |2x - 1| \tag{5.8}$$

dla wszystkich $x \in [0, 1]$ nazywamy funkcją namiotową.

Funkcja namiotowa dana wzorem (5.8) jest oczywiście ciągła. Dlatego wykazanie chaotyczności układu dynamicznego $([0, 1], T)$ jest prostsze, gdyż dowodu wymaga jedynie tranzytywność dynamiki T .

Pomocny tu jest następujący lemat.

Lemat 5.5. Dla wszystkich $n \in \mathbb{N}$, $n \geq 2$ zachodzi równość:

$$T^n = T \circ D_2^{n-1}. \tag{5.9}$$

Dowód. Zastosujemy indukcję. Bezpośrednim rachunkiem (z wykorzystaniem wzorów (5.1) i (5.8)) sprawdzamy, że $T^2 = T \circ D_2$, czyli że $T^2(x) = T(D_2(x))$ dla każdego $x \in [0, 1]$. Załóżmy, że $T^n = T \circ D_2^{n-1}$, gdzie $n \geq 2$. Wtedy mamy:

$$T^{n+1} = T \circ T^n = T \circ (T \circ D_2^{n-1}) = T^2 \circ D_2^{n-1} = (T \circ D_2) \circ D_2^{n-1} = T \circ D_2^n,$$

co dowodzi prawdziwości wzoru (5.9). \square

Lemat 5.6. Układ dynamiczny $([0, 1], T)$ spełnia warunek [T].

Dowód łatwo wynika z wniosku 5.3 i z lematu 5.5.

Wniosek 5.6. Układ dynamiczny $([0, 1], T)$ jest chaotyczny.

Dowód. Wystarczy skorzystać z definicji układu chaotycznego. □

Przejdziemy do trzeciego przekształcenia przedziału $[0, 1]$.

Definicja 5.5. Funkcję $V : [0, 1] \rightarrow [0, 1]$ określoną wzorem

$$V(x) = 4x(1 - x) \tag{5.10}$$

dla wszystkich $x \in [0, 1]$ nazywamy funkcją logistyczną.

Funkcja logistyczna dana wzorem (5.10) jest oczywiście ciągła. Dlatego, jak poprzednio, udowodnienie chaotyczności układu dynamicznego $([0, 1], V)$ będzie wymagało jedynie wykazania tranzytywności dynamiki V .

Pomocne tu są następujące lematy.

Lemat 5.7. Funkcja $h : [0, 1] \rightarrow [0, 1]$ określona wzorem

$$h(x) = \sin^2\left(\frac{\pi x}{2}\right) \tag{5.11}$$

dla wszystkich $x \in [0, 1]$ jest ciągła, rosnąca i przekształca przedział $[0, 1]$ wzajemnie jednoznacznie na siebie, a funkcja odwrotna $h^{-1} : [0, 1] \rightarrow [0, 1]$ jest również ciągła.

Lemat 5.8. Zachodzi równość:

$$V = h \circ T \circ h^{-1}, \tag{5.12}$$

gdzie funkcja h jest określona wzorem (5.11), funkcja T jest określona wzorem (5.8), a funkcja V jest określona wzorem (5.10).

Dowód. Ponieważ funkcja h jest wzajemnie jednoznaczna, więc dowodzona równość (5.12) jest równoważna równości $V \circ h = h \circ T$. Niech $x \in [0, \frac{1}{2}]$. Ze wzoru (5.8) łatwo wynika, że wtedy jest $T(x) = 2x$ i mamy:

$$\begin{aligned} V(h(x)) &= 4h(x)(1-h(x)) = 4\sin^2\left(\frac{\pi x}{2}\right)\left(1-\sin^2\left(\frac{\pi x}{2}\right)\right) = \\ &= 4\sin^2\left(\frac{\pi x}{2}\right)\cos^2\left(\frac{\pi x}{2}\right) = \sin^2(\pi x), \end{aligned}$$

a z drugiej strony

$$h(T(x)) = \sin^2\left(\frac{\pi T(x)}{2}\right) = \sin^2(\pi x),$$

więc $V(h(x)) = h(T(x))$. Gdy $x \in [0, \frac{1}{2}]$, rozumiemy podobnie. \square

Korzystając ze wzoru (5.12) możemy łatwo wyznaczać kolejne iteracje funkcji logistycznej, gdyż, jak łatwo spostrzec,

$$V^n = h \circ T^n \circ h^{-1} \tag{5.13}$$

dla wszystkich $n \in \mathbb{N}$. Wobec ciągłości funkcji h i h^{-1} bez trudu możemy teraz wykazać następujący lemat.

Lemat 5.9. Układ dynamiczny $([0, 1], V)$ spełnia warunek $[T]$.

Wniosek 5.7. Układ dynamiczny $([0, 1], V)$ jest chaotyczny.

Dowód. Wynika to z definicji układu chaotycznego. \square

Badane w tym paragrafie funkcje mogą się wydawać (i są) wybrajmi w sposób przypadkowy. Pewnym usprawiedliwieniem tak szczegółowego studiowania ich własności jest fakt (zob. np. [5], [13]), że zachowanie się tych dynamik jest w pewnym sensie „typowe” dla dużo szerszej klasy dynamik działających w przedziale na osi liczbowej.

6 Izometrie i podobieństwa w przestrzeniach euklidesowych

W tym paragrafie zajmujemy się pewnymi szczególnymi typami przekształceń $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ przestrzeni euklidesowej w przestrzeń euklidesową. Symbol $\mathbf{0}$ będzie oznaczać wektor zerowy przestrzeni \mathbb{R}^n , $\mathbf{0} = (0, 0, \dots, 0)$. Symbol \mathbf{I}_n będzie oznaczać macierz jednostkową stopnia n ,

$$\mathbf{I}_n = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{bmatrix}.$$

Niech $p \in \mathbb{R}^n$ i niech $r \in \mathbb{R}$, $r \neq 0$. *Jednokładnością* o środku w punkcie p i skali r nazywamy przekształcenie $J_{p,r} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ spełniające dla wszystkich $x \in \mathbb{R}^n$ warunek:

$$J_{p,r}(x) - p = r(x - p).$$

Mamy więc:

$$J_{p,r}(x) = (1 - r)p + rx \quad (6.1)$$

dla każdego $x \in \mathbb{R}^n$.

W przypadku gdy p jest wektorem zerowym, wzór (6.1) przyjmuje postać

$$\mathbf{J}_{\mathbf{0},r}(x) = rx,$$

jeżeli więc $x = (x_1, x_2, \dots, x_n)$, $y = (y_1, y_2, \dots, y_n)$ i $y = \mathbf{J}_{\mathbf{0},r}(x)$, to

$$\begin{cases} y_1 = rx_1, \\ y_2 = rx_2, \\ \dots \\ y_n = rx_n. \end{cases}$$

Przekształcenie $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ przestrzeni euklidesowej \mathbb{R}^n w przestrzeń euklidesową \mathbb{R}^m ($n, m \in \mathbb{N}$) nazywamy *izometrią*, jeżeli

$$\|F(x) - F(y)\| = \|x - y\|$$

Macierz A izometrii spełnia równości $A \cdot A^T = \mathbf{I}_n$ oraz $A^T \cdot A = \mathbf{I}_n$, czyli

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \cdot \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 \end{bmatrix},$$

$$\begin{bmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{bmatrix} \cdot \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Macierz kwadratową A stopnia n taką, że $A \cdot A^T = A^T \cdot A = \mathbf{I}_n$ nazywamy macierzą ortogonalną. Każda macierz ortogonalna jest nieosobliwa, tzn. $\det A \neq 0$ (gdzie $\det A$ jest wyznacznikiem macierzy A). Co więcej, $\det A = 1$ lub $\det A = -1$.

Jeżeli we wzorze (Iz) jest $w = \mathbf{0}$, to izometrię nazywamy *obrotem* przestrzeni \mathbb{R}^n wokół punktu $\mathbf{0}$. Obrót nazywamy *właściwym*, jeżeli $\det A = 1$, zaś *niewłaściwym*, gdy $\det A = -1$. Tak więc obrót $R_{\mathbf{0},A} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ przestrzeni \mathbb{R}^n wokół punktu $\mathbf{0}$ dany jest wzorem:

$$[R_{\mathbf{0},A}(x)] = A \cdot [x] \quad (6.3)$$

dla każdego $x \in \mathbb{R}^n$, gdzie A jest macierzą ortogonalną stopnia n .

Niech $w \in \mathbb{R}^n$ będzie ustalonym wektorem. *Translacją* (lub przesunięciem) przestrzeni \mathbb{R}^n o wektor w nazywamy przekształcenie $T_w : \mathbb{R}^n \rightarrow \mathbb{R}^n$ opisane wzorem:

$$T_w(x) = x + w \quad (6.4)$$

dla każdego $x \in \mathbb{R}^n$. Każda translacja jest izometrią. W postaci (6.2) (zob. fakt 6.1) translację T_w opisuje wzór

$$[T_w(x)] = \mathbf{I}_n \cdot [x] + [w],$$

macierz translacji jest więc macierzą jednostkową.

Nietrudno sprawdzić, że złożenie translacji przestrzeni \mathbb{R}^n o wektor u z translacją przestrzeni \mathbb{R}^n o wektor w ($u, w \in \mathbb{R}^n$) jest translacją tej przestrzeni o wektor $u + w$, czyli zachodzi równość

$$T_w \circ T_u = T_{u+w}.$$

Ponadto izometria odwrotna do translacji przestrzeni \mathbb{R}^n o wektor $w \in \mathbb{R}^n$ jest translacją tej przestrzeni o wektor $-w$, czyli

$$T_w^{-1} = T_{-w}.$$

Translacja $T_0 : \mathbb{R}^n \rightarrow \mathbb{R}^n$ przestrzeni \mathbb{R}^n o wektor zerowy jest przekształceniem tożsamościowym, $T_0(x) = x$ dla każdego $x \in \mathbb{R}^n$.

Możemy teraz uogólnić pojęcie obrotu przestrzeni euklidesowej. Ustalmy punkt $p \in \mathbb{R}^n$. *Obrotom* przestrzeni \mathbb{R}^n wokół punktu p nazywamy każde przekształcenie $R_{p,A} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ postaci:

$$R_{p,A} = T_p \circ R_{0,A} \circ T_{-p}, \quad (6.5)$$

gdzie $R_{0,A}$ jest dowolnym obrotem przestrzeni \mathbb{R}^n wokół punktu 0 .

Zajmiemy się bardziej szczegółowo izometriami płaszczyzny euklidesowej \mathbb{R}^2 . Można w tym przypadku opisać wszystkie macierze ortogonalne. Mianowicie, jeżeli A jest macierzą ortogonalną stopnia 2, to

$$A = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix} \quad \text{lub} \quad A = \begin{bmatrix} \cos \alpha & \sin \alpha \\ \sin \alpha & -\cos \alpha \end{bmatrix},$$

gdzie $\alpha \in \mathbb{R}$ jest pewną liczbą rzeczywistą. (Na odwrót, macierze mające powyższą postać są macierzami ortogonalnymi dla każdego $\alpha \in \mathbb{R}$).

Łatwo można sprawdzić, że

$$\det \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix} = 1, \quad \det \begin{bmatrix} \cos \alpha & \sin \alpha \\ \sin \alpha & -\cos \alpha \end{bmatrix} = -1.$$

Zgodnie z definicją obrotu właściwego, obrót właściwy płaszczyzny \mathbb{R}^2 wokół punktu 0 dany jest wzorem (6.3), gdzie

$$A = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix}$$

jest macierzą obrotu i $\alpha \in \mathbb{R}$. Taką izometrię będziemy nazywać *obrotem* płaszczyzny \mathbb{R}^2 o kąt α wokół punktu $\mathbf{0}$, i oznaczać symbolem $R_{\mathbf{0},\alpha}$

Zatem jeżeli $x, y \in \mathbb{R}^2$, $x = (x_1, x_2)$, $y = (y_1, y_2)$ i $y = R_\alpha(x)$, to

$$[R_\alpha(x)] = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad (6.6)$$

a stąd

$$\begin{cases} y_1 = x_1 \cos \alpha - x_2 \sin \alpha, \\ y_2 = x_1 \sin \alpha + x_2 \cos \alpha. \end{cases}$$

Podobnie, translację $T_w : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ płaszczyzny o wektor $w = (w_1, w_2)$ opisują równania

$$\begin{cases} y_1 = x_1 + w_1, \\ y_2 = x_2 + w_2. \end{cases}$$

Niech $K \subset \mathbb{R}^2$ będzie ustaloną prostą. *Symetrią osiową* (będziemy mówić krótko: symetrią) płaszczyzny euklidesowej \mathbb{R}^2 względem prostej k nazywamy przekształcenie $S_K : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ przyporządkowujące każdemu punktowi $x \in \mathbb{R}^2$ punkt $y = S_K(x)$ w taki sposób, że

- (a) punkty x i y leżą na prostej $K^\perp \subset \mathbb{R}^2$ prostopadłej do prostej K ,
- (b) punkt p będący punktem przecięcia prostych K i K^\perp jest środkiem odcinka o końcach x i y .

W prosty sposób można opisać symetrię osiową w przypadku, gdy prosta K jest osią OX układu współrzędnych na płaszczyźnie \mathbb{R}^2 . Mianowicie jeżeli $x, y \in \mathbb{R}^2$, $x = (x_1, x_2)$, $y = (y_1, y_2)$ i $y = S_{OX}(x)$, to

$$\begin{cases} y_1 = x_1, \\ y_2 = -x_2. \end{cases}$$

Analogicznie, dla symetrii osiowej S_{OY} względem osi OY mamy:

$$\begin{cases} y_1 = -x_1, \\ y_2 = x_2. \end{cases}$$

Wyprowadzimy teraz wzory opisujące symetrię względem dowolnej prostej na płaszczyźnie euklidesowej. Wybierzmy punkt $x \in \mathbb{R}^2$, $x = (x_1, x_2)$ i przypuśćmy, że prosta K ma równanie:

$$K : az_1 + bz_2 + c = 0. \quad (6.7)$$

Prosta K^\perp jest zbiorem punktów postaci $z = x + t(a, b)$, gdzie $t \in \mathbb{R}$ (gdyż wektor normalny (a, b) prostej K jest wektorem kierunkowym prostej K^\perp). W szczególności współrzędne punktu $p = (p_1, p_2)$ będącego punktem przecięcia prostych K i K^\perp są dane wzorami

$$\begin{cases} p_1 = x_1 + t_0a, \\ p_2 = x_2 + t_0b, \end{cases} \quad (6.8)$$

gdzie $t_0 \in \mathbb{R}$. Wstawiając te współrzędne do równania (6.7) otrzymujemy:

$$a(x_1 + t_0a) + b(x_2 + t_0b) + c = 0,$$

a stąd

$$t_0 = \frac{-ax_1 - bx_2 - c}{a^2 + b^2}.$$

Wstawiając to do równań (6.8) otrzymujemy po prostych rachunkach:

$$\begin{cases} p_1 = \frac{b^2x_1 - abx_2 - ac}{a^2 + b^2}, \\ p_2 = \frac{-abx_1 + a^2x_2 - bc}{a^2 + b^2}, \end{cases}$$

a ponieważ p jest środkiem odcinka o końcach x i y (gdzie $y = S_K(x)$, $y = (y_1, y_2)$ jest obrazem punktu x w symetrii S_K), zatem $p_1 = \frac{x_1 + y_1}{2}$, $p_2 = \frac{x_2 + y_2}{2}$, czyli $y_1 = 2p_1 - x_1$ i $y_2 = 2p_2 - x_2$. Ostatecznie

$$\begin{cases} y_1 = \left(\frac{b^2 - a^2}{a^2 + b^2}\right)x_1 + \left(\frac{-2ab}{a^2 + b^2}\right)x_2 + \left(\frac{-2ac}{a^2 + b^2}\right), \\ y_2 = \left(\frac{-2ab}{a^2 + b^2}\right)x_1 + \left(\frac{a^2 - b^2}{a^2 + b^2}\right)x_2 + \left(\frac{-2bc}{a^2 + b^2}\right). \end{cases}$$

Powyzsze równości można zapisać w postaci:

$$[S_K(x)] = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \frac{b^2 - a^2}{a^2 + b^2}, & \frac{-2ab}{a^2 + b^2} \\ \frac{-2ab}{a^2 + b^2}, & \frac{a^2 - b^2}{a^2 + b^2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} \frac{-2ac}{a^2 + b^2} \\ \frac{-2bc}{a^2 + b^2} \end{bmatrix}. \quad (6.9)$$

Łatwo można sprawdzić, że macierz

$$A = \begin{bmatrix} \frac{b^2 - a^2}{a^2 + b^2}, & \frac{-2ab}{a^2 + b^2} \\ \frac{-2ab}{a^2 + b^2}, & \frac{a^2 - b^2}{a^2 + b^2} \end{bmatrix}$$

jest ortogonalna, więc (zob. fakt 6.1) każda symetria osiowa S_K jest izometrią.

Przejdziemy teraz do określenia ważnej klasy przekształceń przestrzeni euklidesowej \mathbb{R}^n .

Definicja 6.1. Niech $r \in \mathbb{R}$, $r \neq 0$. Podobieństwem o skali r w przestrzeni euklidesowej \mathbb{R}^n nazywamy każde przekształcenie $P : \mathbb{R}^n \rightarrow \mathbb{R}^n$ postaci:

$$P = F_1 \circ F_2 \circ \dots \circ F_m$$

(gdzie $m \in \mathbb{N}$), przy czym dokładnie jedno spośród przekształceń F_j , $j \in \{1, \dots, m\}$, jest jednokładnością $J_{p,r}$ o skali r i o dowolnym środku $p \in \mathbb{R}$, zaś pozostałe są izometriami w przestrzeni \mathbb{R}^n .

Zachodzi następujący fakt.

Fakt 6.2. Niech $P : \mathbb{R}^n \rightarrow \mathbb{R}^n$ będzie podobieństwem o skali $r \neq 0$. Istnieją wtedy: macierz ortogonalna A stopnia n oraz wektor $w \in \mathbb{R}^n$ takie, że

$$[P(x)] = (rA) \cdot [x] + [w] \quad (6.10)$$

dla każdego $x \in \mathbb{R}^n$. Na odwrót, każde przekształcenie przestrzeni \mathbb{R}^n w siebie powyższej postaci jest podobieństwem o skali r . (Macierz rA nazywamy macierzą podobieństwa P).

Każde podobieństwo $P : \mathbb{R}^n \rightarrow \mathbb{R}^n$ jest funkcją różnowartościową i przekształca przestrzeń \mathbb{R}^n na całą przestrzeń \mathbb{R}^n .

Jeżeli zatem $A = [a_{ij}]$, $i, j \in \{1, \dots, n\}$, $x \in \mathbb{R}^n$ i $y = P(x)$, $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n)$, to

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} ra_{11} & ra_{12} & \cdots & ra_{1n} \\ ra_{21} & ra_{22} & \cdots & ra_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ ra_{n1} & ra_{n2} & \cdots & ra_{nn} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}.$$

Zauważmy, że każda izometria jest podobieństwem (wystarczy we wzorze (6.10) przyjąć $r = 1$).

7 Samopodobieństwo. Fraktale topologiczne

Niech (X, d) będzie przestrzenią metryczną. Przypomnijmy (zob. par. 2), że przekształcenie $T : X \rightarrow X$ nazywamy kontrakcją, jeżeli istnieje stała $L \in \mathbb{R}$ taka, że $0 \leq L < 1$ i nierówność

$$d(T(x), T(y)) \leq Ld(x, y) \quad (7.1)$$

zachodzi dla wszystkich punktów $x, y \in X$. Liczbę L nazywamy stałą kontrakcji T .

Jeżeli $T : X \rightarrow X$ jest dowolnym przekształceniem (niekoniecznie będącym kontrakcją) i $x_0 \in X$, to punkt x_0 nazywamy *punktem stałym* przekształcenia T , jeżeli $T(x_0) = x_0$.

Będziemy rozważać przestrzenie zupełne. Poniższe twierdzenie jest znane jako twierdzenie Banacha o punkcie stałym (spotyka się też nazwę: „Zasada odwzorowań zwężających”), zob. np. [6].

Twierdzenie 7.1. *Jeżeli (X, d) jest przestrzenią metryczną zupełną, to każda kontrakcja $T : X \rightarrow X$ ma dokładnie jeden punkt stały $x_0 \in X$. Dla każdego $x \in X$ ciąg iteracyjny $(T^n(x))$, $n \in \mathbb{N}_0$, jest zbieżny i zachodzi równość:*

$$\lim_{n \rightarrow \infty} T^n(x) = x_0.$$

Jeżeli $L \in [0, 1)$ jest stałą kontrakcji T , to dla każdego $x \in X$ mamy:

$$d(T^n(x), x_0) \leq \left(\frac{L^n}{1-L} \right) d(x, T(x)).$$

Dowód. Niech (X, d) będzie przestrzenią metryczną zupełną oraz załóżmy, że $T : X \rightarrow X$ jest kontrakcją ze stałą L , $0 \leq L < 1$. Wybierzmy dowolny punkt $x \in X$. Nierówność (7.1) daje:

$$d(T(x), T^2(x)) = d[T(x), T(T(x))] \leq Ld(x, T(x)),$$

$$d(T^2(x), T^3(x)) = d[T(T(x)), T(T^2(x))] \leq L^2 d(x, T(x)),$$

$$d(T^3(x), T^4(x)) = d[T(T^2(x)), T(T^3(x))] \leq L^3 d(x, T(x)),$$

i stosując indukcję łatwo stwierdzamy, że dla każdego $n \in \mathbb{N}$ jest

$$d(T^n(x), T^{n+1}(x)) \leq L^n d(x, T(x)). \quad (7.2)$$

Wybermy dowolne liczby $n, p \in \mathbb{N}$. Metryka d spełnia warunek trójkąta: $d(u, w) \leq d(u, v) + d(v, w)$ dla wszystkich $u, v, w \in X$. Stąd i z (7.2) mamy:

$$d(T^n(x), T^{n+1}(x)) \leq L^n d(x, T(x)),$$

$$\begin{aligned} d(T^n(x), T^{n+2}(x)) &\leq d(T^n(x), T^{n+1}(x)) + d(T^{n+1}(x), T^{n+2}(x)) \leq \\ &\leq L^n d(x, T(x)) + L^{n+1} d(x, T(x)) = L^n (1 + L) d(x, T(x)), \end{aligned}$$

$$\begin{aligned} d(T^n(x), T^{n+3}(x)) &\leq d(T^n(x), T^{n+2}(x)) + d(T^{n+2}(x), T^{n+3}(x)) \leq \\ &\leq L^n (1 + L) d(x, T(x)) + L^{n+2} d(x, T(x)) = L^n (1 + L + L^2) d(x, T(x)), \end{aligned}$$

więc ponownie stosując indukcję stwierdzamy, że

$$d(T^n(x), T^{n+p}(x)) \leq L^n \left(\sum_{k=0}^{p-1} L^k \right) d(x, T(x)) \leq L^n \left(\sum_{k=0}^{\infty} L^k \right) d(x, T(x)),$$

czyli

$$d(T^n(x), T^{n+p}(x)) \leq \left(\frac{L^n}{1-L} \right) d(x, T(x)), \quad (7.3)$$

gdyż szereg $\sum_{k=0}^{\infty} L^k$ jest szeregiem geometrycznym zbieżnym i $\sum_{k=0}^{\infty} L^k = \frac{1}{1-L}$.

Ponieważ ciąg (L^n) , $n \in \mathbb{N}$, dąży do zera przy $n \rightarrow \infty$, więc także

$$\lim_{n \rightarrow \infty} \left(\frac{L^n}{1-L} \right) d(x, T(x)) = 0,$$

a to oznacza, że dla każdej liczby $\varepsilon > 0$ istnieje $n_0 \in \mathbb{N}$ takie, że dla wszystkich $n \geq n_0$ jest

$$\left(\frac{L^n}{1-L} \right) d(x, T(x)) < \varepsilon. \quad (7.4)$$

Jeżeli więc $n, m \in \mathbb{N}$, $n \geq n_0$, $m \geq n_0$ i np. $n < m$, to $m = n + p$ i z (7.3) oraz (7.4) otrzymujemy:

$$d(T^n(x), T^m(x)) < \varepsilon.$$

(Jeżeli $n > m$, rozumiemy podobnie). Oznacza to, że ciąg $(T^n(x))$, $n \in \mathbb{N}$, spełnia warunek Cauchy'ego, a ponieważ przestrzeń (X, d) jest z założenia zupełna, więc ciąg ten jest zbieżny. Niech

$$x_0 = \lim_{n \rightarrow \infty} T^n(x).$$

Wobec ciągłości przekształcenia T jest

$$T(x_0) = T\left(\lim_{n \rightarrow \infty} T^n(x)\right) = \lim_{n \rightarrow \infty} T(T^n(x)) = \lim_{n \rightarrow \infty} T^{n+1}(x) = x_0,$$

czyli x_0 jest punktem stałym przekształcenia T . Przekształcenie to nie może mieć innych punktów stałych, bo jeżeli np. $T(y_0) = y_0$ to z nierówności (7.1) otrzymujemy:

$$d(x_0, y_0) = d(T(x_0), T(y_0)) \leq Ld(x_0, y_0),$$

więc ponieważ $0 \leq L < 1$, to musi być $d(x_0, y_0) = 0$, czyli $x_0 = y_0$.

Na koniec przechodząc w nierówności (7.3) do granicy przy $p \rightarrow \infty$ dostajemy dla każdego $n \in \mathbb{N}$:

$$\lim_{p \rightarrow \infty} d(T^n(x), T^{n+p}(x)) = d(T^n(x), x_0) \leq \left(\frac{L^n}{1-L}\right) d(x, T(x)).$$

Dowód jest zakończony. □

Niech teraz $(X, d) = (\mathbb{R}^N, d)$ będzie przestrzenią euklidesową N -wymiarową.

Definicja 7.1. Niech A będzie niepustym podzbiorem przestrzeni euklidesowej \mathbb{R}^n i niech $x \in \mathbb{R}^n$. Odległość punktu x od zbioru A określa wzór:

$$\text{dist}(x, A) = \inf_{a \in A} d(x, a) = \inf_{a \in A} \|x - a\|.$$

Odległość punktu x od zbioru A w przestrzeni euklidesowej jest zawsze skończoną liczbą nieujemną i, jak nietrudno wykazać, $\text{dist}(x, A) = 0$ wtedy i tylko wtedy, gdy $x \in \bar{A}$. Zauważmy, że w takim razie odległość punktu x od zbioru A jest dodatnia wtedy i tylko wtedy, gdy $x \notin \bar{A}$.

Fakt 7.1. Niech A będzie niepustym podzbiorem przestrzeni euklidesowej \mathbb{R}^n . Funkcja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ określona dla $x \in \mathbb{R}^n$ wzorem

$$f(x) = \text{dist}(x, A)$$

spełnia warunek Lipschitza ze stałą 1, jest więc funkcją ciągłą (zob. fakt 2.8).

Ważne własności przestrzeni euklidesowych wyrażają dobrze znane stwierdzenia.

Fakt 7.2. (a) Przestrzeń euklidesowa jest zupełna.

(b) W przestrzeni euklidesowej zbiór jest zwarty wtedy i tylko wtedy, gdy jest domknięty i ograniczony.

Niech A będzie niepustym podzbiorem przestrzeni euklidesowej \mathbb{R}^n . Wybierzmy dowolną liczbę $r \geq 0$. Kulą domkniętą o środku w zbiorze A i promieniu r nazywać będziemy zbiór

$$K[A; r] = \{x \in \mathbb{R}^n : \text{dist}(x, A) \leq r\}.$$

(Używa się także nazwy: otoczka zbioru). Kulą otwartą o środku w zbiorze A i promieniu r nazywać będziemy zbiór

$$K(A; r) = \{x \in \mathbb{R}^n : \text{dist}(x, A) < r\}.$$

Kula otwarta o środku w zbiorze A i dodatnim promieniu jest zawsze zbiorem otwartym, natomiast kula domknięta o środku w zbiorze A i dodatnim promieniu nie musi być zbiorem domkniętym. Jeżeli jednak $A \subset \mathbb{R}^n$ jest niepustym zbiorem zwartym i $r \geq 0$, to $K[A; r]$ jest zbiorem domkniętym, a nawet zwartym.

W dalszym ciągu tego paragrafu symbolem \mathcal{K}_n będziemy oznaczać rodzinę wszystkich zwartych i niepustych podzbiorów przestrzeni euklidesowej \mathbb{R}^n .

Definicja 7.2. Niech $X, Y \in \mathcal{K}_n$. Odległością Hausdorffa zbiorów X i Y nazywamy liczbę

$$d_H(X, Y) = \max \left\{ \max_{x \in X} \text{dist}(x, Y), \max_{y \in Y} \text{dist}(y, X) \right\}.$$

Zauważmy, że powyższy wzór jest poprawny, gdyż dla ustalonego zbioru $Y \in \mathcal{K}_n$ funkcja $f(x) = \text{dist}(x, Y)$ jest ciągła (zob. fakt 7.1), więc w zbiorze zwanym $X \in \mathcal{K}_n$ jest ona ograniczona i osiąga swoje kresy (zob. fakt 2.7). Dlatego też

$$\sup_{x \in X} \text{dist}(x, Y) = \max_{x \in X} \text{dist}(x, Y)$$

i, z podobnych powodów,

$$\sup_{y \in Y} \text{dist}(y, X) = \max_{y \in Y} \text{dist}(y, X).$$

Przypomnimy jeszcze jedną ważną własność dotyczącą podzbiorów dowolnej przestrzeni metrycznej (a więc w szczególności podzbiorów przestrzeni euklidesowej).

Fakt 7.3. *W przestrzeni metrycznej suma dowolnej rodziny zbiorów otwartych jest zbiorem otwartym. Natomiast iloczyn dowolnej rodziny zbiorów domkniętych jest zbiorem domkniętym.*

Lemat 7.1. *Funkcja d_H określona w definicji 7.2 jest metryką w rodzinie \mathcal{K}_n wszystkich zwartych i niepustych podzbiorów przestrzeni euklidesowej R^n . (Metryka d_H zwana jest metryką Hausdorffa).*

Dowód. Nieujemność funkcji d_H jest oczywista. Jeżeli $X \in \mathcal{K}_n$, to $\text{dist}(x, X) = 0$ dla każdego $x \in X$, więc

$$d_H(X, X) = \max \left\{ \max_{x \in X} \text{dist}(x, X), \max_{x \in X} \text{dist}(x, X) \right\} = \max \{0, 0\} = 0.$$

Niech $X, Y \in \mathcal{K}_n$ i przypuścimy, że $d_H(X, Y) = 0$. Wtedy

$$\max_{x \in X} \text{dist}(x, Y) = 0 = \max_{y \in Y} \text{dist}(y, X).$$

Oznacza to, że dla każdego $x \in X$ jest $\text{dist}(x, Y) = 0$ i dla każdego $y \in Y$ jest $\text{dist}(y, X) = 0$. Zbiory X, Y są z założenia domknięte. Z pierwszej z tych równości wynika, że $x \in Y$ dla każdego $x \in X$, zaś z drugiej równości wynika, że $y \in X$ dla każdego $y \in Y$. Więc $X = Y$.

Jeżeli $X, Y \in \mathcal{K}_n$, to równość $d_H(X, Y) = d_H(Y, X)$ jest oczywista.

Pozostaje do udowodnienia nierówność trójkąta:

$$d_H(X, Z) \leq d_H(X, Y) + d_H(Y, Z)$$

dla wszystkich zbiorów $X, Y, Z \in \mathcal{K}_n$. Ustalmy dowolnie wybrany punkt $z \in Z$. Istnieje punkt $y_0 \in Y$ taki, że $\|y_0 - z\| = \text{dist}(z, Y)$, gdyż funkcja $f(y) = \|y - z\|$ jest ciągła, więc w zbiorze zwartym Y jest ona ograniczona i osiąga swoje kresy. Z podobnych powodów istnieje $x_0 \in X$ taki, że $\|x_0 - y_0\| = \text{dist}(y_0, X)$. Nierówność trójkąta dla normy daje: $\|z - x_0\| \leq \|x_0 - y_0\| + \|y_0 - z\|$, czyli

$$\|z - x_0\| \leq \text{dist}(y_0, X) + \text{dist}(z, Y),$$

a ponieważ

$$\text{dist}(y_0, X) \leq \max_{y \in Y} \text{dist}(y, X) \leq d_H(X, Y), \quad \text{dist}(z, Y) \leq \max_{z' \in Z} \text{dist}(z', Y) \leq d_H(Y, Z),$$

więc

$$\|z - x_0\| \leq d_H(X, Y) + d_H(Y, Z).$$

Z drugiej strony $\text{dist}(z, X) = \inf_{x \in X} \|z - x\| \leq \|z - x_0\|$. Wykazaliśmy, że dla każdego $z \in Z$ zachodzi nierówność

$$\text{dist}(z, X) \leq d_H(X, Y) + d_H(Y, Z).$$

W podobny sposób pokazujemy, że dla każdego $x \in X$

$$\text{dist}(x, Z) \leq d_H(X, Y) + d_H(Y, Z).$$

Stąd,

$$d_H(X, Z) = \max \left\{ \max_{x \in X} \text{dist}(x, Z), \max_{z \in Z} \text{dist}(z, X) \right\} \leq d_H(X, Y) + d_H(Y, Z).$$

□

Lemat 7.2. *Jeżeli $(A_n)_{n \in \mathbb{N}}$ jest zstępującym (tzn, takim, że $A_{n+1} \subset A_n$ dla każdego $n \in \mathbb{N}$) ciągiem zbiorów z rodziny \mathcal{K}_n , to*

$$(a). \bigcap_{n=1}^{\infty} A_n \neq \emptyset.$$

(b) Zbiór $\bigcap_{n=1}^{\infty} A_n$ należy do rodziny \mathcal{K}_n .

$$(c) \lim_{k \rightarrow \infty} d_H \left(A_k, \bigcap_{n=1}^{\infty} A_n \right) = 0.$$

Dowód. (a) Wybierzmy z każdego ze zbiorów A_n (w sposób dowolny) punkt x_n . Otrzymujemy w ten sposób ciąg (x_n) taki, że $x_n \in A_n$ dla każdego $n \in \mathbb{N}$. Ponieważ wszystkie zbiory A_n zawierają się w zbiorze A_1 (ciąg jest zstępujący), więc (x_n) jest ciągiem punktów ze zbioru zwartego A_1 i można z niego wybrać podciąg zbieżny (x_{n_k}) (zob. par. 2). Niech $x = \lim_{k \rightarrow \infty} x_{n_k}$. Pokażemy, że $x \in A_n$ dla każdego $n \in \mathbb{N}$. Przypuśćmy, że tak nie jest, czyli, że istnieje liczba naturalna m , dla której $x \notin A_m$. Zbiór A_m jest domknięty (gdyż jest zwarty), więc zachodzi nierówność

$$\text{dist}(x, A_m) = d > 0. \quad (7.5)$$

Ciąg wskaźników (n_k) jest rosnący, zatem istnieje $k_0 \in \mathbb{N}$ takie, że $n_k > m$ dla wszystkich $k \geq k_0$. Wszystkie zbiory A_{n_k} , gdzie $k \geq k_0$, są podzbiorem zbioru A_m , więc z (7.5) otrzymujemy:

$$\text{dist}(x, A_{n_k}) \geq d > 0.$$

Ale $x_{n_k} \in A_{n_k}$, tym bardziej więc $\|x_{n_k} - x\| \geq d > 0$ dla wszystkich $k \geq k_0$. Jest to sprzeczne z faktem, że $x = \lim_{k \rightarrow \infty} x_{n_k}$.

Pokazaliśmy, że $x \in \bigcap_{n=1}^{\infty} A_n$, więc zbiór $\bigcap_{n=1}^{\infty} A_n$ jest niepusty.

(b) Zbiór $\bigcap_{n=1}^{\infty} A_n$ jest domknięty jako iloczyn rodziny zbiorów domkniętych A_n (zob. fakt 7.3). Jest to zbiór ograniczony jako podzbiór zbioru ograniczonego A_1 . Zatem (zob. fakt 7.2) $\bigcap_{n=1}^{\infty} A_n$ jest zbiorem zwartym. Na mocy punktu (a) zbiór ten jest niepusty, czyli jest elementem rodziny \mathcal{K}_n .

(c) Niech $C = \bigcap_{n=1}^{\infty} A_n$. Ponieważ $C \subset A_n$, więc

$$d_H(A_n, C) = \max_{x \in A_n} \text{dist}(x, C) \quad (7.6)$$

dla wszystkich $n \in \mathbb{N}$. Przypuśćmy, że dowodzona równość (c) nie zachodzi. Istnieje zatem liczba $r > 0$ i taki podciąg (A_{n_k}) ciągu (A_n) , że dla każdego $k \in \mathbb{N}$ mamy: $d_H(A_{n_k}, C) > r$. Wobec równości (7.6), w każdym zbiorze A_{n_k} ($k \in \mathbb{N}$) możemy znaleźć punkt x_k taki, że

$$\text{dist}(x_k, C) > r. \quad (7.7)$$

Ciąg (x_k) jest ciągiem punktów ze zbioru zwartego A_1 (gdyż wszystkie zbiory A_{n_k} zawierają się w tym zbiorze), więc można wybrać z niego podciąg zbieżny (x_{k_j}) . Niech $x = \lim_{j \rightarrow \infty} x_{k_j}$. Wobec ciągłości funkcji $x \rightarrow \text{dist}(x, C)$ z nierówności (7.7) wynika, że $\text{dist}(x, C) \geq r > 0$. Z drugiej strony rozumując podobnie jak w punkcie (a) możemy dowieść, że $x \in \bigcap_{n=1}^{\infty} A_n$, czyli $x \in C$ i $\text{dist}(x, C) = 0$. Otrzymana sprzeczność oznacza, że zachodzi dowodzona równość. \square

Twierdzenie 7.2. . *Przestrzeń metryczna (\mathcal{X}_n, d_H) jest przestrzenią zupełną.*

Dowód. Niech (A_n) będzie ciągiem spełniającym warunek Cauchy'ego w przestrzeni (\mathcal{X}_n, d_H) . Mamy wykazać, że ciąg ten jest zbieżny. Przyjmijmy dla każdego $n \in \mathbb{N}$

$$B_n = \overline{\bigcup_{k=n}^{\infty} A_k}, \quad B = \bigcap_{n=1}^{\infty} B_n.$$

Pokażemy, że $B \in \mathcal{X}_n$ i $\lim_{n \rightarrow \infty} A_n = B$ w przestrzeni (\mathcal{X}_n, d_H) , czyli, że $\lim_{n \rightarrow \infty} d_H(B, A_n) = 0$. Wszystkie zbiory A_n , a w konsekwencji także wszystkie zbiory B_n zawierają się w pewnym ograniczonym zbiorze z przestrzeni \mathbb{R}^n , są więc zbiorami ograniczonymi. Istotnie, ponieważ ciąg (A_n) spełnia warunek Cauchy'ego, więc istnieje liczba $n_0 \in \mathbb{N}$ taka, że dla wszystkich $n, m \geq n_0$ jest $d_H(A_n, A_m) < 1$. Stąd mamy dla każdego $n \in \mathbb{N}$:

$$A_n \subset \overline{A_n} \subset \overline{\bigcup_{k=1}^{\infty} A_k} = \overline{\bigcup_{k=1}^{n_0} A_k} \cup \overline{\bigcup_{k=n_0}^{\infty} A_k} \subset \bigcup_{k=1}^{n_0} A_k \cup K[A_{n_0}; 1] =$$

$$= \bigcup_{k=1}^{n_0} \overline{A_k \cup K[A_{n_0}; 1]} = \bigcup_{k=1}^{n_0} A_k \cup K[A_{n_0}; 1],$$

gdyż zbiory A_k i $K[A_{n_0}; 1]$ są domknięte.

Zbiory B_n są z definicji domknięte, więc są zwarte (zob. fakt 7.2). Bez trudu możemy sprawdzić, że ciąg (B_n) jest ciągiem zstępującym. Na podstawie lematu 7.2 mamy: $B \in \mathcal{K}_n$ oraz

$$\lim_{k \rightarrow \infty} d_H(B_k, B) = 0. \quad (7.8)$$

Wybermy dowolną liczbę $\varepsilon > 0$. Wobec równości (7.8) istnieje $n_1 \in \mathbb{N}$ takie, że dla $n \geq n_1$ mamy $d_H(B_n, B) < \varepsilon$, czyli $B_n \subset K[B, \varepsilon]$. Z warunku Cauchy'ego wynika istnienie takiej liczby $n_2 \in \mathbb{N}$, że dla $n, k \geq n_2$ mamy $d_H(A_n, A_k) < \varepsilon$, a więc $A_k \subset K[A_n, \varepsilon]$. Niech $n_0 = \max\{n_1, n_2\}$. Dla $n \geq n_0$ będzie wtedy

$$A_n \subset \overline{\bigcup_{k=n}^{\infty} A_k} = B_n \subset K[B, \varepsilon]$$

oraz

$$B \subset \overline{\bigcup_{k=n}^{\infty} A_k} \subset K[A_n, \varepsilon].$$

Stąd wynika, że $d_H(A_n, B) \leq \varepsilon$. Wykazaliśmy więc, że $\lim_{n \rightarrow \infty} A_n = B$ w przestrzeni (\mathcal{K}_n, d_H) , co kończy dowód. \square

Lemat 7.3. *Jeżeli $F : \mathbb{R}^N \rightarrow \mathbb{R}^N$ jest przekształceniem ciągłym, zaś $A \in \mathcal{K}_N$ jest zbiorem niepustym i zwartym, to jego obraz $F(A)$ w przekształceniu F jest zbiorem niepustym i zwartym.*

Dowód. Wybiermy dowolny ciąg (y_n) , $n \in \mathbb{N}$, punktów ze zbioru $F(A)$. Mamy pokazać, że z ciągu tego można wybrać podciąg zbieżny, którego granica należy do $F(A)$.

Dla każdego $n \in \mathbb{N}$, $y_n = F(x_n)$, gdzie $x_n \in A$. Ciąg (x_n) jest ciągiem punktów ze zwartego zbioru A , istnieje więc podciąg (x_{n_k}) , $k \in \mathbb{N}$, zbieżny do pewnego punktu $x_0 \in A$. Niech $y_0 = F(x_0)$. Wtedy $y_0 \in F(A)$ i wobec ciągłości przekształcenia F mamy:

$$\lim_{k \rightarrow \infty} y_{n_k} = \lim_{k \rightarrow \infty} F(x_{n_k}) = F\left(\lim_{k \rightarrow \infty} x_{n_k}\right) = F(x_0) = y_0.$$

Dowód został zakończony. □

Niech $\{F_1, F_2, \dots, F_m\}$ będzie rodziną przekształceń ciągłych $F_k : \mathbb{R}^N \rightarrow \mathbb{R}^N$ w przestrzeni euklidesowej \mathbb{R}^N . Jeżeli $A \in \mathcal{K}_N$, to każdy spośród zbiorów $F_k(A)$, gdzie $k \in \{1, \dots, m\}$, jest zwarty. Zauważmy, że w takim razie zbiór

$$\bigcup_{k=1}^m F_k(A)$$

także jest zwarty. Jest on bowiem domknięty, a ponieważ każdy ze zbiorów $F_k(A)$ jest ograniczony, to ich suma także jest zbiorem ograniczonym. Zbiór domknięty i ograniczony w przestrzeni euklidesowej \mathbb{R}^N jest zwarty.

Definicja 7.3. Niech $\{F_1, F_2, \dots, F_m\}$ będzie rodziną przekształceń ciągłych $F_k : \mathbb{R}^N \rightarrow \mathbb{R}^N$ w przestrzeni euklidesowej \mathbb{R}^N . Operatorem Barnsleya-Hutchinsona wyznaczonym przez rodzinę $\{F_1, F_2, \dots, F_m\}$ nazywamy funkcję $\mathcal{H}_{F_1, F_2, \dots, F_m} : \mathcal{K}_N \rightarrow \mathcal{K}_N$ przyporządkowującą każdemu zbiorowi $X \in \mathcal{K}_N$ zbiór

$$\mathcal{H}_{F_1, F_2, \dots, F_m}(X) = \bigcup_{k=1}^m F_k(X). \quad (7.9)$$

Zauważmy, że układ $(\mathcal{K}_N, \mathcal{H}_{F_1, F_2, \dots, F_m})$ jest układem dynamicznym w sensie paragrafu 2, którego przestrzenią fazową jest rodzina \mathcal{K}_N .

Wszystkie powyższe spostrzeżenia dotyczą także kontrakcji, ponieważ w przestrzeni (\mathbb{R}^N, d) każda kontrakcja jest przekształceniem ciągłym. Jeżeli $W : \mathbb{R}^N \rightarrow \mathbb{R}^N$ jest kontrakcją ze stałą L , to warunek (7.1) przybiera postać:

$$\|W(x) - W(y)\| \leq L \|x - y\| \quad (7.10)$$

dla wszystkich $x, y \in \mathbb{R}^N$.

Lemat 7.4. Jeżeli $W : \mathbb{R}^N \rightarrow \mathbb{R}^N$ jest kontrakcją ze stałą L , to dla dowolnego zbioru niepustego $A \subset \mathbb{R}^N$ i każdego punktu $x \in \mathbb{R}^N$ zachodzi nierówność:

$$\text{dist}(W(x), W(A)) \leq L \cdot \text{dist}(x, A).$$

Dowód. Niech $A' = W(A)$, $x' = W(x)$. Dla każdego $a' \in A'$ mamy

$$\text{dist}(x', A') = \inf_{y \in A'} \|x' - y\| \leq \|x' - a'\|.$$

Jeżeli jednak $a' \in A'$, to $a' = W(a)$ dla pewnego $a \in A$. Ze wzoru (7.10) dostajemy:

$$\|x' - a'\| = \|W(x) - W(a)\| \leq L\|x - a\|.$$

Gdy a' przebiega zbiór A' , to a przebiega zbiór A , wobec czego

$$\text{dist}(x', A') \leq L \cdot \inf_{a \in A} \|x - a\|,$$

czyli $\text{dist}(x', A') \leq L \cdot \text{dist}(x, A)$. □

Udowodnimy teraz podstawowe twierdzenie w rozważanej teorii. Niech d_H oznacza metrykę Hausdorffa w przestrzeni \mathcal{K}_N .

Twierdzenie 7.3. *Niech $\{W_1, W_2, \dots, W_m\}$ będzie rodziną kontrakcji $W_k : \mathbb{R}^N \rightarrow \mathbb{R}^N$ w przestrzeni euklidesowej \mathbb{R}^N . Dla każdego $k \in \{1, \dots, m\}$ niech r_k będzie stałą kontrakcji W_k . Wtedy operator Barnsleya-Hutchinsona wyznaczony przez rodzinę $\{W_1, W_2, \dots, W_m\}$ jest kontrakcją w przestrzeni metrycznej (\mathcal{K}_N, d_H) , ze stałą*

$$r = \max\{r_1, r_2, \dots, r_m\}.$$

Dowód. Wybierzmy dowolne zbiory $X, Y \in \mathcal{K}_N$ i niech $\mathcal{H}_{W_1, W_2, \dots, W_m}$ będzie operatorem Barnsleya-Hutchinsona wyznaczonym przez rodzinę $\{W_1, W_2, \dots, W_m\}$. Wprowadźmy oznaczenia:

$$X' = \mathcal{H}_{W_1, W_2, \dots, W_m}(X), \quad Y' = \mathcal{H}_{W_1, W_2, \dots, W_m}(Y).$$

Mamy wykazać, że $d_H(X', Y') \leq r d_H(X, Y)$. Wybierzmy dowolny punkt $x' \in X'$. Ze wzoru (7.9) wynika, że $x' \in W_k(X)$ dla pewnego wskaźnika $k \in \{1, \dots, m\}$. Wobec tego $x' = W_k(x_k)$, gdzie $x_k \in X$. Z lematu 7.4 wynika, że

$$\text{dist}(x', W_k(Y)) \leq r_k \cdot \text{dist}(x_k, Y) \leq r \cdot \text{dist}(x_k, Y) \leq r \cdot \max_{x \in X} \text{dist}(x, Y), \quad (7.11)$$

gdyż $r_k \leq r$. Z drugiej strony

$$\text{dist}(x', Y') \leq \text{dist}(x', W_k(Y)),$$

ponieważ $W_k(Y) \subset Y'$ na mocy określenia zbioru Y' i wzoru (7.9). Ostatnia nierówność wraz z (7.11) daje

$$\text{dist}(x', Y') \leq r \cdot \max_{x \in X} \text{dist}(x, Y),$$

a ponieważ punkt $x' \in X'$ był wybrany dowolnie, więc także

$$\max_{x' \in X'} \text{dist}(x', Y') \leq r \cdot \max_{x \in X} \text{dist}(x, Y).$$

W analogiczny sposób pokazujemy, że

$$\max_{y' \in Y'} \text{dist}(y', X') \leq r \cdot \max_{y \in Y} \text{dist}(y, X).$$

Wobec tego mamy:

$$\begin{aligned} d_H(X', Y') &= \max \left\{ \max_{x' \in X'} \text{dist}(x', Y'), \max_{y' \in Y'} \text{dist}(y', X') \right\} \\ &\leq \max \left\{ r \cdot \max_{x \in X} \text{dist}(x, Y), r \cdot \max_{y \in Y} \text{dist}(y, X) \right\} \\ &= r \max \left\{ \max_{x \in X} \text{dist}(x, Y), \max_{y \in Y} \text{dist}(y, X) \right\} = r d_H(X, Y), \end{aligned}$$

co należało udowodnić. □

Wniosek 7.1. Niech $\{W_1, W_2, \dots, W_m\}$ będzie rodziną kontrakcji $W_k : \mathbb{R}^N \rightarrow \mathbb{R}^N$ w przestrzeni euklidesowej \mathbb{R}^N . Dla każdego $k \in \{1, \dots, m\}$ niech r_k będzie stałą kontrakcji W_k . Wtedy istnieje dokładnie jeden zbiór $A \in \mathcal{H}_N$ będący punktem stałym operatora Barnsleya-Hutchinsona wyznaczonego przez rodzinę $\{W_1, W_2, \dots, W_m\}$, tj. taki, że

$$\mathcal{H}_{W_1, W_2, \dots, W_m}(A) = \bigcup_{k=1}^m W_k(A) = A.$$

Dla każdego zbioru $X \in \mathcal{K}_N$ ciąg iteracyjny $(\mathcal{H}_{W_1, W_2, \dots, W_m}^n(X))$, $n \in \mathbb{N}_0$, jest zbieżny w przestrzeni metrycznej (\mathcal{K}_N, d_H) i zachodzi równość:

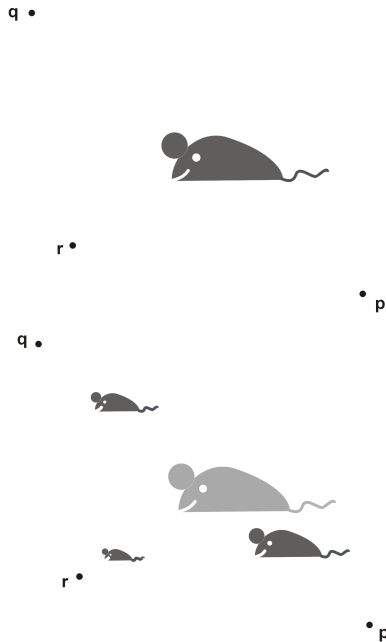
$$\lim_{n \rightarrow \infty} \mathcal{H}_{W_1, W_2, \dots, W_m}^n(X) = A.$$

Ponadto dla każdego zbioru $X \in \mathcal{K}_N$ mamy:

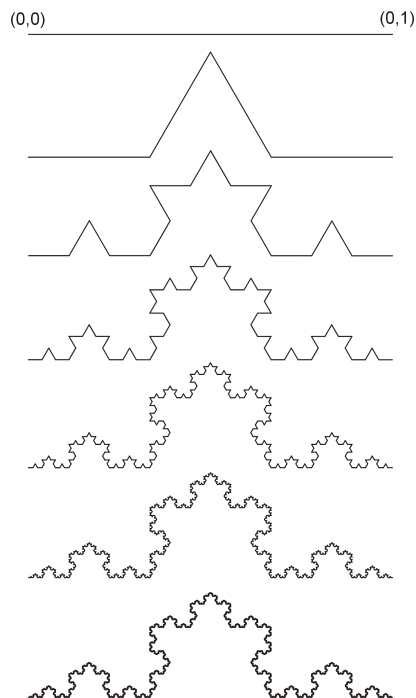
$$d_H(\mathcal{H}_{W_1, W_2, \dots, W_m}^n(X), A) \leq \left(\frac{r^n}{1-r}\right) d_H(X, \mathcal{H}_{W_1, W_2, \dots, W_m}(X)),$$

gdzie $r = \max\{r_1, r_2, \dots, r_m\}$.

Dowód. Wystarczy skorzystać z twierdzenia Banacha o punkcie stałym (Twierdzenie 6.1). □



Rysunek 1: Działanie operatora Barnsleya-Hutchinsona na zbiór zwarty na płaszczyźnie



Rysunek 2: Kolejne przybliżenia krzywej von Kocha. Zbiorem wyjściowym jest odcinek A o końcach $(0,0)$ i $(0,1)$

Definicja 7.4. Niech $\{W_1, W_2, \dots, W_m\}$ będzie rodziną kontrakcji $W_k : \mathbb{R}^N \rightarrow \mathbb{R}^N$ w przestrzeni euklidesowej \mathbb{R}^N . Jedyne zbiór niepusty i zwarty A będący punktem stałym operatora Barnsleya-Hutchinsona $\mathcal{H}_{W_1, W_2, \dots, W_m}$ wyznaczonego przez rodzinę $\{W_1, W_2, \dots, W_m\}$, o którym mowa we wniosku 7.1, nazywamy atraktorem Barnsleya układu dynamicznego $(\mathcal{H}_N, \mathcal{H}_{W_1, W_2, \dots, W_m})$.

Często mamy do czynienia z sytuacjami, w których kontrakcje z rodziny $\{W_1, W_2, \dots, W_m\}$ są podobieństwami w przestrzeni \mathbb{R}^N (zob. par. 6). Mamy łatwy do wykazania lemat.

Lemat 7.5. W przestrzeni euklidesowej \mathbb{R}^N podobieństwo $P : \mathbb{R}^N \rightarrow \mathbb{R}^N$ o skali $r \neq 0$ jest kontrakcją wtedy i tylko wtedy, gdy $0 < |r| < 1$.

Definicja 7.5. Niech $\{P_1, P_2, \dots, P_m\}$ będzie rodziną podobieństw $P_k : \mathbb{R}^N \rightarrow \mathbb{R}^N$ w przestrzeni euklidesowej \mathbb{R}^N i niech $\mathcal{H}_{P_1, P_2, \dots, P_m}$ będzie operatorem Barnsleya-Hutchinsona wyznaczonym przez rodzinę $\{P_1, P_2, \dots, P_m\}$, Mówimy, że układ dynamiczny $(\mathcal{K}_N, \mathcal{H}_{P_1, P_2, \dots, P_m})$ spełnia warunek zbioru otwartego, jeżeli P_k jest kontrakcją dla każdego $k \in \{1, \dots, m\}$ oraz istnieje taki ograniczony, niepusty zbiór otwarty $G \subset \mathbb{R}^N$, że

$$\bigcup_{k=1}^m P_k(G) \subset G$$

i zbiory $P_k(G)$ dla $k \in \{1, \dots, m\}$ są parami rozłączne.

Definicja 7.6. Niech $\{P_1, P_2, \dots, P_m\}$ będzie rodziną podobieństw $P_k : \mathbb{R}^N \rightarrow \mathbb{R}^N$ w przestrzeni euklidesowej \mathbb{R}^N taką, że układ dynamiczny $(\mathcal{K}_N, \mathcal{H}_{P_1, P_2, \dots, P_m})$ (gdzie $\mathcal{H}_{P_1, P_2, \dots, P_m}$ jest operatorem Barnsleya-Hutchinsona wyznaczonym przez tę rodzinę) spełnia warunek zbioru otwartego. Wtedy atraktor Barnsleya A rozważanego układu nazywamy zbiorem samopodobnym lub fraktalem topologicznym.

Wymiarem samopodobieństwa zbioru A nazywamy liczbę $s \in \mathbb{R}$ będącą rozwiązaniem równania

$$\sum_{k=1}^m |r_k|^s = 1,$$

gdzie, dla każdego $k \in \{1, \dots, m\}$, r_k jest skalą podobieństwa P_k .

Jest wiele sposobów zdefiniowania pojęcia wymiaru zbioru w przestrzeni euklidesowej w sposób odpowiedni dla zbiorów o „fraktalnym” charakterze. Rozważany tutaj wymiar samopodobieństwa zbioru jest bardzo szczególnym przypadkiem. Podstawowe znaczenie w tej dziedzinie ma tzw. wymiar Hausdorffa (dla samopodobnych podzbiorów przestrzeni \mathbb{R}^N pokrywa się on z wymiarem samopodobieństwa), zob. np. [3], [4], [7]. Czytelnicy zainteresowani technikami stosowanymi przy określaniu wymiaru zbiorów Julia na płaszczyźnie zespolonej odsyłamy do pozycji [1] i [14].

Literatura

- [1] D.S. Alexander, *A History of Complex Dynamics from Schröder to Fatou and Julia*, Vieweg & Sohn Verlagsgesellschaft, Braunschweig, Wiesbaden 1994.
- [2] Bau-Sen Du, A simple proof of Sharkovsky's Theorem, *Amer. Math. Monthly*, 111, 2004.
- [3] K.J. Falconer, *The geometry of fractal sets*, Cambridge University Press, Cambridge 1985.
- [4] K.J. Falconer, *Techniques in fractal geometry*, John Wiley and Sons, Chichester 1997.
- [5] M.J. Feigenbaum, Quantitative universality for a class of nonlinear transformations, *Journal of Statistical Physics*, 19, 1978.
- [6] K. Goebel, W.A. Kirk, *Topics in metric fixed point theory*, Cambridge University Press, Cambridge 1990.
- [7] J.E. Hutchinson, Fractals and self similarity, *Indiana University Mathematics Journal*, 30, 1981.
- [8] J. Kudrewicz, *Fraktale i chaos*, WNT, Warszawa 1993.
- [9] K. Kuratowski, *Topology*, Academic Press, New York, 1966.
- [10] A. Lasota, *Układy dynamiczne na miarach*, Wydawnictwo Uniwersytetu Śląskiego, Katowice 2008.
- [11] T.Y. Li, J.A. Yorke, Period three implies chaos, *Amer. Math. Monthly*, 82, 1975.
- [12] B.B.Mandelbrot, *The Fractal Geometry of Nature* W. H. Freeman & Co., San Francisco 1982.

-
- [13] J.D. Murray, *Wprowadzenie do biomatematyki*, Wydawnictwo Naukowe PWN, Warszawa 2006.
- [14] J. Myjak, T. Szarek, On the Hausdorff dimension of Cantor-like set with overlaps, *Chaos, Solitons and Fractals*, 18, 2003.
- [15] J. von Neumann, *Theory of self-reproducing automata*, University of Illinois Press, Urbana II, 1966.
- [16] E. Ott, *Chaos w układach dynamicznych*, WNT, Warszawa 1987.
- [17] H.-O. Peitgen, H. Jürgens, D. Saupe, *Fraktale. Granice chaosu*, PWN, Warszawa 2002.
- [18] R. Penrose, *Nowy umysł cesarza*, Wydawnictwo Naukowe PWN, Warszawa 1996.
- [19] W. Rudin, *Podstawy analizy matematycznej*, Wydawnictwo Naukowe PWN, Warszawa 1996.
- [20] T.M. Sękowski, *Zagadnienia matematycznej teorii chaosu*, Wydawnictwo Uniwersytetu Marii Curie-Skłodowskiej, Lublin 2007.
- [21] A.N. Szarkowski, Coexistence of the cycles of a continuous mapping of the line into itself, *Ukrainski Matematyczny Żurnal*, 16, 1965.
- [22] P. Właż, *Analiza automatów komórkowych*, rozprawa doktorska, Wydawnictwo UMCS, Lublin 1997.