Marek Szydłowski*
Paweł Tambor**

* Centrum Układów Złożonych
Uniwersytet Jagielloński
** Centrum Kopernika Badań Interdyscyplinarnych
Uniwersytet Jana Kochanowskiego w Kielcach

* Theory of Complex Systems Department
Jagiellonian University
** Copernicus Center Interdisciplinary Studies
Jan Kochanowski University in Kielce

EMERGENTNY I UNIWERSALNY CHARAKTER PRAWA ROZKŁADU ZIPFA W NAUCE

The Emergent and Universal Nature of Zipf's Law in Science

Słowa kluczowe: prawo Zipfa, relacje potęgowe, złożoność, emergencja.

Streszczenie

Prawo rozkładu Zipfa znajduje zastosowanie w bardzo wielu i często odległych od siebie dziedzinach rzeczywistości fizycznej, biologicznej czy też społeczno-ekonomicznej. Nasze rozważania umieszczamy w kontekście dyskusji filozoficznej na temat złożoności badanych układów fizycznych, która to złożoność daje się jednak efektywnie wyrazić przez proste relacje potęgowe. Pokazujemy, że pewne wielkości charakterystyczne dla relacji opisywanych przez układy fizyczne mają charakter relacji potęgowych. Prawo Zipfa jest emergentne z praw losowego zachowania układu rządzonych przez prawa wyprowadzane z teorii procesów stochastycznych. Z naszej analizy wynika, że prawo Zipfa świadczy o emergencji o charakterze metodologicznym.

Key words: Zipf's law, power-law relations, complexity, emergence.

Abstract

In the paper we discuss Zipf's law and its variety of applications in the broad range of scientific disciplines from physics, biology, music to sociology and economics. Our analyses are placed in the context of the philosophical debate on complexity and emergence in science. We show that Zipf's law has the form of power-law scaling and in its nature is emergent and derivable from a more fundamental theory (the theory of stochastic processes). During this reconstruction we are dealing with a methodological type of emergence.

Wstęp

W pracy przedstawiamy argumentację na poparcie tezy, że fakt opisu świata w kategoriach prawa Zipfa jest efektem takich własności samego świata, że jesteśmy w stanie ująć w postaci prostych formuł dość złożone zachowania układów fizycznych[1]. Wskazujemy również na emergentny charakter prawa Zipfa, dla którego poszukujemy struktur głębszych z poziomu dolnego.

Prawo Zipfa jest świadectwem skuteczności opisu świata w jego różnych skalach i obszarach w kategoriach prostych potęgowych relacji pomiędzy zmiennymi opisującymi zjawiska. Zależność potęgowa nie opiera się na jakiejś fundamentalnej matematycznej strukturze, jaką definiuje Tegmark w swojej koncepcji *Mathematical Universe* (algebry Boole'a). Prawo rozkładu Zipfa może być klasycznym przykładem emergencji, podobnym do emergencji termodynamiki z mechaniki statystycznej. Fakt skuteczności opisu przy pomocy tak niezwykle elementarnego prawa w badaniach trendów na rynkach finansowych, w muzyce, układzie immunologicznym, programach komputerowych, rozkładzie pożarów w miastach czy też grawitacji kwantowej sugeruje, może być także świadectwem pewnych pozaempirycznych kryteriów epistemologicznych, takich jak: prostota, piękno, ekonomia środków poznawczych. Nasza generalna teza jest następująca: prawo Zipfa jest emergentne z praw losowego zachowania układu rządzonych przez prawa wyprowadzane z teorii procesów stochastycznych i jako takie spełnia kryteria emergencji w jej wersji metodologicznej.

Organizacja naszej pracy jest następująca. W części pierwszej dokonujemy krótkiego omówienia prawa Zipfa oraz przeglądu jego przykładowych zastosowań. W części drugiej podajemy argumenty na rzecz emergentnej natury prawa Zipfa oraz wskazujemy na ontologiczny i metodologiczny charakter tej emergencji, sugerując, że może to być prawo wyrażające własność skalowania świata. W części ostatniej formułujemy także główne wnioski wynikające z filozoficznej refleksji nad uniwersalnością prawa Zipfa.

1. Prawo rozkładu Zipfa – kontekst odkrycia

Zacznijmy od próby wydobycia pewnych relacji ilościowych w lingwistyce. Teksty pisane tworzone są przez ludzi, ale dobrze wiemy z badań lingwistycznych, że posiadają pewne złożone struktury na różnych poziomach: na poziomie morfologii słów, składni zdaniowej itd.[2] Rodzi się pytanie, czy teksty mogą po-

[1] Por. M. Tegmark, *The mathematical universe*, „Foundations of Physics" 2008, nr 38, s. 101–150.

[2] Por. J. R. Pierce, *Symbole, sygnały i szum*, Biblioteka Problemów, PWN, Warszawa 1967.

siadać pewne ogólne prawidłowości, które moglibyśmy ująć w relacje ilościowe. W tym celu rozważmy określony tekst czy też korpus tekstów i skonstruujmy listę rankingową słów według malejącej liczby wystąpień w tekście. Jest to najprostszy sposób uporządkowania słów. W ten sposób każdemu słowu zostaje przyporządkowana ranga $r(w)$. Tym samym najczęściej występujące słowo będzie posiadać rangę 1, drugie co do częstości występowania rangę 2 itd. (tab. 1).

Tabela 1

Przykładowe rangi $r(w)$ oraz częstości $c(w)$ dla różnych słów wg *Słownika frekwencyjnego polszczyzny współczesnej*

Ranga	Częstość	Słowo
1	14767	w
2	12473	i
3	11093	się
4	8750	na
5	7878	nie
6	7605	z
7	6004	do
8	5233	to
9	4675	że

Prawo Zipfa: Jeżeli słowo w_1 ma rangę 10 razy większą niż słowo w_2, to słowo w_1 ma częstość 10 razy mniejszą niż słowo w_2. Częstość słowa jest odwrotnie proporcjonalna do jego rangi,

$$c(w) \sim \frac{A}{r(w)}.$$

W latach 30. ubiegłego wieku nauczyciel języka niemieckiego na Uniwersytecie Harvarda odkrył niezwykle interesującą zależność pomiędzy rangą słowa w korpusie językowym a częstością występowania słów. Policzył on zależność częstotliwości występowania słów (liczby wystąpień) w języku angielskim[3] $h(w)$ jako funkcję liczby r określającej rangę. *De facto* dokonał on tego, co nazwalibyśmy histogramem; związek między wielkościami został opisany zależnością hiperboliczną[4]:

---

[3] Gdzie „the” ma rangę 1, „and” – rangę 2, „of” – rangę 3 itd.

[4] Por. G. K. Zipf, *Human behavior and the principle of least effort*, Addison – Wesley Press 1949.

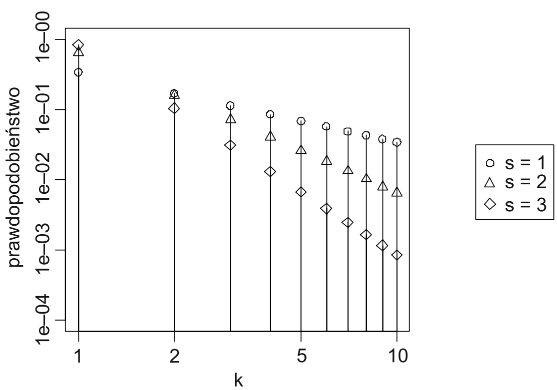
$$h(w) = \frac{A}{r(w)}, \quad h(w) \propto \frac{1}{r(w)}.$$

Ta zależność posiada niezwykle prostą interpretację. Im większa ranga, tym mniejsza częstość. Co to znaczy? Jeśli jakieś słowo, powiedzmy w_1, ma rangę 10 razy większą niż inne, powiedzmy w_2, to słowo w_1 ma częstość 10 razy mniejszą niż słowo w_1.

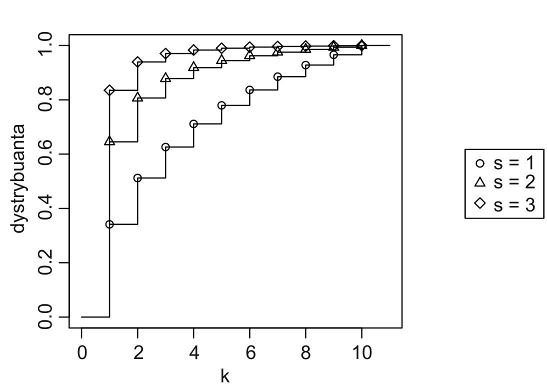
Zależność jest oczywiście równoważna stwierdzeniu, że dla każdego słowa w tekście iloczyn rangi i częstości jest stały. Jeśli N to liczba elementów i k jest ich rangą, to formalnie:

$$f(k,s,N) = \frac{1/k^s}{\sum_{n=1}^{N} \left(1/n^s \right)}$$

gdzie s to wartość eksponenty charakteryzującej rozkład.



Rys. 1. Rozkład Zipfa dla różnych s



Rys. 2. Wykresy dystrybuanty rozkładu Zipfa

Rozkłady ranga–częstość spełniające prawo Zipfa:

- lingwistyka – rozkład występowania słów,
- naukometria – rozkład cytowań artykułów naukowych,
- ekonomia – rozkład dochodów ludności,
- ekonomia – rozkład wielkości miast,
- astronomia – rozmiary meteorytów.

Później okazało się, że własność ta jest uniwersalna, tzn. analogiczne relacje zachodzą dla innych języków, także dla tekstów pisanych w językach etnicznych. Dziś prawo Zipfa należy do kanonu dziedziny ilościowych badań praw występujących w języku, zwanej lingwistyką kwantytatywną.

Reasumując, prawo Zipfa wyraża fakt, że w dowolnym języku *większość* słów jest używana rzadko i gdy uporządkujemy słowa danego języka według częstości ich występowania w wypowiedziach, to częstość n-tego słowa w rankingu jest odwrotnie proporcjonalna do $\frac{1}{n^{\alpha}}$, gdzie α jest bliskie jedności. Prawo to jest rodzajem uniwersalnej nieliniowej relacji językowej; stwierdza ową relację, ale nie daje wyjaśnienia jej natury, podobnie jak w analogiczny sposób prawo grawitacji Newtona nie daje odpowiedzi, jaka jest istota grawitacji, którą odkrył Albert Einstein.

Spójrzmy trochę ogólniej na prawo Zipfa, u podstaw którego leży „czynność" rangowania pewnego zbioru danych. Załóżmy, że posiadamy pewną liczbę n danych (i = 1,2,…,n), dla których ustalamy relację słabego porządku

$$x_1 \geq x_2 \ldots \geq x_2.$$

Wówczas treść prawa Zipfa może być zapisana w formie:

$$x(r) = \frac{C}{r^{\alpha}}$$

gdzie parametr á jest niezwykle bliski jedności. Logarytmując obustronnie zależność (3), uzyskujemy $x(r)$ jako funkcję n na skali logarytmicznej w postaci zależności liniowej:

$$\ln x(r) = \ln C - \alpha \ln r \quad (\textit{ln} \text{ oznacza logarytm przy podstawie } e).$$

Załóżmy dalej, że $x(r)$ jest pewną zmienną losową[5]. Wówczas, ze statystycznego punktu widzenia, prawo Zipfa jest modelem średniej albo ln $x(r)$ jako funkcji liniowej (regresja liniowa) ln (r), tj.:

$$E\,(\log r) = C - \alpha \log r.$$

Jednakże wykres relacji log-log zrangowanych danych nie stanowi rygorystycznego testu. Wobec tego rodzi się pytanie: czy można zaproponować inne

[5] Por. W. Li, *Zipf's law everywhere*, „Glottometrics 2003, nr 5, s. 14–21.

zależności funkcjonalne jako alternatywę lepiej dopasowującą te same dane? W literaturze można odnaleźć wiele alternatyw do prawa Zipfa sformułowanego dla zrangowanych danych, jak chociażby następujący rozkład[6]:

$$x(r) = \frac{C}{r^a B^r} \quad \text{albo } E\ logx(r) = c - alogr - be^{logr},$$

gdzie $b = ln\ B$; czy też wariant rozkładu Weibulla:

$$E\ (ln\ \mathrm{x}(r)) = \mathrm{c} - a\ \ln r - \mathrm{be}^{\beta \ln r},\ \beta - \text{dodatkowa stała.}$$

W literaturze dotyczącej prawa Zipfa toczy się dyskusja na temat odstępstwa – systematycznych odchyleń od linii prostej w log-log wykresie funkcji liniowej[7]. Nie istnieją jednakże żadne dostateczne powody uznania, że prawo Zipfa należałoby zastąpić innym. Jest tak z tego powodu, że inne funkcje dopasowują zrangowane dane lepiej[8]. W rozstrzygnięciu kwestii, jak ma się prawo Zipfa do innych funkcji dopasowujących dane, naturalne wydaje się zastosowanie kryterium Akaike czy też bayesowskiego kryterium BIC[9] .

Dokonamy teraz krótkiego przeglądu zastosowań prawa Zipfa w różnych dziedzinach i podamy literaturę, która może poprowadzić dalej zainteresowanych poszerzeniem wiedzy. Zacznijmy od językoznawstwa, gdzie prawo Zipfa pozostaje najbardziej znanym prawem językowym sformułowanym w ramach tzw. lingwistyki kwantytatywnej.

1.1. Prawo Zipfa w lingwistyce kwantytatywnej

Sformułowanych zostało kilka hipotez, które w zadowalający sposób tłumaczą pojawienie się – emergencję – prawa Zipfa w pewnych kontekstach pozajęzykowych. Natomiast nie istnieje jakiś ogólnie akceptowalny model matematyczny tego zjawiska w odniesieniu do słów w tekstach tworzonych przez ludzi[10].

---

[6] Por. G. U. Yule, *A mathematical theory of evolution based on the conclusions of dr J.C. Willis f.r.s.*, „Philosophical Transactions" 1925, s. 21–87.

[7] Por. J. R. Piqueira, L. H. Monteiro, T. M. Magalhaos, R. T. de Ramos, R. B. Sassi i E. G. Cruz, *Zipf's law organizes a psychiatric ward*, „Journal of Theoretical Biology" 1999, nr 198, s. 439–443.

[8] Por. W. Li, *Zipf's law...*, s. 15.

[9] Wyselekcjonowanie prawa Zipfa z użyciem kryterium prostoty i pewne idee w tym kontekście były dyskutowane w pracach: R. E. Quandt, *Statistical discrimination among alternative hypotheses and some economic regularities*, „Journal of Regional Science" 1964, nr 5, s. 1–23 oraz C. M. Urzua, *A simple and efficient test for Zipf's law*, „Economics Letters" 2000, nr 66, s. 257–260.

[10] Polecamy strony z interdyscyplinarnych bibliografii Wentiana Li (gttp://www.nslij-genetics.org/wli/zipf/) oraz A. Pawłowskiego (http://lingwistyka.uni.wroc.pl/bgl/).

Ogólnie znany jest fakt, że z prawa Zipfa można wyprowadzić tzw. prawo Lotki. W tym celu rozważa się funkcję $F(c)$ jako liczbę typów o częstości c. W ten sposób otrzymujemy *F(1)* – liczbę słów występujących jeden raz, *F(2)* – dwa razy, etc. Przez częstość częstości rozumiemy liczbę słów (typów) o danej częstości. Możemy sporządzić wykres funkcji częstości $F(c)$ jako funkcję *e* w skali logarytmicznej, który dobrze odzwierciedlany się linią prostą. Formalnie: jeżeli $F(c_1) = 10\ F(c_2)$, to $100c_1 = c_2$.

Równoważnym stwierdzeniem jest, że iloczyn częstości $F(c)$ i kwadratu częstości c jest stały albo:

$$F(c) \cong \frac{A}{c(c+1)},$$

gdzie A jest liczbą słów pojawiających się w tekście. Okazuje się, że połowa słownictwa każdego tekstu to słowa pojawiające się tylko raz (tzw. *hapaksy*).

Wyprowadzenie prawa Lotki jest elementarne na bazie wprowadzonych pojęć. Jeśli przypomnimy sobie, że oznaczyliśmy przez $F(c)$ liczbę typów o częstości c, $c(r)$ będzie częstością typu o randze r, a $r(c)$ największą rangą typu o częstości c, to prawo Zipfa $c(r) = \frac{A}{r}$, gdzie [...] – część całkowita danej liczby nie większa od tej liczby.

Stąd z prawa Zipfa $c(r)=0$, wtedy i tylko wtedy, gdy $r > A$. Zatem A to liczba słów występujących w tekście, w ten sposób: $r(c) = \frac{A}{c}$

$$F(c) = r(c) - r(c+1) = \frac{A}{c} - \frac{A}{c+1} = \frac{A}{c(c+1)}$$

Stąd otrzymaliśmy, że prawo Lotki można wyprowadzić z prawa Zipfa. Dodajmy, że prawo Lotki zostało sformułowane dziesięć lat wcześniej od prawa Zipfa i w dodatku w zupełnie odmiennym kontekście rozkładu cytowań prac naukowych[11]. Jego treść jest niezwykle prosta: *liczba autorów cytowanych n-razy jest odwrotnie proporcjonalna do* n^2.

Spotykamy się tutaj z sytuacją, w której prawo Zipfa pojawia się poza lingwistyką. Jest to sytuacja typowa, ponieważ – jak pokażemy dalej – tym prawem opisuje się także rozkład dochodów ludności (prawo Pareto, zasada 80/20), czy też rozkład wielkości miast (prawo Gibrata). Próba ilościowego ujęcia zjawiska korpusu językowego doprowadziła nas do wniosku, że ranga słowa w tekście w przybliżeniu opisywana jest prawem Zipfa, zgodnie z którym ranga słowa jest odwrotnie proporcjonalna do jego częstości.

[11] Por. A. J. Lotka, *The frequency diystribution of scientific productivity*, „J. Washington Academy Sciences” 1926, nr 16, s. 317–324.

Wykryte prawidłowości językowe są słuszne dla bardzo szerokiego zakresu tekstów w różnych językach przy dość nieostrej definicji słowa[12]. W korpusie języka naturalnego, częstotliwość występowania słów jest odwrotnie proporcjonalna do pozycji w rankingu.

Lingwiści podają odstępstwa wykresu ranga–częstość od prawa Zipfa i wskazują na ich systematyczność. Wykres ranga–częstość może zależeć od rodzajów tekstów, dla których badamy tę relację. Zaproponowany został wzór Benoita B. Mandelbrota, w którym zostały *a priori* wprowadzone nowe elementy uzależniające $c(w)$ od badanego tekstu[13].

$$c(w) \cong \left[\left(\frac{\rho + A}{\rho + r(w)} \right)^{B} \right], B > 1.$$

Dla tekstów krótkich $B < 1$, natomiast dla długich $B > 1$. Wzór Mandelbrota jest zbyt prosty, by dobrze dopasowywać wykresy ranga–częstość czy liczba typów–liczba okazów[14].

1.2. Prawo Zipfa w muzyce

Prawo Zipfa jest stosowane w teorii muzyki[15]. Muzycy powołują się na Galena, który pisze: „Piękno nie zawiera się w składnikach, ale w harmonicznym złożeniu całości”[16]. Prawo Zipfa jest ważne w kontekście badań struktur muzycznych, ich hierarchicznej organizacji i stanów emocjonalnych (smutek, gniew, poczucie szczęścia itp.). Manaris wykorzystuje prawo Zipfa, będące odzwierciedleniem pewnych własności skalowania (pojawiających się na wielu polach: od ekologii do muzyki), do klasyfikacji muzyki[17]. Odkrywa się pewne formalne

[12] Por. L. L. Goncalves, L. B. Goncalves, *Fractals power law in literary English*, „Physica A” 2006, nr 360, s. 557–575; G. Gottfried, *Teoria poznania od Kartezjusza do Wittgensteina*, Wydawnictwo WAM, Kraków 2007; H. Xiao, *On aplicalibility of Zipf's law in chinese word frequency distribution*, „Journal of Chinese Language of Computing” 2000, nr 18, s. 33–46; S. Shtrikman, *Some comments on Zipf's law for the Chinese language*, „Journal of Information Science” 1994, nr 20, s. 142–143.

[13] Por. B. B. Mandelbrot, *Structure formalle dee textes et communication*, „Word” 1954, nr 10, s. 1–27.

[14] Por. L. Dębowski, *Zipf's law against the text size: A half-rational model*, „Glottometrics” 2002, nr 4, s. 49–60.

[15] Por. B. Manaris i in., *Music classification and aesthetics*, „Computer Music Journal” 2005, nr 29, s. 55–69.

[16] „Beauty does not consist in the elements, but in the harmonious proposition of the parts” (tłum. autorzy).

[17] Por. D. H. Zanette, *Zipf's law and the creation of musical context*, CoRR, cs.CL/ 0406015 (2004).

struktury fraktalne, które zdaniem L. Solomona mają wpływ np. na binarną strukturę pierwszego z cyklu sześciu Ecossaises W.083 Beethovena i samopodobieństwo wykorzystanych tam motywów[18].

W ten sposób, przy pomocy metod ilościowych, odkrywane są znane już w starożytności związki muzyki z naturą i matematyką. Poszukuje się tym samym obiektywnych wzorców piękna muzycznego, dostrzegając struktury samopodobne, dla których charakterystyczne są relacje potęgowe.

W roku 1990 H. J. Hsu i A. J. Hsu odkryli niezwykle interesującą zależność typu potęgowego pomiędzy liczbą kolejnych nut odległych od siebie o i półtonów od tejże wielkości i[19]. Jeśli F oznacza względną częstość wystąpienia interwału o długości i półtonów ($i = 0$ dla prymy, $i = 1$ dla małej sekundy, ..., $i = 8$ dla oktawy), to empirycznie można ustalić zależność $F \propto i^D$, gdzie D jest pewnym parametrem. Chociaż autorzy tego *explicite* nie stwierdzili, należy ją uznać za przejaw prawa Zipfa[20] . Parametr D zależy od badanego utworu i zawiera się pomiędzy 1,34 dla Toccaty fis-mol Bacha (BWV 910) poprzez 1,73 dla sonaty F-dur Mozarta (KV 533) do 2,42 dla Inwencji nr 1 C-dur Bacha (BWV 772). Autorzy znajdują wyjątki w postaci utworów Stockhausena, ale muzyka ta nie podpada pod reguły muzyki klasycznej.

Własność samopodobieństwa ze względu na zmienne przestrzenne posiadają klasyczne fraktale, ale takie skalowanie może również zachodzić w zmiennej czasowej – i to jest właśnie przypadek muzyki. Odkryty związek w muzyce ma znaczenie dla automatycznego komponowania muzyki. W roku 1975 Richard Voss i John Clark odkryli korelacje nie tyle między dwoma kolejnymi dźwiękami (charakteryzujące lokalne własności muzyki), ile globalne własności całych utworów w kategoriach widma mocy[21] .

Jak wiadomo, sygnał zależny od czasu możemy rozłożyć na sumę drgań harmonicznych o pewnych częstościach f_K i odpowiadających im amplitudach (rozkład Fouriera). Widmo mocy $S_V(f)$ sygnału zawiera informacje o udziale określonych mod częstości w pełnym widmie, powiedzmy dla fali dźwiękowej. Jest to informacja, ile energii tej fali przypada na drgania o określonej częstości f.

Zaobserwowano uniwersalność zależności potęgowej $S_V(f) \propto \frac{1}{f^{\beta}}, \beta \cong 1$ w często bardzo odległych sytuacjach[22]. Voss i Clarke wykazali, że muzyka wykazuje szum typu $1/f$ dla prawie wszystkich melodii. Kompozycje atonalne Stoc-

---

[18] I. Linstedt, *Fraktale i muzyka*, „Ruch Muzyczny" 2009, nr 7, s. 6.

[19] Por. K. J. Hsü, A. J. Hsü, *Fractal geometry of music*, „Proceedings of the National Academy of Science" 1990, nr 87 (3), s. 938–941.

[20] Por. D. Wolf, *Noise in physical system*, Springer Heidelberg, New York 1978.

[21] Por. R. F. Voss i J. Clarke, *1/f noise in music and speech*, „Nature" 1975, nr 258, s. 317; R. F. Voss, J. Clarke, *1/f noise in music from 1/f noise*, „Journal of Acoustical Society of America" 1978, nr 63, s. 258–263.

[22] Por. D. Wolf, op. cit.

khausena, Cartera są białym szumem ($\beta = 0$), natomiast utwory średniowieczne, a także Beethovena, Straussa czy Beatlesów charakteryzują się szumem typu $1/f$. Ponieważ szum jest typowy dla zjawisk przyrodniczych, Wolf powiada, że muzyka imituje charakterystyczny sposób, w jaki zjawiska przyrodnicze przebiegają w czasie. Zarówno muzyka, jak i szum są pośrednie pomiędzy białym szumem losowym ($\beta = 0$), a klasycznym brownowskim ruchem losowym ($\beta = 2$). Własność samopodobieństwa muzyki przejawia się w tym, że najmniejsza faza dla utworu klasycznego jest podobna do całego utworu[23]. Widzimy więc, że chociaż prawo Zipfa oddaje lokalne własności muzyki, badania widma mocy odkrywają jej głęboką własność globalną wyrażoną przez prostą relację potęgową.

1.3. Inne przykłady aplikacji prawa Zipfa

Oto najbardziej spektakularne aplikacje prawa Zipfa (wymieniamy je bez szczegółowej analizy, podając jedynie referencje):

1. Zależność między liczbą mieszkańców a liczbą określającą miejsce na liście rankingowej miast uporządkowanych według wielkości jest zależnością potęgową[24].
2. Liczba przedsiębiorstw o obrotach zawartych w określonym przedziale rośnie jak odwrotność miejsca na liście uporządkowanej według wartości tychże obrotów[25].
3. Liczba trzęsień ziemi rośnie od największych do najsłabszych według zależności potęgowej.
4. Funkcja rozkładu galaktyk w gromadach ma charakter potęgowy; podobnie rozkład rozmiarów galaktyk, rozmiaru planet, satelitów.
5. Erupcje wulkaniczne i ich rozmiary są zgodne z rozkładem Zipfa, to samo dotyczy rozmiarów wysp.
6. Proteiny i sieci metaboliczne posiadają własność niezmienniczości względem skali – topologię, dla której charakterystyczne są rozkłady Zipfa.
7. W przyrodzie występują tzw. relacje alometryczne. To one sprawiają, że nie istnieją krasnoludki czy też góry o wysokości 20 km. Energia biologiczna i jej transport wykazują własności skalowania (metabolizm vs. masa) dla rozmiarów aż do 27 rzędu.

---

[23] Por. Z.-Y. Su, T. Wu, *Music walk, fractals geometry in music*, „Physica A" 2007, nr 380, s. 418–428; M. Beltran del Rio, G. G. Cocho, G. G. Naumis, *Universality in tail of musical note rank distribution*, „Physica A" 2008, nr 387, s. 5552–5560.

[24] M. Gell-Mann, *Jaguar i kwark*, CiS, Warszawa 1996; B. B. Mandelbrot, *Structure formalle dee textes et communication*, „Word" 1954, nr 10, s. 1–27.

[25] Por. B. B. Mandelbrot, op. cit.

8. Sieci genetyczne i rozkłady rozmiarów gatunków wykazują własność rozkładu Zipfa; podobnie sieci rozkładów portów lotniczych względem ich rozmiarów.

9. Prawo Zipfa jest odkrywane w programach komputerowych[26].

10. To samo prawo opisuje rozkład fraktalnych dziur (*voids*) w wielkoskalowej strukturze Wszechświata[27].

11. Prawo Zipfa jest odkrywane w trendach ewolucyjnych w rynkach finansowych[28].

12. Każde otwarcie w szachach rozpoczyna ciąg dalszych ruchów, które mogą być reprezentowane przez graf, którego węzły są sytuacjami szachisty, natomiast krawędzie dozwolonymi ruchami z każdej pozycji. Autorzy badają częstości ruchów otwarcia i pokazują, że ich rozkład podlega uniwersalnemu prawu Zipfa[29].

2. Emergentny charakter prawa Zipfa

W artykule posługujemy się pojęciem emergencji, które wymaga szczególnego wyjaśnienia. Bardzo trudno sformułować spójną teorię emergencji, ponieważ istnieje wiele odmian emergencji, szeroki jest też zakres zastosowań tego pojęcia w nauce i filozofii nauki. O emergencji mówimy zarówno w kontekście dyskusji nad relacją umysł–ciało, jak i w teorii układów złożonych, samoorganizacji, chemii fizycznej czy nawet matematyce. Najczęściej podłożem do emergentyzmu jest dyskusja nad redukcjonizmem, fizykalizmem lub holizmem w filozofii nauki. Wydaje się, że między wyodrębnionymi poziomami rzeczywistości lub opisu można wskazać na istnienie pewnej hierarchii: wyróżnia się poziom niższy – bazowy i poziom wyższy, na którym pojawiają się własności nazywane emergentnymi.

Historycznie rzecz biorąc, za początek idei emergencji w filozofii można uznać prace J. S. Milla (*A system of logic*, 1843) i G. H. Lewesa. Mill związał ideę emergentyzmu w kontekście kazualnym ze wskazaniem na dwa możliwe rodzaje związków przyczynowych między złożoną całością a jej elementami składowymi. Otóż, albo można potraktować tę całość jako sumę części, albo na poziomie całości pojawiają się cechy nieobecne na poziomie składników. Szcze-

[26] Por. H. Zhang, *Discovering power law in computer programs*, „Information Processing and Management" 2009, nr 45, s. 477–483.

[27] Por. J. C. Gaite, *Zipf's law for fractal voids and a new void-finder*, „The European Physical Journal B" 2005, nr 47, s. 93–98.

[28] P. V. Balakrishnan, J. M. Miller, S. G. Shanker, *Power law and evolutionary trends in stock markets*, „Economic Letters" 2008, nr 98, s. 194–200.

[29] Por. B. Blasius, R. Tonjes, *Zipfs law in the popularity distribution of chess openings*, „Physics Review Letters" 2009, nr 103, s. 218701.

gólnych przykładów takiej sytuacji dostarcza choćby chemia. Sód i chlor, brane pod uwagę osobno, posiadają zupełnie inne własności niż chlorek sodu, który tworzą (NaCl). Całość, która powstaje jako prosta suma elementów (skutek jest sumą przyczyn), nazwana została przez Milla skutkiem homopatycznym. Odpowiednio skutek, który nie jest sumą przyczyn składowych, jest określany jako heteropatyczny. Lewes nazywa takie skutki emergentnymi w odróżnieniu od pierwszych, które określa mianem „rezultantów"[30].

Wyróżnionym kontekstem, w którym rozważa się istnienie efektów emergentnych, jest fakt istnienia hierarchii poziomów: dolnego (tożsamego ze składnikami danego układu) i górnego, który objawia skutki nieredukowalne do własności bazowych. Jak widzimy, emergencja jest tu traktowana synchronicznie, trzeba by raczej mówić o poziomie emergentnym, skutkach, własnościach emergentnych (jak nowe prawa, własności, itd.) lub po prostu o emergentach. Za klasyczny okres w historii pojęcia emergencji można uznać tzw. emergentyzm brytyjski (lata 20. ubiegłego wieku – prace S. Alexandra, C.L. Morgana i C.D. Broada[31]). Emergentyzm w wersji Alexandra nazywany jest ewolucyjnym i traktuje emergencję diachronicznie[32]. Bardziej współczesne odmiany emergencji wskazują na jej walor interakcyjny: własność emergentna w układzie złożonym może zostać wyjaśniona poprzez wskazanie na typ wzajemnych oddziaływań między jego częściami.

Jedna z trudności związanych ze sformułowaniem jednolitej koncepcji emergencji polega na tym, że istnieje wielość tzw. jednostek emergencji. Poza tym rozróżnia się konteksty:

a) ontologiczny – gdy mówimy o emergencji własności, struktury itp.;

b) epistemologiczny – gdy wiedzy o własnościach pewnego układu złożonego nie jesteśmy w stanie uzyskać poprzez analizę własności jego składników oraz

c) metodologiczny – gdy prawa na poziomie emergentnym nie da się bez specjalnych założeń wywieść z praw obowiązujących na poziomie bazowym.

Przy badaniu emergentnego charakteru praw potęgowych istotny jest dla nas właśnie kontekst metodologiczny.

Emergentny charakter prawa Zipfa zilustrujemy na przykładzie lingwistyki kwantytatywnej. Poszukujemy odpowiedzi na pytanie, dlaczego prawo Zipfa obowiązuje dla tak wielu rozkładów ranga–częstość. Proponowane są różne hipotezy wyjaśnienia prawa Zipfa, tj. wydedukowania go z bardziej fundamentalnego poziomu. Przy pewnej świadomości, że niektóre hipotezy mogą okazać się błęd-

[30] Por. R. Poczobut, *System – struktura – emergencja*, (w:) M. Heller, J. Mączka (red.), *Struktura i emergencja*, Biblos, Tarnów 2008, s. 88.

[31] Por. S. Alexander, *Space, time, and deity*, Macmillan Alexander, London 1920; C. Morgan, *Emergent evolution*, Holt, New York 1923; C. Broad, *The minds and its place in nature*, Routledge and Kegan Paul, London 1925.

[32] Por. R. Poczobut, op. cit.

ne bądź może istnieć wiele niezależnych hipotez, konstruowane są matematyczne modele tekstu oparte na założeniach takich jak: 1) pewien probabilistyczny model tekstu; 2) tekst traktowany jako wynik optymalizacji pewnych kryteriów; 3) słowa w tekście dają się wyodrębnić przy pomocy pewnego algorytmu.

Wymieńmy dla ilustracji kilka propozycji[33]. Pierwszą jest model „małpy na klawiaturze". W tej koncepcji zakłada się, że tekst jest ciągiem liter i odstępów. Tekst uzyskujemy, wciskając losowo klawisze klawiatury[34]. Z tych minimalistycznych założeń uzyskujemy to, co trzeba, żeby lista rankingowa słów spełniła prawo Zipfa–Mandelbrota. Jako wady modelu wymienia się to, że wszystkie słowa o długości n mają podobną częstość; proces stochastyczny w tym modelu ma własność ergodyczności (ustalone słowo ma bardzo podobną częstość w dostatecznie długim tekście); zróżnicowanie słownikowe tekstów jest banalne.

Kolejnym modelem jest tzw. losowanie ze sprzężeniem zwrotnym. Model został zaproponowany przez noblistę w dziedzinie ekonomii Herberta A. Simona (1916–2001)[35]. Oparty jest na kilku założeniach odnośnie do tekstu: 1) tekst jest ciągiem słów i jest otrzymywany poprzez losowanie kolejnych słów; 2) gdy losujemy słowo z n-tej pozycji tekstu, to prawdopodobieństwo wylosowania słowa, które zostało już wcześniej wylosowane, jest proporcjonalne do jego częstości na pozycjach od 1 do $(n-1)$ (prawo Gibrata), natomiast prawdopodobieństwo wylosowania nowego słowa jest niezerowe.

Innym modelem wykorzystującym pojęcie entropii jest model Mandelbrota, oparty na minimalizacji średniej długości słowa na bit. Podstawowe założenia modelu są następujące: tekst nie jest wynikiem procesu losowego, lecz optymalizacji; przy ustalonej liczbie okazów słów w tekście rozkład częstości słów maksymalizuje iloraz H/c, gdzie H jest entropią informacyjną słowa w tekście, c – średnią długością słowa w tekście; długość słowa o randze r jest proporcjonalna do $\ln r$.

Z założeń uzyskujemy, że lista rankingowa słów spełnia prawo Zipfa–Mandelbrota. Entropia słowa w tekście[36]:

$$H = -\sum_{r=1}^{\infty} \frac{\frac{c(r)}{N} \log c(r)}{N} \quad ,$$

[33] Por. L. Dębowski, op. cit.

[34] Por. G. A. Miller, *Some effects of intermiittent silence*, „American Journal of Psychology" 1957, nr 70, s. 311–314.

[35] Por. H. A. Simon, *On a class a skew distribution function*, „Biometrica" 1955, nr 42, s. 425–440; G. Udny Yule, *A mathematical theory of evolution*, „Philosophical Transactions of the Royal Society of London" 1925, nr 213.

[36] B. B. Mandelbrot, op. cit.

a średnia długość:

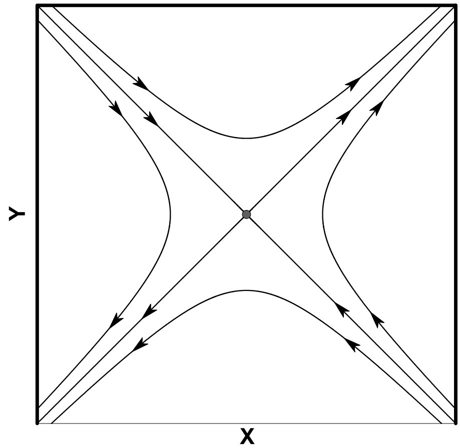
$$c \propto \sum_{r=1}^{\infty} \frac{c(r)}{N} \log r \,.$$

Żadna z prezentowanych dotychczas koncepcji nie wyjaśnia, czym jest słowo. Nie jest uwzględniany fakt, że słowa są jednostkami definiowanymi przez znaczenie. Wyjaśnienia opierają się na ukrytym założeniu, że słowa są elementami pewnego języka danego *a priori* itd.

3. Relacje potęgowe emergentną własnością świata

Zwróćmy uwagę, że tezę o uniwersalności danego prawa potęgowego można uzasadniać na dwa sposoby. Po pierwsze, twierdząc, że np. prawo Zipfa opisuje stan stacjonarny procesu dynamicznego, zmierzającego do powstania struktury, w przestrzeni fazowej będącego punktem siodłowym. W takim ujęciu własność opisana w kategoriach relacji potęgowych może być interpretowana jako własność układu samego w sobie.

Inną propozycją jest podejście epistemologiczne, w którym wskazuje się na szczególne własności naszego poznania. Załóżmy, że umysł, rozpoznając zjawisko, nie wie, w jakiej skali ono się rozgrywa i stara się w pierwszej kolejności rozpoznać je tak, aby nie posiadało swojej skali, czyli tak, aby jego opis nie zależał od wyboru specyficznej skali. Przywiązanie do skal wydaje się być ściśle związane z naszą ludzką egzystencją, podczas gdy sam poznający umysł nie musi czuć skal czasowych czy przestrzennych. Jeśli przyjąć takie założenie, to relacje, które próbuje ustalić pomiędzy zmiennymi opisującymi układ, powinny być niezmiennicze (nie zmieniać swojej formy) przy przeskalowaniach zmiennej niezależnej i zmiennych zależnych. Jeśli założyć dla prostoty, że mamy do czynienia z procesem dynamicznym opisywanym przez układ dynamiczny (deterministyczny model procesu dynamicznego) w przestrzeni fazowej (x, y), to łatwo pokazać, że zależność pomiędzy zmiennymi stanu układu będzie miała koniecznie charakter potęgowy. Natomiast w przestrzeni stanów układu (przestrzeni fazowej) ścieżki ewolucyjne układu odpowiadają tzw. punktowi siodłowemu. Na rysunku 3 (zwanym portretem fazowym) widzimy punkt równowagi zlokalizowany w początku układu, jedną parę trajektorii wychodzących z niego (separatrysy wychodzące) i jedną parę trajektorii wchodzących do niego (separatrysy wchodzące). Te proste separują pewne obszary, w których zależności pomiędzy zmiennymi są hiperbolami opisywanymi poprzez rozwiązania potęgowe. Przedstawiono także siodło dwuwymiarowe, ale w ogólności może to być siodło wielowymiarowe. Potęgowe zachowania rozwiązań $y(x)$ w pobliżu siodła są uniwersalne.



Rys. 3. Portret fazowy siodła dwuwymiarowego
Ewolucja układu jest reprezentowana przez krzywe fazowe. Są one w przestrzeni (x, y) opisywane przez relacje potęgowe. W szczególnym przypadku, gdy stałe a i b są różnych znaków, niezmiennicza względem skali wielkość jest reprezentowana przez krzywe we współrzędnych x, y, zwane krzywymi fazowymi.

Z drugiej strony wskazujemy, że nasze poznanie, chcąc uchwycić to, co istotne w procesie dynamicznym opisywanym przez układ dynamiczny $\frac{dx}{dt} = P(x, y), \frac{dy}{dt} = Q(x, y)$, poszukuje w pierwszym przybliżeniu relacji określającej wielkość niezmienniczą względem skali. Załóżmy, że poszukujemy relacji niezmienniczej względem zmiany skali:

$$x \to ax = x'; \quad y \to by = y'; \quad t \to ct = t'.$$

$$\frac{dx'}{dt'} = P(x', y'); \quad \frac{dy'}{dt'} = Q(x', y').$$

Równanie niezmiennika wyznaczamy z relacji:

$$\frac{dx}{ax} = \frac{dy}{by} = \frac{dt}{ct},$$

w ten sposób uzyskujemy prawo:

$$y \propto x^{\frac{b}{a}}.$$

Widzimy, że uzyskane prawo, wyrażone w kategoriach relacji potęgowych, znajduje uzasadnienie na gruncie własności naszego poznania. Umysł skaluje rzeczywistość, by potem określić równanie dla związku między wielkościami x i y; ta relacja ma charakter relacji potęgowej. Uważamy, że powyższy wniosek ma dla naszej dyskusji kapitalne znaczenie w związku z klasycznym w epistemologii problemem wiedzy. Odwołajmy się do jej klasycznego ujęcia jako prawdziwego i uzasadnionego przekonania. Otóż, jeśli udało nam się zrekonstruować pewien mechanizm emergencji relacji o cechach relacji typu potęgowego, świadczy to o tym, że mamy do czynienia z jakimś rodzajem pozaempirycznej intuicji, którą daje się uchwycić w pierwszej kolejności w postaci mechanizmu (jak wyżej).

Ponieważ w pracy jako wyróżnione traktujemy podejście ontologiczne i metodologiczne (nie epistemologiczne), kluczowe pozostaje wyjaśnienie, dlaczego proste prawa potęgowe opisują ogromną złożoność modelowanych procesów. Można oczywiście uważać to za fakt będący świadectwem matematyczności świata, ale tutaj pojawia się problem, jak wyjaśnić przypadki odstępstwa od opisu przez prawa potęgowe, a w szczególności prawo Zipfa, tzn. jak wyjaśnić, że te prawa są jedynie przybliżeniem badanej rzeczywistości. Z filozoficznego punktu widzenia jest to poszukiwanie poziomu bardziej fundamentalnego, z którego wyprowadzimy tę zależność potęgową.

Prawo Zipfa jest emergentne z praw losowego zachowania układu rządzonych przez prawa wyprowadzane z teorii procesów stochastycznych, jak pokazali to ostatnio Lin i Loeb[37]. Złożone układy opisujące oddziaływanie ludzi z otoczeniem, które wyrażają się za pomocą prostych relacji Zipfa, można wyprowadzić z praw losowego wzrostu skupisk: „Wyprowadzamy prosty model statystyczny, który tłumaczy wszystkie z tych praw skalowania, bazując na jednej wspólnej zasadzie pociągającej za sobą przypadkowy przestrzenny wzrost grup ludzkich we wszystkich skalach. Za pomocą tego modelu dokonuje się ważnych nowych predykcji dotyczących rozprzestrzeniania się chorób lub innych zjawisk społecznych”[38].

Z kolei Corominas-Murtra, Hanel i Thurner wskazują na to, że pewne procesy stochastyczne prowadzą do prawa Zipfa. W zakończeniu konkludują: „prawo Zipfa wyłania się (*emerguje*) jako prosta konsekwencja łamania symetrii kierunku w procesach stochastycznych”[39].

---

[37] Por. H. Lin, A. Loeb, *A unifying theory for scaling laws of human populations*, arXive:1501.00738, 2015.

[38] Tekst oryginalny: “We derive a simple statistical model that explains all of these scaling laws based on a single unifying principle involving the random spatial growth of clusters of people on all scales. The model makes important new predictions for the spread of diseases and other social phenomena”.

[39] Por. B. Corominas-Murtra, R. Hanel i S Thurner, *Understanding Zipf 's law with playing dice: history-dependent stochastic processes with collapsing sample-space have power-law rank distributions*, arXive:1407.2775, 2014.

Reasumując, uniwersalne prawo Zipfa znajduje wyjaśnienie w prawach teorii procesów stochastycznych. Wyjaśnienie znajduje również sam fakt jego uniwersalności. Można powiedzieć, że obecność prawidłowości potęgowych typu prawa Zipfa jest świadectwem zjawiska emergencji metodologicznej w nauce globalnej w tym samym sensie, jak termodynamika jest świadectwem emergencji z praw mechaniki statystycznej.