Liczby normalne i wybrane ich własności

Karol GRYSZKA, Kraków*

Jest to tekst związany z odczytem wygłoszonym na LXI Szkole Matematyki Poglądowej, *Matematyczne zmiany*, Wola Ducka, luty 2020.

Redakcja

Liczby normalne

Wybierzemy dowolną liczbę rzeczywistą, niech dla przykładu będzie to $\sqrt{2}$. Rozważmy następnie jej rozwinięcie dziesiętne, to jest

$$\sqrt{2} = 1{,}414213562373095048801688724209698078569671875376948073176\ldots$$

i zapytajmy o częstotliwość występowania, na przykład, cyfry 0. W powyższym przykładzie jest ich 6 na 57 wskazanych w rozwinięciu cyfr znaczących, a więc jest to proporcja bliska 1 : 10. Wszystkich cyfr jest 10, więc proporcja ta jest bliska takiej, jakiej oczekiwalibyśmy w losowej liczbie o 57 cyfrach. Czy to przypadek? Co byśmy otrzymali, gdybyśmy rozważyli nie 57, a milion, bilion, googol cyfr? Czy częstotliwość występowania cyfry 0 również byłaby bliska 10%?

W teorii liczb wyróżniane są tak zwane liczby *normalne*, a więc takie liczby, w których, kolokwialnie mówiąc, częstość występowania coraz dłuższych bloków cyfr w rozwinięciu dziesiętnym jest taka, jaką ma liczba z losowymi cyframi w rozwinięciu dziesiętnym. W szczególności,w liczbie normalnej cyfra 0 występuje średnio raz na 10 cyfr, a blok cyfr 154 średnio raz na 1000 bloków trzycyfrowych. Jeśli zaś okaże się, że na przykład blok cyfr 1234 występuje średnio tylko raz na milion razy, to taka liczba nie będzie normalna. Dokonamy teraz formalizacji tej intuicji.

Losowość oznacza, że każda cyfra w rozwinięciu wybierana jest niezależnie od pozostałych i z prawdopodobieństwem równym $\frac{1}{10}$.

Niech $w = a_1 \ldots a_k$ będzie ciągiem cyfr w systemie o podstawie $b > 1$ (to jest każde $a_i \in \{0, 1, \ldots, b-1\}$) – wtedy liczbę $|w| := k$ nazwiemy długością ciągu w. Niech ponadto dla $x \in [0, 1]$ dane będzie jej rozwinięcie w systemie o podstawie b,

$$x = \sum_{i=1}^{+\infty} \frac{x_i}{b^i},$$

gdzie $x_i \in \{0, 1, \ldots, b-1\}$. Będziemy również pisać $x = 0, x_1x_2x_3\ldots$ i stosować następującą konwencję: jeśli $\hat{x}$ jest ciągiem (skończonym lub nieskończonym) cyfr zapisanym w postaci $\hat{x} = x_1x_2x_3x_4\ldots$, to liczbę $0, x_1x_2x_3\ldots$ będziemy oznaczać przez $0, \hat{x}$. Dla ustalonego $n > 0$ definiujemy *częstotliwość występowania ciągu w* w rozwinięciu x do pozycji n:

$$F(w, x, n) = \frac{\#\{1 \leqslant i \leqslant n : \ x_i = a_1, \ldots, x_{i+k-1} = a_k\}}{n}.$$

Przechodząc do granicy, otrzymujemy wielkość

$$F(w, x) = \lim_{n\to+\infty} F(w, x, n).$$

Na przykład jeśli $\hat{x} = 3141515$, to $x = 0,3141515$. Ponadto dla $w = 15$ i $n = 4$ otrzymujemy $F(w, x, n) = 1/4$, dla $n = 7$ zaś $F(w, x, n) = 2/7$. Jeśli teraz $y = 1/7$, to $F(1, y) = 1/6$, ale $F(3, y) = 0$.

Jest to asymptotyczna częstość występowania ciągu w w x. Łatwo zauważyć, że taka granica nie zawsze musi istnieć (Czytelnik z łatwością wskaże taką liczbę x, dla której nie istnieje $F(1, x)$).

Definicja 1. *Liczbę $x \in [0, 1]$ nazywamy* ***normalną*** *w systemie o podstawie b, jeśli dla dowolnego w zachodzi $F(w, x) = b^{-|w|}$.*

Przedstawiona wyżej koncepcja liczby normalnej pochodzi od Borela [3] i w tym roku obchodzimy okrągłą rocznicę 111 lat od jej powstania. W dalszej części artykułu podamy jeszcze kilka innych własności tych liczb odkrytych przez Borela.

Zgodnie z powyższą definicją należy sprawdzić wszystkie możliwe skończone ciągi w. Zadanie to jest dla „losowej" liczby praktycznie niemożliwe do wykonania, okazuje się jednak, że istnieją nietrudne konstrukcje prowadzące do uzyskania liczb normalnych. Z drugiej strony, losowo wybrana liczba, jeśli ma być normalna, musi być liczbą niewymierną.

*Uniwersytet Pedagogiczny w Krakowie, karol.gryszka@up.krakow.pl

Obserwacja 2. *Jeśli $x \in [0,1]$ jest normalna, to $x \notin \mathbb{Q}$.*

Dowód. Jeśli x jest wymierna, to jej rozwinięcie jest skończone lub od pewnego momentu okresowe. Jeśli jest skończone, to nie ma czego dowodzić. Niech więc $x = 0, x'_1 \dots x'_k(x_1 x_2 \dots x_n)$ będzie tym rozwinięciem. Wybieramy $w = x_1 x_2 \dots x_n y$, gdzie $y \equiv x_1 + 1 \mod b$. Łatwo spostrzec, że dla dowolnego n mamy $F(w, x, n) = 0$. □

Zauważmy, że przez utożsamienie ciągu $\hat{x} \in \{0, 1, \dots, b-1\}^{\mathbb{N}}$ z liczbą $0, \hat{x}$ możemy definicję liczby normalnej przenieść bezpośrednio na definicję ciągu normalnego. Definicja taka jest identyczna z powyższą, jedyną różnicą jest typ obiektu, jaki jest definiowany (ciąg lub liczba).

W 1933 roku D. G. Champernowne [4] określił następujący ciąg liczbowy, który jest zestawieniem kolejnych liczb naturalnych w porządku rosnącym:

$$01234567891011121314151617181920\dots.$$

Ciąg ten nazywamy ciągiem Champernowne'a. Służy on do zdefiniowania stałej Champernowne'a:

$$c := 0,01234567891011121314151617181920\dots.$$

Champernowne wykazał, że c jest liczbą normalną w systemie dziesiętnym, tym samym jednocześnie wskazując jedną z pierwszych metod na konstrukcję liczb normalnych – zestawienie ze sobą liczb w pewnym porządku. Dokładnie taką ideę zastosował w 1935 roku Besicovitch [2] i wykazał, że liczba

$$0,149162536496481100121144\dots$$

będąca zestawieniem kwadratów kolejnych liczb naturalnych jest normalna. W 1946 roku z kolei Copeland i Erdős [5] dowodzą, że zestawiając liczby pierwsze

$$0,23571113171923293137414347 4953\dots,$$

również otrzymujemy liczbę normalną. Jednym z ciekawszych wyników tego typu jest rezultat Nakai i Shiokawy z 1992 roku [7].

Twierdzenie 3. *Niech $f : \mathbb{R} \to \mathbb{R}$ będzie niestałym wielomianem o tej własności, że $f(x) > 0$ dla $x > 0$. Wtedy liczba*

$$0, \lfloor f(1) \rfloor \lfloor f(2) \rfloor \lfloor f(3) \rfloor \lfloor f(4) \rfloor \lfloor f(5) \rfloor \dots,$$

gdzie $\lfloor x \rfloor$ oznacza cechę (część całkowitą) x, jest normalna.

Jeśli $f(x) = x$, to twierdzenie sprowadza się (po dodaniu 0) do stałej Champernowne'a. Dla $f(x) = x^2$ otrzymujemy stałą Besicovitcha. Wybierając teraz $f(x) = x^{10}$ otrzymujemy stałą

$$0,1102459049104857697656 25\dots,$$

która również jest liczbą normalną.

Jedną z ciekawszych metod wyznaczania liczb normalnych w danej bazie są ciągi de Bruijna. Niech dany będzie zbiór B cyfr w systemie o podstawie (bazie) b.

Definicja 4. ***Ciągiem de Bruijna rzędu** n w zbiorze B nazywamy ciąg o długości $b^n + n - 1$ o tej własności, że każdy ciąg w długości n pojawia się w nim dokładnie jeden raz.*

Ciągi de Bruijna są ciągami o minimalnej długości, które spełniają postulowaną w definicji własność. Nie będziemy opisywać metody wyznaczania takiego ciągu dla danego rzędu, ograniczymy się tutaj jedynie podania dwóch przykładów:

1. jeśli $b = 2$ i $n = 2$, to odpowiednim ciągiem de Bruijna jest 00110,
2. w przypadku $b = 3$ i $n = 2$ tym ciągiem jest 0011221020.

Oczywiście, powyższe przykłady to nie jedyne możliwe ciągi de Bruijna rzędu 2. Łatwo zauważyć, że jest ich co najmniej $b!$ – wystarczy w tym celu permutować cyfry.

W jaki sposób otrzymać z takiego ciągu liczbę normalną? Okazuje się, że jeżeli $b > 2$, to ciąg de Bruijna rzędu n można „przedłużyć” do ciągu rzędu $n+1$, tym samym więc ma sens określenie takiego ciągu $\hat{d} = d_1 d_2 d_3 \dots$, dla którego każdych $b^n + n - 1$ początkowych wyrazów jest ciągiem de Bruijna rzędu n. Prowadzi to do zdefiniowania liczby

$$d = 0, \hat{d} = 0, d_1 d_2 d_3 \dots$$

w pewnym ustalonym systemie pozycyjnym b. Takich liczb można, oczywiście, zdefiniować wiele, niemniej wszystkie one dzielą wspólną cechę.

Twierdzenie 5. ([1]) *Liczba d zdefiniowana powyżej jest normalna.*

Nie jest zaskoczeniem, że istnieje wiele liczb, które nie są normalne. Ponadto istnieje prosta konstrukcja, która prowadzi do tak zwanej *liczby całkowicie nienormalnej*, to znaczy takiej liczby, która nie jest normalna w żadnym systemie pozycyjnym. Ściślej – w dowolnym systemie pozycyjnym i dla dowolnego w wartość $F(w,x)$ nie istnieje. Rozważmy mianowicie funkcję

$$f : \mathbb{N} \setminus \{0,1\} \to \mathbb{N}$$

daną relacją:

$$f(n) = \begin{cases} 4, & n = 2, \\ n^{\frac{f(n-1)}{n-1}}, & n \geqslant 3. \end{cases}$$

Niech teraz dana będzie liczba

$$\alpha := \prod_{n=2}^{+\infty} (1 - f(n)^{-1}).$$

$f(3) = 3^2 = 9.$

$f(4) = 4^3 = 64$

$f(5) = 5^{16} = 152\ 587\ 890\ 625$

$f(6) = 6^{30\ 517\ 578\ 125}$

$\approx 4,46 \cdot 10^{23\ 747\ 291\ 576}$

Wtedy

$$\alpha = 0,6562499999956991\underbrace{9999\ldots9999}_{23\ 747\ 291\ 559}8528404201690728\ldots.$$

Liczba ta posiada w różnych odległych pozycjach rozwinięcia dziesiętnego bardzo długie ciągi dziewiątek. Efektem tego jest nieistnienie granic $F(w,\alpha)$. Liczba ta jest w szczególności liczbą Liouville'a (zainteresowanego Czytelnika odsyłamy do [6]).

Jak dużo jest liczb normalnych? Na to pytanie łatwo jest odpowiedzieć w języku teorii mnogości.

Twierdzenie 6. *Liczb normalnych jest nieprzeliczalnie wiele.*

Dowód – poglądowo. Wybieramy takie pozycje p_n w rozwinięciu stałej Champernowne'a c, które spełniają zależność $p_1 = 1$ oraz

$$p_{n+1} \geqslant 2^{2^n} p_n$$

oraz na pozycji p_n rozpoczyna się sekwencja cyfr odpowiedzialna za wszystkie bloki o określonej ilości cyfr.

Na przykład, $p_1 = 1$ implikuje, że $p_2 \geqslant 4$, a więc można przyjąć $p_2 = 11$ (gdyż na pozycji 11 w ciągu stałej c rozpoczynają się bloki dwucyfrowe). Teraz $p_3 \geqslant 256 \cdot 11 = 2816$, zaś liczby od 0 do 999 mają łącznie

$$10 + 90 \cdot 2 + 900 \cdot 3 = 2890$$

cyfr, można więc przyjąć $p_3 = 2891$ i jest to pozycja rozpoczynająca sekwencję cyfr bloków czterocyfrowych liczb.

Mając już wybrane pozycje, „rozciągamy" c tak, aby miejsca wskazane przez p_n były chwilowo miejscami pustymi. W te miejsca wstawiamy teraz jakąkolwiek kombinację cyfr od 0 do 9 i otrzymujemy stałą $\tilde{c}$.

Ponieważ miejsca, na których umieszczamy dodatkowe cyfry są rzadkie (pozycje rosną szybciej niż wykładniczo), nie zmienia to wartości $F(w,\tilde{c})$ dla wszystkich w. Ponadto swoboda w dodawaniu cyfr pozwala określić nieprzeliczalnie wiele $\tilde{c}$. Kończy to dowód twierdzenia. $\square$

Odpowiedź teoriomongościowa nie sprawia problemu. Spróbujmy teraz odpowiedzieć na nieco trudniejsze pytanie – ile jest liczb normalnych w ujęciu teorii miary (definicję przypominamy na marginesie).

Jeśli X jest ustalonym zbiorem, to rodzinę zbiorów $\mathcal{B}$ nazywamy *σ-algebrą*, jeśli:
- $\emptyset \in \mathcal{B}$,
- $A \in \mathcal{B} \Rightarrow X \setminus A \in \mathcal{B}$,
- $A_1, A_2, \ldots \in \mathcal{B} \Rightarrow A_1 \cup A_2 \cup \ldots \in \mathcal{B}$.

Jeśli teraz $\mathcal{B}$ jest σ-algebrą, to funkcję $P : \mathcal{B} \to [0,+\infty]$ nazywamy *miarą*, gdy:
- $P(\emptyset) = 0$,
- jeśli $A_1, A_2, \ldots$ są zbiorami rozłącznymi, to $P(A_1 \cup A_2 \cup \ldots) = P(A_1) + P(A_2) + \ldots$.

Zbiory $A \in \mathcal{B}$ nazywamy mierzalnymi. Jeśli $P(X) = 1$, to miarę P nazywamy *probabilistyczną*. Zbiór A jest *miary zero* (*pełnej miary*), gdy $P(A) = 0$ ($P(A) = 1$).

Na odcinku $[0,1]$ liczby wymierne są rzadkością – z prawdopodobieństwem równym 0 losowo wybrana liczba z tego odcinka jest ilorazem dwóch liczb całkowitych. Z drugiej strony trójkowy zbiór Cantora, a więc zbiór takich liczb z odcinka $[0,1]$, których rozwinięcie w systemie trójkowym posiada tylko zera i dwójki, jest zbiorem nieprzeliczalnym o mierze zerowej, a więc ponownie losowo wybrana liczba z przedziału z zerową szansą jest elementem zbioru Cantora. Jak jest zatem z liczbami normalnymi? Odpowiedź na to pytanie uzyskamy z użyciem aparatu pochodzącego od układów dynamicznych.

Układy dynamiczne i twierdzenie ergodyczne

Na początku zauważmy, że jeżeli liczba x jest normalna, to część ułamkowa $\{10^n \cdot x\}$ jest również liczbą normalną dla dowolnego n. Innymi słowy – mnożenie przez 10 liczby normalnej i ucięcie wyniku do części ułamkowej nie zmienia „normalności liczby", a sama operacja może być rozumiana jako zapomnienie pierwszej cyfry rozwinięcia dziesiętnego tej liczby. Jeszcze inaczej, rozważmy zbiór $10^{\mathbb{N}}$ wszystkich ciągów nieskończonych rozwinięć dziesiętnych wszystkich liczb z zakresu $[0,1]$. Jeśli jakaś liczba ma rozwinięcie skończone, uzupełniamy je

Rozważając zbiór $10^{\mathbb{N}}$ numerujemy kolejne wyrazy ciągu od 1, to jest zakładamy w tym napisie, że $0 \notin \mathbb{N}$.

zerami uzyskując ciąg nieskończony. Niech teraz σ będzie odpowiednikiem operacji $\{10 \cdot x\}$ na ciągu. Wtedy

$$\sigma : 10^{\mathbb{N}} \to 10^{\mathbb{N}}$$

działa następująco:

$$\sigma(x_1 x_2 x_3 x_4 \ldots) = x_2 x_3 x_4 \ldots.$$

Jest to tak zwany ***operator przesunięcia „w lewo”***. W takim kontekście operację $\{10 \cdot x\}$ na liczbie możemy utożsamić z operacją $\sigma(x)$ na ciągu zbudowanym z kolejnych cyfr rozwinięcia dziesiętnego. Co więcej, operacje te są ze sobą powiązane tożsamością

$$0,\sigma(x) = \{10 \cdot 0,x\},$$

a więc są ze sobą *topologicznie sprzężone*.

Odwzorowania ciągłe $F : X \to X$ oraz $G : Y \to Y$ są topologicznie sprzężone, gdy istnieje homeomorfizm $\psi : X \to Y$ (a więc bijekcja ciągła, której odwrotna jest również funkcją ciągłą i taka, że $\psi \circ F = G \circ \psi$).

Przypomnijmy przed kontynuacją artykułu kilka podstawowych pojęć z topologii oraz układów dynamicznych. Jeśli (X, d) jest przestrzenią metryczną oraz $F : X \to X$ jest odwzorowaniem ciągłym, to parę (X, F) nazywamy *układem dynamicznym*. Możemy wtedy również pisać krótko: F jest układem dynamicznym. Jeśli $x \in X$, to zbiór $\{x, F(x), F^2(x), \ldots\}$ nazywamy *orbitą* punktu x i oznaczamy przez $orb(x)$. Zbiór $A \subset X$ jest *gęsty* w X, gdy każdy element $x \in X$ jest granicą ciągu elementów ze zbioru A. Układ dynamiczny (X, F) jest *tranzytywny*, jeśli istnieje w nim taki punkt, którego orbita jest zbiorem gęstym. Taki punkt nazywa się wtedy *tranzytywnym*.

Przypomnijmy, *metryką* nazywamy taką funkcję $d : X \times X \to [0, +\infty)$, która spełnia następujące aksjomaty:

1. $d(x, x) = 0$ dla wszystkich $x \in X$,
2. $d(x, y) = d(y, x)$ dla wszystkich $x, y \in X$,
3. $d(x, y) \leqslant d(x, z) + d(z, y)$ dla wszystkich $x, y, z \in X$.

Wtedy parę (X, d) nazywamy przestrzenią metryczną. Powiemy ponadto, że $x_n \to x$ w (X, d), gdy $d(x_n, x) \to 0$ (zbieżność ciągu liczbowego).

Na zbiorze $10^{\mathbb{N}}$ można określić metodę mierzenia odległości między dwoma nieskończonymi ciągami. Niech $x = (x_0, x_1, x_2, \ldots)$ oraz $y = (y_1, y_2, \ldots)$ i definiujemy

$$d(x, y) := \begin{cases} 0, & x = y, \\ 2^{-k}, & k = \min\{n : x_n \neq y_n\}. \end{cases}$$

Dowodzi się, że d jest metryką na zbiorze $10^{\mathbb{N}}$. Intuicja stojąca za takim sposobem mierzenia odległości jest dość prosta – porównujemy dwa ciągi x oraz y tak długo, aż pojawi się pierwsza pozycja, na której one się różnią. Numer k-tej pozycji jest następnie wykładnikiem w liczbie 2^{-k}, określającej odległość tych ciągów. Jest jasne, że im k jest większe, a więc im dłuższy jest początkowy segment cyfr, na którym oba ciągi są równe, tym odległość jest mniejsza i tym samym, ciągi są „bliższe”.

Definicja 7. *Układ dynamiczny* $(10^{\mathbb{N}}, \sigma)$ *nazywamy* ***pełnym shiftem na*** 10 ***symbolach***.

W powyższej definicji oraz wszystkich rozważaniach liczba 10 może zostać zastąpiona przez dowolny abstrakcyjny zbiór n-elementowy A – mówimy wtedy o pełnym shifcie na n symbolach ze zbioru A, a zbiór A nazywamy wtedy *alfabetem*.

W teorii układów dynamicznych ważną rolę odgrywa jej gałąź zwana teorią ergodyczną. Jej istotą jest wyposażenie układu dynamicznego (a więc odwzorowania ciągłego na pewnej przestrzeni metrycznej, a ogólniej również na przestrzeni topologicznej) w miarę, a więc funkcję mówiącą o tym, jak mierzyć zbiory oraz jak duże one mogą być. Przykładem miary jest długość odcinka (miara Lebesgue'a), ale również może nią być liczność zbioru (jeśli zbiór jest nieskończony, to jego miara jest równa ∞).

Nie będziemy przypominać formalnej definicji miary Lebesgue'a. Jej intuicją jest to, iż mierzy ona długość odcinka. Ponadto każdy zbiór przeliczalny jest miary 0 (choć nie jest prawdą fakt odwrotny!).

Rozważmy teraz odwzorowanie $T : [0, 1] \to [0, 1]$ oraz niech $A \in \mathcal{B}([0, 1])$, czyli A będzie zbiorem borelowskim, to jest takim podzbiorem odcinka $[0, 1]$, który można otrzymać przez operację przeliczalnej sumy, przeliczalnego przekroju lub dopełnienia przedziałów otwartych. Dowodzi się, że zbiory borelowskie tworzą σ-algebrę oraz mogą być użyte do zdefiniowania miary. Sama definicja jest dużo bardziej złożona od strony formalnej, powyższa intuicja jest jednak w zupełności wystarczająca. W szczególności z tej intuicji wynika następująca obserwacja.

Obserwacja 8. *Przedziały domknięte są borelowskie. Zbiory jednoelementowe są borelowskie. Każdy zbiór przeliczalny jest borelowski.*

Definicja 9. *Ustalmy $n \geqslant 0$, $x \in [0,1]$, $A \in \mathcal{B}([0,1])$ oraz rozważmy miarę probabilistyczną $m(x,n)$ zadaną wzorem*

$$m(x,n)(A) := \frac{\#\{0 \leqslant j < n \ : \ T^j(x) \in A\}}{n}.$$

*Jest to n-ta **miara empiryczna punktu x przekształcenia** T.*

Wyżej zdefiniowana miara określa, z jaką częstością orbita danego punktu do elementu $T^{n-1}(x)$ „odwiedza” zbiór A. W szczególności ma sens rozważenie granicy takich częstości dla $n \to +\infty$. Zauważmy istotne podobieństwo takiej definicji z definicją liczby normalnej.

Definicję miary empirycznej można przenieść w naturalny sposób na inne odwzorowania oraz przestrzenie i zbiory borelowskie (σ-algebry). W tak abstrakcyjnym ujęciu możemy teraz rozważyć dowolną miarę probabilistyczną P na przestrzeni X. Jeśli teraz $x \in X$ oraz $A \in \mathcal{B}(X)$, to można zadać naturalne pytanie: jak dobrze miara empiryczna punktu x przybliża daną miarę probabilistyczną P? Postawmy definicję, której sformułowanie będzie miało charakter intuicyjny.

Definicja 10. ***Punkt x przekształcenia jest generyczny** dla miary P w układzie dynamicznym (X,F), jeśli $m(x,n)$ wraz ze wzrostem n coraz lepiej przybliża miarę P.*

Definicja formalna jest następująca: *punkt x jest generyczny*, jeśli ciąg miar $\{m(x,n)\}_{n\in\mathbb{N}}$ zbiega do P w $*$-słabej topologii, a więc

$$\int_X |f| dm(x,n) \to \int_X |f| dP$$

dla dowolnej ciągłej i ograniczonej funkcji f.

Jeśli P jest długością odcinka (miarą Lebesgue'a) i $T : [0,1] \to [0,1]$, to generyczność punktu x oznacza w szczególności, że jeśli na przykład

$$A = \left[\frac{10^3}{2^{21}}, \frac{10^3+1}{2^{21}}\right],$$

to odsetek liczb $T^j(x)$ należących do A jest tym bliższy długości przedziału A, im więcej punktów z orbity punktu x będziemy rozważać. Ale to oczywiście wybrany przykład. Generyczność zakłada ponadto, że taka własność jest prawdziwa dla dowolnego zbioru $A \in \mathcal{B}([0,1])$.

Przykład 11. Jeśli $T : [0,1] \to [0,1]$ dane jest wzorem $T(x) = \{10 \cdot x\}$, to punkty $x = \frac{1}{2}$ oraz $y = \frac{2}{3}$ nie są generyczne dla miary Lebesgue'a. Istotnie, jeśli $A = [\frac{1}{2}, 1]$, to $m(x,n)(A) = \frac{1}{n}$ oraz jeśli $B = [0, \frac{1}{2}]$, to $m(y,n)(B) = 0$.

Układy dynamiczne z miarą mają wiele istotnych własności, których nie sposób opisać w tak krótkim artykule. Na potrzeby tego artykułu oraz sformułowania zasadniczego twierdzenia wprowadzimy jeszcze dwa ważne pojęcia tej dziedziny.

Definicja 12. *Powiemy, że układ dynamiczny (X,F) **zachowuje miarę** P, jeśli $P(A) = P(F^{-1}(A))$ dla dowolnego zbioru mierzalnego A. Mówimy również wtedy, że P jest F-niezmiennicza. Miarę P nazywamy ergodyczną względem F, jeśli warunek $F^{-1}(A) = A$ implikuje $P(A) = 0$ lub $P(A) = 1$.*

Przypomnijmy jeszcze definicję funkcji mierzalnej.

Definicja 13. *Funkcję $f : \mathbb{R} \to \mathbb{R}$ nazywamy **mierzalną**, jeśli dla dowolnego a $f^{-1}((-\infty, a])$ jest zbiorem z $\mathcal{B}(\mathbb{R})$.*

Zauważmy, że każda funkcja ciągła jest mierzalna (gdyż wtedy przeciwobraz zbioru domkniętego $(-\infty, a]$ jest zbiorem domkniętym zgodnie z definicją ciągłości, a zbiory domknięte są zbiorami borelowskimi, czyli mierzalnymi). Funkcje $F(x) = \{10 \cdot x\}$ oraz $G(x) = \operatorname{sgn}(x)$ są mierzalne (co Czytelnik może sprawdzić samodzielnie).

Sformułujemy teraz głównie twierdzenie tego artykułu.

Twierdzenie 14 (twierdzenie ergodyczne Birkhoffa [8]). *Jeśli $f : X \to \mathbb{R}$ jest funkcją mierzalną oraz P jest miarą ergodyczną i niezmienniczą względem F, to dla P-prawie wszystkich $x \in X$ spełniony jest warunek*

$$\frac{1}{n}\sum_{j=0}^{n-1} f(F^j(x)) \to \frac{1}{P(X)} \int_X f dP.$$

Określenie „P-prawie wszystkich” oznacza, że jeśli rozważymy wszystkie punkty $x \in X$ spełniające warunek z tezy twierdzenia, to tworzą one zbiór pełnej miary.

Interpretacja twierdzenia ergodycznego jest następująca. Załóżmy, że dany jest pewien proces fizyczny, w którym przestrzeń stanów X ewoluuje w czasie dyskretnym i ewolucja ta opisana jest przez funkcję $F : X \to X$. Ponadto dana jest funkcja $f : X \to \mathbb{R}$, która jest funkcją pomiaru (zwraca dla stanu $x \in X$ wartość $f(x) \in \mathbb{R}$). Ustalmy stan x. Interesuje nas teraz średnia wartość pomiaru tego stanu oraz jego ewolucji w czasie (lewa strona, czyli poprzednik implikacji w twierdzeniu ergodycznym). Jeżeli dokonamy pomiaru w dużej liczbie ewolucji tego stanu, to spodziewamy się, że tak uśredniony pomiar będzie bliski średniej wartości funkcji pomiaru liczonej dla całej przestrzeni (prawa strona, czyli następnik implikacji). Twierdzenie Birkhoffa mówi, że tak właśnie jest.

Ponieważ F jest obrotem, to przypadek $F^{-1}(A) = A$ zachodzi tylko wtedy, gdy $A = \emptyset$ lub $A = \mathbb{S}^1$. A zatem A musi być albo zbiorem miary zero albo pełnej miary, co dowodzi ergodyczności miary. Oczywiście miara ta jest niezmiennicza względem obrotu, tym samym założenia twierdzenia ergodycznego są spełnione dla miary.

Przykład 15. Niech $\mathbb{S}^1$ będzie okręgiem na płaszczyźnie zespolonej oraz $F : \mathbb{S}^1 \to \mathbb{S}^1$ dana jest wzorem $F(x) = xe^{2\pi i\alpha}$, gdzie $\alpha \notin \mathbb{Q}$. Niech teraz f będzie funkcją charakterystyczną odcinka $[a, b]$ na okręgu, gdzie przez odcinek $[a, b]$ na okręgu rozumiemy zbiór

$$[a, b] := \{z \in \mathbb{C} \ : \ z = e^{2i\pi\phi}, \phi \in [a, b]\}.$$

Miara Lebesgue'a λ na okręgu zdefiniowana jest jako $\lambda([a, b]) = b - a$ jest ergodyczna względem odwzorowania F oraz F zachowuje miarę λ. Wtedy

$$\frac{1}{n}\sum_{j=0}^{n-1} f(F^j(x)) = \frac{\#\{0 \leqslant j < n \ : \ F^j(x) \in [a, b]\}}{n}$$

oraz z twierdzenia ergodycznego:

$$\frac{\#\{0 \leqslant j < n \ : \ F^j(x) \in [a, b]\}}{n} \to \int_{\mathbb{S}^1} f d\lambda = \lambda([a, b]) = b - a.$$

Innymi słowy – odsetek tych punktów, które pod działaniem obrotu na okręgu o kąt α zawitają do pewnego przedziału $[a, b]$ będzie zbiegał do długości tego przedziału.

W szczególności wynika z tego, że dla dowolnej wartości kąta w zakresie liczb π, e, $\sqrt{3}$, $7 - \sqrt{44}$, częstotliwość „odwiedzania" odcinka $[0, \frac{1}{7}]$ jest taka sama i równa długości tego odcinka, to jest $\frac{1}{7}$.

Zauważmy bardzo duże podobieństwo tego przykładu z definicją liczby normalnej. Rozważając mianowicie inną funkcję możemy udowodnić następujące twierdzenie.

Twierdzenie 16 (Borel [3]). *Prawie wszystkie liczby są normalne (w systemie dziesiętnym).*

Dowód. Niech $x \in [0, 1]$ oraz ustalmy ciąg w długości $|w| = k$. Rozważmy miarę Lebesgue'a $P = \lambda$ oraz $F(x) = \{10 \cdot x\}$. Niech ponadto

$$f(x) = \begin{cases} 1, & x \in [0, w; 0, w + 10^{-k}), \\ 0, & \text{w p.p.} \end{cases}$$

Wszystkie założenia twierdzenia ergodycznego są spełnione oraz lewa strona twierdzenia ergodycznego to $m(x, n) = F(w, x, n)$, prawa zaś to $10^{-|w|}$. Zatem dla prawie wszystkich x zachodzi

$$F(w, x, n) \to 10^{-|w|}.$$

□

Istnieje znacznie mocniejsza wersja twierdzenia Borela.

Twierdzenie 17 (Borel [3]). *Prawie wszystkie liczby są normalne jednocześnie we wszystkich systemach pozycyjnych o podstawach będących liczbami naturalnymi począwszy od* 2.

Dowód. Niech $N_n \subset [0, 1]$ będzie zbiorem wszystkich liczb normalnych w systemie o podstawie $n \geqslant 2$ oraz niech $M_n = [0, 1] \setminus N_n$. Należy teraz udowodnić, że

$$\lambda\left(\bigcap_{n\geqslant 2} N_n\right) = 1.$$

Z poprzedniego twierdzenia Borela wynika, że $\lambda(N_n) = 1$ dla wszystkich $n \geqslant 2$, przeto $\lambda(M_n) = 0$. Wtedy z własności miary wynika, że

$$\lambda\left(\bigcup_{n\geqslant 2} M_n\right) \leqslant \sum_{n\geqslant 2} \lambda(M_n) = 0.$$

Stąd

$$\lambda\left(\bigcap_{n\geqslant 2} N_n\right) = 1 - \lambda\left(\bigcup_{n\geqslant 2} M_n\right) = 1.$$

□

Z powyższego twierdzenia wynika, że jeśli tylko wybierzemy dowolną liczbę x, to z prawdopodobieństwem równym 1 ta liczba jest normalna we wszystkich systemach pozycyjnych o podstawie całkowitej i większej od 1. Z drugiej jednak strony nie posiadamy wielu narzędzi, by istotnie pokazać, że losowo wybrana liczba (lub ulubiona przez Czytelnika liczba z przedziału $[0,1]$) jest normalna w choć jednej wybranej bazie. Prostym tego przykładem są części ułamkowe liczb π, e oraz $\sqrt{2}$ – ich normalność jest do dziś problemem otwartym. Podobnie znacznie szersze jest pytanie o to, czy wszystkie niewymierne liczby będące pierwiastkami wielomianu o współczynnikach całkowitych są normalne. I na to pytanie nie znamy odpowiedzi.

Twierdzenie ergodyczne ma wiele innych i ciekawych zastosowań. Można z jego pomocą na przykład określić, jakie jest prawdopodobieństwo, że pierwszą cyfrą potęgi dwójki jest 4 i że jedynka jest najczęściej występującą pierwszą cyfrą takiej potęgi (wynik dla 4 to $\log_{10}\frac{5}{4} \approx 0,0969$, dla 1 zaś to $\log_{10} 2 \approx 0,301$). Można też rozważyć podobny problem dla drugich i kolejnych cyfr, wreszcie można podobne pytanie zadań w kontekście rozwinięcia liczby rzeczywistej w ułamek łańcuchowy (zobacz [8]).

Na koniec tej sekcji zauważmy jeszcze, że liczby normalne są jednocześnie punktami generycznymi miary Lebesgue'a dla przekształcenia F opisanego w dowodzie twierdzenia Borela. W dalszej części zobaczymy, jaką inną cechę posiadają punkty generyczne czyli liczby normalne, tym razem jednak przyjrzymy się aspektowi czysto topologicznemu.

Punkty tranzytywne

Rozważmy ponownie układ dynamiczny $(10^{\mathbb{N}}, \sigma)$. Rozpatrzmy ciąg Champernowne'a

$$c = 0123456789101112131415161718192021\ldots.$$

Ostatnim problemem, jakim zajmiemy się w tym artykule jest próba zbadania dynamiki takiego ciągu. Rozumiemy przez to próbę opisania, jak wygląda orbita ciągu c. W szczególności pytaniem, które nas interesuje, jest gęstość tej orbity.

Twierdzenie 18. *Zbiór* $\mathrm{orb}(c)$ *jest gęsty w* $10^{\mathbb{N}}$.

Zanim przedstawimy (poglądowy) dowód, wprowadźmy kilka oznaczeń dla uproszczenia zapisu. Jest to notacja na wyrazach lub ciągach, nie na liczbach!

- $0^1 = 0, 0^2 = 00, 0^3 = 000, \ldots,$
- $(123)^1 = 123, (123)^2 = 123123, (123)^3 = 123123123, \ldots,$
- $(123)^\infty = 123123123123\ldots$ (powtarzanie nieskończenie wiele razy).

Jeśli teraz na przykład $x = (0123456789)^\infty$, to

$$\begin{aligned} a^{(1)} &= 01234567890^\infty, \\ a^{(2)} &= (0123456789)^2 0^\infty, \\ a^{(3)} &= (0123456789)^3 0^\infty, \\ a^{(k)} &= (0123456789)^k 0^\infty \end{aligned}$$

definiuje ciąg zbieżny do x (dlaczego)? Przechodzimy teraz do dowodu twierdzenia.

Dowód. Ustalmy $x \in 10^{\mathbb{N}}$. Niech w_1 będzie blokiem 10 początkowych wyrazów ciągu x. Z normalności ciągu c blok w_1 pojawia się pewnym miejscu c, powiedzmy w miejscu k_1. Wobec tego

$$d(\sigma^{k_1}(c), x) \leqslant 2^{-10},$$

gdyż porównywane ciągi mają na pewno pierwszych 10 wyrazów identycznych.

Rozważmy teraz blok 10^2 początkowych wyrazów ciągu x. Niech będzie to blok w_2, który pojawia się w c nieskończenie wiele razy. Na pewno więc pojawia się na miejscu $k_2 > 10^2 k_1$. Wtedy też

$$d(\sigma^{k_2}(c), x) \leqslant 2^{-10^2}.$$

Postępując teraz analogicznie konstruujemy przez indukcję taki ciąg (k_n), dla którego spełnione są warunki:

- $k_n > 10^n k_{n-1}$,
- $d(\sigma^{k_n}(c), x) \leqslant 2^{-10^n}$.

Wprost z konstrukcji wynika teraz, że

$$d(\sigma^{k_n}(c), x) \to 0$$

gdy $n \to +\infty$. Tym samym wykazaliśmy gęstość zbioru $orb(c)$. □

Bardziej formalnie: dla dowolnego x oraz dowolnego $\varepsilon > 0$ istnieje takie $n > 0$, że $d(\sigma^n(c), x) < \varepsilon$.

Powyższe twierdzenie można ująć słownie następująco: ustalając dowolnie wybrany ciąg x, obojętnie jakie posiadający wyrazy, zawsze jesteśmy w stanie znaleźć taki punkt orbity ciągu Champernowne'a, który jest dowolnie blisko danego ciągu. Zauważmy, że sam dowód nie korzysta istotnie z tego, w jaki sposób ciąg zadający liczbę normalną został opisany. Wystarczyła nam wiedza co do tego, że jest to ciąg normalny, a więc każda sekwencja cyfr musi wystąpić w nim nieskończenie wiele razy, a więc na dowolnie dalekim miejscu ciągu. Tym samym możemy powyższe twierdzenie uogólnić następująco.

Twierdzenie 19. *Niech d będzie dowolnym ciągiem z $10^{\mathbb{N}}$ i takim, dla którego liczba $0,d$ jest normalna. Wtedy orbita ciągu d jest gęsta w $10^{\mathbb{N}}$. W szczególności, prawie wszystkie elementy zbioru $10^{\mathbb{N}}$ są punktami tranzytywnymi w $(10^{\mathbb{N}}, \sigma)$.*

Literatura

[1] Becher V., Carton O. (2018) *Normal Numbers and Computer Science*. In: Berthé V., Rigo M. (eds) Sequences, Groups, and Number Theory. Trends in Mathematics. Birkhäuser.

[2] Besicovitch, A. S. (1935), /emphThe asymptotic distribution of the numerals in the decimal representation of the squares of the natural numbers, Mathematische Zeitschrift, 39: 146–156, doi:10.1007/BF01201350

[3] Borel, E. (1909) *Les probabilités dénombrables et leurs applications arithmétiques*, Rendiconti del Circolo Matematico di Palermo, 27: 247–271, doi:10.1007/BF03019651

[4] Champernowne, D. G. (1933),*The construction of decimals normal in the scale of ten*, Journal of the London Mathematical Society, **8** (4): 254–260, doi:10.1112/jlms/s1-8.4.254.

[5] Copeland, A. H.; Erdős, P. (1946), *Note on normal numbers*, Bulletin of the American Mathematical Society, 52 (10): 857–860, doi:10.1090/S0002-9904-1946-08657-7

[6] Martin, G. (2001). Absolutely abnormal numbers. The American Mathematical Monthly, 108 (8), 746–754.

[7] Nakai, Y.; Shiokawa, I. (1992), *Discrepancy estimates for a class of normal numbers*, Acta Arithmetica, 62 (3): 271–284, doi:10.4064/aa-62-3-271-284

[8] Pasquinelli, I., *Birkhoff Ergodic Theorem and Applications*, notatki do wykładu, http://www.maths.dur.ac.uk/users/irene.pasquinelli/dyn/C3L5.pdf (dostęp 18.05.2020).