

Rachunek prawdopodobieństwa a fizyka

Mirosław LACHOWICZ, Andrzej PALCZEWSKI, Warszawa

1. Wprowadzenie — Klasyczna definicja prawdopodobieństwa

Wykład ten chcemy poświęcić przedstawieniu wpływu fizyki na rozwój rachunku prawdopodobieństwa. Dokładniej rzecz ujmując, chcemy w tym wykładzie pokazać, jak powstająca w drugiej połowie XIX w. fizyka statystyczna, stawiająca sobie jako cel wyjaśnienie na gruncie molekularno-cząsteczkowej teorii budowy materii podstaw termodynamiki, zderzyła się z zupełnym brakiem odpowiedniego aparatu matematycznego. Jak wreszcie, w drodze burzliwych dyskusji, w których padały zarzuty najcięższego kalibru, doprowadziła do zasadniczej rewizji podstaw rachunku prawdopodobieństwa, co w ostateczności zaowocowało powstaniem aksjomatycznej definicji prawdopodobieństwa sformułowanej przez Kołmogorowa.

Należy jednak zaznaczyć, że inspirujący wpływ fizyki statystycznej na prace Kołmogorowa nie jest widoczny tak bezpośrednio, jak np. wpływ mechaniki kwantowej na rozwój analizy funkcjonalnej.

Jak powszechnie wiadomo, rachunek prawdopodobieństwa wyrósł z prostych problemów, które pojawiły się przy grach hazardowych. Legendarną postacią, której przypisuje się zainicjowanie tych badań jest Antoine Gombaud chevalier de Mééré (1607–84). Jak podają niektóre źródła de Mééré zapoczątkował grę, w której należało obstawiać za, lub przeciw, wypadnięciu przynajmniej jednej *szóstki* w 4 kolejnych rzutach kością. De Mééré wyliczył, że prawdopodobieństwo tego, że w 4 rzutach nie wypadnie ani razu *szóstka* wynosi $(\frac{5}{6})^4 = \frac{625}{1296}$. Stąd prawdopodobieństwo wypadnięcia chociaż jednej *szóstki* jest $1 - \frac{625}{1296} = \frac{671}{1296} = 0.51775$. Ponieważ prawdopodobieństwo to jest większe od $\frac{1}{2}$, de Mééré obstawiając za wypadnięciem *szóstki* dość często wygrywał i wkrótce nie mógł znaleźć partnerów gotowych grać przeciwko niemu. Chcąc uatrakcyjnić grę zaproponował nowe zasady: rzucać należy dwiema kośćmi do wypadnięcia *podwójnej szóstki*. Pragnąc zachować swoją przewagę, de Mééré starał się obliczyć, ile należy wykonać rzutów, aby prawdopodobieństwo wypadnięcia *podwójnej szóstki* w grze dwiema kośćmi było takie samo jak wypadnięcie pojedynczej *szóstki* w grze jedną. Rozumowanie de Mééré było następujące: szansa otrzymania jednej *szóstki* jest sześciokrotnie większa niż szansa otrzymania dwóch *szóstek*, a zatem liczba rzutów powinna być sześciokrotnie większa od 4. Korzystając z tego rozumowania de Mééré zaproponował grę, w której obstawiał wypadnięcie *podwójnej szóstki* w 24 rzutach dwiema kośćmi. Niestety, w przeciwieństwie do poprzedniej gry, częściej przegrywał niż wygrywał. W tej sytuacji zwrócił się do znakomitego matematyka tych czasów, a swojego znajomego, Blaise Pascala (1623–62) z prośbą o wyjaśnienie powstałego paradoksu. Pascal przeprowadził odpowiednie obliczenia wykazując, że prawdopodobieństwo wypadnięcia co najmniej jednej *podwójnej szóstki* w 24 rzutach wynosi $1 - (\frac{35}{36})^{24} = 0.4914$ i dopiero przy 25 rzutach staje się większe od $\frac{1}{2}$ (dokładniej jest równe 0.50553).

Przytoczona historia znamy z wymiany korespondencji między B. Pascallem a P. de Fermatem. Należy jednak podkreślić, że de Mééré, choć niewątpliwie zwrócił uwagę wielkich matematyków na problemy prawdopodobieństwa, nie był pierwszym, który interesował się tymi zagadnieniami. Przed nim problemami gry w kości zajmowali się Cardano, Tartaglia, a nawet Galileusz. Ponadto należy przypuszczać, że de Mééré był bardziej zainteresowany teoretyczną niż praktyczną stroną zagadnienia.

Bogaty przegląd paradoksów rachunku prawdopodobieństwa znajdzie czytelnik w książce G. Székely *Paradoxes in probability theory and mathematical statistics*.

Paradoks petersburski polega na tym, że występująca tu zmienna losowa nie ma wartości oczekiwanej. Szerzej o tym paradoksie i jego rozwiązaniu pisze W. Feller w *Wzrycie do rachunku prawdopodobieństwa* tom I.

Nie chcąc prezentować tutaj całej historii rozwoju rachunku prawdopodobieństwa, powiedzmy tylko, że takie i podobne paradoksy wstrząsały matematyką przez cały wiek XVII i XVIII. Nie omijały one nawet najlepszych umysłów tamtej epoki. Przykładem może tu być tzw. *paradoks petersburski* dyskutowany przez Daniela Bernoulliego (1700–82). Gra w tym przypadku polegała na rzucaniu monetą do czasu wypadnięcia *orła*. Jeśli następowało to w r -tym rzucie, to gracz otrzymywał 2^r dukatów. Pytanie, które prowadziło do paradoksu brzmiało: ile należy zapłacić za prawo uczestniczenia w grze, aby była to „gra uczciwa”. Paradoksy te wskazywały na luki istniejące w podstawach budowanej nauki. Nie zmienia to jednak faktu, że wiek XVIII przyniósł burzliwy rozwój teorii prawdopodobieństwa i wielkie jej sukcesy: prawo wielkich liczb Bernoulliego i twierdzenie graniczne de Moivre-Laplace'a. Z punktu widzenia tego wykładu ważna jest również zaproponowana przez Pierre de Laplace'a (1749–1827) definicja prawdopodobieństwa, która przetrwała aż do początku XX w. i którą jeszcze teraz posługuje się szkolny wykład rachunku prawdopodobieństwa.

Definicja. Prawdopodobieństwo $P(A)$ zdarzenia A równa się stosunkowi liczby zdarzeń elementarnych sprzyjających zajściu zdarzenia A do liczby wszystkich zdarzeń elementarnych. Zakłada się przy tym, że wszystkie zdarzenia elementarne są jednakowo prawdopodobne.

Ta, tzw. klasyczna definicja prawdopodobieństwa, spotkała się w ciągu XIX w. z wielostronną krytyką. Jednym ze źródeł tej krytyki było ostatnie zdanie definicji. Zacytujmy w tym celu uwagę uczynioną przez Henri Poincarégo (1854–1912): „*Jak zdefiniować, że wszystkie zdarzenia elementarne są jednakowo prawdopodobne? Definicja matematyczna nie jest w tym przypadku możliwa; będziemy musieli przy każdym zastosowaniu podawać dodatkowe warunki, omawiać, że takie to a takie zdarzenia będziemy uważali za jednakowo prawdopodobne. Te uzupełnienia nie są zupełnie dowolne, jednak wymykają się matematycznej analizie, bo potem jak zostały zrobione nie poddaje się ich weryfikacji.*” Problem jest więc pewnego rodzaju błędnym kołem: chcąc zdefiniować prawdopodobieństwo należy określić zdarzenia jednakowo prawdopodobne, a żeby powiedzieć, że dwa zdarzenia są jednakowo prawdopodobne trzeba wiedzieć, co to jest prawdopodobieństwo.

Sytuacja ta nie jest beznadziejna we wszystkich przypadkach. Jest wiele takich problemów, gdzie zdarzenia elementarne wykazują pewną symetrię i ich jednakowe prawdopodobieństwo może być uzasadnione właśnie tą symetrią.

Drugi rodzaj niedostatków przedstawionej definicji jest związany z faktem, że w zasadzie definicję tę można stosować tylko do tego, co współcześnie nazywamy rozkładami skończonymi, tj. do zmiennych losowych przyjmujących tylko skończoną liczbę wartości. Wobec takiej definicji teoria prawdopodobieństwa była zupełnie nieprzydatna w wielu zastosowaniach praktycznych.

W szczególności okazała się ona zupełnie bezradna wobec problemów rozkładów ciągłych, którymi za sprawą Maxwella, Boltzmana i Gibbsa zaczęła się zajmować fizyka w drugiej połowie XIX w. Przytoczmy tu za Norbertem Wienerem rozumowanie, które pokazuje jawnie na czym polega trudność:

„*Jeśli strzelam do tarczy pociskiem o wymiarach punktu geometrycznego, to prawdopodobieństwo, że trafię w pewien określony punkt tarczy będzie na ogół równe zeru, jakkolwiek nie jest wcale niemożliwe, że mi się to uda, co więcej, w każdym przypadku rzeczywiście trafiam w pewien punkt, choć prawdopodobieństwo tego zdarzenia jest równe zeru. Zdarzenie, którego prawdopodobieństwo jest równe jedności, polegające na tym, że trafię w jakiś punkt tarczy składa się zatem ze zdarzeń o prawdopodobieństwie równym zeru.*”

Oczywiście trudno winić za to twórcę klasycznej definicji Laplace'a. Rozwiązanie problemu, który opisał Wiener, wymaga zrozumienia idei zbiorów nieprzeliczalnych, które to pojęcie powstało, za sprawą Cantora, dopiero w końcu XIX w.

2. Boltzmann i narodziny fizyki statystycznej

Baza, na której powstała fizyka statystyczna, była molekularna teoria budowy materii. Przekonanie o takiej naturze materii pojawiło się w czasach starożytnych (już Demokryt twierdził, że materia składa się z *atomów*) jednak dopiero w XIX w. prace Clausiusa i Meyera przyczyniły się do przyjęcia tego poglądu na gruncie fizyki. W 1859 r. James Clerk Maxwell (1831–79) opierając się na modelu molekularnym opisał materię znajdującą się w spoczynku. Kontynuatorem tych idei był Ludwig Boltzmann (1844–1906), który usiłował wyprowadzić termodynamiczne własności gazów z ich mechanicznego modelu. W tym celu zakładał, że gaz złożony jest z wielkiej liczby atomów, które jak kule bilardowe poruszają się swobodnie w przestrzeni, zderzają i znowu swobodnie pędzą dalej. Jeśli więc wyobrazimy sobie gaz jako zbiór wielkiej liczby cząstek (atomów), powiedzmy N , których ruch jest rządzone prawami mechaniki (np. równaniami Newtona), to ruch ten opisuje układ $6N$ równań różniczkowych. Oczywiście taki model miał dla Boltzmana niewielką korzyść praktyczną (nawet przy obecnej technice komputerowej nie jesteśmy w stanie rozwiązać tego układu równań dla większości praktycznych problemów). Zamiast tego rozpatrywał on sześciowymiarową przestrzeń fazową położeń i pędów (nazwiemy tę przestrzeń przestrzenią fazową μ). Każda z N cząstek reprezentowana jest przez pewien punkt w tej przestrzeni. Boltzmann zakładał, że cząstki są nierozróżnialne i nie interesował się ich dokładnymi położeniami w przestrzeni μ , lecz gęstością rozkładu punktów, które reprezentują cząstki. Jeśli wprowadzimy funkcję opisującą ten rozkład, oznaczmy ją za Boltzmannem przez f , to cała dynamika układu N cząstek sprowadza się do opisu zmian w czasie funkcji $f = f(t, q, p)$, gdzie t jest czasem, q – wektorem położenia,

a p – wektorem pędu. Funkcja f zawiera całą informację o układzie i tak np.

$$\int_{\mathbb{R}^3} f(t, q, p) dp$$

jest (po pomnożeniu przez masę cząsteczkową) gęstością masy gazu.

Analogicznie

$$\int_{\mathbb{R}^3} p f(t, q, p) dp$$

jest pędem w punkcie q , a

$$\int_{\mathbb{R}^3} |p|^2 f(t, q, p) dp$$

energiami (inaczej temperaturą) w punkcie q .

Nim przejdziemy do dalszego opisu wyników Boltzmana zauważmy czym są jego konstrukcje wyrażone w języku współczesnej teorii prawdopodobieństwa. Po pierwsze, przejście od układu N cząstek do funkcji rozkładu f oznacza przejście z N do nieskończoności (!). Po drugie, funkcja f jest gęstością rozkładu prawdopodobieństwa. Oczywiście tak nazywa się te pojęcia obecnie, gdyż w czasach Boltzmana nawet nie istniał aparat matematyczny umożliwiający jego wprowadzenie (nie istniała teoria miary !!!).

Tymczasem Boltzmann nie poprzestał na wprowadzeniu funkcji f , ale posługując się heurystycznymi rozważaniami znalazł równanie opisujące ewolucję w czasie tej funkcji (równanie Boltzmana – 1872). Istotną rolę w rozważaniach Boltzmana odgrywała hipoteza o molekularnym chaosie, sformułowana w 1857 r. przez Clausiusa, a mówiąca o statystycznej niezależności stanów cząstek. Nadała ona opisowi Boltzmana charakter statystyczny. Nie będziemy tutaj zajmowali się samym równaniem Boltzmana, ale omówimy pewien fakt, który jest bezpośrednią konsekwencją tego równania – sławne *twierdzenie H* Boltzmana.

Twierdzenie H (Boltzmana). Zdefiniujmy funkcję

$$H(t) = \int_{\mathbb{R}^3 \times \mathbb{R}^3} f(t, q, p) \log f(t, q, p) dq dp.$$

Jeśli $f(t, q, p)$ jest rozwiązaniem równania Boltzmana, to

$$\frac{dH}{dt} \leq 0.$$

Sformuławszy to twierdzenie Boltzmann zauważył związek, jaki zachodzi między funkcją $H(t)$ a występującą w termodynamice entropią S

$$S = -kH,$$

gdzie k jest tzw. stałą Boltzmana, a entropia S jest liczona na jednostkę objętości.

W przekonaniu Boltzmana zbudowana teoria była wspaniałym sukcesem kinetyczno-molekularnej teorii budowy materii. Oto zbudowana w oparciu o klasyczną mechanikę teoria doprowadziła do uzyskania drugiej zasady termodynamiki mówiącej o nieodwracalności przemian jednych form energii w inne.

Jedno z możliwych sformułowań drugiej zasady termodynamiki mówi, że entropia układu rośnie w czasie.

W rzeczywistości to, co Boltzmann uważał za największy sukces swojej teorii, było najzacieklej atakowane przez jej przeciwników. W krótkim czasie powstały dwa paradoksy związane z *twierdzeniem H*. Pierwszym z nich był *paradoks odwracalności* sformułowany przez Kelvina, a następnie Loschmidta (1876). Opiera się on na fakcie, że równania klasycznej mechaniki są odwracalne w czasie. Jeśli dokona się transformacji $t \rightarrow -t$ to równania nie ulegną zmianie.

Tymczasem po tej transformacji funkcja H nie może maleć. Jasne więc, że *twierdzenie H* i równanie Boltzmana, z którego to twierdzenie wynika, nie może być konsekwencją czysto mechanicznego modelu. Drugi paradoks pochodzi

od Ernesta Zermelo (1871–1953) i nazywany jest *paradoksem powracalności* (1896). Paradoks ten opiera się na znanym twierdzeniu Poincarégo, które mówi, że każdy zachowawczy i zamknięty układ dynamiczny powraca, po dostatecznie długim czasie, do dowolnie małego otoczenia prawie każdego stanu początkowego. Zatem funkcja H , która początkowo maleje powinna następnie rosnać, aby przy powrocie w pobliże punktu, z którego zaczęła się ewolucja, przyjmować wartości bliskie wartościom początkowym. Ponieważ funkcja H cały czas maleje, pojawia się sprzeczność z wnioskiem z twierdzenia Poincarégo.

Pomiędzy Boltzmannem a jego oponentami toczyła się bardziej walka na słowa niż na argumenty naukowe. Boltzmann do końca życia był przekonany o słuszności swojej teorii nie umiejąc jednak w sposób zadowalający odeprzeć stawianych mu zarzutów. Wyjaśnienie skąd biorą się wspomniane paradoksy zostało podane dopiero wiele dziesiątków lat później. Nie uprzedzając dalszego wykładu powiedzmy już teraz, że jest ono związane z właściwym interpretowaniem pojęć probabilistycznych, które Boltzmann wprowadził w sposób intuicyjny. Ale nawet dzisiaj, ponad sto lat od czasu odkrycia Boltzmann, nie możemy powiedzieć, że wiemy jak w pełni poprawnie matematycznie wyprowadzić *twierdzenie H* z mechanicznego modelu gazu.

3. Gibbs i probabilistyczna interpretacja entropii

Mając świadomość trudności, które napotkał Boltzmann, Josiah Gibbs (1859–1903) dokonał innego podejścia do problemu wyjaśnienia podstaw termodynamiki. Przedstawimy ten opis używając dla ułatwienia języka współczesnej probabilistyki. Jest to pewne odstępstwo od wierności na rzecz zrozumiałości, jednakże oryginalne rozumowania Gibbsa, choć znacznie bardziej intuicyjne, w zasadzie przebiegały analogicznie.

Rozważmy układ N cząstek. Tym razem zbudujemy dla nich przestrzeń fazową Γ , $6N$ -wymiarową, której każdy punkt reprezentuje stan układu. W przypadku obserwacji makroskopowych nie mamy ani możliwości, ani potrzeby ustalania stanu układu w każdej chwili. Interesuje nas zaledwie kilka makroskopowych własności, np. znamy z pewną dokładnością energię całkowitą tego układu. Taki warunek spełnia nieskończenie wiele stanów układu. Dlatego wyobraźmy sobie nie pojedynczy układ, ale nieskończoną liczbę kopii tego układu, znajdujących się we wszystkich możliwych stanach spełniających dane warunki. Zbiór wszystkich takich kopii nazywa się *zespołem*. Odpowiada mu pewien rozkład prawdopodobieństwa w przestrzeni Γ . Zespół, dla którego rozkład prawdopodobieństwa na powierzchni stałej energii ma stałą gęstość, nazywa się *zespołem mikrokanonicznym*.

Korzystając z tego modelu przedstawmy boltzmanowską interpretację entropii. Niech $\omega(E)$ oznacza miarę obszaru zajmowanego przez zespół mikrokanoniczny odpowiadający energii E , wtedy

$$S = k \log \omega(E).$$

Gibbs porzucił ideę indywidualnego układu i rozważał układy rozłożone w całej przestrzeni Γ . Niech $D = D(t, q_1, \dots, q_N, p_1, \dots, p_N)$ oznacza gęstość tego rozkładu w chwili czasu t . W otoczeniu powierzchni stałej energii postąpić możemy następująco. Bierzymy cienką powłokę $[E, E + \Delta E]$ i dzielimy ją na skończoną liczbę komórek $\Delta_1, \Delta_2, \dots, \Delta_n$. W każdej komórce definiujemy średnią gęstość

$$P_i = P_i^* = \frac{1}{|\Delta_i|} \int_{\Delta_i} D(t, q_1, \dots, q_N, p_1, \dots, p_N) dq_1 \dots dq_N dp_1 \dots dp_N.$$

Za Gibbsem entropię zdefiniujemy następującym wzorem

$$\eta_t = \sum_i |\Delta_i| P_i^* \log P_i^* = \int P_i \log P_i dq_1 \dots dq_N dp_1 \dots dp_N.$$

Twierdzenie o entropii (Gibbsa).

$$\eta_t \leq \eta_0,$$

gdzie η_0 jest entropią odpowiadającą stanowi początkowemu.

Jest rzeczą oczywistą, że entropię Gibbsa można związać z entropią termodynamiczną zależnością

$$S(t) = -k\eta_t.$$

Ważną cechą twierdzenia Gibbsa jest to, że jego dowód nie wymaga odwoływania się do równania Boltzmannna. Zauważmy na koniec, że podejście Gibbsa, które też daje kinetyczno-molekularne podstawy termodynamiki odwołuje się do jeszcze bardziej zaawansowanych pojęć probabilistyki. W jego interpretacji ewolucja układu jest bowiem pewnym procesem stochastycznym.

4. Aksjomatyczna definicja prawdopodobieństwa

W początku XX w. było zupełnie jasne, że nie wszystko jest w porządku z logicznymi podstawami rachunku prawdopodobieństwa. W szczególności było oczywiste, że aparat używany przez fizykę statystyczną daleko wybiega poza to, co matematycy potrafili zdefiniować. Pierwszym, który na tę sytuację zwrócił wyraźnie uwagę, był Poincaré, sam aktywnie zaangażowany w dyskusję z Boltzmannem. W swoich pracach Poincaré skoncentrował się na logicznych niedostatkach klasycznej definicji prawdopodobieństwa podkreślając, jak to zauważyliśmy w pierwszej części tego wykładu, że definicja ta prowadzi do błędnego koła. Nic więc dziwnego, że w okresie tym rozpoczęto prace nad nowymi podstawami rachunku prawdopodobieństwa. Ponieważ był to jednocześnie okres wielkiego triumfu podejścia aksjomatycznego (aksjomatyzacja arytmetyki przez Peano, aksjomatyzacja geometrii przez Hilberta), stąd także dla rachunku prawdopodobieństwa starano się stworzyć podstawy aksjomatyczne.

Rzecz jasna, rozwój nowoczesnego rachunku prawdopodobieństwa nie byłby jednak możliwy bez postępu w innych działach matematyki. Przełomowe było tutaj dokonane przez Cantora odkrycie zbiorów nieprzeliczalnych. Właściwego aparatu dostarczył znacznie później, bo na przełomie XIX i XX w., Lebesgue tworząc teorię miary zbiorów. Dopiero uzbrojeni w tę teorię mogli kolejni matematycy stworzyć nowe podstawy rachunku prawdopodobieństwa opierając się na intuicjach analogicznych do tych, które leżały u podstaw definicji klasycznej.

Pierwsze próby aksjomatyzacji rachunku prawdopodobieństwa pojawiły się w Rosji. W Petersburgu istniała bardzo silna szkoła probabilistów, której przedstawicielami byli m.in. Czebyszew, Markow i Lapunow. Pewną słabością tej znakomitej szkoły była jej izolacja od postępów w fizyce. Niemniej jednak to właśnie tam Siergiej N. Bernstein (1880–1968) stworzył pierwszą aksjomatykę rachunku prawdopodobieństwa. Niezależnie od Bernsteina niemiecki uczyony Richard von Mises (1883–1953) zaproponował własne aksjomaty. Niewątpliwą zasługą obu tych aksjomatyzacji było wprowadzenie do teorii tego, co dzisiaj nazywamy przestrzenią zdarzeń i uświadomienie, że prawdopodobieństwo jest jedynie pewną funkcją na tej przestrzeni. Należy także podkreślić, że o ile Bernstein tworzył swoje aksjomaty nie odwołując się do doświadczeń nauk przyrodniczych, to von Mises budując swoją teorię kwestionował wręcz jej samodzielne matematyczne istnienie. Dla von Misesa prawdopodobieństwo jest tam, gdzie jest wielka liczba niezależnych eksperymentów, a teoria jest potrzebna do tego, aby poprawnie oliczać wartości oczekiwane tych eksperymentów.

Prawdziwy przełom w podstawach teorii prawdopodobieństwa nastąpił jednak dopiero wtedy, gdy uświadomiono sobie związki tej teorii z teorią funkcji rzeczywistych. Stało się to za sprawą prac Aleksandra Chinczyna (1894–1959) i Andrieja Kołmogorowa (1903–87) na temat prawa wielkich liczb. Odkryto wtedy analogię między prawem wielkich liczb a zbieżnością funkcji według miary oraz mocnym prawem wielkich liczb a zbieżnością funkcji prawie wszędzie. Po tych odkryciach Kołmogorow przystąpił do pracy nad logicznymi podstawami

U Misesa tę przestrzeń nazywa się *przestrzenią próbek*, który to termin występuje jeszcze obecnie w wielu wydawanych w Polsce książkach o rachunku prawdopodobieństwa.

To bazujące głównie na fizyce podejście von Misesa sprawiło, że jego koncepcja prawdopodobieństwa była używana przez fizyków jeszcze wiele lat po tym, gdy powstała nowoczesna aksjomatyka.

rachunku prawdopodobieństwa. Znał on wtedy prace obu swoich prekursorów Bernsteina i von Misesa, miał świadomość analogii między miarą zbioru a prawdopodobieństwem zdarzenia, całą a wartością oczekiwaną, ortogonalnością funkcji a niezależnością zdarzeń losowych. Jego praca przyniosła aksjomatyzację rachunku prawdopodobieństwa, która w ciągu następnych lat stała się powszechnie akceptowana i stanowi do dziś podstawę tej teorii matematycznej. Przytoczmy na zakończenie tę definicję.

Definicja przestrzeni probabilistycznej (Kolmogorow - 1933). Niech dana będzie przestrzeń zdarzeń Ω . Każde zdarzenie jest pewnym podzbiorem Ω .

1. Zbiór wszystkich zdarzeń \mathbf{B} tworzy algebra zbiorów, tzn. jeśli A i B należą do \mathbf{B} , to należą tam także zbiory $A \cup B$, $A \cap B$, A' i B' (A' oznacza uzupełnienie zbioru A).
2. Ω i zbiór pusty \emptyset należą do \mathbf{B} .
3. Każdemu elementowi A zbioru \mathbf{B} przypisuje się pewną liczbę nieujemną $P(A)$ nazywaną prawdopodobieństwem zdarzenia A .
4. $P(\Omega) = 1$.
5. Jeśli $A \cap B = \emptyset$, to

$$P(A \cup B) = P(A) + P(B).$$

6. Jeśli mamy ciąg zstępujący zbiorów

$$A_1 \supset A_2 \supset \dots \supset A_n \supset \dots$$

mających puste przecięcie, to

$$\lim_{n \rightarrow \infty} P(A_n) = 0.$$

Na zakończenie, jako anegdotę powiedzmy, że rachunek prawdopodobieństwa odwdziczył się fizyce za wkład jaki wniosła w jego rozwój. Dla fizyków końca XIX i początków XX w. konieczność używania metod statystycznych wynikała jedynie z braku dostatecznej wiedzy o układzie (nie znamy położenia i pędów wszystkich cząstek) oraz niedostatków mocy obliczeniowej (rozwiązanie $6N$ równań). Podobnie myślało również wielu matematyków (np. Poincaré). Tymczasem narodziny mechaniki kwantowej zasadniczo zmieniły ten obraz. Teraz już nie było położenia i pędów cząstek. Istniało jedynie prawdopodobieństwo znalezienia cząstki we wskazanym obszarze przestrzeni fazowej μ . W efekcie więc same cząstki zostały zastąpione przez odpowiednie gęstości prawdopodobieństwa, a równania fizyki stały się równaniami opisującymi te gęstości. Że nie był to łatwy przełom dla fizyki, może świadczyć fakt długiej walki z probabilistyczną interpretacją teorii kwantów. Einstein właściwie do śmierci nie pogodził się w pełni z tą interpretacją. Dowodzić tego może jego słynne zdanie „*Nie wierzę, żeby Bóg grał w kości.*”

Dodatek. Wyjaśnienie paradoksów fizyki statystycznej

Omawiając zastrzeżenia jakie podnoszono wobec teorii Boltzmanna obiecałmy współczesne wyjaśnienie występujących paradoksów. Należy jednak pamiętać, że w całej ogólności teoria Boltzmanna nie została dopracowana matematycznie nawet obecnie. Stąd też chcąc wyjaśnić paradoksy nie będziemy rozpatrywali gazu rzeczywistego, ale pewien uproszczony model, który z jednej strony ma wszystkie niezbędne cechy modelu kompletnego, z drugiej zaś jest dostatecznie prosty, aby wszystkie rozumowania można było doprowadzić do końca. Model ten nosi w literaturze nazwę *kołowego modelu Kaca*. Rozważmy na okręgu układ n równomiernie rozłożonych punktów. Wśród tych punktów m tworzy wyróżniony zbiór S . W każdym z n punktów umieszczona jest kula o jednym z dwóch kolorów: biała lub czarna. Wszystkie kule poruszają się wykonując równocześnie przeskok do sąsiedniego punktu w kierunku przeciwnym do ruchu wskazówek zegara, przy czym kule zmieniają swój kolor według następującej reguły: jeśli kula startuje z punktu należącego do zbioru S , to zmienia swój kolor na przeciwny, jeśli startuje z punktu nie należącego do S , to nie zmienia koloru. Zadanie nasze polega na tym, aby znając początkowy rozkład kolorów, określić ten rozkład po pewnej (dostatecznie dużej) liczbie ruchów.

Opis tego modelu, jak również wszystkie przytoczone rachunki, zostały zaczerpnięte z książki Marka Kaca *Kilka zagadnień stochastycznych fizyki i matematyki*.

Wprowadźmy następujące oznaczenia: niech $N_c(t)$ i $N_b(t)$ będą odpowiednio liczbami czarnych i białych kul w chwili t . Ponumerujmy punkty na okręgu od 1 do n i niech p oznacza jedną z liczb zawartych między 1 a n . Następnie niech

$$\varepsilon_p = \begin{cases} -1, & \text{jeśli } p \text{ należy do } S, \\ +1, & \text{jeśli } p \text{ nie należy do } S, \end{cases}$$

$$\eta_p(t) = \begin{cases} +1, & \text{jeśli w punkcie } p \text{ w chwili } t \text{ jest czarna kula,} \\ -1, & \text{jeśli w punkcie } p \text{ w chwili } t \text{ jest biała kula.} \end{cases}$$

Ewolucję w czasie naszego układu można teraz zapisać równaniem

$$(1) \quad \eta_p(t+1) = \varepsilon_{p-1} \eta_{p-1}(t),$$

ponieważ kula znajdująca się w punkcie p w chwili $t+1$ przysłała tam z punktu $p-1$ zmieniając kolor, lub nie, zależnie od tego, czy punkt $p-1$ był w zbiorze S , czy też nie. Korzystając z wzoru (1) otrzymujemy

$$\eta_p(t) = \eta_{p-t}(0) \varepsilon_{p-1} \varepsilon_{p-2} \dots \varepsilon_{p-t}.$$

Jeśli teraz założymy, że w chwili $t=0$ wszystkie kule są czarne, czyli $\eta_{p-t}(0) = 1$ dla każdego p i t , to

$$\sum_p \eta_p(t) = \sum_p \varepsilon_{p-1} \varepsilon_{p-2} \dots \varepsilon_{p-t}.$$

Wartość sumy po lewej stronie tej równości to nadwyżka kul czarnych nad kulami białymi

$$N_c(t) - N_b(t) = \sum_p \varepsilon_{p-1} \varepsilon_{p-2} \dots \varepsilon_{p-t}.$$

Równanie (1) opisuje dynamikę danego modelu, a zatem może być uznane za analog równań klasycznej mechaniki. Podobnie jak klasyczna mechanika, rozważany model jest w pełni odwracalny i ma własność powracalności. Jak łatwo zauważyć, jeżeli zmienimy kierunek ruchu po okręgu (co odpowiada transformacji $t \rightarrow -t$), to wrócimy do punktu wyjścia. Zatem faktycznie model jest odwracalny. Powracalność jest jeszcze łatwiejsza do zaobserwowania, ponieważ model jest okresowy o okresie $2n$. Rzeczywiście, po n krokach każda kulka powróci do swojego początkowego położenia obiegwszy okrąg. W tym czasie przejdzie ona przez m punktów zbioru S . Po przejściu przez m punktów kulka dowolnego koloru będzie miała kolor $(-1)^m \eta_p$. Stąd po dwukrotnym obejściu okręgu każda kulka wróci do swojego początkowego koloru.

Model jest, oczywiście, w pełni deterministyczny. Naśladować idee Boltzmanna, można do modelu wprowadzić „losowość”. W tym celu rozważmy wszystkie możliwe zbiory S o tej własności, że mają one dokładnie m elementów (mamy $\binom{n}{m}$ takich zbiorów) i założmy jednakowe prawdopodobieństwo wybrania każdego z tych zbiorów. Możemy teraz obliczyć wartość średnią wyrażenia $\frac{1}{n}(N_c(t) - N_b(t))$ po wszystkich możliwych zbiorach S . Nie będziemy przytaczali rachunków, które prowadzą do wyliczenia tej średniej. Wynik jest następujący

$$\left\langle \frac{N_c(t) - N_b(t)}{n} \right\rangle \approx (1 - 2\sigma)^t,$$

przy założeniu, że $\sigma = \frac{m}{n}$ jest ustalone i mniejsze od $\frac{1}{2}$ oraz $n, m \rightarrow \infty$. Widzimy, że z czasem liczby białych i czarnych kulek będą się zbliżały wykładniczo, niezależnie od tego jaki był stan początkowy. Dla danego modelu stwierdzenie to jest analogiem *twierdzenia H Boltzmanna*.

Widać więc, że model ten doskonale ilustruje matematyczne źródło nieodwracalności, które zdecydowanie nie ma „mechanicznego” charakteru, lecz powstaje na skutek uśrednienia po zbiorach S oraz przejścia do „granicznego układu nieskończonego”. Jest to pełna analogia z modelem Boltzmanna, w którym „losowość” została wprowadzona przez założenie molekularnego chaosu. Idąc dalej możemy obliczyć wariancję zmiennej $\frac{1}{n}(N_c(t) - N_b(t))$ i przekonać się, że jest ona rzędu $1/\sqrt{n}$. Można zatem powiedzieć, że wyrównywanie się liczby kul białych i czarnych będzie obserwowane, ale tylko dla czasów małych w porównaniu z czasem powrotu dla tego zagadnienia (tzn. dla $t \ll 2n$).

Na koniec oddajmy głos samemu twórcy modelu Markowi Kacowi: „Zastanówmy się bliżej nad tym, co to wszystko znaczy. Przypuśćmy, że chcemy wykreślić przebieg funkcji $\frac{1}{2}(N_c(t) - N_b(t))$ w zależności od t . Ile otrzymamy krzywych? Otrzymamy $\binom{n}{m}$ krzywych, albowiem dla każdego wyboru zbioru S otrzymuje się inną krzywą. Każda z tych krzywych związana jest z jednym zbiorem i jest periodyczna o okresie $2n$. Ustalmy sobie teraz t i popatrzmy na te krzywe tylko w odniesieniu do wybranego punktu. Przypuśćmy, że n jest bardzo duże w porównaniu z t , np. niech n będzie rzędu 10^{23} , a t — rzędu 10^6 . Otóż w chwili t każda z tych krzywych ma pewną wartość; wszystkie te wartości koncentrują się wokół $(1 - 2\sigma)^t$. Gdybyśmy nakreślili krzywą wykładniczą $(1 - 2\sigma)^t$, wówczas zauważylibyśmy, że w każdym ustalonym punkcie t olbrzymia większość rozważanych krzywych skupia się wokół tej krzywej wykładniczej.”