

Termodynamika i Fizyka Statystyczna R

Przewodnik dla ostrożnych

Wersja: 15 maja 2026

1 Wstęp: Z rozpaczy ku fizyce statystycznej

„Ludwig Boltzmann, who spent much of his life studying statistical mechanics, died in 1906, by his own hand. Paul Ehrenfest, carrying on the work, died similarly in 1933. Now it is our turn to study statistical mechanics. Perhaps it will be wise to approach the subject cautiously.”

— David Goodstein, *States of Matter*

Powyższy cytat, choć podszyty czarnym humorem, doskonale oddaje ciężar gatunkowy dziedziny, którą będziemy się zajmować. Dlaczego mechanika statystyczna doprowadzała wybitne umysły do takiej frustracji? Aby to zrozumieć, musimy spojrzeć na ograniczenia tradycyjnego opisu fizycznego.

Mechanika klasyczna i kwantowa odnoszą ogromne, wręcz spektakularne sukcesy w odkrywaniu i analizie praw rządzących ruchem pojedynczych ciał czy cząstek. Potrafimy z niezwykłą precyzją wyliczyć trajektorię rzuconego kamienia, ruch planet, czy stany energetyczne atomu wodoru. Jednak wraz ze wzrostem liczby cząstek w rozpatrywanym układzie, problem drastycznie się komplikuje, i to na tyle, że klasyczne metody analityczne po prostu zawodzą.

Warto przypomnieć, że już słynny problem trzech ciał oddziałujących grawitacyjnie nie posiada ogólnego rozwiązania analitycznego w postaci zwartej (tzw. *closed-form solution*). Owszem, w 1912 roku fiński matematyk Karl F. Sundman zdołał udowodnić istnienie ścisłego rozwiązania w postaci zbieżnego szeregu nieskończonego. Z praktycznego i fizycznego punktu widzenia jest ono jednak całkowicie bezużyteczne – zbiega ono tak wolno, że aby uzyskać sensowną precyzję dla krótkiego odcinka czasu, należałoby zsumować miliony wyrazów tego szeregu. Co więcej, pod koniec XIX wieku Henri Poincaré udowodnił, że układ trzech ciał jest fundamentalnie chaotyczny. Oznacza to ekstremalną wrażliwość na najdrobniejsze zmiany warunków początkowych. Trajektorie ciał splątują się ze sobą w skomplikowany, niepowtarzalny i niemożliwy do długoterminowego przewidzenia sposób.

A mówimy tu przecież zaledwie o trzech oddziałujących obiektach! A co, jeśli tych ciał jest nie trzy, lecz tryliony? Czy pozostaje nam tylko załamać ręce i pogрузić się w rozpacz?

1.1 Bariera obliczeniowa, czyli dlaczego komputery nas nie ocalą

Pozostają nam rzecz jasna jeszcze metody numeryczne. Naiwne podejście podpowiada: weźmy po prostu równania Newtona, wsadźmy je do potężnego komputera i pozwólmy maszynie rozwiązać je krok po kroku. Niestety, nawet one nie pozwalają na opis układów o rozmiarach makroskopowych.

Warto w tym miejscu spojrzeć prawdzie w oczy: w tej walce najpotężniejsze komputery świata są równie bezradne co kartka i ołówek. Dziś dysponujemy superkomputerami klasy eksa-skalowej (wykonującymi ponad 10^{18} operacji zmiennoprzecinkowych na sekundę, czyli eksaflopy). Obecnymi rekordzistami są giganty takie jak **Frontier** (znajdujący się w Oak Ridge National Laboratory w USA) czy **Aurora** (Argonne National Laboratory). Ich kolosalna moc obliczeniowa

nie opiera się na klasycznych procesorach, lecz na zaprzęgnięciu do pracy dziesiątek tysięcy potężnych akceleratorów graficznych (GPU) liczących wszystko równolegle.

Co to oznacza w praktyce dla fizyka? Największe z możliwych obecnie symulacji dynamiki molekularnej na takich maszynach potrafią śledzić ruch rzędu biliona (10^{12}) atomów. Brzmi to dumnie, ale czy rozwiązuje nasz problem? Absolutnie nie. Bilion atomów to wciąż zaledwie mikroskopijny pyłek, znikomy ułamek mikromola materii. Co gorsza, średni czas między oddziaływaniem cząstek, np. w gazie, jest rzędu nanosekund, a żeby symulacja numeryczna była w ogóle stabilna, krok czasowy musi być ułamkiem femtosekundy (10^{-15} s). To oznacza, że po tygodniach ciągłego działania największych klastrów obliczeniowych świata na pełnych obrotach, potrafimy prześledzić ewolucję takiego układu zaledwie przez ułamki mikrosekundy. Wynik jest przytłaczająco mizerny – ewolucja zaledwie *jednego mola* gazu (10^{24} cząstek) przez *jedną sekundę* wymagałaby mocy obliczeniowej o jakieś 25 rzędów wielkości większej niż ta, którą dysponuje dzisiaj cała ludzkość.

1.2 Zbój Gębon i demon Laplace’a

Jeżeli jednak przez moment wyobrazilibyśmy sobie, że taki nadnaturalny komputer istnieje, potrafi wyliczyć trajektorie cząstki po cząstce w szklance wody i wyrzuci z siebie w mgnieniu oka biliardy liczb opisujących konfiguracje układu w każdej kolejnej chwili czasu... moglibyśmy ze zdumieniem stwierdzić, że te potoki liczb **nikogo tak naprawdę nie interesują**.

Nikt nie zadałby pytania: „gdzie dokładnie znajduje się cząsteczka nr 156 893 211 w czwartej mikrosekundzie?”. Pytania, które byśmy zadawali, byłyby zupełnie innej natury: jaka jest gęstość cząstek tu czy tam? Gdzie średnio wędrują cząstki? Jakie wywierają ciśnienie? Czyli nie tyle interesowałoby nas zachowanie pojedynczych indywiduów, ile pewne własności kolektywne, uśrednione, makroskopowe.

Dramat zbója Gębona

Sytuacja ta do złudzenia przypomina los zbójcy Gębona z opowiadania Stanisława Lema „*Jak Trurl i Klapaucjusz demona drugiego rodzaju stworzyli, aby zbójcę Gębona pokonać*”. Gębon, łaknący potęgi płynącej z Wszelkiej Informacji, został ukarany przez genialnych konstruktorów maszyną (Demonem Drugiego Rodzaju), która z chaotycznego ruchu cieplnego atomów powietrza odczytywała absolutnie wszystkie prawdziwe fakty. Zbójca momentalnie został pogrzebany pod górami papierowych taśm zawierających nieskończoną ilość bezużytecznych dla niego danych:

„(...) wydawało się zbójowi, że zaraz już dowie się rzeczy niesłychanych, takich, które mu oczy na Istotę Bytu otworzą, więc wczytywał się we wszystko, co leciało spod brylancika (...) a tymczasem czytał o tym, jak się wiję arlebardzkie wiją i że córka króla Petrycego z Labaudii zwała się Garbunda, i co jadł na drugie śniadanie Fryderyk II, król bładawców, nim wojnę wypowiedział Gwendolinom, i ile powłok elektronowych liczyłby sobie atom termionolium, gdyby taki pierwiastek był możliwy, i jakie są wymiary dziurki tylnej małego ptaszka, zwanego kurkucielem, którego na swych rozamforach malują Marłajowie Wabędzcy, jak również o trzech smakach poliwonnych szlamu oceanicznego na Wodocji Przyzroczonej, i o kwiatku Łubuduku, który myśliwych staromalfandzkich wali siarczyście na odlew, świtem wzruszony, i jak wyprowadzić wzór na dostawę kąta podstawy wieloboku, ikoseadrem zwanego, i kto był jubilerem Fafucjusza, rzeźnika mańkuta Buwantów, i ile pism filatelistycznych będzie wychodziło w roku siedemdziesięcioletnim na Morkonaucji, i gdzie znajduje się trupek Cybrycji Kraśnopiętej, którą gwoździem przebił po pijanemu niejaki Malkonder, i czym się różni Maciąg od Naciągu, a także kto ma najmniejszą w Kosmosie pielownicę wzdłużną i dlaczego pchły smoczkotyłkie mchu jeść nie chcą, i na czym polega gra zwana Balansyer

Zadni Ściągany, i ile było ziarenek lwichwostu w tej kupce, co ją Abrukwian Po-listny nogą trącił, kiedy się pośliznął na ósmym kilometrze szosy albacjerskiej w Dolinie Wzduchów Szedziwych.”

Dokładnie to samo spotkałoby nas, gdybyśmy dysponowali komputerem śledzącym ruch i parametry każdego pojedynczego atomu. Utonęlibyśmy w oceanie absolutnie prawdziwych, lecz fizycznie bezużytecznych danych.

Innym krytycznym elementem w tych rozważaniach jest chaos deterministyczny. Najpierw musielibyśmy do naszego wysnionego komputera załadować gigantyczny plik z warunkami początkowymi wszystkich atomów. Musiałyby to być warunki znane z absolutnie nieskończoną precyzją. Rzeczywistość jest bowiem bezlitosna: w układach wielociałowych, już po czasach rzędu kilku czasów między zderzeniami, dwie trajektorie różniące się w warunkach początkowych choćby o niewyobrażalnie mały ułamek, pójdą dwiema zupełnie innymi drogami (tzw. efekt motyla w skali mikro).

Wszystko to prowadzi nas do nieuniknionego wniosku: **nie można po prostu „wyliczyć świata”**. Pełna informacja o mikro stanie układu w ogólności nie jest dla nas dostępna – to jest to, o czym mówiliśmy na samym początku wykładu. Taka absolutna wiedza deterministyczna jest dostępna wyłącznie dla hipotetycznych bytów, takich jak słynny demon pojawiający się w Esejach Laplace’a z roku 1814: „Wyobraźmy sobie istotę, która w danej chwili czasu znalazłaby wszystkie siły działające w świecie i położenia wszystkich jego składników. Jeśli istota ta byłaby wystarczająco potężna, by poddać te dane analizie, mogłaby zawrzeć w jednym równaniu ruch największych ciał Wszechświata i najmniejszego atomu, dla takiej istoty nic nie byłoby niepewne a przyszłość tak jak i przeszłość rozpościerałaby się przed jej oczami.” Daleko nam jednak do demona Laplace’a. Mamy drastycznie ograniczone możliwości pomiarowe oraz obliczeniowe i musimy sobie jakoś, skromnie, po ludzku z tym radzić.

1.3 Paradygmat statystyczny

Wszystko to sprawia, że musimy definitywnie zrezygnować z podejścia indywidualistycznego na rzecz opisu uśrednionego, statystycznego, wyrażonego w języku rachunku prawdopodobieństwa. Przyznajemy się do fundamentalnej niewiedzy: nie będziemy wiedzieli, dokąd lecą poszczególne atomy. Zadowolimy się wiedzą o tym, ile ich *średnio* poleciało tu i tam, oraz jakie są prawdopodobieństwa wystąpienia różnorodnych efektów.

Paradoksalnie, to zrzeczenie się precyzji na poziomie mikro pozwala nam na precyzyjny opis na poziomie makro. W ten sposób będziemy w stanie zrozumieć i przewidzieć najróżniejsze zjawiska – od tak prozaicznych (że kiedy ściskamy balon z gazem, to on się ogrzewa, albo że jeżeli ogrzewamy jakiś materiał, to on się rozszerza), aż po zjawiska kwantowe i egzotyczne, jak kondensacja Bosego-Einsteina (BEC).

Co ciekawe, takie statystyczne podejście zrodzone z fizyki może być (i z wielkim powodzeniem jest!) stosowane nie tylko do opisu martwej materii w laboratoriach. Odnajdujemy je wszędzie tam, gdzie duża liczba oddziałujących ze sobą jednostek tworzy zjawiska kolektywne.

- **W psychologii społecznej i socjologii** funkcjonuje zasada: „Nie wiem, jak zachowa się konkretna jednostka – ale znam psychologię tłumu”. Reakcja jednego człowieka na bodziec może być nieprzewidywalna, ale zachowanie stutysięcznego tłumu na stadionie poddaje się już dość rygorystycznym prawom statystyki.
- **W fizyce ruchu drogowego** mamy do czynienia z podobnym ewenementem. Chociaż nie potrafimy przewidzieć zachowania i ruchu poszczególnych kierowców na autostradzie (bo zawsze trafią się narwańcy, piraci drogowi, niedzielni „wolniacy” i osoby zamyślane), to jednak za pomocą zaledwie kilku założeń o ich *średnim zachowaniu* jesteśmy w stanie niezwykle skutecznie symulować ruch drogowy całych metropolii. Potrafimy przewidywać

powstawanie korków, czy też modelować powstawanie fal na przemian zagęszczonego i rozrzedzonego ruchu drogowego (tzw. *stop-and-go waves*). Zresztą samo powstawanie i propagacja korków to bardzo ciekawy przykład zjawiska emergentnego (kolektywnego), który pojawia się dopiero wtedy, kiedy mamy do czynienia z wieloma pojazdami – odpowiednikami „atomów” na drodze – i który jest absolutnie niemożliwy do przewidzenia na podstawie znajomości mechaniki i ruchu pojedynczych samochodów. No bo jak, na litość boską, wpasć na to analizując ruch samotnego samochodu na pustej drodze, że w gęstym ruchu ulicznym tzw. korki fantomowe (korki powstające bez wyraźnej przyczyny, np. wypadku) **propagują się zawsze do tyłu z prędkością około 15 km/h**? Co ciekawe, prędkość propagacji tej fali zagęszczenia w tył jest jak najwyraźniej stała i uniwersalna dla prawie wszystkich szos na świecie, niezależnie od tego, czy to amerykańska autostrada, czy niemiecki Autobahn!

- Zupełnie **podobnie zachowują się piesi** w tłumie. Zjawiska ucieczki, problematyka paniki, zatorów w wąskich przejściach ewakuacyjnych – to wszystko fizyka statystyczna w najczystszej postaci. Piesi traktowani są jako samobieżne cząstki oddziałujące na siebie siłami odpychania społecznego, co pozwala inżynierom i architektom projektować bezpieczniejsze stadiony i stacje metra.

To właśnie siła statystyki. Nie znając szczegółów, poznajemy całościowy obraz. W kolejnej części przejdziemy do formalnego aparatu matematycznego, który nam to umożliwi – do rachunku prawdopodobieństwa.

Część II

Rachunek prawdopodobieństwa

*„He deals the cards to find the answer
The sacred geometry of chance
The hidden law of a probable outcome
The numbers lead a dance.”*

— Sting, *Shape of My Heart*

1 Pojęcia podstawowe

Większość z nas uważa, że intuicyjnie rozumie, czym jest prawdopodobieństwo. Przecież to proste – jaka jest szansa na to, że w rzucie monetą wypadnie reszka? Odpowiadamy bez wahania: $1/2$. Ale co właściwie przez to rozumiemy?

Mamy tu do dyspozycji dwie główne możliwości interpretacyjne:

1. Chodzi o wynik pojedynczego rzutu (Podejście subiektywne / informacyjne)

Mówi się wtedy, że prawdopodobieństwo P opisuje po prostu **stan naszej wiedzy** na temat wyniku jeszcze niewykonanego doświadczenia. Takie ujęcie natychmiast naraża się na mocny zarzut o subiektywizm („każdy obserwator wie coś innego”). Odpiera się go jednak tłumacząc, że wcale nie chodzi o subiektywne widzimisię badacza, lecz o to, co obiektywnie *można wiedzieć*, przeprowadzając dane doświadczenie dokładnie według zadanego przepisu.

2. Chodzi o częstość wystąpienia orła w długiej serii rzutów (Podejście częstościowe)

Jeśli nie podoba nam się abstrakcyjny opis „stanu wiedzy”, możemy oprzeć się na empirii. Rzucaliśmy monetą wystarczająco dużo razy i tak po prostu wyszło. To podejście prowadzi nas do **definicji częstościowej**, która mówi, że prawdopodobieństwo to granica stosunku liczby zdarzeń sprzyjających (N_r) do całkowitej liczby prób (N):

$$p_r = \lim_{N \rightarrow \infty} \frac{N_r}{N}$$

Niestety, i ta ścieżka kryje dwa poważne zarzuty:

- Każda fizyczna seria prób w naturze jest w rzeczywistości **skończona**. Nigdy nie osiągniemy granicy w nieskończoności.
- Skąd w ogóle wiemy, że zdarzenia w tej serii są ze sobą faktycznie niezależne?

Co gorsza, gdy zapytamy, co mówi teoria na temat częstości występowania danego zdarzenia w długiej, ale wciąż skończonej serii doświadczeń, okaże się, że teoria ta opisuje to... znowu używając języka prawdopodobieństwa. Wpadamy tu w zgrabne, błędne koło.

Co ciekawe, w historii nauki nie brakowało śmiałków (lub ludzi skrajnie znudzonych), którzy postanowili sprawdzić podejście częstościowe empirycznie:

- **John Kerrich** – angielski matematyk, który na początku II wojny światowej wykładał w Kopenhadze. Aresztowany przez Niemców, spędził wojnę w obozie jenieckim na Jutlandii. Z nudów przeprowadzał eksperymenty probabilistyczne. Rzucił monetą 10 tysięcy razy i otrzymał 5067 orłów.
- **Karl Pearson** – słynny statystyk, miał jeszcze więcej czasu i samozaparcia. Wykonał 24 000 rzutów, z czego 12 012 zakończyło się wyrzuceniem orła.

O monetach warto wiedzieć jeszcze kilka rzeczy. O ile rzuconej monety nie da się łatwo oszukać (w przeciwieństwie do kości do gry, które można obciążyć), o tyle sytuacja zmienia się dramatycznie, jeśli monetą **zakręcimy jak bączkiem**. Polscy matematycy z Siedlec, Tomasz Gliszczyński i Włodzimierz Zawadowski (o których w 2002 roku pisał nawet brytyjski *The*

Times), kazali swoim studentom zakręcić 250 razy nową belgijską monetą o nominale 1 Euro. Wynik? Król Belgii (awers) pokazał się 140 razy, a reszka tylko 110. Strona z twarzą króla była po prostu cięższa.

Niezmiernie ważna jest też sama technika rzutu. Persi Diaconis – niegdyś profesjonalny magik, a dziś wybitny profesor matematyki na Uniwersytecie Stanforda – udowodnił, że potrafi rzucić monetą w taki sposób, aby za każdym razem wylądowała na żądanej stronie. Udowodnił, że rzut monetą to nie magia ani czysta losowość, lecz zagadnienie z zakresu mechaniki klasycznej.

Budowanie intuicji: Eksperyment Alicji i Boba

Aby uniknąć pułapek intuicji, w mechanice statystycznej musimy zbudować precyzyjny, matematyczny model zjawisk losowych. Zanim jednak podamy aksjomaty, prześledźmy proces ich powstawania na bogatszym przykładzie.

Wyobraźmy sobie następującą sytuację: Alicja dzwoni do Boba i podaje mu wyniki swoich rzutów dwiema kostkami do gry. Ściślej mówiąc, podaje mu tylko *sumy* wyrzuconych oczek:

$$4, 7, 5, 2, 6, 11, 7, 9, 9, 6 \dots$$

Na razie jest to mało ciekawy ciąg liczb. Ale po dłuższym powtarzaniu eksperymentu Bob zauważa pewną prawidłowość – dostajemy ciąg, w którym najczęściej wychodzi 6, 7 i 8, a najrzadziej 2 i 12. Tym z Was, którzy grali w Catana, ten fakt jest na pewno dobrze znany. Jak to opisać i sformalizować? Najpierw doprecyzujemy, co Boba w ogóle interesuje. Niech $\mathcal{A}$ oznacza **zbiór interesujących nas zdarzeń**. To mogą być proste zdarzenia, np.:

- S_2 (suma wynosi 2),
- S_6 (suma wynosi 6),
- S_9 (suma wynosi 9).

Ale mogą to być również zdarzenia złożone, traktowane jako operacje na zbiorach, np. suma logiczna $A = S_2 \cup S_6 \cup S_9$, albo zdarzenie $B =$ „cokolwiek, byle nie 2 i 12”. Możemy mieć też w tym zbiorze istotne przypadki skrajne:

- **zdarzenie niemożliwe** (np. suma oczek wynosi 13). Z definicji chcemy, by $P(\text{niemożliwego}) = 0$.
- **zdarzenie pewne** (np. suma oczek to liczba całkowita). Z definicji $P(\text{pewnego}) = 1$.

Teraz chcielibyśmy każdemu zdarzeniu ze zbioru $\mathcal{A}$ przypisać prawdopodobieństwo. Będzie to funkcja przyporządkowująca zdarzeniom liczby rzeczywiste:

$$P : \mathcal{A} \ni S \rightarrow P(S) \in \mathbb{R}$$

Rozsądnie jest też przyjąć addytywność prawdopodobieństw: alternatywa zdarzeń to w języku zbiorów ich suma ($A \cup B$), więc chcemy, aby prawdopodobieństwo alternatywy zdarzeń wykluczających się było równe sumie ich prawdopodobieństw.

Zauważmy fundamentalną rzecz: każde zdarzenie (np. S_5) analizowane przez Boba jest tak naprawdę uwarunkowane niewidocznymi dla niego, mikroskopowymi stanami kostek. Wprowadźmy zbiór **wszystkich możliwych par liczb** określających pojedyncze wyniki na obu kostkach. Niech Ω oznacza zbiór tych par:

$$\Omega = \{(i, j) \mid i, j \in \{1, 2, \dots, 6\}\}$$

Elementy tego bazowego zbioru nazywamy **zdarzeniami elementarnymi**.

Każde makroskopowe zdarzenie S to po prostu podzbiór zbioru Ω . Zauważmy jednak, że nie wszystkie zdarzenia elementarne są bezpośrednio identyfikowalne dla Boba! Np. konkretny rzut (2, 3) należy do zdarzenia makroskopowego S_5 , ponieważ:

$$S_5 = \{(1, 4), (2, 3), (3, 2), (4, 1)\}$$

Gdybyśmy założyli (co dla idealnych kostek jest rozsądne), że wszystkie z 36 możliwych zdarzeń elementarnych (i, j) są równoprawdopodobne, to z zasady addytywności dostalibyśmy, że prawdopodobieństwo każdego z nich to $1/36$. Stąd łatwo wyznaczyć prawdopodobieństwa interesujących Boba sum:

$$P(S_5) = \frac{4}{36} \quad \text{bo sprzyjają mu pary: } (1, 4), (2, 3), (3, 2), (4, 1)$$

$$P(S_7) = \frac{6}{36} \quad \text{bo sprzyjają mu pary: } (1, 6), (6, 1), (2, 5), (5, 2), (4, 3), (3, 4)$$

1.1 Formalizm matematyczny i aksjomatyka Kołmogorowa

Ogólny schemat tego, co intuicyjnie zrobiliśmy przed chwilą, stanowi fundament formalnego rachunku prawdopodobieństwa. Wprowadzamy w nim trzy filary, tworzące tzw. przestrzeń probabilistyczną $(\Omega, \mathcal{A}, P)$:

1. Przestrzeń próbek (Zbiór zdarzeń elementarnych) Ω

W każdym modelu występuje zbiór Ω , zawierający wszystkie możliwe wyniki danego eksperymentu losowego na poziomie najbardziej podstawowym. Dla dwóch rzutów monetą byłoby to $\Omega = \{(O, O), (O, R), (R, R), (R, O)\}$.

2. Zbiór zdarzeń losowych (σ -algebra) $\mathcal{A}$

Zdarzenia losowe (to, co nas interesuje) to podzbiory zbioru Ω . Wyodrębniamy zbiór $\mathcal{A}$, który zawiera te podzbiory. Aby matematyka działała poprawnie, od rodziny zbiorów $\mathcal{A}$ (zwanej σ -algebrą) wymagamy spełnienia trzech warunków:

1. $\Omega \in \mathcal{A}, \emptyset \in \mathcal{A}$ (zdarzenie pewne i niemożliwe zawsze należą do naszych rozważań).
2. $A, B \in \mathcal{A} \implies A \cup B, A \cap B, A \setminus B \in \mathcal{A}$ (zbiór jest zamknięty na podstawowe operacje logiczne: sumę, iloczyn, różnicę).
3. $A_n \in \mathcal{A}, n = 1, 2, \dots \implies \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$ (nieskończona suma zdarzeń również jest zdarzeniem).

Ten ostatni, nieco tajemniczy punkt (który dzięki prawom De Morgana gwarantuje również zamkniętość na nieskończone przecięcia zbiorów) zapewnia nam *ciągłość prawdopodobieństwa*. Możemy na to patrzeć na dwa równoważne sposoby:

- **Ciągłość z dołu:** dla rosnącego ciągu zdarzeń $(A_1 \subset A_2 \subset \dots)$ chcemy, by $P(A_n) \rightarrow P(A)$, gdzie $A = \bigcup_{n=1}^{\infty} A_n$.
- **Ciągłość z góry:** dla malejącego ciągu zdarzeń $(A_1 \supset A_2 \supset \dots)$ chcemy, by $P(A_n) \rightarrow P(A)$, gdzie $A = \bigcap_{n=1}^{\infty} A_n$.

Dlatego właśnie nieskończone sumy (jak i iloczyny) zdarzeń muszą należeć do naszej σ -algebry, aby granice te miały matematyczny sens.

3. Miara prawdopodobieństwa P

Prawdopodobieństwo to funkcja przypisująca poszczególnym zdarzeniom z algebry $\mathcal{A}$ liczby rzeczywiste, spełniające aksjomaty Kołmogorowa:

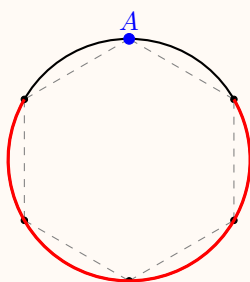
1. Nieujemność: Dla każdego $A \in \mathcal{A}$, zachodzi $P(A) \geq 0$.
2. Unormowanie: $P(\Omega) = 1$ oraz $P(\emptyset) = 0$.
3. Przeliczalna addytywność: Jeśli ciąg zdarzeń $A_1, A_2, \dots \in \mathcal{A}$ składa się ze zdarzeń parami rozłącznych ($A_i \cap A_j = \emptyset$ dla $i \neq j$), to:

$$P\left(\bigcup_n A_n\right) = \sum_n P(A_n)$$

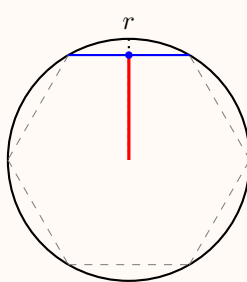
Ten ostatni aksjomat jest niezwykle ważny w praktyce. Gwarantuje on na przykład, że jeżeli podzielimy całą naszą przestrzeń zdarzeń elementarnych Ω na nieskończenie wiele (ale przeliczalnie wiele, np. ponumerowanych liczbami naturalnymi) rozłącznych kawałeczków A_n , to ich suma to po prostu Ω , a suma ich prawdopodobieństw sumuje się do jedności: $\sum_{n=1}^{\infty} P(A_n) = P(\Omega) = 1$. Bez tego nie moglibyśmy sensownie opisywać rozkładów na nieskończonych, ale dyskretnych zbiorach (np. gdy pytamy o liczbę atomów, które uległy rozpadowi – może to być $0, 1, 2, \dots$ w nieskończoność).

Przerywnik: Paradoks Bertranda

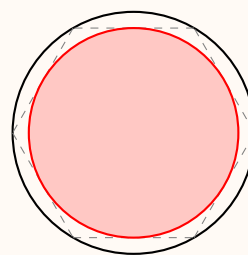
Aby pokazać, że matematyczny rygor nie jest sztuką dla sztuki, rozważmy klasyczny problem z 1889 roku. Losowo wybieramy cięciwę okręgu o promieniu r . Jakie jest prawdopodobieństwo, że jej długość jest **większa niż promień okręgu**?



Metoda 1: $p = \frac{4}{6} = \frac{2}{3}$



Metoda 2: $p = \frac{\sqrt{3}}{2}$



Metoda 3: $p = \frac{3}{4}$

Rysunek: Trzy różne rozwiązania paradoksu Bertranda (dla warunku, że długość cięciwy jest większa od promienia okręgu r , co geometrycznie odpowiada bokowi wpisanego sześciokąta). Kolorem czerwonym oznaczono zbiór zdarzeń sprzyjających dla każdego schematu losowania.

- **Odpowiedź nr 1:** Wybieramy losowo punkt na okręgu i z niego prowadzimy cięciwę do drugiego, losowo wybranego punktu na obwodzie. Rozważając sześciokąt foremny wpisany w okrąg (którego bok ma długość dokładnie r), zauważamy, że drugi punkt musi spaść na 4 z 6 łuków utworzonych przez wierzchołki sześciokąta. Prawdopodobieństwo wynosi $p = 4/6 = 2/3$.
- **Odpowiedź nr 2:** Załóżmy, że losujemy odległość środka cięciwy od środka okręgu – wartość z przedziału $(0, r)$. Aby cięciwa była dłuższa niż bok naszego sześciokąta (r), jej odległość od środka okręgu musi być mniejsza od wysokości trójkąta równobocznego budującego ten sześciokąt, czyli $r\sqrt{3}/2$. Ponieważ wybór jest jednorodny na odcinku r , daje nam to $p = \frac{r\sqrt{3}/2}{r} = \sqrt{3}/2 \approx 0.866$.
- **Odpowiedź nr 3:** Wybieramy losowo punkt wewnątrz całego okręgu, który staje się środkiem naszej cięciwy. Aby cięciwa była dłuższa od r , wybrany punkt musi leżeć wewnątrz mniejszego okręgu o promieniu $r\sqrt{3}/2$. Stosunek pól tych dwóch okręgów wynosi $3/4 = 0.75$.

Morał: Która odpowiedź jest poprawna? Wszystkie! Paradoks obnaża brutalną prawdę: zadanie probabilistyczne nie ma najmniejszego sensu bez jasnego sprecyzowania mechanizmu (schematu) losowania, a zatem wybory zbioru zdarzeń elementarnych. Pytanie „które zdarzenia są równoprawdopodobne?” jest ważniejsze niż same obliczenia.

1.2 Zdarzenia niezależne i prawdopodobieństwo warunkowe

Nasze przewidywania tego, co się wydarzy, zależą krytycznie od posiadanej informacji. I tak np. prawdopodobieństwo tego, że wracając wieczorem do domu zostaniemy pobici, mocno rośnie, gdy dowiemy się, że tego właśnie dnia Legia będzie grała mecz.

Jako mniej drastyczny przykład weźmy rzut dwiema monetami z przestrzenią próbek $\Omega = \{(O, O), (O, R), (R, O), (R, R)\}$. Niech zdarzenie $A = \{(O, O)\}$ (wypadły dwa orły), a zdarzenie B to „wypadła przynajmniej jedna reszka” czyli $\{(O, R), (R, O), (R, R)\}$. Jeśli wiemy, że zaszło zdarzenie B , to zdarzenie A staje się absolutnie niemożliwe. Stąd prawdopodobieństwo zdarzenia A pod warunkiem zajścia B wynosi zero.

Prowadzi to do definicji **prawdopodobieństwa warunkowego**:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Inny klasyczny przykład: w pokoju jest 100 osób. 20 z nich ma brodę, 25 to osoby czarnoskóre, a 5 to czarnoskóre osoby z brodą. Jakie jest prawdopodobieństwo, że wylosowana osoba ma brodę, *jeśli wiemy*, że jest czarnoskóra? $P(\text{brodą}|\text{czarnoskóra}) = 5/25 = 1/5$.

Jeżeli jednak zdobyta informacja nie zmienia absolutnie nic, mówimy o **zdarzeniach niezależnych**. Zdarzenia A i B są niezależne, jeśli zajście B nie wpływa na szansę zajścia A :

$$P(A|B) = P(A)$$

Podstawiając to do definicji prawdopodobieństwa warunkowego, otrzymujemy kluczowy wzór na iloczyn zdarzeń niezależnych:

$$P(A \cap B) = P(A) \cdot P(B)$$

Np. szansa na wyrzucenie dwóch szóstek w dwóch rzutach niezależnymi kostkami to po prostu $1/6 \cdot 1/6 = 1/36$.

Brak zrozumienia pojęcia niezależności to źródło najpopularniejszych przesądów:

- „Po dwóch dziewczynkach teraz musi być chłopiec, żeby statystyka się wyrównała.” (Natura nie ma pamięci).
- Naiwna strategia gry w ruletkę: „Poczekam na długą serię czerwonych i wtedy na pewno wypadnie czarne.” (Kulka też nie ma pamięci!).
- Stary dowcip o matematyku, który po fali zamachów bombowych zaczął wносить na pokład samolotu własną bombę. Wyliczył bowiem, że szansa, iż na pokładzie znajdzie się jedna bomba, wynosi 1 na 10 tysięcy, ale szansa na to, że w jednym samolocie znajdą się *dwie* bomby, wynosi już 1 na 100 milionów. (Błąd leży w niezależności – jego własna bomba nie zmniejsza prawdopodobieństwa cudzej bomby!).

Prawdopodobieństwo na sali sądowej

Z założeniem o niezależności trzeba niezmiernie uważać! Oto dwa mrożące krew w żyłach, autentyczne przykłady nadużycia rachunku prawdopodobieństwa przez oskarżycieli, które zrujnowały ludzkie życia.

Przypadek Josepha Trujillo (1946)

W USA zastrzelono oficera policji. O zabójstwo oskarżono Josepha Trujillo. Na jego ubraniu znaleziono 11 włókienek materiału odpowiadających materiałowi z ubrania ofiary. Prokurator stwierdził arbitralnie, że prawdopodobieństwo przypadkowej zgodności jednego włókna to 0.1. Mnożąc to przez siebie (zakładając niezależność!), przekonywał ławę przysięgłych, że dla 11 włókien szansa pomyłki to $0.1^{11} = 0.00000000001$. Co więcej, w pistolecie oskarżonego i w broni zbrodni znaleziono 5 cech wspólnych, m.in.:

- pistolety nieautomatyczne (prokurator założył $p = 0.5$),
- kaliber .38 ($p = 0.2$),
- lekko skrzywiona lufa ($p = 0.2$).

Pomnożenie tych „prawdopodobieństw” dało znów absurdalnie małą liczbę. Błąd polegał na tym, że te zdarzenia **nie były niezależne**. Włókna z tego samego płaszcza są ze sobą silnie skorelowane (jak spadnie jedno, to i drugie), a rewolwery kalibru .38 rzadziej bywały bronią automatyczną niż małe kalibry, więc cechy broni też były zależne. Niestety, Trujillo został skazany i stracony.

Sprawa małżeństwa Collinsów (1964)

W Kalifornii starsza kobieta została uduszona w alejce. Świadek zeznał, że z miejsca zbrodni wybiegła blondynka z włosami spiętymi w kucyki, wsiadła do żółtego samochodu prowadzonego przez czarnoskórego mężczyznę z wąsami i brodą. Krótco potem policja zatrzymała Janet i Malcolma Collinsów, którzy odpowiadali rysopisowi i posiadali żółtego Forda Lincolna. Prokurator, zachłyśnięty pseudomatematyką, zaprezentował ławie przysięgłych następujące prawdopodobieństwa:

- blondynka: $p = 1/3$
- kucyki: $p = 1/10$
- małżeństwo o różnych kolorach skóry: $p = 1/1000$
- żółty samochód: $p = 1/10$
- mężczyzna z wąsami: $p = 1/4$
- czarnoskóry z brodą: $p = 1/10$

Mnożąc te liczby, oskarżyciel otrzymał wynik 1 do 12000000, argumentując, że szansa, by winny był ktoś inny niż Collinsowie, jest bliska zeru. Ława przysięgłych skazała ich na wyroki wieloletniego więzienia (w pierwotnych doniesieniach groziła im kara śmierci).

Kilka lat później Sąd Najwyższy Kalifornii w głośnym wyroku *People v. Collins* zmasakrował te argumenty i anulował wyrok. Dlaczego? Po pierwsze, liczby były całkowicie wyssane z palca. Po drugie i najważniejsze – **cechy te absolutnie nie były niezależne!** Fakt, że mężczyzna nosi brodę, jest silnie skorelowany z posiadaniem przez niego wąsów. Fakt, że obserwujemy osobę o ciemnej skórze z brodą, w tamtym czasie nie był niezależny od faktu, że jest to osoba o określonym statusie demograficznym. Co więcej, eksperci matematyczni powołani przez obronę pokazali, że stosując nawet te wadliwe wartości prokuratora do populacji kilku milionów mieszkańców aglomeracji Los Angeles, istniało około 41% szans, że w mieście znajduje się *jeszcze co najmniej jedna taka para*. Skazanie na podstawie samej statystyki było matematyczną zbrodnią.

2 Zmienne losowe i ich rozkłady

*„Roll of the dice, take a spin of the wheel
Out of your hands now, so, how do you feel?
Do you feel lucky or do you feel scared?
Take what luck brings and be devil-may-care.”*

— Iron Maiden, *The Angel and the Gambler*

2.1 Zmienna losowa: pomost między abstrakcją a liczbą

Zanim przejdziemy dalej, zróbmy krótkie przypomnienie. Nasza matematyczna konstrukcja prawdopodobieństwa opiera się na trzech filarach: przestrzeni próbek (zbiór zdarzeń elementarnych) Ω , zbiorze zdarzeń losowych $\mathcal{A}$ i mierze prawdopodobieństwa P .

Jednak w fizyce rzadko interesuje nas to, czy moneta upadła orłem, czy reszką, albo jak dokładnie ułożyły się włókna w rzucie kością. Chcemy operować na konkretnych wielkościach fizycznych: energiach, prędkościach, położeniach. Potrzebujemy narzędzia, które przetłumaczy abstrakcyjne zdarzenia na język liczb rzeczywistych. Tym narzędziem jest **zmienna losowa**.

Zmienna losowa to funkcja określona na zbiorze zdarzeń elementarnych, przyjmująca wartości rzeczywiste:

$$X : \Omega \rightarrow \mathbb{R}$$

Przykładowo: zdarzeniu „wypadła reszka” przyporządkowujemy liczbę 1, a „wypadł orzeł” liczbę 0. Albo inaczej: wylosowanej w loterii osobie (zdarzenie elementarne to konkretny człowiek) przyporządkowujemy jej wzrost w centymetrach.

Uwaga! Ważny wymóg formalny

Zmienna losowa nie może być zupełnie dowolną funkcją. Aby cała nasza misterna maszyna probabilistyczna działała, będziemy od niej wymagali, aby dla każdego przedziału $[a, b] \subset \mathbb{R}$ zbiór postaci:

$$A_{ab} = \{\omega \in \Omega : a \leq X(\omega) \leq b\}$$

należał do naszej algebry zdarzeń $\mathcal{A}$. Krótko mówiąc: przeciwobrazy odcinków w $\mathbb{R}$ muszą być elementami zbioru zdarzeń losowych.

Dlaczego to takie ważne? Ponieważ prawdopodobieństwo P potrafimy liczyć *tylko* dla elementów ze zbioru $\mathcal{A}$! Jeśli jakiś przeciwobraz nie trafi do $\mathcal{A}$, nie będziemy w stanie odpowiedzieć na pytanie „jakie jest prawdopodobieństwo, że zmienna X przyjmie wartość z tego przedziału?”. Spójrzmy na **przykład patologiczny**: Niech nasza przestrzeń to $\Omega = \{1, 2, 3\}$, zaś algebra (bardzo uboga) to $\mathcal{A} = \{\{1\}, \{2, 3\}, \emptyset, \Omega\}$. Mamy też prawdopodobieństwa: $P(\{1\}) = 1/6$, $P(\{2, 3\}) = 5/6$. (Pamiętajmy, że zbiór pusty zawsze należy do $\mathcal{A}$, $P(\emptyset) = 0$).

Zdefiniujmy teraz funkcję $X : \Omega \rightarrow \mathbb{R}$ poprzez:

$$X(1) = 2, \quad X(2) = 3, \quad X(3) = 4.$$

Czy to jest zmienna losowa? **Nie!** Sprawdźmy przeciwobraz przedziału $(3.5, 4.5]$. Szukamy takich ω , dla których $3.5 < X(\omega) \leq 4.5$. Jedynek takim argumentem jest $\omega = 3$. Zatem przeciwobraz to zbiór $\{3\}$. Niestety, zbioru $\{3\}$ nie ma w naszej algebrze $\mathcal{A}$! Nie potrafimy odpowiedzieć na pytanie, z jakim prawdopodobieństwem zmienna przyjmuje wartość 4.

Widząc powyższy problem patologiczny, można by zadać bardzo logiczne pytanie: „Czy to nie oznacza, że zbiór zdarzeń losowych $\mathcal{A}$ musi po prostu zawierać absolutnie wszystkie możliwe podzbiory Ω , żebyśmy mieli spokój?”.

Otóż **nie!** Ten formalny wymóg wcale nie zmusza nas do powiększania algebry $\mathcal{A}$. Zmusza nas do czegoś zupełnie innego: do **dopasowania zmiennej losowej do dostępnej informacji** (w żargonie matematycznym mówimy, że zmienna losowa musi być *mierzalna* względem danej σ -algebry).

Wróćmy do naszego patologicznego przykładu. Zbiór $\mathcal{A} = \{\{1\}, \{2, 3\}, \emptyset, \Omega\}$ reprezentuje sytuację, w której nasz przyrząd pomiarowy potrafi odróżnić wynik 1 od reszty, ale **nie potrafi** rozróżnić 2 od 3 (wyniki te zawsze występują jako nierozdzielny pakiet $\{2, 3\}$).

Jeżeli spróbujemy zdefiniować funkcję X , która daje inny wynik dla 2 (np. $X(2) = 3$) i inny dla 3 (np. $X(3) = 4$), to funkcja ta po prostu wymaga większej precyzji, niż oferuje nasza algebra. Dlatego *nie jest* zmienną losową.

Gdybyśmy jednak zdefiniowali inną funkcję Y , np. reprezentującą wygraną w grze, w której

za wyrzucenie 1 dostajemy 100 zł, a za wyrzucenie 2 lub 3 płacimy 50 zł kary:

$$Y(1) = 100, \quad Y(2) = -50, \quad Y(3) = -50$$

To taka funkcja **jest pełnoprawną zmienną losową** w tej bardzo ubogiej algebrze! Sprawdźmy: przeciwobraz wartości -50 to zbiór $\{2, 3\}$, który bezpiecznie leży w naszym $\mathcal{A}$. Zmienna losowa musi po prostu „szanować” ograniczenia informacyjne naszego zbioru zdarzeń i nie może rozróżniać stanów, których nasza algebra nie widzi.

Ta kluczowa własność przeciwobrazów pozwala nam na zdefiniowanie fundamentalnego pojęcia – **dystrybuanty zmiennej losowej**:

$$F_X(x) = P(\{\omega \in \Omega : -\infty < X(\omega) \leq x\}) \equiv P(X \leq x)$$

Zapis $P(X \leq x)$ to oczywiście skrót myślowy. Oznacza on „prawdopodobieństwo tych wszystkich zdarzeń elementarnych ω , dla których wartość zmiennej losowej $X(\omega)$ jest mniejsza bądź równa x ”.

Jeżeli dystrybuenta jest funkcją absolutnie ciągłą (i tym samym różniczkowalną prawie wszędzie), pozwala ona łatwo wyznaczyć tzw. **gęstość prawdopodobieństwa**, która jest po prostu jej pochodną:

$$\rho_X(x) = \frac{dF_X(x)}{dx}$$

2.2 Rozkład dyskretny i ciągły

Jeśli nasza zmienna losowa przyjmuje tylko pewien dyskretny ciąg wartości (np. $x_1, x_2, x_3 \dots$), to sytuacja staje się bardzo prosta. Aby w pełni opisać jej rozkład, wystarczy podać ciąg prawdopodobieństw:

$$p_k = P(X = x_k)$$

(Jeszcze raz przypominam: to zapis skrócony. Chodzi o prawdopodobieństwo tych wszystkich zdarzeń z Ω , dla których X przyjmuje konkretnie wartość x_k). Oczywiście prawdopodobieństwa te nie mogą być ujemne ($p_i \geq 0$) i muszą sumować się do jedności: $\sum_i p_i = 1$.

W języku analizy matematycznej możemy zapisać gęstość takiego dyskretnego rozkładu przy pomocy delty Diraca:

$$\rho_X(x) = \sum_i p_i \delta(x - x_i)$$

Z kolei dystrybuenta staje się funkcją “schodkową”, opisywaną za pomocą funkcji skokowej (thety) Heaviside’a (gdzie $\theta(x) = 0$ dla $x \leq 0$ oraz $\theta(x) = 1$ dla $x > 0$):

$$F_X(x) = \sum_i p_i \theta(x - x_i)$$

Przykład: Rzut kostką. Zmienna losowa to liczba wyrzuconych oczek ($X \in \{1, 2, 3, 4, 5, 6\}$). Gęstość rozkładu to sześć szpilek Diraca (każda o wadze $1/6$), a dystrybuenta to równe schodki pnące się w górę co każde oczko aż do wartości 1.

Kluczowe przykłady rozkładów dyskretnych spotykane w fizyce:

1. **Rozkład dwumianowy:** Mamy N niezależnych prób. W każdej próbie mamy tylko dwie możliwości (np. sukces/porażka, orzeł/reszka) z prawdopodobieństwami odpowiednio p oraz $1 - p$. Zmienną losową n jest łączna liczba odniesionych sukcesów (np. liczba wyrzuconych orłów w N rzutach). Prawdopodobieństwo wyrzucenia dokładnie n orłów to:

$$p_n = \binom{N}{n} p^n (1 - p)^{N-n}$$

2. Rozkład Poissona:

$$p_n = e^{-\lambda} \frac{\lambda^n}{n!}$$

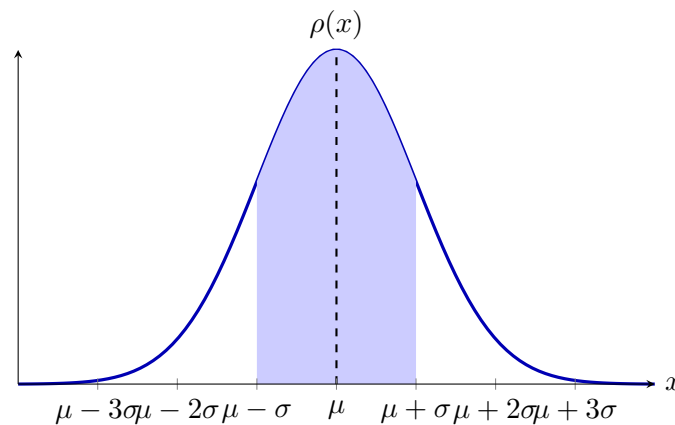
Jest to model opisujący prawdopodobieństwo wystąpienia dokładnie n rzadkich zdarzeń w określonym czasie t (np. liczba rozpadów promieniotwórczych w próbce w ciągu sekundy). Zakładamy, że kolejne wystąpienia są od siebie niezależne, a średnia spodziewana liczba zjawisk w tym czasie wynosi właśnie λ .

Zbiór wartości zmiennej losowej może być również **ciągły** – np. dzienny opad deszczu w Rabce, położenie cząstki w pudle, czy prędkość atomu w gazie.

Królem wszystkich rozkładów ciągłych jest oczywiście **Rozkład Gaussa (Normalny)**:

$$\rho(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Gdzie μ to wartość oczekiwana, a σ to odchylenie standardowe determinujące szerokość „dzwonu”.



Rysunek 1: Gęstość ciągłego rozkładu normalnego (Gaussa). Zacieniowany obszar reprezentuje prawdopodobieństwo znalezienia wartości w przedziale jednego odchylenia standardowego od średniej.

2.3 Momenty rozkładu

Aby scharakteryzować rozkład w sposób syntetyczny, wprowadzamy pojęcie wartości oczekiwanej. Wartością oczekiwaną funkcji $f(X)$ zmiennej losowej X nazywamy liczbę:

$$\langle f(x) \rangle = \int_{-\infty}^{\infty} f(x) \rho(x) dx$$

(Dla rozkładów dyskretnych całka zamienia się oczywiście w sumę).

W szczególności interesują nas tzw. momenty rozkładu:

- $\langle x \rangle = \int x \rho(x) dx$ – średnia (wartość oczekiwana zmiennej), czyli **pierwszy moment**.
- $\langle x^k \rangle = \int x^k \rho(x) dx$ – **k -ty moment zwykły**.
- $\langle (x - \langle x \rangle)^k \rangle$ – **k -ty moment centralny** (mierzony względem średniej).
- $\langle (x - \langle x \rangle)^2 \rangle = \langle x^2 \rangle - \langle x \rangle^2 \equiv \sigma^2$ – **wariancja**, czyli drugi moment centralny. Opisuje, jak bardzo wyniki rozrzucane są wokół średniej.

Uwaga – nie każdy rozkład ma skończone momenty!

Istnieją rozkłady patologiczne, np. rozkład Cauchy’ego (opisujący m.in. kształt linii widmowych), dla którego średnia i wariancja są matematycznie nieokreślone (całki są rozbieżne), co sprawia sporo problemów nieostrożnym badaczom.

Paradoks Petersburski (Daniel Bernoulli, 1730)

Aby pokazać, że sama średnia (wartość oczekiwana) nie opisuje w pełni ludzkiego pojmowania szansy, rozważmy słynny paradoks, który spędzał sen z powiek osiemnastowiecznym matematykom.

Gramy przeciwko kasynu. Krupier rzuca monetą. Jeśli wypadnie orzeł – dostajemy 1 rubla, a gra się kończy. Jeśli reszka – gramy dalej. Jeśli orzeł wypadnie w drugim rzucie – dostajemy 2 ruble i kończymy. W trzecim – 4 ruble, w czwartym – 8, i tak dalej. Ogólnie: jeśli pierwszy orzeł wypadnie w n -tym rzucie, nasza wygrana wynosi 2^{n-1} rubli.

Zadajmy standardowe pytanie: jaka jest matematyczna wartość oczekiwana (średnia) naszej wygranej? Prawdopodobieństwo, że wygramy w 1. rzucie to $1/2$. Że w drugim (ciąg Reszka, Orzeł) to $1/4$. W n -tym rzucie to $1/2^n$.

Liczymy średnią wygraną:

$$\langle \text{wygrana} \rangle = \left(\frac{1}{2} \cdot 1 \right) + \left(\frac{1}{4} \cdot 2 \right) + \dots + \left(\frac{1}{2^n} \cdot 2^{n-1} \right) + \dots$$

$$\langle \text{wygrana} \rangle = \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \dots = \infty$$

Wartość oczekiwana naszej wygranej to nieskończoność! A teraz kluczowe pytanie z teorii gier: **Ile bylibyś w stanie zapłacić wpisowego, żeby wziąć udział w tej grze?**

Naiwnie, z punktu widzenia matematyki, gra jest dla nas opłacalna (spodziewamy się zysku), o ile wpisowe jest mniejsze niż oczekiwana wygrana. Skoro oczekiwana wygrana jest nieskończona, powinniśmy zastawić dom, sprzedać nerkę i wziąć gigantyczny kredyt, byle tylko móc zagrać. A jednak żaden rozsądny człowiek nie zapłaciłby za wejście do tej gry więcej niż 10, może 20 rubli.

Dlaczego? Pomocnicze pytanie: są dwie gry. W jednej masz prawdopodobieństwo jedna miliarda na wygranę miliarda złotych. W drugiej masz prawdopodobieństwo $1/2$ na wygranę 2 złotych. Wartość oczekiwana jest ta sama: około 1 złotego. W którą grę zagrasz chętniej? Ludzie instynktownie odrzucają nieskończone małe prawdopodobieństwa potężnych wygranych, ponieważ szansa na odzyskanie wpisowego staje się nierealna. (Później doprowadziło to do powstania pojęcia „użyteczności krańcowej” w ekonomii).

Znajomość tego mechanizmu to podstawa **strategii d’Alemberta (martyngału)** w ruletce: stawiamy na czerwone. Przegrywamy? Podwajamy stawkę. Przegrywamy znowu? Znowu podwajamy. Teoretycznie, pierwsza wygrana pokryje nam wszystkie dotychczasowe straty z nawiązką. Problem w tym, że bardzo szybko docieramy do potężnych kwot (jak w Paradoksie Petersburskim), a kasyna mają limity maksymalnego zakładu, co brutalnie weryfikuje naiwne zastosowanie wartości oczekiwanej w praktyce.

2.4 Wielowymiarowe zmienne losowe

W fizyce rzadko mamy do czynienia z tylko jedną zmienną losową. Cząstka w gazie porusza się w trzech wymiarach, więc jej prędkość to wektor losowy składający się z trzech składowych. Załóżmy więc, że mamy więcej niż jedną zmienną losową: $X_1, X_2, X_3, \dots, X_n$, określone na tej samej przestrzeni zdarzeń Ω .

Ich **łącznie dystrybuantę** definiujemy jako:

$$F_{X_1 \dots X_n}(x_1, \dots, x_n) = P(X_1 < x_1, \dots, X_n < x_n)$$

Pamiętajmy o interpretacji tego wyrażenia: jest to prawdopodobieństwo części wspólnej (przecięcia) zdarzeń. Innymi słowy, to prawdopodobieństwo, że zmienna X_1 przyjmie wartość mniejszą od x_1 , i **jednocześnie** zmienna X_2 będzie mniejsza od x_2 , itd. W przypadku dyskretnym przykładem może być rzut dwiema kostkami: interesuje nas wynik na pierwszej (X_1) i na drugiej (X_2). W przypadku ciągłym – np. łączny rozkład wysokości męża i żony w losowo wybranych małżeństwach w Polsce.

Analogicznie do przypadku jednowymiarowego, gęstość prawdopodobieństwa to po prostu wielokrotna pochodna mieszana dystrybucyj:

$$\rho_{\mathbf{X}}(x_1, \dots, x_n) = \frac{\partial^n F_{\mathbf{X}}(x_1, \dots, x_n)}{\partial x_1 \dots \partial x_n}$$

Często znając pełen rozkład, interesuje nas tylko jedna zmienna (np. prędkość tylko w osi x , ignorując resztę). Otrzymujemy wtedy tzw. **rozkłady marginalne (brzegowe)**, całkując (lub sumując) po wszystkich pozostałych zmiennych, których chcemy się „pozbyć”:

$$\rho_1(x_1) = \int_{-\infty}^{\infty} \rho(x_1, x_2) dx_2$$

Momenty wielowymiarowe definiujemy analogicznie, licząc wielokrotne całki: $\langle x_1 \rangle$, $\langle x_2 \rangle$, czy wreszcie momenty mieszane określające korelacje, np. $\langle x_1 x_2 \rangle$.

Korzystając z poznanego wcześniej pojęcia niezależności zdarzeń losowych, możemy zdefiniować **niezależność zmiennych losowych**. Mówimy, że zmienne X_1 i X_2 są niezależne, jeżeli dla dowolnych przedziałów $A_1, A_2 \subset \mathbb{R}$ ich przeciwobrazy są zdarzeniami niezależnymi. W praktyce (w języku gęstości prawdopodobieństwa) oznacza to, że łączna gęstość faktoryzuje się na iloczyn gęstości brzegowych:

$$X_1 \text{ i } X_2 \text{ są niezależne} \iff \rho(x_1, x_2) = \rho_1(x_1)\rho_2(x_2)$$

Dowód jest natychmiastowy - z jednej strony $P(A_1 \cap A_2) = \int_{A_1} dx_1 \int_{A_2} dx_2 \rho(x_1, x_2)$, z drugiej $P(A_1)P(A_2) = \int_{A_1} dx_1 \rho_1(x_1) \int_{A_2} dx_2 \rho_2(x_2)$. Oba te wyrażenia powinny być równe (z definicji niezależności A_1 i A_2), a skoro ma to zachodzić dla dowolnych A_1 i A_2 to funkcje podcałkowe powinny być równe, a zatem $\rho(x_1, x_2) = \rho_1(x_1)\rho_2(x_2)$.

Dla zmiennych niezależnych wartość oczekiwana iloczynu jest iloczynem wartości oczekiwanych: $\langle x_1 x_2 \rangle = \langle x_1 \rangle \langle x_2 \rangle$ (ponieważ całka podwójna się po prostu rozбивa na dwie niezależne).

2.5 Przykład fizyczny: Rozkład Maxwella (1860) – oryginalne wyprowadzenie

To, co teraz zrobimy, to jedno z najpiękniejszych wyprowadzeń w historii fizyki. James Clerk Maxwell w 1860 roku postawił sobie za cel znalezienie rozkładu prędkości w gazie nieoddziałujących ze sobą cząsteczek w równowadze termicznej.

Maxwell przyjął tylko dwa, dość naturalne założenia:

1. **Niezależność składowych:** Prędkości w poszczególnych kierunkach kartezjańskich (v_x, v_y, v_z) są wzajemnie niezależnymi zmiennymi losowymi. Łączne prawdopodobieństwo to iloczyn prawdopodobieństw składowych:

$$p(v_x, v_y, v_z) = f(v_x)f(v_y)f(v_z)$$

2. **Izotropia przestrzeni:** W gazie w równowadze nie ma żadnego wyróżnionego kierunku. Jeśli obrócimy układ współrzędnych, rozkład nie powinien się zmienić. Gęstość musi zależeć wyłącznie od modułu prędkości (lub jej kwadratu $v^2 = v_x^2 + v_y^2 + v_z^2$), a nie od kierunku:

$$p(v_x, v_y, v_z) = \Psi(v_x^2 + v_y^2 + v_z^2)$$

Zestawiając oba założenia, dostajemy równanie funkcyjne:

$$f(v_x)f(v_y)f(v_z) = \Psi(v_x^2 + v_y^2 + v_z^2)$$

Rozpatrzmy teraz przypadek, gdy cząstka porusza się tylko wzdłuż osi X, czyli $v_y = v_z = 0$. Oznaczmy dla wygody $f(0) = c$. Mamy wtedy:

$$f(v_x) \cdot c \cdot c = \Psi(v_x^2 + 0 + 0) \implies c^2 f(v_x) = \Psi(v_x^2)$$

Zatem $f(v_x) = \Psi(v_x^2)/c^2$. Wstawiając to z powrotem do głównego równania, otrzymujemy:

$$\frac{\Psi(v_x^2)}{c^2} \frac{\Psi(v_y^2)}{c^2} \frac{\Psi(v_z^2)}{c^2} = \Psi(v_x^2 + v_y^2 + v_z^2)$$

$$\Psi(v_x^2)\Psi(v_y^2)\Psi(v_z^2) = c^6 \Psi(v_x^2 + v_y^2 + v_z^2)$$

Dla ułatwienia zapisu zmienmy zmienne: niech $\alpha = v_x^2$, $\beta = v_y^2$, $\gamma = v_z^2$. Równanie przyjmuje postać:

$$\Psi(\alpha)\Psi(\beta)\Psi(\gamma) = c^6 \Psi(\alpha + \beta + \gamma)$$

Zróżniczkujemy teraz to wyrażenie stronami po zmiennej α :

$$\Psi'(\alpha)\Psi(\beta)\Psi(\gamma) = c^6 \Psi'(\alpha + \beta + \gamma)$$

Podzielmy to nowe równanie przez równanie wyjściowe:

$$\frac{\Psi'(\alpha)}{\Psi(\alpha)} = \frac{\Psi'(\alpha + \beta + \gamma)}{\Psi(\alpha + \beta + \gamma)}$$

Podstawmy $\alpha = \gamma = 0$:

$$\frac{\Psi'(\beta)}{\Psi(\beta)} = \frac{\Psi'(0)}{\Psi(0)}$$

Zauważmy, że lewa strona tej równości zależy od β , a prawa nie, a tak może być tylko wtedy, gdy obie strony są równe stałej!

$$\frac{\Psi'(\beta)}{\Psi(\beta)} = \frac{\Psi'(0)}{\Psi(0)} \equiv -\lambda,$$

gdzie ujemny znak wybieramy przezornie, wiedząc, że prawdopodobieństwo dla nieskończonych prędkości musi spadać do zera. Otrzymaliśmy proste równanie różniczkowe: $\frac{d\Psi}{\Psi} = -\lambda d\beta$, które całkujemy:

$$\ln \Psi(\beta) = -\lambda\beta + C \implies \Psi(\beta) = Ae^{-\lambda\beta}$$

Stąd:

$$\Psi(\alpha + \beta + \gamma) = \frac{A^3}{c^6} e^{-\lambda(\alpha + \beta + \gamma)}$$

Ostatecznie nasza łączna gęstość prawdopodobieństwa prędkości wynosi:

$$p(v_x, v_y, v_z) = Be^{-\lambda(v_x^2 + v_y^2 + v_z^2)}$$

Pozostaje nam jedynie wyznaczyć stałą normalizacyjną B (całka po całej przestrzeni pędów musi wynosić 1) oraz parametr fizyczny λ . Wiemy z analizy, że całka Gaussa to:

$$\int_{-\infty}^{\infty} e^{-\lambda v_x^2} dv_x = \sqrt{\frac{\pi}{\lambda}}$$

Co dla trzech wymiarów daje $B = \left(\frac{\lambda}{\pi}\right)^{3/2}$.

Parametr λ możemy powiązać ze średnią energią kinetyczną cząsteczek. Obliczając wartość oczekiwaną kwadratu prędkości $\langle v^2 \rangle$ (np. różniczkując po parametrze), otrzymujemy:

$$\frac{m\langle v^2 \rangle}{2} = \frac{3m}{4\lambda}$$

Wybiegając nieco w przód do termodynamiki – z zasady ekwipartycji energii wiemy, że średnia energia kinetyczna zależy od temperatury T :

$$\frac{m\langle v^2 \rangle}{2} = \frac{3}{2}kT$$

Zestawiając te dwa wzory, otrzymujemy ostatecznie parametr $\lambda = \frac{m}{2kT}$. W ten sposób abstrakcyjna matematyka połączyła się z mikroskopowym ruchem gazu i makroskopową temperaturą.

2.6 Funkcje zmiennych losowych

Często znamy rozkład zmiennej losowej X (o gęstości $\rho_X(x)$), ale w eksperymencie mierzymy wielkość Y , która jest pewną jej funkcją: $Y = f(X)$. Pytanie brzmi: jaka jest gęstość prawdopodobieństwa $\rho_Y(y)$ nowej zmiennej?

Mechanizm jest niezwykle elegancki i opiera się na delcie Diraca. Obliczamy prawdopodobieństwo, że nowa zmienna przyjmie wartość ze zbioru A :

$$P(Y \in A) = \int_A \rho_Y(y) dy = \int_{f^{-1}(A)} \rho_X(x) dx$$

Zamiast całkować po skomplikowanym przeciwobrazie $f^{-1}(A)$, możemy całkować po całej przestrzeni używając dystrybucji Diraca jako matematycznego „sita”, które wyłapie tylko te wartości x , dla których $f(x) = y$:

$$P(Y \in A) = \int_A dy \int_{-\infty}^{\infty} dx \rho_X(x) \delta(y - f(x))$$

Porównując lewą i prawą stronę, natychmiast odczytujemy wzór na nową gęstość:

$$\rho_Y(y) = \int_{-\infty}^{\infty} \rho_X(x) \delta(y - f(x)) dx = \langle \delta(y - f(X)) \rangle_{\rho_X}$$

Ten potężny wzór można łatwo uogólnić na przypadek wielowymiarowych zmiennych losowych $\mathbf{X} = (X_1, X_2, \dots, X_n)$ oraz wielowymiarowych funkcji $\mathbf{Y} = \mathbf{f}(\mathbf{X})$:

$$\rho_{\mathbf{Y}}(\mathbf{y}) = \int d\mathbf{x} \rho_{\mathbf{X}}(\mathbf{x}) \delta(\mathbf{y} - \mathbf{f}(\mathbf{x})) = \langle \delta(\mathbf{y} - \mathbf{f}(\mathbf{X})) \rangle_{\rho_{\mathbf{X}}}$$

Przykład: Suma dwóch niezależnych zmiennych losowych

Mamy dwie rzutki, dwie kostki lub dwie cząstki gazu. Mamy zmienne niezależne X_1 i X_2 o zadanych rozkładach $\rho_1(x_1)$ oraz $\rho_2(x_2)$. Szukamy rozkładu gęstości ich **sumy**: $Y = X_1 + X_2$.

Używamy naszego wzoru z deltą Diraca (pamiętając, że dla zmiennych niezależnych łączna gęstość to iloczyn $\rho_1 \cdot \rho_2$):

$$\rho_Y(y) = \langle \delta(y - X_1 - X_2) \rangle_{\rho_1 \times \rho_2} = \int \int dx_1 dx_2 \delta(y - x_1 - x_2) \rho_1(x_1) \rho_2(x_2)$$

Całkujemy po zmiennej x_2 . Delta Diraca „wybiera” za nas wartość $x_2 = y - x_1$:

$$\rho_Y(y) = \int_{-\infty}^{\infty} dx_1 \rho_1(x_1) \rho_2(y - x_1)$$

Wniosek jest fundamentalny: gęstość prawdopodobieństwa sumy niezależnych zmiennych losowych jest splotem ich indywidualnych gęstości. To właśnie ten fakt leży u podstaw słynnego Centralnego Twierdzenia Granicznego.

Część III

Procesy nierównowagowe

1 Ruchy Browna i błędzenie losowe

*„And you may find yourself
In another part of the world
And you may find yourself
Behind the wheel of a large automobile
And you may find yourself in a beautiful house
With a beautiful wife
And you may ask yourself
Well ... how did I get here?” — Talking
Heads, Once in a Lifetime*

1.1 Kropla mleka i siła życiowa

Na wykładzie mieliśmy okazję spojrzeć przez mikroskop na kroplę mleka rozcieńczoną w wodzie. Drobne kuleczki tłuszczu zawieszone w płynie nieustannie, chaotycznie przemieszczały się we wszystkich kierunkach. Ten nieustanny taniec mikroskopijnych cząstek to właśnie **ruchy Browna**. Zjawisko to było odkryte w 1827 roku przez szkockiego botanika Roberta Browna, który badał pod mikroskopem małe cząstki pyłku kwiatowego zawieszone w wodzie. Ponieważ pyłek to element rozrodczy roślin, Brown sądził, że ten nieustanny ruch to przejaw mitycznej „siły życiowej” (*vis vitalis*) – legendarnego tchnienia, które miało odróżniać materię ożywioną od nieożywionej. Brown przeprowadził z pyłkiem imponującą liczbę eksperymentów. Zauważył, że ruchy te stają się szybsze wraz ze wzrostem temperatury płynu. Co więcej, pokazał, że ten ruch nigdy nie ustaje – próbki zamknięte w szafie nawet na kilka lat po wyjęciu wciąż wykazywały tę samą, nieprzerwaną aktywność.

Prawdziwe zrozumienie tego zjawiska przyszło dopiero na początku XX wieku (w latach 1905–1906), głównie za sprawą prac Alberta Einsteina oraz Mariana Smoluchowskiego. Udo wodnili oni, że ruchy Browna są efektem wynikającym bezpośrednio z faktu, że woda składa się z niewidzialnych atomów. Niezliczone miliardy maleńkich cząsteczek wody nieustannie i ze wszystkich stron zderzają się z dużą, zawieszoną w niej cząstką pyłku.

W tym miejscu pojawia się jednak bardzo naturalne pytanie: skoro cząstka pyłku jest bombardowana ze wszystkich stron przez miliardy cząsteczek wody, to czy średnio rzecz biorąc nie uderza w nią tyle samo cząsteczek z lewej, co z prawej strony? Dlaczego więc cząstka w ogóle się przemieszcza, a nie stoi w miejscu, wibrując jedynie wokół środka?

Marian Smoluchowski odpierał on te zarzuty, porównując ruch cząstki Browna do grającego w kasynie hazardzisty:

„Jest to taki sam błąd rozumowania, jak gdyby człowiek uprawiający grę hazardową (np. rzucanie kostki) sądził, że nigdy większej straty ani też większego zysku mieć nie będzie, niż wynosi stawka na jeden rzut. Wiemy dobrze, że szczęście i nieszczęście zwykle niezupełnie się równoważą; że im dłużej gra trwa, tym większa jest przeciętna suma albo wygrana, albo stracona”.

Załóżmy, że w każdej rundzie rzucamy monetą: jeśli wypadnie orzeł, gracz zyskuje 1 jednostkę kapitału (+1 z $p = 1/2$), a jeśli reszka – traci 1 jednostkę (−1 z $p = 1/2$). Średnio rzecz biorąc, wygrana z takiej gry to zero. Ale jeśli spojrzymy na rozkład kapitału w czasie, zobaczymy, że im dłużej nasz hazardzista gra, tym bardziej jego kapitał odchyła się od wartości początkowej. Kwadrat odchylenia stanu konta naszego gracza od stanu początkowego $(N_t - N_0)^2$ rośnie pro-

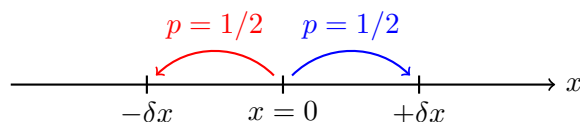
porcjonalnie do liczby rozegranych partii n . Oznacza to, że typowe odchylenie od średniej (czyli statystyczna nadwyżka wygranych nad przegranymi) rośnie jak pierwiastek z liczby prób: $\sqrt{n}$.

Aby przenieść tę hazardową analogię na grunt fizyki, musimy dokonać prostego tłumaczenia pojęć. Rola „rzutu monetą” przypada pojedynczym uderzeniom niewidzialnych molekuł cieczy, a „kapitałem” gracza staje się aktualne położenie cząstki pyłku względem punktu startu. Uderzenie z jednej strony to nasza wygrana, uderzenie z przeciwnej – strata. Zgodnie z prawami statystyki, nawet przy idealnej symetrii, z upływem czasu cząstka będzie doświadczać coraz większych fluktuacji. Oznacza to, że co rusz pojawiać się będą statystyczne **nadwyżki** uderzeń z jednego kierunku, które zaczną chaotycznie przesuwac pyłek w przestrzeni.

Sprawdźmy, czy te statystyczne nadwyżki mają w ogóle znaczenie w skali fizycznej. Ziarenko zawieszony jest uderzany przez molekuły cieczy średnio $n = 10^{20}$ razy w ciągu zaledwie jednej sekundy. Zgodnie z regułą opisaną wyżej, przeciętna nadwyżka uderzeń z jednej strony nad drugą wyniesie w tym czasie pierwiastek z tej gigantycznej liczby zderzeń, czyli około $\sqrt{10^{20}} = 10^{10}$ (dziesięć miliardów)! Nawet jeżeli pęd przekazany przez pojedynczą molekułę jest znikomo mały, to sumaryczny efekt dziesięciu miliardów „nadmiarowych” uderzeń z jednego kierunku jest wystarczająco potężny, by popychać cząstkę pyłku w widoczny pod mikroskopem sposób.

1.2 Matematyczny model błędzenia przypadkowego w 1D

Zbudujmy prosty, jednowymiarowy model tego zjawiska. Rozważmy cząstkę na osi x , która w chwili $t = 0$ startuje z punktu $x = 0$.



Wprowadzamy trzy proste założenia:

1. Każda cząstka w pojedynczym kroku czasowym δt wykonuje skok w prawo lub w lewo o ustaloną odległość δx . Dokładne wartości δt oraz δx zależą od fizycznych własności cieczy: jej lepkości, temperatury i rozmiaru samej cząstki.
2. Prawdopodobieństwo skoku w lewo lub w prawo jest równe i wynosi $1/2$. Co kluczowe, kierunek skoku **nie zależy od tego, co działo się wcześniej** – brakuje tu jakiegokolwiek pamięci, a kolejne kroki są zdarzeniami w pełni niezależnymi.
3. Różne cząstki w cieczy przemieszczają się niezależnie od siebie.

Jakie są konsekwencje takiego modelu dla ruchu po n krokach czasowych? Położenie po n -tym kroku opisuje równanie:

$$X(n) = X(n-1) + \Delta X$$

gdzie zmienna losowa (skok) ΔX przyjmuje wartość $+\delta x$ z $p = 1/2$ oraz $-\delta x$ również z $p = 1/2$.

1. Średnie przesunięcie

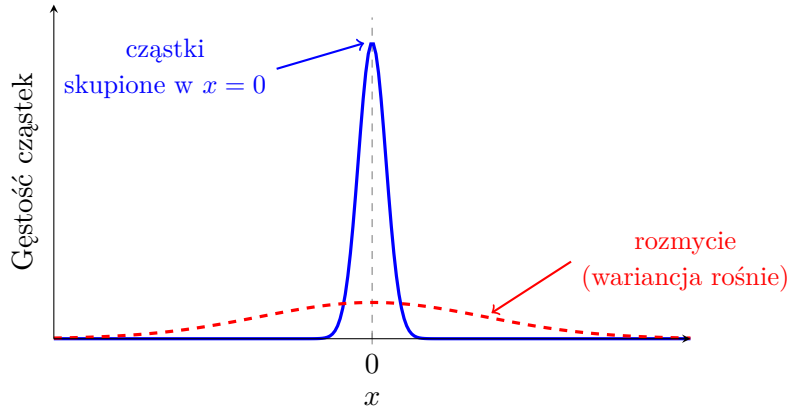
Wyśredniujmy nasze równanie stronami:

$$\langle X(n) \rangle = \langle X(n-1) \rangle + \langle \Delta X \rangle$$

Ponieważ $\langle \Delta X \rangle = 0$, otrzymujemy:

$$\langle X(n) \rangle = \langle X(n-1) \rangle = \langle X(n-2) \rangle = \dots = \langle X(0) \rangle = 0$$

Zgodnie z przewidywaniami, średnie przesunięcie cząstki jest równe 0. Gdybyśmy zamiast jednej cząstki wypuszczali całą ich gromadę z $x = 0$ i przyjęli założenie że ich ruchy wzajemnie na siebie nie wpływają, to powyższy wynik oznaczałby, że choć to skupisko cząstek z upływem czasu rozmywa się na boki, środek paczki cały czas pozostaje w punkcie $x = 0$.



2. Średnie przesunięcie kwadratowe

Podnieśmy równanie ruchu do kwadratu:

$$X^2(n) = (X(n-1) + \Delta X)^2 = X^2(n-1) + 2\Delta X \cdot X(n-1) + (\Delta X)^2$$

i znów wyśredniujmy stronami:

$$\langle X^2(n) \rangle = \langle X^2(n-1) \rangle + 2\langle \Delta X \cdot X(n-1) \rangle + \langle (\Delta X)^2 \rangle$$

W tym miejscu korzystamy z założenia o braku pamięci. Skoro kierunek nowego skoku ΔX jest niezależny od przeszłego położenia $X(n-1)$, to wartość oczekiwana ich iloczynu jest iloczynem wartości oczekiwanych. A skoro $\langle \Delta X \rangle = 0$, to cały człon mieszany znika:

$$2\langle \Delta X \cdot X(n-1) \rangle = 2 \cdot \langle \Delta X \rangle \cdot \langle X(n-1) \rangle = 0$$

Wartość oczekiwana kwadratu pojedynczego skoku wynosi natomiast $\langle (\Delta X)^2 \rangle = (\delta x)^2$. Zostaje nam prosta rekurencja:

$$\langle X^2(n) \rangle = \langle X^2(n-1) \rangle + (\delta x)^2$$

Zwijając to równanie dla n kroków aż do zera (zakładając $\langle X^2(0) \rangle = 0$), dostajemy:

$$\langle X^2(n) \rangle = n(\delta x)^2$$

Ostatni krok to zamiana liczby kroków n na fizyczny czas. Skoro każdy krok trwa δt , to całkowity czas $t = n\delta t$, skąd $n = t/\delta t$. Podstawiając to do równania:

$$\langle X^2(t) \rangle = \left[\frac{(\delta x)^2}{\delta t} \right] t$$

Dla wygody wprowadzamy w tym miejscu wielkość zwaną **stałą dyfuzji** D , zdefiniowaną jako $D = \frac{(\delta x)^2}{2\delta t}$. W ten sposób otrzymujemy najważniejszy wzór na dyfuzję w jednym wymiarze:

$$\langle x^2 \rangle = 2Dt$$

W przypadku przestrzeni trójwymiarowej wariancje z trzech wzajemnie niezależnych osi po prostu się dodają, co daje wynik 3 razy większy:

$$\langle r^2 \rangle = \langle x^2 \rangle + \langle y^2 \rangle + \langle z^2 \rangle = 6Dt$$

1.3 Czy dyfuzja jest efektywna?

Aby określić rozmiar „rozmywania” się paczki cząstek po upływie zadanego czasu, patrzymy na odchylenie standardowe położenia, będące pierwiastkiem z wariancji:

$$\sqrt{\langle x^2 \rangle} = \sqrt{2Dt}$$

Sprawdźmy teraz, jak ten statystyczny mechanizm transportu radzi sobie w różnych rzędach wielkości, przyjmując wartość stałej dyfuzji dla małych cząsteczek w wodzie w temperaturze pokojowej wynoszącą $D \sim 10^{-5} \text{ cm}^2/\text{s}$. Z przekształcenia wzoru wiemy, że czas dyfuzji to mniej więcej $t \approx x^2/2D$.

- **Skala mikro (wymiary bakterii):** Dla odległości rzędu $x = 10^{-4} \text{ cm}$, czas wynosi $t \approx (10^{-4})^2/(2 \cdot 10^{-5}) \approx 5 \cdot 10^{-4} \text{ s}$. Zaledwie 0.5 milisekundy! W mikroskopijnym świetle pojedynczych komórek transport dyfuzyjny jest bardzo efektywny.
- **Skala próbki chemicznej:** Jeśli chcemy, aby nasza substancja przebyła $x = 1 \text{ cm}$, czas wzrasta lawinowo do $t \approx 1^2/(2 \cdot 10^{-5}) = 50\,000 \text{ sekund}$. To oznacza aż ~ 14 godzin. Dyfuzja w tej skali staje się procesem niezwykle mało efektywnym.
- **Skala makroskopowa (pokój):** Powszechnie uważa się, że zapachy w pokoju rozchodzą się dzięki mechanizmowi dyfuzji. Policzmy to! Dla dystansu 5 m (500 cm), dyfuzja wylicza nam czas na przebycie tego dystansu równy dobremu kilku miesiącom. W rzeczywistości, jeżeli po chwili czujemy na drugim końcu pokoju zapach perfum, nie jest to absolutnie zasługa powolnej i ślepej dyfuzji, lecz mechanizmu **konwekcji** – makroskopowych, fizycznych prądów mieszających powietrze w pomieszczeniu.

1.4 Równanie dyfuzji: opis ciągły

Zamiast śledzić chaotyczną trajektorię pojedynczej cząstki, możemy przejść do opisu ewolucji prawdopodobieństwa znalezienia cząstki w danym punkcie x w chwili t , czyli $P(x, t)$. Wychodząc z naszego dyskretnego modelu błędzenia, cząstka może znaleźć się w punkcie x w chwili $t + \delta t$ tylko wtedy, gdy wcześniej (w chwili t) była w sąsiednim węźle i wykonała odpowiedni skok. Równanie bilansu prawdopodobieństwa ma postać:

$$P(x, t + \delta t) = \frac{1}{2}P(x - \delta x, t) + \frac{1}{2}P(x + \delta x, t)$$

Odejmując obustronnie $P(x, t)$, dzieląc przez krok czasowy δt i wyciągając czynnik przed nawias po prawej stronie, od razu formujemy klasyczne ilorazy różnicowe:

$$\frac{P(x, t + \delta t) - P(x, t)}{\delta t} = \frac{(\delta x)^2}{2\delta t} \left[\frac{P(x - \delta x, t) - 2P(x, t) + P(x + \delta x, t)}{(\delta x)^2} \right]$$

Dokonując przejścia do granicy ciągłej ($\delta t \rightarrow 0, \delta x \rightarrow 0$), dyskretne prawdopodobieństwo $P(x, t)$ zastępujemy ciągłą gęstością prawdopodobieństwa $\rho(x, t)$ (formalnie $P(x, t) \approx \rho(x, t)\delta x$, przy czym stały czynnik δx upraszcza się po obu stronach równania). Nasze ułamki stają się z definicji odpowiednio: pierwszą pochodną gęstości po czasie oraz jej drugą pochodną przestrzenną. Rozpoznając w czynniku przed nawiasem wyprowadzoną wcześniej stałą dyfuzji $D = \frac{(\delta x)^2}{2\delta t}$, natychmiast otrzymujemy różniczkowe **równanie dyfuzji**:

$$\frac{\partial \rho}{\partial t} = D \frac{\partial^2 \rho}{\partial x^2}$$

Przeszliśmy w ten sposób od mikroskopowego chaosu do gładkiego, deterministycznego równania na pole koncentracji. Rozwiążmy to równanie przy założeniu, że w chwili $t = 0$ wszystkie cząstki startują z początku układu współrzędnych:

$$\rho(x, 0) = \delta(x)$$

Rozwiązaniem dla tak zadanego warunku początkowego jest tzw. **rozwiązanie fundamentalne** (funkcja Greena), mające postać rozmywającego się w czasie rozkładu Gaussa:

$$\rho(x, t) = \frac{1}{\sqrt{4\pi Dt}} \exp\left(-\frac{x^2}{4Dt}\right)$$

Porównując ten wynik ze standardową postacią gęstości rozkładu normalnego ($\propto \exp\left(-\frac{x^2}{2\sigma^2}\right)$), natychmiast odczytujemy wariancję:

$$2\sigma^2 = 4Dt \quad \implies \quad \sigma^2 = 2Dt$$

W ten sposób zatoczyliśmy pełne koło. Rozwiązując różniczkowe równanie ciągłe, otrzymaliśmy wariancję położenia identyczną z wynikiem $\langle X^2 \rangle = 2Dt$, który wcześniej wyprowadziliśmy czysto algebraicznymi metodami dla dyskretnego błędzenia przypadkowego.

1.5 Procesy Markowa i równowaga szczegółowa

Analizując błędzenie przypadkowe, poczyniliśmy jedno fundamentalne założenie: prawdopodobieństwo tego, w którą stronę cząstka skoczy w kolejnym kroku, nie zależy w żaden sposób od jej przeszłości. Przyszłość zależy *wyłącznie* od jej obecnego położenia.

Procesy losowe posiadające tę własność „braku pamięci” nazywamy **procesami Markowa**. Modele oparte na dyskretnych stanach i dyskretnych krokach czasowych z tą własnością to **łańcuchy Markowa**. Z fizycznego punktu widzenia oznacza to, że układ bardzo szybko „zapomina” o swoich warunkach początkowych wskutek chaotycznych zderzeń.

Wyobraźmy sobie, że nasz łańcuch Markowa ewoluuje przez bardzo długi czas, aż osiągnie **stan stacjonarny** (równowagowy). Oznacza to, że rozkład prawdopodobieństwa znalezienia układu w danym makroście przestaje się zmieniać w czasie: $P(n, t+1) = P(n, t) \equiv P_{\text{eq}}(n)$.

Aby utrzymać ten stacjonarny rozkład, średnia liczba przejść (strumień prawdopodobieństw) wychodzących z dowolnego stanu musi być równa liczbie przejść do niego wchodzących. Najprostszym sposobem na zagwarantowanie takiego bilansu jest żądanie, aby bilansowały się procesy zachodzące między *każdą parą pojedynczych stanów*. Koncepcję tę nazywamy **zasadą równowagi szczegółowej** (*detailed balance*).

Zgodnie z nią, strumień prawdopodobieństwa ze stanu n do $n+1$ jest dokładnie równoważony przez strumień powrotny:

$$P_{\text{eq}}(n) \cdot W(n \rightarrow n+1) = P_{\text{eq}}(n+1) \cdot W(n+1 \rightarrow n) \quad (1.1)$$

Uwaga: równowaga szczegółowa jako narzędzie

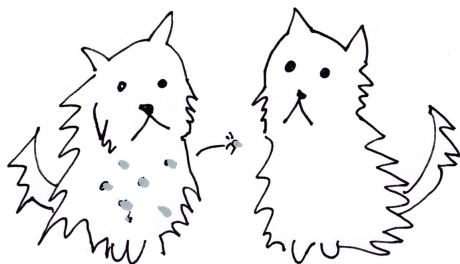
Zasada równowagi szczegółowej jest warunkiem *wystarczającym* na stacjonarność rozkładu, ale nie koniecznym – rozkład stacjonarny mógłby istnieć nawet bez niej (np. przy cyklicznej cyrkulacji prawdopodobieństwa $A \rightarrow B \rightarrow C \rightarrow A$). Stosujemy ją w fizyce statystycznej jako niezwykle potężne i wygodne *narzędzie*: jeżeli uda nam się znaleźć rozkład, który ją spełnia, to mamy absolutną pewność, że jest to rozkład stacjonarny.

2 Model Ehrenfestów (psy i pchły)

„They say jump
Jump, they say
They say jump
Jump, they say”
— David Bowie, *Jump They Say*

Model zaproponowany w 1907 roku przez Paula i Tatiangę Ehrenfestów wybitny matematyk Mark Kac nazwał „prawdopodobnie jednym z najbardziej pouczających modeli w całej fizyce”. Pokazuje on w sposób krystalicznie przejrzysty, jak mikroskopowa odwracalność i losowość prowadzą do makroskopowej nieodwracalności i równowagi termodynamicznej – a wszystko to za pomocą obrazu, który każdy może sobie wyobrazić: dwóch psów i przeskakujących między nimi pcheł.

Wyobraźmy sobie zatem dwa psy: Azora i Burka, na których żeruje łącznie N rozróżnialnych pcheł (patrz Ry. 2). Stan (makrostan) układu w dowolnej chwili określamy przez liczbę pcheł znajdujących się na Azorze, którą oznaczamy jako n . Siłą rzeczy, na Burku przebywa w tym samym czasie $N - n$ pcheł.



Rysunek 2: Azor, Burek i ich pchły w chwili początkowej (wszystkie na Azorze).

W każdym dyskretnym kroku czasowym wybieramy losowo (z jednakowym prawdopodobieństwem $1/N$) jedną z N pcheł, która natychmiast przeskakuje na drugiego psa. Zastanówmy się, jak zmienia się stan układu w pojedynczym kroku. Zakładając, że obecnie na Azorze mamy n pcheł, w kolejnym kroku stan ten może zmienić się tylko o ± 1 :

- **Skok w górę** ($n \rightarrow n+1$): Zdarzy się tylko wtedy, gdy wylosowana zostanie pchła siedząca obecnie na Burku. Ponieważ takich pcheł jest $N - n$, prawdopodobieństwo tego przejścia wynosi:

$$W(n \rightarrow n+1) = \frac{N-n}{N}$$

- **Skok w dół** ($n \rightarrow n-1$): Zdarzy się, jeśli wylosujemy pchłę siedzącą na Azorze. Takich pcheł jest n , więc prawdopodobieństwo wynosi:

$$W(n \rightarrow n-1) = \frac{n}{N}$$

Zastanówmy się nad fizyczną intuicją tego procesu. Jeśli początkowo większość z N pcheł znajduje się na Azorze, to w pierwszych chwilach liczba n będzie gwałtownie maleć na korzyść (czy raczej niekorzyść) Burka. Po dostatecznie długim czasie układ osiągnie stan równowagi, w którym liczba pcheł na każdym z psów będzie fluktuować wokół $N/2$. Co istotne, rozmiar tych fluktuacji (mierzony względną dyspersją $\sigma/\langle n \rangle$) silnie zależy od układu: dla małego N układ będzie gwałtownie przeskakiwać między skrajnymi stanami, natomiast w granicy makroskopowej (bardzo duże N) rozkład skoncentruje się niezwykle ciasno wokół wartości średniej.

W kontekście tego modelu postawimy sobie trzy fundamentalne pytania:

1. Jaki jest równowagowy rozkład prawdopodobieństwa $P_{\text{eq}}(n)$ znalezienia dokładnie n pcheł na Azorze po bardzo długim czasie?
2. Jak w funkcji czasu ewoluuje średnia liczba pcheł $\langle n(t) \rangle$?
3. Ile średnio kroków trzeba odczekać, aby wszystkie pchły samorzutnie powróciły na Azora?

2.1 Rozkład równowagowy przez równowagę szczegółową

Szukamy rozkładu stacjonarnego $P_{\text{eq}}(n)$, tj. takiego, który nie zmienia się w czasie. Naszą strategią będzie skorzystanie z omówionej wcześniej zasady równowagi szczegółowej. Podstawiając prawdopodobieństwa przejścia, otrzymujemy relację rekurencyjną:

$$P_{\text{eq}}(n) \cdot \frac{N-n}{N} = P_{\text{eq}}(n+1) \cdot \frac{n+1}{N}, \quad (2.1)$$

skąd:

$$\frac{P_{\text{eq}}(n+1)}{P_{\text{eq}}(n)} = \frac{N-n}{n+1}. \quad (2.2)$$

Rozwijamy tę rekurencję krok po kroku, zaczynając od $P_{\text{eq}}(0) \equiv a$:

$$P_{\text{eq}}(1) = \frac{N}{1} \cdot a, \quad (2.3)$$

$$P_{\text{eq}}(2) = \frac{N-1}{2} \cdot P_{\text{eq}}(1) = \frac{N(N-1)}{1 \cdot 2} \cdot a, \quad (2.4)$$

$$P_{\text{eq}}(n) = \frac{N \cdot (N-1) \cdots (N-n+1)}{n!} \cdot a = \binom{N}{n} \cdot a. \quad (2.5)$$

Z warunku normalizacji $\sum_{n=0}^N P_{\text{eq}}(n) = 1$ i tożsamości $\sum_{n=0}^N \binom{N}{n} = 2^N$ wyznaczamy stałą $a = 2^{-N}$. Ostatecznie:

$$\boxed{P_{\text{eq}}(n) = \frac{1}{2^N} \binom{N}{n}}. \quad (2.6)$$

Jest to **rozkład dwumianowy** z parametrem $p = 1/2$. Ma on wartość oczekiwaną $\langle n \rangle = N/2$ i odchylenie standardowe $\sigma = \sqrt{N}/2$. Względna szerokość rozkładu maleje jak $\sigma/\langle n \rangle = 1/\sqrt{N}$ – dla makroskopowej liczby cząstek rozkład jest ekstremalnie wąski i z astronomiczną pewnością obserwujemy stan $n \approx N/2$.

Warto przy tym zauważyć, że wynik ten ma głębokie uzasadnienie kombinatoryczne: rozkład dwumianowy z $p = 1/2$ to po prostu rozkład jednostajny po wszystkich 2^N mikroskopowych konfiguracjach (przypisaniach każdej pchły do jednego z psów). Każda konfiguracja mikroskopowa jest jednakowo prawdopodobna – to zasada równouprawnienia a priori stanów mikroskopowych.

2.2 Relaksacja wartości średniej

Zapytajmy teraz, jak wartość oczekiwana $\langle n(t) \rangle$ ewoluuje w czasie, startując z warunkiem początkowym $n(0) = N$ (wszystkie pchły na Azorze).

W każdym kroku n zmienia się o $+1$ z prawdopodobieństwem $\frac{N-n}{N}$ lub o -1 z prawdopodobieństwem $\frac{n}{N}$, więc:

$$\langle n(t+1) \rangle - \langle n(t) \rangle = (+1) \cdot \frac{N - \langle n \rangle}{N} + (-1) \cdot \frac{\langle n \rangle}{N} = 1 - \frac{2\langle n \rangle}{N}, \quad (2.7)$$

skąd:

$$\langle n(t+1) \rangle = \langle n(t) \rangle \left(1 - \frac{2}{N}\right) + 1. \quad (2.8)$$

Aby rozwiązać rekurencję wprowadzamy nową zmienną opisującą **nadwyżkę** pcheł na Azorze ponad wartość równowagową:

$$m(t) = \langle n(t) \rangle - \frac{N}{2}. \quad (2.9)$$

Podstawiając $\langle n(t) \rangle = m(t) + N/2$ do równania rekurencyjnego:

$$m(t+1) + \frac{N}{2} = \left(m(t) + \frac{N}{2}\right) \left(1 - \frac{2}{N}\right) + 1. \quad (2.10)$$

Upraszczając:

$$m(t+1) = m(t) \left(1 - \frac{2}{N}\right), \quad (2.11)$$

co daje

$$m(t) = m(0) \left(1 - \frac{2}{N}\right)^t. \quad (2.12)$$

Warunek początkowy $n(0) = N$ daje $m(0) = N/2$, zatem wracając do oryginalnej zmiennej:

$$\boxed{\langle n(t) \rangle = \frac{N}{2} + \frac{N}{2} \left(1 - \frac{2}{N}\right)^t.} \quad (2.13)$$

Z granicy $t \rightarrow \infty$ od razu widać, że układ dąży do wartości równowagowej $N/2$, co stanowi potwierdzenie spójności rachunku. Dla dużych N korzystamy z przybliżenia $\ln(1-2/N) \approx -2/N$, co daje wykładniczą relaksację:

$$\langle n(t) \rangle \approx \frac{N}{2} \left(1 + e^{-2t/N}\right), \quad (2.14)$$

ze charakterystycznym czasem relaksacji $t_{\text{eq}} = N/2$. Im więcej pcheł, tym wolniejsza relaksacja – co ma głęboki sens fizyczny: większy układ potrzebuje więcej kroków, by “wymieszać” się do równowagi.

2.3 Czas powrotu do stanu początkowego

Pytanie trzecie jest filozoficznie najgłębsze: jak długo trzeba czekać, żeby wszystkie N pcheł znalazło się z powrotem na Azorze?

Jest to pytanie o **średni czas powrotu** (*mean recurrence time*) do stanu n w łańcuchu Markowa. Fundamentalne twierdzenie teorii łańcuchów Markowa mówi, że dla ergodycznego łańcucha skończonego średni czas powrotu do stanu n wynosi:

$$T(n) = \frac{1}{P_{\text{eq}}(n)} = \frac{2^N}{\binom{N}{n}}. \quad (2.15)$$

Dla stanu początkowego $n = N$ (wszystkie pchły na Azorze):

$$T(N) = \frac{2^N}{\binom{N}{N}} = 2^N. \quad (2.16)$$

Dla $N = 20000$ daje to $T(N) = 2^{20000} \text{ s} \approx 10^{6000} \text{ lat}$ – czas gigantyczny, nawet w porównaniu z wiekiem Wszechświata (10^{10} lat). Jak ujął to Richard Feynman: „*It is just unlikely. It would never happen in a million years*”.

Kontrast staje się jeszcze wyraźniejszy, gdy zapytamy o czas powrotu do stanu *równowowego* $n = N/2$. Korzystając ze wzoru Stirlinga $k! \approx \sqrt{2\pi k} \left(\frac{k}{e}\right)^k$, szacujemy mianownik:

$$\binom{N}{N/2} = \frac{N!}{(N/2)!(N/2)!} \approx \frac{2^N}{\sqrt{\pi N/2}}, \quad (2.17)$$

skąd:

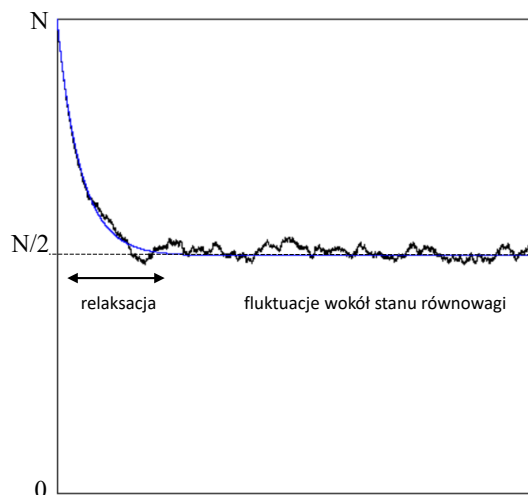
$$T(N/2) = \frac{2^N}{\binom{N}{N/2}} \approx \sqrt{\frac{\pi N}{2}}. \quad (2.18)$$

Dla $N = 20000$ dostajemy $T(N/2) = \sqrt{10000\pi} \text{ s} \approx 177 \text{ s}$ – niecałe trzy minuty.

Wniosek: mikroodwracalność a makronieodwracalność

Zestawienie $T(N/2) \approx \sqrt{\pi N/2}$ kroków i $T(N) = 2^N$ kroków doskonale ilustruje sedno fizyki statystycznej. Jak pisał Marian Smoluchowski: *jeżeli stan początkowy charakteryzuje się długim czasem powrotu, to ewolucja układu będzie miała cechy procesu nieodwracalnego; jednak dla stanów początkowych o krótkim czasie powrotu nie ma sensu mówić o nieodwracalności*. Makroskopowa nieodwracalność nie wynika zatem z żadnego mikroskopowego zakazu – każda reguła ewolucji jest w pełni odwracalna.

Rysunek 3 schematycznie pokazuje pełną historię układu startującego z $n(0) = N$.



Rysunek 3: Ewolucja liczby pcheł na Azorze w czasie. W fazie relaksacji $\langle n(t) \rangle$ spada wykładniczo do $N/2$; film tej ewolucji wygląda wyraźnie inaczej od odwróconego – układ jest makroskopowo nieodwracalny. W stanie równowagowym n_A fluktuuje losowo wokół $N/2$ z amplitudą $\sim \sqrt{N}/2$; tutaj film odwrócony jest nieodróżnialny od oryginalnego.

Rysunek ten unaocznia kluczowe rozróżnienie wprowadzone przez Smoluchowskiego: *nieodwracalność jest cechą fazy relaksacji, nie stanu równowagowego*. Gdy układ już osiągnie stan równowagowy i oscyluje wokół $N/2$, nie ma żadnego sposobu, żeby patrząc na trajektorię stwierdzić, “w którą stronę” biegnie czas.

Psy a gaz

Warto jeszcze zauważyć, że otrzymany rozkład równowagowy

$$P_{\text{eq}}(n) = \binom{N}{n} \frac{1}{2^N} \quad (2.19)$$

powinien przypominać nam dobrze znany wynik z innego kontekstu. Na ćwiczeniach rozważaliśmy N cząstek gazu zamkniętych w pudle o objętości V , w którym wydzieliliśmy podobieństwo V_0 . Prawdopodobieństwo znalezienia dokładnie n cząstek w lewej połowie wynosi:

$$P(n) = \binom{N}{n} \left(\frac{V_0}{V}\right)^n \left(1 - \frac{V_0}{V}\right)^{N-n}. \quad (2.20)$$

Jest to rozkład dwumianowy z parametrem $p = V_0/V$. Podstawiając $V_0 = V/2$, czyli $p = 1/2$, dostajemy dokładnie nasz rozkład psów Ehrenfestów. To nie jest przypadek – w obu przypadkach mamy N niezależnych, rozróżnialnych obiektów, z których każdy z jednakowym prawdopodobieństwem $1/2$ należy do jednego z dwóch zbiorów.

Rozkład dwumianowy z $p = 1/2$ ma wartość oczekiwaną $\langle n \rangle = N/2$ i wariancję $\sigma_n^2 = N/4$. Kluczową wielkością opisującą względną szerokość fluktuacji jest **dyspersja względna**:

$$\frac{\sigma_n}{\langle n \rangle} = \frac{\sqrt{N/4}}{N/2} = \frac{1}{\sqrt{N}}. \quad (2.21)$$

Fluktuacje względne maleją jak $1/\sqrt{N}$ – zgodnie ze słynnym **prawem pierwiastka**, które jest fundamentalnym wynikiem dla układów złożonych z wielu niezależnych składników. Dla makroskopowej liczby cząstek $N \sim 10^{23}$ względne odchylenie od równowagi wynosi $\sim 10^{-12}$, co jest zupełnie nieobserwowalne. Tak mocnego zapchlenia się (na szczęście) nie obserwuje, więc w przypadku psów Ehrenfestów fluktuacje są trochę większe – $N = 100$ pcheł typowa fluktuacja wynosi 10%, dla $N = 10000$ – zaledwie 1%.

3 Entropia jako miara informacji

„Bawimy się
Bawimy ładnie lub nie
Nieważne
Bo nico niełatwo rozdrażnić
Ono tam jest
A imię jego na E
Paradne
Coś jak Etiopia lecz ładniej”

— Lech Janerka, *Dobranoc*

Stan makroskopowy układu fizycznego może być zrealizowany przez gigantyczną liczbę mikrostanów. Makroskopowo wiemy o układzie wszystko, ale mikroskopowo nie mamy pojęcia, w jakiej dokładnie konfiguracji w danej chwili się on znajduje. Sięgnijmy więc teraz do teorii informacji, aby nauczyć się mierzyć tę niewiedzę.

Jak precyzyjnie zmierzyć ilość brakującej informacji? Wyobraźmy sobie prostą grę: ukryliśmy kulkę w jednym z $N = 8$ identycznych pudełek. Ile optymalnie zadanych pytań typu „tak/nie” (wyszukiwanie binarne) potrzeba, by ją znaleźć? Naiwną strategią byłoby pytanie o każde pudełko po kolei („Czy jest w pudełku nr 1?”, „A może nr 2?”). W najgorszym scenariuszu, aby zyskać pewność, musimy wtedy zadać $N - 1$ pytań. Inteligentniejsza strategia

1. „Czy kulka jest w pudełkach 1-4?” (zostają 4 możliwości)
2. „Czy jest w pudełkach 1-2?” (zostają 2 możliwości)
3. „Czy to pudełko nr 1?” (zostaje 1 możliwość – znamy odpowiedź).

Dla 8 pudełek potrzebujemy dokładnie 3 pytań, co wynika z prostej relacji $2^3 = 8$. Z każdym pytaniem dzielimy przestrzeń stanów przez dwa. Uogólniając to na dowolną liczbę N równoprawdopodobnych stanów, miara informacji I potrzebnej do usunięcia naszej niewiedzy wynosi:

$$I = k \log N \quad (3.1)$$

gdzie k to stała dobierana w zależności od wybranej jednostki (wybór logarytmu przy podstawie 2 i $k = 1$ daje informację w **bitach**; w fizyce używa się stałej Boltzmanna k_B i logarytmu naturalnego).

Dlaczego logarytm?

Wybór logarytmu nie jest przypadkowy. Gwarantuje on **addytywność** informacji dla układów niezależnych. Jeśli szukamy kulki w jednym z N pudełek oraz niezależnie monety w jednym z

M pudełek, łączna liczba kombinacji to $N \cdot M$. Oczekujemy, że całkowita informacja potrzebna do ich zlokalizowania będzie po prostu sumą informacji cząstkowych: $I_{N \cdot M} = I_N + I_M$. Tę kluczową własność posiada właśnie funkcja logarytmiczna: $\log(N \cdot M) = \log N + \log M$.

3.1 Szpieg i kolorowe pudełka

Uogólnijmy teraz tę sytuację na nierówne prawdopodobieństwa. Załóżmy, że nasze 8 pudełek jest pokolorowane: 4 są czarne, 3 białe, a 1 jest zielone. Kulka nadal ląduje w jednym z nich całkowicie losowo.

Wysyłamy na zwiady szpiega. Szpiegowi nie udaje się przeniknąć do pomieszczenia z samymi pudełkami, ale udaje mu się ustalić, że kulka znajduje się w czarnym pudełku”. Szpieg przesyła nam ten komunikat. Pytanie brzmi: ile informacji przekazał nam szpieg w swojej wiadomości?

Przed otrzymaniem tego meldunku mieliśmy 8 równoprawdopodobnych możliwości – aby wskazać kulkę, brakowało nam $k \log 8$ informacji. Po usłyszeniu hasła „czarne”, nasz zbiór możliwości zawęził się do 4 pudełek – nadal brakuje nam $k \log 4$ pytań (informacji) do pełnego sukcesu. Różnica tych wielkości to informacja, którą *zyskaliśmy* dzięki działaniom szpiega:

$$I_{\text{czarne}} = k \log 8 - k \log 4 = k \log \frac{8}{4} = k \log 2 \quad (3.2)$$

Gdyby szpieg wyczytał w depeszy słowo „białe”, zyskalibyśmy $I_{\text{białe}} = k \log 8 - k \log 3 = k \log \frac{8}{3}$.

Ogólnie, jeśli pudełek koloru i jest dokładnie N_i spośród wszystkich N pudełek, informacja zyskana po komunikacie o kolorze i wynosi $k \log(N/N_i)$. Ponieważ prawdopodobieństwo trafienia kulki do pudełka koloru i przed rozpoczęciem poszukiwań wynosiło po prostu $p_i = N_i/N$, otrzymujemy fundamentalny wzór na ilość informacji niesioną przez zdarzenie:

$$I_i = -k \log p_i \quad (3.3)$$

Im rzadsze (mniej prawdopodobne) zdarzenie, z tym większym zaskoczeniem je przyjmujemy i tym więcej informacji dostarcza nam komunikat o jego zajściu.

3.2 Szpieg na horyzoncie, czyli entropia Shannona

Wyobraźmy sobie teraz nieco inną sytuację. Widzimy naszego szpiega majaczącego na horyzoncie. Biegnie w naszą stronę z zaklejoną kopertą zawierającą wykradziony meldunek o kolorze. Jeszcze do nas nie dotarł i nie mamy pojęcia, jaki kolor wymieni w raporcie. Możemy jednak zapytać: jaka jest wartość oczekiwana informacji, którą za chwilę od niego dostaniemy?

Z prawdopodobieństwem p_i szpieg przyniesie nam wiadomość o kolorze i , co dostarczy nam I_i informacji. Średnia (oczekiwana) ilość informacji, jaką zyskamy po otwarciu koperty, to uśredniona po wszystkich możliwościach wartość I_i :

$$S = \langle I \rangle = \sum_i p_i I_i \quad (3.4)$$

Podstawiając nasz wzór na I_i , dostajemy słynne równanie:

$$S = -k \sum_{i=1}^M p_i \log p_i \quad (3.5)$$

To właśnie jest **entropia informacyjna** wprowadzona przez Claude’a Shannona w 1948 roku jako fundament teorii informacji. Zauważmy, że wielkość ta nie mówi, ile informacji *dostaliśmy* po otwarciu koperty (to mierzy I_i). Entropia mierzy, ile informacji *spodziewamy się dostać*, zanim kopertę w ogóle otworzymy. Jest to zatem miara naszej początkowej niewiedzy.

Dlaczego „entropia”?

Kiedy w 1948 roku Claude Shannon sformułował matematyczne podstawy teorii informacji, zastanawiał się, jak nazwać odkrytą przez siebie miarę ($S = -k \sum p_i \log p_i$). Wahał się między terminami takimi jak „informacja” czy „niepewność”. Z prośbą o radę zwrócił się do Johna von Neumanna, którego spotkał w instytucjonalnej kantynie. Von Neumann po bez wahania udzielił mu następującej rady:

„Powinieneś nazwać to entropią, z dwóch powodów. Po pierwsze, twoja funkcja niepewności była już wcześniej używana w mechanice statystycznej pod tą właśnie nazwą, więc ma ugruntowaną pozycję. Po drugie, i co znacznie ważniejsze, nikt tak naprawdę nie wie, czym w ogóle jest entropia, więc w każdej dyskusji będziesz miał przewagę.”

3.3 Własności entropii informacyjnej

Kształt rozkładu prawdopodobieństwa jednoznacznie determinuje wartość entropii. Przeanalizujmy dwa skrajne przypadki:

- **Rozkład równomierny (maksymalna niewiedza):** Gdy wszystkie z M możliwych zdarzeń są jednakowo prawdopodobne ($p_i = 1/M$), wzór redukuje się do:

$$S = -k \sum_{i=1}^M \frac{1}{M} \log \frac{1}{M} = -k \cdot M \cdot \frac{1}{M} (-\log M) = k \log M \quad (3.6)$$

Otrzymujemy w ten sposób maksymalną możliwą wartość entropii. Ponieważ nie wiemy absolutnie nic o tym, które zdarzenie nastąpi, każda wiadomość o wyniku niesie maksimum informacji. Fakt, że to właśnie rozkład płaski maksymalizuje entropię przy zadanym warunku normalizacji ($\sum p_i = 1$), można udowodnić stosując metodę mnożników Lagrange’a.

- **Stan pewny (pełna informacja):** Gdy z góry wiemy, że na pewno zajdzie konkretne zdarzenie (np. $p_1 = 1$, a pozostałe $p_{i>1} = 0$), entropia znika. Stosując standardową w teorii informacji konwencję graniczną $\lim_{x \rightarrow 0^+} x \log x = 0$, otrzymujemy:

$$S = -k \cdot 1 \cdot \log 1 - \sum_{i>1} k \cdot 0 \cdot \log 0 = 0 \quad (3.7)$$

Entropia wynosi zero. Skoro mamy całkowitą pewność co do wyniku przed jego ogłoszeniem, sama wiadomość nie niesie najmniejszego elementu zaskoczenia i w związku z tym nie dostarcza żadnej nowej informacji.

Entropia jako miara “niewiedzy”

Entropia S mierzy nasze *przeciętne zaskoczenie* wynikiem losowania – albo równoważnie, naszą *niewiedzę* o tym, co nastąpi. Im bardziej płaski rozkład prawdopodobieństwa (im mniej wiemy), tym większa entropia. Im bardziej skupiony rozkład (im więcej wiemy), tym entropia mniejsza, aż do zera w granicy pewności.

Wkrótce zobaczymy, że entropia Boltzmanna znana z termodynamiki, $S = k_B \ln \Omega$, jest szczególnym przypadkiem tego wzoru – dla rozkładu równomiernego po Ω mikroskopowych stanach układu.

4 Entropia Gibbsa psów Ehrenfesta

„Entropy, it’s a matter of course

Entropy, energy at all levels

Entropy, from it you cannot divorce.”

— Bad Religion, *Entropy*

4.1 Dwa poziomy opisu i dwie entropie

W modelu psów Ehrenfestów mamy do czynienia z dwiema różnymi warstwami opisu probabilistycznego, które warto wyraźnie rozróżnić.

Poziom mikroskopowy. Każda pchła ma swoją tożsamość – możemy je ponumerować. Mikrostan układu to konkretne przypisanie każdej pchły do jednego z psów. Oznaczamy przez $P_{\text{mic}}(t)$ prawdopodobieństwo, że układ w chwili t jest w danym mikro stanie.

Poziom makroskopowy. Obserwujemy jedynie, ile pcheł siedzi na Azorze – nie wiemy, które to są. Makrostan opisuje liczba $m \in \{0, 1, \dots, N\}$. Oznaczamy przez $P_m(t)$ prawdopodobieństwo makrostanu m w chwili t . Kluczową rolę gra liczba mikrostanów odpowiadających makrostanowi m :

$$G_m = \binom{N}{m}, \quad (4.1)$$

gdyż tyle jest sposobów wyboru m pcheł spośród N , które mają siedzieć na Azorze.

Zakładamy, że wszystkie mikro stany w danym makro stanie są jednakowo prawdopodobne (zasada równouprawnienia a priori). Prawdopodobieństwo konkretnego mikrostanu należącego do makrostanu m wynosi zatem:

$$p_{i \in m} = \Pi_m \equiv \frac{P_m(t)}{G_m}. \quad (4.2)$$

Josiah Willard Gibbs (1839–1903)

Jeden z trojga ojców mechaniki statystycznej, obok Maxwella i Boltzmanna. W swoim monumentalnym dziele *Elementary Principles in Statistical Mechanics* (1902) jako pierwszy sformułował mechanikę statystyczną w języku ansamblów i rozkładów prawdopodobieństwa – a nie pojedynczych trajektorii. Entropia nosząca jego imię jest własnością całego rozkładu, nie pojedynczego stanu.

4.2 Entropia Gibbsa

Entropię Gibbsa definiujemy jako entropię informacyjną Shannona rozkładu po *mikro stanach*:

$$S_G(t) = -k \sum_i p_i \log p_i, \quad (4.3)$$

gdzie suma biegnie po wszystkich 2^N mikro stanach. Wyrażamy ją przez zmienne makroskopowe. Każdy makrostan m zawiera G_m mikrostanów, każdy z prawdopodobieństwem $\Pi_m = P_m/G_m$:

$$S_G(t) = -k \sum_{m=0}^N G_m \cdot \Pi_m \log \Pi_m = -k \sum_{m=0}^N G_m \cdot \frac{P_m}{G_m} \log \frac{P_m}{G_m} \quad (4.4)$$

czyli:

$$S_G(t) = -k \sum_{m=0}^N P_m \log P_m + k \sum_{m=0}^N P_m \log G_m. \quad (4.5)$$

Pierwszy składnik to entropia rozkładu po *makro stanach* – miara naszej niewiedzy o tym, w którym makro stanie jest układ. Drugi składnik to średnia entropia Boltzmanna – średnia (po makro stanach) liczba $k \log G_m$ mikrostanów odpowiadających danemu makrostanowi.

4.3 Entropia Boltzmanna

Entropia Boltzmanna przypisuje pojedynczemu makrostanowi m wartość:

$$S_B(m) = k \log G_m = k \log \binom{N}{m}. \quad (4.6)$$

Jest to własność *jednego* makrostanu – nie całego rozkładu. Obserwujemy jeden makrostan, wyobrażamy sobie G_m równoprawdopodobnych mikrostanów, które go realizują, i obliczamy S_B .

Obliczmy S_B dla stanu równowagi. Wiemy, że:

$$n = \langle n \rangle|_{t \rightarrow \infty} = \frac{N}{2}, \quad (4.7)$$

skąd:

$$S_B^{\text{eq}} = k \log \binom{N}{N/2} = k \log \frac{N!}{(N/2)!(N/2)!}. \quad (4.8)$$

Dla dużych N stosujemy wzór Stirlinga $\log N! \approx N \log N - N$:

$$S_B^{\text{eq}} \approx k \left(N \log N - N - 2 \cdot \frac{N}{2} \log \frac{N}{2} + 2 \cdot \frac{N}{2} \right) = k N \log 2. \quad (4.9)$$

4.4 Entropia Gibbsa rośnie monotonicznie – dowód

Pokażemy, że $S_G(t+1) \geq S_G(t)$, co jest odpowiednikiem II zasady termodynamiki w naszym modelu.

Równanie master. Ewolucja rozkładu $P_m(t)$ opisana jest równaniem:

$$P_m(t+1) = \frac{m+1}{N} P_{m+1}(t) + \left(1 - \frac{m-1}{N}\right) P_{m-1}(t). \quad (4.10)$$

Przepiszemy je w języku $\Pi_m = P_m/G_m$. Korzystając z tożsamości:

$$\frac{m+1}{N} \cdot \frac{G_{m+1}}{G_m} = \frac{m+1}{N} \cdot \frac{N-m}{m+1} = \frac{N-m}{N}, \quad (4.11)$$

oraz

$$\left(1 - \frac{m-1}{N}\right) \cdot \frac{G_{m-1}}{G_m} = \frac{N-m+1}{N} \cdot \frac{m}{N-m+1} = \frac{m}{N}, \quad (4.12)$$

otrzymujemy:

$$\Pi_m(t+1) = \frac{m}{N} \Pi_{m-1}(t) + \left(1 - \frac{m}{N}\right) \Pi_{m+1}(t). \quad (4.13)$$

Jest to ważona średnia dwóch wartości Π z sąsiednich stanów, ze współczynnikami $\alpha_1 = m/N$ i $\alpha_2 = 1 - m/N$ spełniającymi $\alpha_1 + \alpha_2 = 1$.

Funkcja $\varphi(x) = x \log x$ jest wypukła (bo $\varphi''(x) = 1/x > 0$ dla $x > 0$). Sieczna funkcji wypukłej leży powyżej jej wykresu, a zatem

$$\varphi(\alpha_1 x_1 + \alpha_2 x_2) \leq \alpha_1 \varphi(x_1) + \alpha_2 \varphi(x_2), \quad \alpha_i \geq 0, \quad \sum_i \alpha_i = 1. \quad (4.14)$$

Stosując to do równania (4.13):

$$\Pi_m(t+1) \log \Pi_m(t+1) \leq \frac{m}{N} \Pi_{m-1}(t) \log \Pi_{m-1}(t) + \left(1 - \frac{m}{N}\right) \Pi_{m+1}(t) \log \Pi_{m+1}(t). \quad (4.15)$$

Mnożymy przez G_m i sumujemy po m od 0 do N :

$$\sum_{m=0}^N G_m \Pi_m(t+1) \log \Pi_m(t+1) \leq \sum_{m=0}^N G_m \left[\frac{m}{N} \Pi_{m-1}(t) \log \Pi_{m-1}(t) + \left(1 - \frac{m}{N}\right) \Pi_{m+1}(t) \log \Pi_{m+1}(t) \right]. \quad (4.16)$$

Rozbijmy prawą stronę na dwie sumy i przesuniemy w nich indeksy. W pierwszej sumie wprowadzamy nową zmienną $n = m - 1$ (co wymaga odrzucenia znikającego wyrazu dla $m = 0$), a w drugiej $n = m + 1$ (odrzucając znikający wyraz dla $m = N$):

$$\text{RHS} = \sum_{n=0}^{N-1} \left(G_{n+1} \frac{n+1}{N} \right) \Pi_n \log \Pi_n + \sum_{n=1}^N \left(G_{n-1} \frac{N-n+1}{N} \right) \Pi_n \log \Pi_n. \quad (4.17)$$

Korzystając z tożsamości dla współczynników dwumianowych, $G_{n+1} \frac{n+1}{N} = G_n \left(1 - \frac{n}{N}\right)$ oraz $G_{n-1} \frac{N-n+1}{N} = G_n \frac{n}{N}$, możemy z powrotem sprowadzić obie sumy do granic od 0 do N (gdyż dla skrajnych wyrazów dodawane współczynniki i tak wynoszą zero):

$$\text{RHS} = \sum_{n=0}^N G_n \left(1 - \frac{n}{N}\right) \Pi_n \log \Pi_n + \sum_{n=0}^N G_n \frac{n}{N} \Pi_n \log \Pi_n = \sum_{n=0}^N G_n \Pi_n(t) \log \Pi_n(t). \quad (4.18)$$

Stąd otrzymujemy:

$$-\sum_m G_m \Pi_m(t+1) \log \Pi_m(t+1) \geq -\sum_m G_m \Pi_m(t) \log \Pi_m(t), \quad (4.19)$$

co po podstawieniu definicji S_G daje entropię stale rosnącą (lub stałą w stanie stacjonarnym):

$$\boxed{S_G(t+1) \geq S_G(t)}. \quad (4.20)$$

4.5 Entropia Gibbsa w stanie równowagowym

W stanie równowagi termodynamicznej wszystkie z 2^N możliwych mikrostanów układu są jednakowo prawdopodobne. Prawdopodobieństwo znalezienia układu w dowolnym konkretnym mikrostanie i wynosi zatem $p_i = 1/2^N$.

Podstawiając to bezpośrednio do definicji entropii Gibbsa, otrzymujemy natychmiast:

$$S_G^{\text{eq}} = -k \sum_{i=1}^{2^N} p_i \log p_i = -k \sum_{i=1}^{2^N} \frac{1}{2^N} \log \left(\frac{1}{2^N} \right) = -k \cdot 2^N \cdot \frac{1}{2^N} (-N \log 2) = kN \log 2. \quad (4.21)$$

Otrzymany wynik zgadza się z entropią Boltzmanna S_B^{eq} obliczoną wcześniej. Zwróćmy jednak uwagę na fundamentalną różnicę pojęciową między tą wartością a entropią Boltzmanna. Entropia Gibbsa zależy od rozkładu na *wszystkich* 2^N mikrostanach. Tymczasem entropia Boltzmanna z definicji zlicza mikrostany należące tylko do makrostanu równowagowego ($n = N/2$):

$$S_B^{\text{eq}} = k \log \binom{N}{N/2}. \quad (4.22)$$

Dlaczego zatem w granicy termodynamicznej obie entropie dają ten sam wynik rzędu $kN \log 2$? Wynika to z faktu, że dla bardzo dużych N makrostan równowagowy całkowicie dominuje przestrzeń stanów. Rozkład jest tak niewyobrażalnie ostro wypiętrzony wokół $N/2$, że logarytm z liczby stanów w samym tylko pikcie równowagowym (S_B^{eq}) jest w praktyce równy logarytmowi z liczby wszystkich możliwych stanów w całym układzie (S_G^{eq}).

Boltzmann vs Gibbs: dwie filozofie entropii

Obie entropie przyjmują tę samą wartość w stanie równowagi, ale poza nim różnią się w sposób zasadniczy.

Entropia Boltzmanna $S_B(m) = k \log G_m$ jest własnością *jednego makrostanu*. Obserwujemy konkretny układ w stanie m i wyobrażamy sobie liczbę mikrostanów go realizujących. Cena tej prostoty jest taka, że naruszenie II zasady termodynamiki jest nominalnie możliwe – wskutek losowych fluktuacji układ może tymczasowo trafić do stanu o mniejszej liczbie mikrostanów (jego entropia spadnie). Oczywiście, dla układów makroskopowych są one astronomicznie rzadkie, a układ po chwili i tak powróci do stanu o wyższej entropii.

Entropia Gibbsa S_G jest własnością *całego rozkładu* $\{P_m(t)\}$ – dotyczy zespołu statystycznego (ansamblu) lub naszego stanu wiedzy, a nie pojedynczej realizacji układu. Jej monotoniczność $S_G(t+1) \geq S_G(t)$ staje się twardym prawem matematycznym, ale płacimy za to pewną cenę: entropia ta „nie widzi” chwilowych fluktuacji pojedynczego układu.

5 Mikro- i makrostany: Skąd się bierze nieodwracalność?

*„There’s nothing to it
It’s just a matter of time
Everything put together
Sooner or later falls apart.”*

— Paul Simon, *Everything Put Together Falls Apart*

5.1 Eksperyment myślowy: Dwa filmy

Wyobraźmy sobie dwa filmy wideo przedstawiające rozprężanie się gazu w zamkniętym pudełku. W obu przypadkach cząsteczki startują z lewej połowy naczynia i zaczynają chaotycznie odbijać się od ścianek. Różnica polega wyłącznie na skali – na pierwszym filmie śledzimy zaledwie pięć cząsteczek, podczas gdy na drugim jest ich milion.

Prawdziwa przepaść między tymi dwoma światami ujawnia się dopiero wtedy, gdy spróbujemy odtworzyć oba filmy wstecz. Film z pięcioma cząsteczkami puszczony od tyłu nie wzbudzi naszych podejrzeń i będzie wyglądał zupełnie naturalnie. Cząsteczki odbijają się od siebie, po czym w wyniku czystego przypadku na moment znów gromadzą się w lewej połowie naczynia. Dla tak małych układów ciągle, stosunkowo duże fluktuacje gęstości są na porządku dziennym.

Gdy jednak puścimy wstecz drugi film, natychmiast uznamy go za niefizyczny absurd. Zobaczymy wtedy, jak równomiernie rozproszony gaz nagle, sam z siebie, gromadzi się w lewej połowie naczynia. Oczywiście, z punktu widzenia mikroskopowych praw mechaniki takie zdarzenie nie jest zakazane, jednak prawdopodobieństwo tak gigantycznej fluktuacji jest znikomo małe. To właśnie dlatego w prawdziwym życiu nie musimy się obawiać, że cząsteczki powietrza w sali wykładowej nagle zgromadzą się w jednym kącie, dusząc wszystkich słuchaczy – czas oczekiwania na takie wydarzenie znacznie przekracza wiek Wszechświata.

Ostatecznie, o ile w układzie pięciu cząstek parametry fizyczne (takie jak gęstość) będą nieustannie się zmieniać, o tyle milion cząstek po krótkim czasie tworzy uśrednioną, wypełniającą całą przestrzeń „zupę”, w której wizualnie nie zachodzą już żadne zmiany. Zjawisko to ilustruje kluczowy fakt: tam, gdzie cząsteczek jest dużo, procesy fizyczne nabierają zupełnie nowej, emergentnej cechy – stają się **nieodwracalne**.

5.2 Statystyka rozprężającego się gazu

Aby to zjawisko ubrać w ramy matematyczne, przypomnijmy rozpatrywany na ćwiczeniach przykład gazu złożonego z N nieoddziałujących cząstek w pojemniku o objętości V , w którym

wydzielamy myślowo podobojętość V_0 . Jeśli odczekaliśmy wystarczająco długo, aby gęstość gazu stała się jednorodna, to rozkład prawdopodobieństwa znalezienia dokładnie n cząstek w objętości V_0 jest dany klasycznym rozkładem dwumianowym:

$$P(n) = \binom{N}{n} \left(\frac{V_0}{V}\right)^n \left(1 - \frac{V_0}{V}\right)^{N-n}$$

Średnia liczba cząstek w tym obszarze wynosi po prostu $\langle n \rangle = N \frac{V_0}{V}$.

Najważniejszym wnioskiem z analizy tego rozkładu jest zachowanie fluktuacji. Dla rozkładu dwumianowego wariancja to $Np(1-p)$. Dzieląc ją przez kwadrat średniej i pierwiastkując, otrzymujemy względne odchylenie standardowe. Dla małego wycinka ogromnego układu ($V_0 \ll V$) upraszcza się ono do postaci:

$$\frac{\sigma_n}{\langle n \rangle} \approx \frac{1}{\sqrt{\langle n \rangle}}$$

Jest to słynne **prawo pierwiastkowe**, charakterystyczne dla niezależnych układów statystycznych. Stanowi ono matematyczny dowód na to, dlaczego w skali makro nie doświadczamy dziwnych zjawisk: ze wzrostem układu fluktuacje stają się relatywnie coraz mniejsze.

Dla uproszczenia przyjmijmy teraz, że interesują nas równe połówki naczynia ($V_0/V = 1/2$). JAK już wspominaliśmy, przypadek ten mapuje się na model Ehrenfestów - cząsteczka gazu wybierająca lewą lub prawą połowę naczynia to dokładnie to samo, co pchła wybierająca psa Azora lub Burka. Rozkład prawdopodobieństwa przyjmuje znaną nam już doskonale z poprzedniego podrozdziału postać:

$$P(n) = \frac{1}{2^N} \binom{N}{n}$$

Zamiast więc powtarzać całe rozumowanie od nowa, możemy śmiało przenieść nasze wnioski z modelu pcheł bezpośrednio na grunt termodynamiki gazu.

Wyobraźmy sobie, że obserwujemy ten gaz, interesując się tylko jego zachowaniem w skali makro. Zamiast śledzić położenia i pędy wszystkich molekuł, mierzymy jedynie „liczbę cząstek w prawej części naczynia”.

Wprowadza to fundamentalne dla statystyki rozróżnienie:

- **Stan mikroskopowy (mikrostan) gazu:** Scharakteryzowany przez podanie położení i pędów wszystkich (a więc $\sim 10^{23}$) cząsteczek. Odpowiada on jednemu konkretnemu punktowi w $6N$ -wymiarowej przestrzeni fazowej.
- **Stan makroskopowy (makrostan):** Opisany przez podanie tylko tych wielkości, które mogą być wyznaczone przez pomiary makroskopowe (np. objętość, ciśnienie, liczba cząstek w badanej połowie).

Przy czym, jak widzieliśmy na przykładzie psów i pcheł, **każdemu makrostanowi odpowiada ogromna liczba mikrostanów** (może być on realizowany na bardzo dużo sposobów).

Makrostany są dostępne naszym ludzkim zmysłom. Natomiast pełna informacja o mikrostanach w ogólności nie jest nam dostępna – dysponuje nią tylko hipotetyczny byt zwany demonem Laplace’a. Co ciekawe, jako że równania mikroskopowe Newtona są w pełni odwracalne w czasie, demon ten w ogóle nie dostrzega nieodwracalności! Pełna informacja o stanie świata w jego oczach jest zachowana, bo zawsze może on precyzyjnie odwrócić prędkości wszystkich cząstek i wrócić do stanu początkowego. Według słów samego Laplace’a: „dla takiej istoty nic nie byłoby niepewne a przyszłość tak jak i przeszłość rozpościerałaby się przed jej oczami.”

5.3 Czas powrotu i stan równowagi

Wróćmy do naszego gazu. Wyobraźmy sobie, że kręcimy film, robiąc zdjęcia klatka po klatce, powiedzmy co sekundę. Musimy tu poczynić jedno kluczowe założenie: odstęp czasu między

kolejnymi klatkami jest dłuższy niż tzw. czas korelacji układu. Dzięki temu cząsteczki zdążają się wielokrotnie zderzyć i przetasować, a układ „zapomni” swój poprzedni stan. Dopiero wtedy każda klatka filmu stanowi dla nas nową, statystycznie niezależną próbkę (losowanie) z dostępnej przestrzeni stanów.

Przez większość czasu na filmie będziemy obserwować, że $n_1 \approx n_2 \approx N/2$, choć rzadko będzie to dokładnie połowa. Ponieważ cząsteczki nieustannie przelatują przez środek naczynia, mówimy, że liczba cząstek n_1 fluktuuje. Zazwyczaj fluktuacje te są bardzo małe ($|n_1 - N/2| \ll N/2$).

Nikt nie spodziewa się, że wszystkie cząsteczki gazu nagle zgromadzą się w jednej części naczynia. Oszacujmy prawdopodobieństwo wystąpienia takiej skrajnej fluktuacji (czyli pustki w lewej połowie, $n_1 = 0$). Korzystając z naszego równania (będącego odpowiednikiem ucieczki wszystkich pcheł na jednego psa), wynosi ono $P(0) = \frac{1}{2^N}$.

Skoro każda klatka to niezależne losowanie, to aby uchwycić taką sytuację, musielibyśmy obejrzeć średnio 2^N klatek filmu. Dla makroskopowego układu $N \approx 10^{24}$ zajęłoby to nam:

$$2^{10^{24}} = (2^{10})^{10^{23}} \approx (10^3)^{10^{23}} = 10^{3 \cdot 10^{23}} \text{ s}$$

Byłby to więc raczej długi seans, biorąc pod uwagę, że wiek Wszechświata to zaledwie $T_u \approx 10^{17}$ s. Nawet dla zupełnie mikroskopijnego układu złożonego z $N = 60$ cząstek dostajemy czas oczekiwania 10^{18} s, czyli dziesięć wieków Wszechświata!

Wniosek jest bezlitosny: tak duże odstępstwa od stanu $n_1 \approx n_2$ nie zachodzą w praktyce prawie nigdy.

Zróbmy teraz następujące doświadczenie: za pomocą tłoka zbieramy wszystkie cząstki w lewej części naczynia, po czym tłok usuwamy. Spodziewamy się, że nierównomierność ruchu spowoduje intensywny przepływ na pustą stronę. Jak będzie zachowywać się wartość oczekiwana $\langle n(t) \rangle$ w lewej połowie?

Odpowiedź już znamy - to dokładnie to samo zjawisko, które rozpatrywaliśmy dla psów Ehrenfesta. Usunięcie tłoka to fizyczny odpowiednik sytuacji, w której wszystkie N pcheł umieściliśmy początkowo na jednym psie. Jak już udowodniliśmy w tamtym modelu, wartość oczekiwana $\langle n(t) \rangle$ będzie szybko maleć (stan makroskopowy zmienia się w czasie!), aż zrówna się z drugą połową ($n_1 \approx n_2 \approx N/2$). Wtedy makrostan przestanie się zauważalnie zmieniać w czasie, osiągając **stan równowagi termodynamicznej**.

Podsumowując:

- Stan równowagi to makrostan, który nie zmienia się w czasie (z wyjątkiem przypadkowych mikroskopijnych fluktuacji) – odpowiada on $n = N/2$.
- Dzięki prawu odwrotności pierwiastka ($1/\sqrt{n}$) zauważamy, że to właśnie okolice stanu równowagi są realizowane przez przytłaczającą większość możliwych mikrostanów.

Aby unaocnić sobie, jak niewyobrażalnie duża jest ta „przytłaczająca większość” mikrostanów, spróbujmy oszacować, jaka część układów przebywa blisko stanu równowagi. Gdy N jest bardzo duże, zmienna n staje się w praktyce ciągła. Na podstawie Centralnego Twierdzenia Granicznego wiemy, że nasz dyskretny rozkład dwumianowy dąży do ciągłego rozkładu Gaussa z dyspersją $\sigma = \frac{1}{2}\sqrt{N}$. Przechodzimy w ten sposób do gęstości prawdopodobieństwa:

$$\rho(n) = \sqrt{\frac{2}{\pi N}} \exp\left(-2 \frac{(n - N/2)^2}{N}\right)$$

Całkowanie tej gęstości prawdopodobieństwa wokół niewielkiego otoczenia połowy cząstek, powiedzmy w granicach od $(1 - a)\frac{N}{2}$ do $(1 + a)\frac{N}{2}$ (gdzie a to ułamek określający szerokość naszego okna tolerancji), prowadzi do wyniku:

$$\text{Erf}\left(a\sqrt{\frac{N}{2}}\right)$$

gdzie Erf to funkcja błędu Gaussa, zdefiniowana jako $\text{Erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-z^2} dz$. Z własności matematycznych tej funkcji wiemy, że dla dużych argumentów dąży ona asymptotycznie do jedynki w następujący sposób:

$$\text{Erf}(x) \rightarrow 1 - \frac{e^{-x^2}}{\sqrt{\pi}x}$$

Wstawiając konkretne liczby, widzimy potęgę tego zjawiska:

- dla bardzo małego układu $N = 10^3$ i odchylenia $a = 0.9$, prawdopodobieństwo znalezienia układu w tym szerokim oknie wynosi ok. $1 - 10^{-3}$,
- ale dla wciąż fizycznie mikroskopijnego $N = 10^7$ i bardzo wąskiego okna $a = 0.999$ (dopuszczamy wahania rzędu zaledwie 0,1%), wynosi ono aż $1 - 2 \cdot 10^{-3}$.

Wraz ze wzrostem N , rozkład ten niezwykle gwałtownie zwęża się względem swojej wartości średniej. Oznacza to, że dla makroskopowej liczby cząstek stany równowagowe zajmują o całe rzędy wielkości większy procent przestrzeni fazowej niż stany nierównowagowe. Nic więc dziwnego, że jeśli przyjmiemy, że trajektoria układu błąka się przypadkowo w przestrzeni fazowej (odwiedzając dostępne stany z równym prawdopodobieństwem), to niezwykle szybko wyjdzie z bardzo ciasnego obszaru nierównowagowego i na zawsze „utonie” w gigantycznym morzu stanów równowagi.

5.4 Paradoks Loschmidta

Podsumowując: gaz zgromadzony w rogu rozpręża się w sposób nieodwracalny i przechodzi do stanu równowagi ($N/2$ cząstek w połówkach). Przejście to polega na zmianie z makrostanu realizowanego przez małą liczbę mikrostanów, do makrostanu o astronomicznie dużej ich liczbie. Przykłady zjawisk nieodwracalnych to wszystko to, co nie jest nudne: ludzie, życie (zjawiska odwracalne to np. ruch planet).

Budzi to jednak kontrowersje, ujęte w tzw. **argumencie Loschmidta**: przecież na poziomie mikroskopowym równania Newtona są w pełni odwracalne! Skąd nagle ta asymetria? Jeśli odwrócimy prędkości cząsteczek, to $n(t)$ zacznie znów rosnać do N i gaz zgromadzi się w lewej części naczynia. Otrzymaliśmy zatem coś nieodwracalnego z odwracalnych równań ruchu.

Odpowiedź Boltzmanna na ten zarzut była następująca: „*liczą się nie tylko równania, ale i warunki początkowe*”. Asymetria kryje się właśnie w nich: prawie wszystkie położenia i pędy cząstek stłoczonych w rogu prowadzą po puszczeniu układu do ich rozprężania, podczas gdy tylko bardzo specyficzne, superdokładnie dobrane położenia w całym pudle prowadzą do ponownego sprężania w narożniku.

Paradoks parkowania

Ten argument Boltzmanna dobrze ilustruje tzw. **paradoks parkowania**. Bardzo często, aby zaparkować samochód, zmuszeni jesteśmy ustawić go w wąskiej luce między dwoma innymi autami. Jest to manewr trudny – znacznie trudniejszy niż manewr odwrotny (wyjechanie). A przecież kinematycznie samochód może jechać po tej samej trasie zarówno do przodu, jak i do tyłu, więc skąd ta trudność? Aby to zrozumieć, ponownie wyróżnijmy dwa poziomy opisu:

- **Opis mikroskopowy:** podajemy dokładnie położenie samochodu na jezdni („mikrostan”).
- **Opis makroskopowy:** wyróżniamy tylko dwie sytuacje („makrostany”): (1) samochód zaparkowany i (2) samochód poza miejscem parkowania.

Stan (1) może być zrealizowany przez bardzo niewiele mikrostanów samochodu. Tymczasem stan (2) na ulicy odpowiada bardzo dużej liczbie położzeń (mikrostanów). Z tego powodu przejście z (1) do (2) jest łatwe – wielu możliwym realizacjom odpowiada sukces. Proces odwrotny jest trudniejszy w realizacji: startując z jezdni, musimy precyzyjnie trafić do makrostanu (1),

któremu odpowiada bardzo mało ułożeń samochodu w luce parkingowej.

5.5 Granica termodynamiczna i ekstensywność entropii

Aby spojrzeć na problem gazu z nieco bardziej wyrafinowanej, statystycznej perspektywy, porzucimy na chwilę zgrubny podział na lewą i prawą stronę naczynia. Zamiast tego podzielmy nasze pudło na V identycznych, mikroskopijnych komórek. Założymy przy tym, że są one na tyle małe, iż w każdej z nich może znajdować się maksymalnie jedna cząstka.

W tak skonstruowanym modelu określenie mikrostanu układu sprowadza się do precyzyjnego wskazania, które dokładnie komórki są w danej chwili zajęte. Makrostanem z kolei będzie po prostu całkowita liczba cząstek N uwięzionych w całym pudle. Skoro z dostępnej puli V komórek musimy wybrać te N , w których umieścimy nasze cząstki, całkowita liczba mikrostanów Σ realizujących ten konkretny makrostan jest dana klasycznym symbolem Newtona:

$$\Sigma = \binom{V}{N} = \frac{V!}{N!(V-N)!}$$

Fizyka makroskopowa rzadko zajmuje się jednak układami, w których możemy ręcznie policzyć komórki i molekuly. Aby nasze rozważania miały sens dla rzeczywistych, obserwowalnych gazów, musimy dokonać matematycznego przejścia do tzw. **granicy termodynamicznej**. Formalnie zapisujemy ją jako:

$$N \rightarrow \infty, \quad V \rightarrow \infty, \quad \rho_N = \frac{N}{V} = \text{const.}$$

Fizyczny sens tej granicy jest bardzo intuicyjny: wyobrażamy sobie, że rozmiary naszego układu (objętość) oraz całkowita ilość materii dążą do nieskończoności, ale dzieje się to w sposób w pełni proporcjonalny, dzięki czemu lokalna gęstość układu ρ_N pozostaje przez cały czas stała.

Aby sprawdzić, jak w tej granicy zachowuje się ogromna liczba mikrostanów, najwygodniej jest badać jej logarytm. Dla tak gigantycznych liczb możemy bez obaw skorzystać z przybliżenia Stirlinga ($\log N! \approx N \log N - N$). Aplikując to rozwinięcie do naszej silni i porządkując logarytmy z iloczynów, otrzymujemy:

$$\begin{aligned} \log \Sigma &= V \log V - V - (N \log N - N) - [(V - N) \log(V - N) - (V - N)] \\ &= V [\log V - \rho_N \log(V \rho_N) - (1 - \rho_N) \log(V(1 - \rho_N))] \\ &= V [\log V - \rho_N \log V - \rho_N \log \rho_N - (1 - \rho_N) \log V - (1 - \rho_N) \log(1 - \rho_N)] \end{aligned}$$

Choć ostateczne równanie wygląda dość paskudnie, to jednak łatwo zauważamy, że wszystkie człony zawierające $\log V$ wewnątrz nawiasu kwadratowego sumują się do zera, ponieważ $1 - \rho_N - 1 + \rho_N = 0$. Pozbywając się ich, otrzymujemy zgrabny wynik:

$$\log \Sigma = -V (\rho_N \log \rho_N + (1 - \rho_N) \log(1 - \rho_N))$$

Z matematycznego punktu widzenia możemy ten rezultat (pomijając zaniedbywalne w granicy termodynamicznej reszty wyższego rzędu, tzw. $o(V)$) zapisać w zwartej postaci:

$$\log \Sigma \approx V s$$

gdzie wprowadziliśmy nową zmienną s , zdefiniowaną jako:

$$s = -(\rho_N \log \rho_N + (1 - \rho_N) \log(1 - \rho_N)) \quad (5.1)$$

Co ważne, wykazaliśmy właśnie, że wielkość $\log \Sigma$ jest **parametrem ekstensywnym** – zależy ona wprost proporcjonalnie od całkowitej objętości układu V . Oznacza to, że jeśli weźmiemy dwa

identyczne pudła z gazem i połączymy je w jedno, wartość $\log \Sigma$ wzrośnie dokładnie dwukrotnie (zachowuje się więc dokładnie tak samo jak całkowita masa układu czy jego energia wewnętrzna). Z kolei zdefiniowane przez nas s jest **parametrem intensywnym**, który zależy wyłącznie od ułamkowej gęstości układu, a więc nie ulega zmianie przy skalowaniu rozmiarów pudełka.

Przyjrzyjmy się jednak bliżej samej postaci matematycznej funkcji s :

$$s = -(\rho_N \log \rho_N + (1 - \rho_N) \log(1 - \rho_N))$$

Dla uważnego czytelnika ten zapis powinien wydać się znajomy - to nic innego jak entropia Shannona dla najprostszego rozkładu dwupunktowego. Faktycznie: $\rho_N = N/V$ to po prostu ułamek zajętych komórek, a więc prawdopodobieństwo p , że losowo wybrana komórka z naszej siatki jest pełna. Z kolei $(1 - \rho_N)$ to prawdopodobieństwo zdarzenia przeciwnego - że komórka pozostaje pusta $(1 - p)$. Entropia informacyjna przypadająca na jedną komórkę to zatem $s = -[p \log p + (1 - p) \log(1 - p)]$, czyli dokładnie wzór (5.1).

Skoro $\log \Sigma = V \cdot s$, dochodzimy do wniosku, że logarytm z liczby mikrostanów jest całkowitą entropią informacyjną naszego układu makroskopowego (miarą naszej niewiedzy o tym, w którym dokładnie mikro stanie znajduje się układ). To zderzenie statystyki z teorią informacji prowadzi nas wprost do serca fizyki. Z dokładnością do stałej proporcjonalności (stałej Boltzmanna k_B , która przelicza entropię na dzule na kelwin), wielkość ta definiuje **entropię termodynamiczną**:

$$S = k_B \log \Sigma$$

W ten oto sposób abstrakcyjna miara braku informacji o ułożeniu cząsteczek stała się mierzalną w laboratorium wielkością fizyczną, a zjawisko makroskopowej nieodwracalności zyskało twarde, matematyczne fundamenty.

5.6 Cząstki w komórkach

Rozważmy teraz układ o objętości V przedzielony przegrodą na dwa równe obszary. W lewej części wiemy N_L cząstek, co wymusza, by w prawej znajdowało się ich $N_P = N - N_L$. Całkowita liczba mikrostanów takiego układu jest iloczynem liczb stanów dostępnych dla każdego z podukładów:

$$\Sigma_L = \binom{V/2}{N_L} \binom{V/2}{N - N_L} \quad (5.2)$$

Logarytmując powyższe wyrażenie i przechodząc do granicy termodynamicznej, otrzymujemy sumę wkładów entropowych obu części:

$$\log \Sigma_L = -\frac{V}{2} [\rho_L \log \rho_L + (1 - \rho_L) \log(1 - \rho_L) + \rho_P \log \rho_P + (1 - \rho_P) \log(1 - \rho_P)] \quad (5.3)$$

gdzie lokalne gęstości zdefiniowane są jako $\rho_L = \frac{N_L}{V/2}$ oraz $\rho_P = \frac{N - N_L}{V/2} = 2\rho - \rho_L$ (przy czym $\rho = N/V$ to średnia gęstość całego układu). Zauważmy, że $\frac{\partial \rho_P}{\partial \rho_L} = -1$.

Kluczowe pytanie brzmi: który makro stan (określony przez gęstość ρ_L) jest realizowany przez największą liczbę mikrostanów? Aby to rozstrzygnąć, wyznaczamy pochodną $\log \Sigma_L$ względem ρ_L :

$$\frac{d \log \Sigma_L}{d \rho_L} = -\frac{V}{2} \left[\log \left(\frac{\rho_L}{1 - \rho_L} \right) - \log \left(\frac{\rho_P}{1 - \rho_P} \right) \right] = \frac{V}{2} \log \left(\frac{(1 - \rho_L)(2\rho - \rho_L)}{\rho_L(1 - 2\rho + \rho_L)} \right) \quad (5.4)$$

Analiza znaku pochodnej prowadzi do następujących wniosków:

- Dla $\rho_L < \rho$ pochodna jest dodatnia (liczba mikrostanów rośnie wraz ze wzrostem ρ_L).
- Dla $\rho_L > \rho$ pochodna jest ujemna.
- Dla $\rho_L = \rho$ pochodna wynosi zero, co oznacza osiągnięcie ekstremum (maksimum).

Jak jednak wiemy z wcześniejszych rozważań, samo istnienie matematycznego maksimum to nie wszystko. Kluczowe jest ekstremalne wypikowanie tego rozkładu w granicy termodynamicznej (zapewniane tu przez gigantyczny czynnik $V/2$ przed całym nawiasem). Liczba mikrostanów w nieskończenie bliskim sąsiedztwie podziału gęstości 50%/50% jest tak przytłaczająco większa od reszty, że sumarycznie pochłania ona niemal całe prawdopodobieństwo. Z tego powodu makroskopowe odchylenia od stanu jednorodnego stają się w praktyce niemożliwe do zaobserwowania. Wynika z tego, że po usunięciu więzów (przegrody), układ spontanicznie dąży do stanu o jednorodnym rozkładzie gęstości, ponieważ to on odpowiada największej liczbie konfiguracji mikroskopowych.

6 Dążenie do równowagi

*„I tried so hard and got so far
But in the end, it doesn't even matter
I had to fall to lose it all
But in the end, it doesn't even matter.”*
— Linkin Park, *In the End*

Analizowaliśmy już przypadki, w których **układ izolowany** przechodził ze stanu nierównowagowego do równowagowego, jak przy swobodnym rozprężaniu gazu czy w modelu psów Ehrenfestów. W opisie takich procesów wygodnie jest posługiwać się pojęciem **zespołu statystycznego** (ansamblu). Zamiast śledzić ewolucję jednego konkretnego układu, wyobrażamy sobie bardzo dużą liczbę jego makroskopowo identycznych kopii i badamy rozkład prawdopodobieństwa mikrostanów, to znaczy sprawdzamy, jaka część układów w zespole realizuje dany mikrostan w danej chwili.

Startując z makrostanu nierównowagowego, który realizowany jest przez znacznie mniejszą liczbę mikrostanów niż stan równowagi, i usuwając więzy ograniczające układ, inicjujemy jego ewolucję. W modelu Ehrenfestów takim więzem było początkowe przygotowanie układu: wszystkie pchły siedziały na Azorze i nie mogły przeskakiwać na Burka. W przykładzie gazu analogiczną rolę pełniła przegroda, która zmuszała wszystkie cząstki do pozostawiania w jednej połowie naczynia. Po usunięciu tych więzów układ uzyskuje dostęp do znacznie większej liczby mikrostanów.

W stanie nierównowagowym rozkład prawdopodobieństwa po mikrostanach nie jest jednostajny: niektóre mikrostan są uprzywilejowane przez warunek początkowy, inne zaś mają znacznie mniejsze prawdopodobieństwo. Po usunięciu więzów ansambl zaczyna stopniowo rozlewać się po coraz większej części dostępnej przestrzeni stanów. W modelu Ehrenfestów można to prześledzić jawnie: jeśli n oznacza liczbę pcheł siedzących na Azorze, to liczba mikrostanów realizujących taki makrostan wynosi

$$\Omega(n) = \binom{N}{n},$$

a maksimum tej funkcji wypada dla $n \approx N/2$. Podobnie dla gazu w naczyniu podzielonym myślowo na dwie równe części, jeśli N_L oznacza liczbę cząstek w lewej połowie, to liczba realizujących taki makrostan mikrostanów jest proporcjonalna do

$$\Omega(N_L) = \binom{N}{N_L},$$

a maksimum znów wypada przy $N_L \approx N/2$.

Widzimy więc, że po usunięciu więzów układ nie zmierza ku jakiemuś wyróżnionemu mikrostanowi, lecz ku takim wartościom parametrów makroskopowych, którym odpowiada największa

liczba realizacji mikroskopowych. Jest to boltzmannowski sens dochodzenia do równowagi: układ izolowany przechodzi samorzutnie ku coraz bardziej typowym makrostanom.

Aby jednak przypisać temu obrazowi precyzyjny sens probabilistyczny, potrzebujemy jeszcze założenia dotyczącego samego stanu równowagi. W modelu Ehrenfestów można pokazać ściśle, że asymptotyczny rozkład równowagowy odpowiada sytuacji, w której wszystkie dopuszczalne mikrostan są jednakowo prawdopodobne. W ogólnej teorii statystycznej przyjmujemy tę własność jako **postulat równych prawdopodobieństw *a priori***:

W stanie równowagi układu izolowanego wszystkie dopuszczalne mikrostan, zgodne z nałożonymi więzami makroskopowymi, są jednakowo prawdopodobne.

Dla układu o ustalonych E , N i V zapisujemy to w postaci

$$P_i = \frac{1}{\Sigma(E, N, V)},$$

gdzie $\Sigma(E, N, V)$ oznacza liczbę wszystkich dopuszczalnych mikrotanów.

Celem fizyki statystycznej jest przewidywanie wartości wielkości makroskopowych na podstawie opisu mikroskopowego. Jeśli wielkość A przyjmuje w mikro stanie i wartość A_i , to jej wartość średnia dana jest przez

$$\langle A \rangle = \sum_i P_i A_i,$$

a dla zmiennych ciągłych przez odpowiednią całkę. Dlatego kluczowym zadaniem fizyki statystycznej jest wyznaczenie prawdopodobieństw P_i mikrotanów.

Postulat równych prawdopodobieństw *a priori* wiąże teraz świat mikro ze światem makro. Jeżeli wszystkie dopuszczalne mikrostan są równoprawdopodobne, to prawdopodobieństwo makrostanu jest proporcjonalne do liczby mikrotanów, które go realizują. Dlatego najbardziej prawdopodobny jest makrostan realizowany przez największą liczbę mikrotanów. To właśnie ten makrostan identyfikujemy ze stanem równowagi.

6.1 Fizyka statystyczna a mechanika klasyczna

W poprzednim rozdziale przedstawiliśmy zgrabny obraz dochodzenia układu do stanu równowagi: układ zapomina o swoim początkowym rozkładzie, entropia rośnie, aż w końcu osiąga maksimum, a w stanie równowagi wszystkie mikrostan stają się równoprawdopodobne. Wielkości makroskopowe stają się wtedy łatwe do obliczenia, bo są po prostu średnimi po takim jednostajnym rozkładzie. Jest to obraz pociągający, ale z punktu widzenia mechaniki klasycznej kompletnie fałszywy!

Dlaczego? Po pierwsze, skoro równania Hamiltona są odwracalne w czasie, to dlaczego entropia miałaby samorzutnie rosnać? Istotnie, jak zaraz zobaczymy, w mechanice klasycznej entropia Gibbsa nie rośnie, lecz pozostaje stała. Po drugie, wskutek zachowania objętości w przestrzeni fazowej rozkład nie może po prostu rozplynać się po całej dostępnej przestrzeni stanów. Obraz samorzutnego dochodzenia do równowagi wymaga więc subtelniejszej interpretacji.

Zacznijmy od ewolucji entropii Gibbsa. Niech Γ oznacza punkt w $6N$ -wymiarowej przestrzeni fazowej, a $f(\Gamma, t)$ będzie znormalizowaną gęstością prawdopodobieństwa w tej przestrzeni, taką że $\int f(\Gamma, t) d\Gamma = 1$. Wielkość $f(\Gamma, t) d\Gamma$ oznacza więc prawdopodobieństwo znalezienia układu w chwili t w elemencie objętości $d\Gamma$ wokół punktu Γ . Przez Φ^t oznaczmy przepływ Hamiltona, tzn. odwzorowanie przypisujące punktowi początkowemu Γ_0 punkt

$$\Gamma = \Phi^t(\Gamma_0), \quad (6.1)$$

do którego układ dochodzi po czasie t .

Entropię Gibbsa definiujemy wtedy jako

$$S_G(t) = - \int f(\Gamma, t) \log f(\Gamma, t) d\Gamma. \quad (6.2)$$

Pokażemy teraz, że wielkość ta jest stała w czasie. Wystarczą do tego dwa fakty. Po pierwsze, z twierdzenia Liouville’a wynika, że przepływ hamiltonowski zachowuje objętość w przestrzeni fazowej:

$$d\Gamma = d\Gamma_0. \quad (6.3)$$

Po drugie, gęstość jest stała wzdłuż trajektorii:

$$f(\Phi^t(\Gamma_0), t) = f(\Gamma_0, 0). \quad (6.4)$$

Skąd to się bierze? Weźmy bardzo małą “kroplę” rozkładu wokół punktu Γ_0 . Jej całkowite prawdopodobieństwo

$$dP = f(\Gamma_0, 0) d\Gamma_0 \quad (6.5)$$

jest podczas ewolucji zachowane, bo układy tylko przesuwają się po trajektoriach – żadne punkty nie znikają i nie pojawiają się znikąd. Ta sama kropla po czasie t trafia do otoczenia punktu $\Phi^t(\Gamma_0)$, więc

$$dP = f(\Phi^t(\Gamma_0), t) d\Gamma. \quad (6.6)$$

Ponieważ zaś ruchu hamiltonowskiego $d\Gamma = d\Gamma_0$, dostajemy

$$f(\Phi^t(\Gamma_0), t) = f(\Gamma_0, 0). \quad (6.7)$$

Możemy teraz wykonać zmianę zmiennych $\Gamma = \Phi^t(\Gamma_0)$:

$$\begin{aligned} S_G(t) &= - \int f(\Gamma, t) \log f(\Gamma, t) d\Gamma = - \int f(\Phi^t(\Gamma_0), t) \log f(\Phi^t(\Gamma_0), t) d\Gamma_0 \\ &= - \int f(\Gamma_0, 0) \log f(\Gamma_0, 0) d\Gamma_0 = S_G(0). \end{aligned} \quad (6.8)$$

A więc w mechanice klasycznej entropia Gibbsa nie rośnie. Dynamika Hamiltona tylko przesuwa i deformuje rozkład w przestrzeni fazowej, ale nie niszczy informacji zapisanej w jego strukturze.

Poznanie demoniczne a poznanie ludzkie

Skoro mikroskopowa informacja nie ginie, jak wytłumaczyć drugą zasadę termodynamiki i powszechnie obserwowany wzrost entropii? Znów rozbijamy się tu o dwie różne wizje świata: demona oraz rzeczywistego obserwatora. Ten pierwszy dysponuje absolutną wiedzą o układzie – zna dokładne położenia i pędy wszystkich cząstek, co oznacza, że posługuje się funkcją $f(\Gamma, t)$ o nieograniczonej rozdzielczości. W miarę ewolucji układu, chmura początkowego prawdopodobieństwa w przestrzeni fazowej wygina się, rozciąga i dzieli na nieskończenie cienkie, splątane włókna niczym kropla soku mieszana w szklance wody. Demon bez trudu śledzi wszystkie te zawiłości, a skoro objętość tych filamentów pozostaje stała (twierdzenie Liouville’a), nie odnotowuje on żadnej utraty informacji, a entropia Gibbsa pozostaje w jego oczach stała. Demon nie dostrzega nieodwracalności – dla niego to świat to wyłącznie w pełni odwracalny, skomplikowany zegar.

My jednak nie potrafimy zarejestrować pędów miliardów pojedynczych atomów, przez co nasz obraz układu jest nieuniknienie rozmyty. Dla naszego wzroku cienkie filamenty stają się nierozróżnialne – nakładają się na siebie i mieszają. Obserwator zastępuje pofałdowany, fraktalny rozkład w przestrzeni fazowej jego wartością uśrednioną na skali makrokomórki. To uśrednianie – wynikające wprost z naszej ignorancji – bezpowrotnie niszczy dokładną informację, sprawiając, że chmura prawdopodobieństwa w naszych oczach "rozmazuje się" i w efekcie zajmuje coraz większą dostępną objętość fazową.

Zamiast używać dokładnej (drobnoziarnistej) gęstości f , wprowadźmy zatem opis gruboziarnisty. Dzielimy więc przestrzeń fazową na komórki ω_i o jednakowej objętości $|\omega|$ i w każdej z nich zastępujemy prawdziwą gęstość przez jej średnią:

$$\tilde{f}(\Gamma, t) = \frac{1}{|\omega|} \int_{\omega_i} f(\Gamma', t) d\Gamma', \quad \Gamma \in \omega_i. \quad (6.9)$$

Jest to właśnie **gęstość gruboziarnista**. Odpowiadająca jej entropia ma postać

$$\tilde{S}(t) = - \int \tilde{f}(\Gamma, t) \log \tilde{f}(\Gamma, t) d\Gamma. \quad (6.10)$$

Załóżmy teraz, że w chwili początkowej rozkład nie zawiera jeszcze struktur mniejszych od rozmiaru komórki, czyli

$$\tilde{f}(\Gamma, 0) = f(\Gamma, 0). \quad (6.11)$$

Wtedy na początku obie entropie są równe:

$$\tilde{S}(0) = S(0). \quad (6.12)$$

Teraz potrzebujemy prostego lematu. Jeśli $\tilde{g}(\Gamma)$ jest funkcją stałą na każdej komórce ω_i , to

$$\int f(\Gamma, t) \tilde{g}(\Gamma) d\Gamma = \int \tilde{f}(\Gamma, t) \tilde{g}(\Gamma) d\Gamma. \quad (6.13)$$

Istotnie, jeśli $\tilde{g} = \tilde{g}_i$ na komórce ω_i , to

$$\begin{aligned} \int f_E \tilde{g} d\Gamma &= \sum_i \tilde{g}_i \int_{\omega_i} f(\Gamma, t) d\Gamma \\ &= \sum_i \tilde{g}_i |\omega| \tilde{f}_i(t) \\ &= \sum_i \int_{\omega_i} \tilde{g}_i \tilde{f}_i(t) d\Gamma = \int \tilde{f}_E \tilde{g} d\Gamma. \end{aligned} \quad (6.14)$$

Wybierając w lemacie funkcję $\tilde{g}(\Gamma) = \log \tilde{f}(\Gamma, t)$, która jest stała na każdej komórce, dostajemy

$$\int \tilde{f}(\Gamma, t) \log \tilde{f}(\Gamma, t) d\Gamma = \int f(\Gamma, t) \log \tilde{f}(\Gamma, t) d\Gamma. \quad (6.15)$$

Możemy więc zapisać:

$$\begin{aligned} \tilde{S}(t) - \tilde{S}(0) &= \tilde{S}(t) - S_{\text{drob}}(t) \\ &= \int f(\Gamma, t) \log f(\Gamma, t) d\Gamma - \int \tilde{f}(\Gamma, t) \log \tilde{f}(\Gamma, t) d\Gamma \\ &= \int f(\Gamma, t) \log \frac{f(\Gamma, t)}{\tilde{f}(\Gamma, t)} d\Gamma. \end{aligned} \quad (6.16)$$

Ostatnia całka jest to tzw. dywergencja Kullbacka-Leiblera między rozkładem dokładnym a gruboziarnistym. Jej nieujemność można sprawdzić elementarnie, korzystając z nierówności

$$\log x \geq 1 - \frac{1}{x}, \quad x > 0, \quad (6.17)$$

w której równość zachodzi tylko dla $x = 1$. Po podstawieniu $x = f_E / \tilde{f}_E$ dostajemy

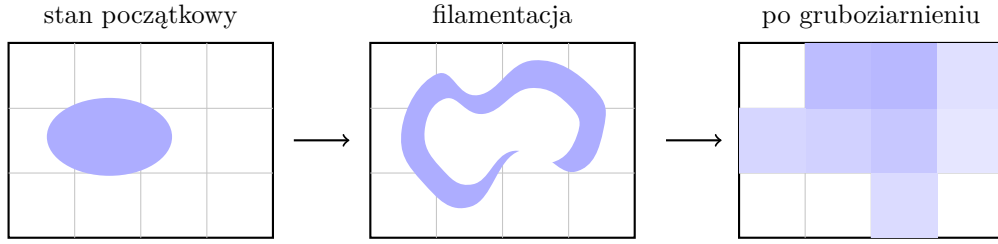
$$f_E \log \frac{f_E}{\tilde{f}_E} \geq f_E - \tilde{f}_E. \quad (6.18)$$

Całkując, otrzymujemy

$$\tilde{S}(t) - \tilde{S}(0) = \int f_E \log \frac{f_E}{\tilde{f}_E} d\Gamma \geq \int (f_E - \tilde{f}_E) d\Gamma = 0, \quad (6.19)$$

bo oba rozkłady są unormowane. A zatem:

$$\boxed{\tilde{S}(t) \geq \tilde{S}(0).} \quad (6.20)$$



Rysunek 4: Schemat wzrostu entropii gruboziarnistej. Dynamika Hamiltona nie zmienia objętości zajmowanej przez rozkład w przestrzeni fazowej, ale rozciąga go i zawija w coraz cieńsze włókna. Po uśrednieniu po komórkach drobna struktura znika, a obserwator makroskopowy widzi coraz bardziej jednorodny obraz.

Równość może zachodzić tylko wtedy, gdy $f(\Gamma, t) = \tilde{f}(\Gamma, t)$ prawie wszędzie, czyli gdy we wnętrzu każdej komórki rozkład pozostał jednorodny. W praktyce ruch hamiltonowski robi coś przeciwnego: rozciąga i zagina początkową „plamę” w przestrzeni fazowej w coraz cieńsze włókna. Drobnociarna gęstość nic na tym nie traci – zachowuje pełną pamięć o tych filamentach. Ale opis grubociarnisty widzi tylko średnią po komórce, więc coraz więcej szczegółowej informacji zostaje bezpowrotnie rozmyte. I właśnie to rozmazywanie, a nie fundamentalna nieodwracalność równań ruchu, stoi za wzrostem entropii makroskopowej.

To geometryczne rozciąganie i zawijanie rozkładu jest najprostszym obrazem **mieszania**. Pojęcie to można sformułować ściślej następująco: jeśli $S \subset C_E$ jest początkowym obszarem, na którym określony jest rozkład, a $V \subset C_E$ dowolnym innym obszarem na powierzchni energii, to przepływ hamiltonowski Φ_t nazywamy **mieszającym**, gdy

$$\frac{\mu(\Phi_t(S) \cap V)}{\mu(V)} \xrightarrow{t \rightarrow \infty} \frac{\mu(S)}{\mu(C_E)}, \quad (6.21)$$

gdzie μ oznacza miarę Liouville’a na C_E . Oznacza to, że po dostatecznie długim czasie fragmenty rozwinętej plamy S trafiają do każdego obszaru V w proporcji wyznaczonej wyłącznie przez miary tych obszarów. W skali makro, układ traci więc pamięć o szczegółach położenia początkowego.

Koktajle a fizyka statystyczna

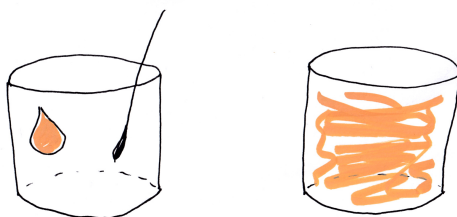
Gibbs, pisząc o mieszaniu, podaje przykład szklanki wody, do której wiano whisky. Początkowo whisky zajmuje tylko niewielką część objętości, lecz po intensywnym wymieszaniu znajdujemy ją wszędzie, w tej samej proporcji (rys. 5). Dobrze wymieszany koktajl to taki, w którym stężenie whisky w każdym elemencie układu jest jednakowe, a więc właśnie taki, o jakim mówi warunek (6.21).

Filamentacja, o której mówiliśmy wcześniej, nie jest jeszcze jedynym sposobem mieszania, ale stanowi naturalny obraz geometryczny tego procesu: coraz cieńsze, splątane pasma rozkładu przenikają każdą makroskopową komórkę, przez co po grubociarnieniu opis staje się coraz bardziej jednorodny.

Przekształcenie kota Arnolda

Klasycznym modelem ilustrującym mieszanie jest **przekształcenie kota Arnolda**. Nie jest to wprawdzie przepływ Hamiltona w czasie ciągłym, lecz dyskretne odwzorowanie zachowujące miarę na torusie $\mathbb{T}^2 = [0, 1) \times [0, 1)$, ale jego własności mają wiele wspólnego z zachowaniem wielocząstkowych układów hamiltonowskich. Przekształcenie zadane jest wzorem

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \pmod{1}. \quad (6.22)$$



Rysunek 5: Mieszanie według Gibbsa: dobrze wymieszany koktajl to taki, w którym procent whisky w każdej podobieństwie jest taki sam.

Wyznacznik macierzy jest równy 1, więc odwzorowanie zachowuje pole powierzchni, czyli dwuwymiarowy odpowiednik miary Liouville'a. Jednocześnie przekształcenie to rozciąga rozkład w jednym kierunku i ściska w drugim. Wartości własne macierzy

$$\lambda_{\pm} = \frac{3 \pm \sqrt{5}}{2}$$

spełniają

$$\lambda_+ > 1, \quad \lambda_- < 1, \quad \lambda_+ \lambda_- = 1,$$

co oznacza, że w każdej iteracji element rozkładu jest wydłużany wzdłuż kierunku niestabilnego i ściskany wzdłuż kierunku stabilnego. Operacja modulo 1 zawiąza następnie obraz z powrotem na torus. W efekcie otrzymujemy właśnie mechanizm rozciągania i fałdowania, który stanowi geometryczne jądro mieszania.

Jeżeli jako rozkład początkowy wybierzemy niewielką plamę albo charakterystyczny obrazek kota, to po kilku iteracjach odwzorowania (6.22) obraz zostaje poszatkowany na cienkie pasma rozciągnięte po całym torusie. Z punktu widzenia opisu drobnociarnistego informacja nie ginie: odwzorowanie jest odwracalne. Jednak po grubociarnieniu obraz szybko staje się praktycznie jednorodny. Przekształcenie kota Arnolda jest więc prostym, lecz bardzo sugestywnym modelem pokazującym, jak odwracalna dynamika może prowadzić do makroskopowego zaniku pamięci o stanie początkowym.

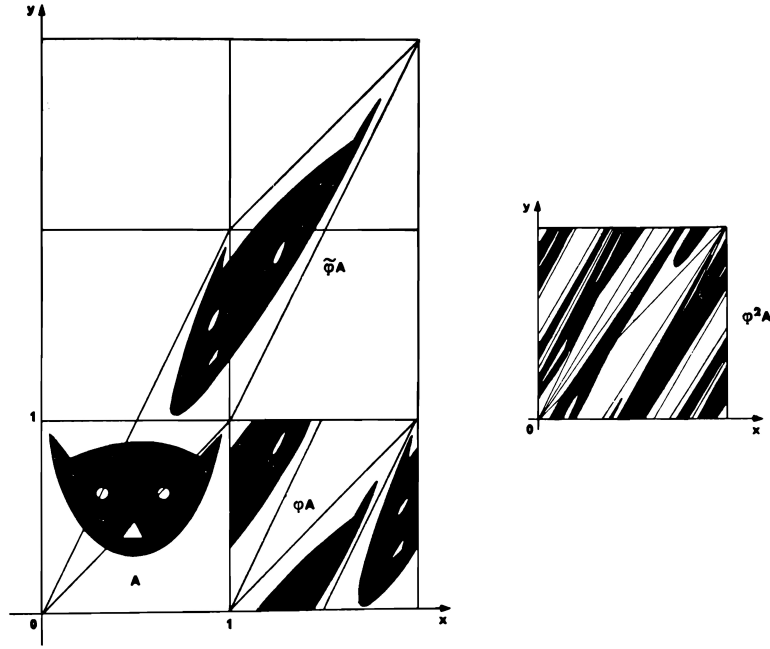
Podsumowanie

W mechanice klasycznej nie ma sprzeczności między odwracalnością równań ruchu a drugą zasadą termodynamiki. Entropia drobnociarnista pozostaje stała, ponieważ dynamika Hamiltona zachowuje objętość w przestrzeni fazowej, a wartość gęstości prawdopodobieństwa jest stała wzdłuż trajektorii. Entropia grubociarnista rośnie natomiast dlatego, że obserwator makroskopowy zastępuje coraz bardziej pofałdowany rozkład jego uśrednieniem po komórkach. Nieodwracalność pojawia się więc dopiero na poziomie opisu, który porzuca informację mikroskopową.

6.2 Ergodyczność i typowość: dlaczego wolno uśredniać po zespole?

W fizyce statystycznej liczymy wielkości makroskopowe jako średnie po rozkładzie na przestrzeni fazowej. W zespole mikrokanonicznym wygląda to tak, jakbyśmy dysponowali całym *zespolem* identycznych kopii układu, rozsianych po powierzchni energii $H(\Gamma_N) = E$, i uśredniali obserwację po wszystkich tych kopiach naraz. Tylko że w rzeczywistym doświadczeniu nie obserwujemy żadnego zespołu. Obserwujemy **jeden** układ, który ewoluuje w czasie. Powstaje więc naturalne pytanie: dlaczego średnia po zespole miałyby cokolwiek mówić o pojedynczym układzie?

Pierwsza, najbardziej naturalna odpowiedź odwołuje się do **ergodyczności**. Zamiast śred-



Rysunek 6: Kot przed i po pojedynczej i podwójnej iteracji przekształcenia (6.22) - oryginalny rysunek z książki V.I. Arnolda “Ergodic Problems of Classical Mechanics”

niej po zespole spróbujmy wziąć średnią po czasie wzdłuż jednej trajektorii. Niech Φ_t oznacza przepływ hamiltonowski na powierzchni energii

$$C_E = \{\Gamma_N : H(\Gamma_N) = E\}. \quad (6.23)$$

Dla dowolnej obserwabli $A(\Gamma_N)$ definiujemy średnią czasową wzdłuż trajektorii startującej z punktu Γ_0 :

$$\bar{A}_t(\Gamma_0) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T A(\Phi_t \Gamma_0) dt, \quad (6.24)$$

o ile granica istnieje. Z drugiej strony odpowiadająca jej średnia mikrokanoniczna ma postać

$$\langle A \rangle_{\text{mc}} = \frac{1}{\omega(E, V, N)} \int A(\Gamma_N) \delta(E - H(\Gamma_N, V)) d\Gamma_N. \quad (6.25)$$

Mówimy, że układ jest **ergodyczny** na powierzchni energii C_E , jeśli dla prawie każdego punktu początkowego $\Gamma_0 \in C_E$ zachodzi

$$\bar{A}_t(\Gamma_0) = \langle A \rangle_{\text{mc}} \quad (6.26)$$

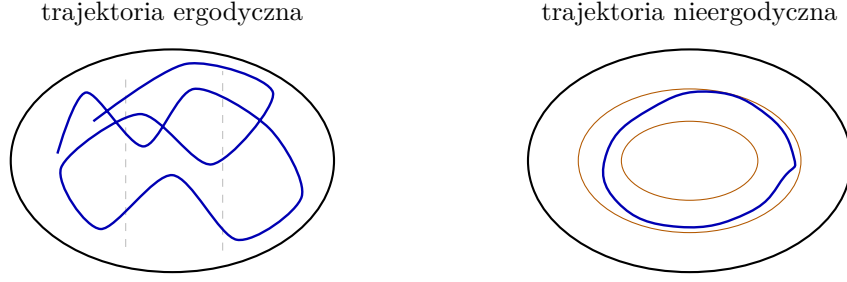
dla każdej rozsądnej obserwabli A . Innymi słowy: średnia czasowa po jednej bardzo długiej trajektorii zgadza się ze średnią po całym zespole. Można udowodnić, że każdy układ mieszający (patrz poprzedni rozdział) jest ergodyczny, lecz twierdzenie odwrotne nie jest prawdziwe.

W szczególności, jeżeli jako obserwabłę weźmiemy funkcję charakterystyczną obszaru $U \subset C_E$:

$$\chi_U(\Gamma_N) = \begin{cases} 1, & \Gamma_N \in U, \\ 0, & \Gamma_N \notin U. \end{cases} \quad (6.27)$$

to średnia czasowa będzie miała prostą interpretację:

$$\bar{\chi}_{U,t}(\Gamma_0) = \lim_{T \rightarrow \infty} \frac{\text{czas spędzony w } U}{T}. \quad (6.28)$$



Rysunek 7: Szkic idei ergodyczności. Po lewej jedna długa trajektoria odwiedza cały dostępny obszar fazowy w sposób reprezentatywny, więc średnia czasowa może pokrywać się ze średnią po zespole. Po prawej ruch pozostaje uwięziony na podzbiorze przestrzeni fazowej, więc taka identyfikacja zawodzi.

Natomiast średnia mikrokanoniczna wynosi

$$\langle \chi_U \rangle_{\text{mc}} = \frac{\omega_U(E)}{\omega(E)} \equiv P_U(E), \quad (6.29)$$

gdzie

$$\omega_U(E) = \int_U \delta(E - H(\Gamma_N, V)) d\Gamma_N \quad (6.30)$$

jest miarą części powierzchni energii leżącej w obszarze U . A zatem dla układu ergodycznego:

$$\lim_{T \rightarrow \infty} \frac{\text{czas spędzony w } U}{T} = \frac{\omega_U(E)}{\omega(E)}. \quad (6.31)$$

Zdanie to można przeczytać po ludzku tak: *ułamek czasu, który pojedynczy układ spędza w obszarze U , jest równy prawdopodobieństwu wylosowania stanu z tego obszaru przy losowaniu jednostajnym na powierzchni energii.*

Na pierwszy rzut oka wygląda to jak dokładnie to, czego potrzebujemy. Jest jednak pewien kłopot. Definicja ergodyczności odwołuje się do granicy $T \rightarrow \infty$, a więc do obserwacji nieskończenie długiej trajektorii. W realnym doświadczeniu nigdy tyle nie czekamy. Co więcej, dla układów makroskopowych czasy odwiedzania różnych zakamarków przestrzeni fazowej stają się niewyobrażalnie długie, więc dosłowna wizja układu “odwiedzającego po kolei” całą powierzchnię energii nie może być rzeczywistym fundamentem fizyki statystycznej.

Właściwy trop jest inny i prowadzi do pojęcia **typowości**. Dla obserwabli makroskopowych nie trzeba zakładać, że układ zdążył zwiedzić całą powierzchnię energii. Wystarczy, że prawie wszystkie mikrostany zgodne z danymi więzami makroskopowymi dają bardzo zbliżoną wartość danej obserwabli. Wtedy pojedynczy układ, o ile znajduje się w stanie typowym, daje wynik praktycznie taki sam jak średnia po zespole. Sens średnich zespołowych bierze się więc nie stąd, że układ wszystko odwiedza, lecz stąd, że *niemal wszystko, co może się zdarzyć, wygląda makroskopowo tak samo*.

Dynamika jest jednak nadal potrzebna. Jak widzieliśmy przy omawianiu gruboziarnienia, ewolucja Hamiltona rozciąga i zawija rozkład w coraz cieńsze filamenty, a w opisie makroskopowym prowadzi to do utraty pamięci o szczegółach warunku początkowego. To właśnie dlatego stany nietypowe są na ogół krótkotrwałe, a układ relaksuje do obszaru stanów typowych.

Z tego punktu widzenia praktycznie użyteczna jest nie tyle ścisła ergodyczność w granicy $T \rightarrow \infty$, ile jej **słaba, operacyjna wersja**: dla istotnych obserwabli makroskopowych i dla czasów długich w porównaniu z czasem relaksacji, ale krótkich w porównaniu z czasem nawrotu, średnie czasowe są bardzo dobrze przybliżone przez średnie zespołowe,

$$\frac{1}{T} \int_0^T A(\Phi_t \Gamma_0) dt \approx \langle A \rangle_{\text{mc}}, \quad \tau_{\text{rel}} \ll T \ll \tau_{\text{recur}}. \quad (6.32)$$

To przybliżenie jest w praktyce całkowicie wystarczające. Nie potrzebujemy bowiem czekać, aż układ “uczciwie” przebiegnie całą powierzchnię energii. Wystarczy, że dla obserwabli makroskopowych niemal wszystkie dopuszczalne mikrostan są do siebie podobne, a dynamika szybko prowadzi układ do tej części przestrzeni fazowej, która jest typowa i przez ogromną większość czasu układ zachowuje się tak, jak przewiduje zespół mikrokanoniczny.

Na koniec warto wspomnieć o jeszcze innym sposobie myślenia, mocno związanym z ujęciem Landaua. Nie odwołuje się on w ogóle do długoczasowej ewolucji pojedynczej trajektorii, lecz do **wielocząstkowości** układu i wynikającego z niej **samośredniowania**. Większość interesujących nas fizycznie obserwabli makroskopowych ma bowiem charakter addytywny: można je zapisać jako sumę wkładów pochodzących z wielu podukładów,

$$A = \sum_{\alpha=1}^M A_{\alpha}. \quad (6.33)$$

W różnych podukładach znajdują się oczywiście nieco inne lokalne konfiguracje cząstek, więc wkłady A_{α} nie są identyczne. Jednak właśnie dlatego, że takich podukładów jest bardzo wiele, suma po nich działa jak skuteczne uśrednienie. W granicy makroskopowej wartość obserwabli w pojedynczej próbce staje się praktycznie równa średniej po zespole, a względne fluktuacje maleją jak $1/\sqrt{M}$ lub $1/\sqrt{N}$. W tym sensie układ makroskopowy *sam* wykonuje potrzebne uśrednienie: przestrzenna suma po ogromnej liczbie różnych lokalnych konfiguracji odtwarza to, co w formalizmie Gibbsa zapisujemy jako średnią po zespole. Jest to więc alternatywa wobec argumentu ergodycznego, i to alternatywa, w której ewolucja czasowa nie odgrywa żadnej zasadniczej roli.

7 Twierdzenie o powracaniu Poincarégo i argument Zermela

A gdyby tak pewnego dnia lub nocy jakiś demon wpelznął za tobą w twą najsamotniejszą samotność i rzekł ci: Życie to, tak jak je teraz przeżywasz i przeżywałeś, będziesz musiał przeżywać raz jeszcze i niezliczone jeszcze razy; i nie będzie nic w niem nowego, tylko każdy ból i każda rozkosz i każda myśl i westchnienie i wszystko niewymownie małe i wielkie twego życia wrócić ci musi, i wszystko w tym samym porządku i następstwie — tak samo ten pająk i ten blask miesiąca pośród drzew i tak samo ta chwila i ja sam. Wieczna klepsydra istnienia odwraca się jeno — a ty z nią, pyłku z pyłu! — Czy nie padłbyś na ziemię i nie zgrzytał zębami i nie przeklął demona, któryby tak mówił?

— Friedrich Nietzsche *Wiedza Radosna* (tłum. L. Staff)

Przechodząc od obrazu Gibbsa do opisu pojedynczego układu w duchu Boltzmann, natrafiamy na dobrze znaną trudność. Gdy Boltzmann sformułował tezę, że entropia układu izolowanego ma tendencję do wzrostu, natychmiast podniesiono przeciw niej dwa klasyczne zarzuty. Pierwszy z nich, sformułowany przez Loschmidta, odwołuje się do odwracalności równań ruchu i był już częściowo omawiany wcześniej. Drugi, wysunięty przez Zermela, opiera się na twierdzeniu o powracaniu Poincarégo i prowadzi do pytania, czy układ może rzeczywiście w sposób trwały przechodzić od makrostanów mniej prawdopodobnych do bardziej prawdopodobnych.

Aby sformułować twierdzenie Poincarégo w przejrzysty sposób, przejdźmy na chwilę od ruchu ciągłego do mapy dyskretniej. Wybierzmy pewien krok czasowy $\tau > 0$ i oznaczmy przez

$$T = \Phi_{\tau} \quad (7.1)$$

odwzorowanie odpowiadające ewolucji układu w czasie τ . Załóżmy, że układ porusza się w pewnym obszarze przestrzeni fazowej D o skończonej objętości oraz że

$$T(D) = D. \quad (7.2)$$

Założenie to jest spełnione na przykład dla układu o ustalonej energii, zamkniętego w skończonym pudle: położenia cząstek są wtedy ograniczone przez ścianki naczynia, a przy zadanej całkowitej energii ograniczone pozostają także ich pędy. Dla dynamiki Hamiltona odwzorowanie T jest ponadto odwracalne i zachowuje objętość fazową.

Twierdzenie Poincarégo o powracaniu. Niech T będzie odwracalnym przekształceniem zachowującym objętość, które przekształca ograniczony obszar D przestrzeni fazowej w siebie. Wtedy dla każdego obszaru $U \subset D$ o dodatniej objętości istnieje punkt $x \in U$ oraz liczba naturalna $n > 0$, takie że

$$T^n(x) \in U. \quad (7.3)$$

Co więcej, nie jest to własność jedynie jakiegoś szczególnego punktu: prawie każdy punkt obszaru U wraca do U , a zbiór punktów, które nigdy do niego nie wracają, ma objętość równą zeru.

Idea dowodu. Najpierw pokażmy istnienie przynajmniej jednego powrotu. Rozpatrzmy kolejne obrazy obszaru U :

$$U = T^0(U), T(U), T^2(U), \dots \quad (7.4)$$

Wszystkie one mają tę samą objętość, ponieważ dynamika Hamiltona objętość zachowuje. Gdyby żadne dwa z tych obszarów się nie przecinały, to mielibyśmy nieskończenie wiele rozłącznych obszarów o tej samej dodatniej objętości upakowanych w skończonym obszarze D . To niemożliwe, a zatem istnieją liczby całkowite $k > l \geq 0$, dla których

$$T^k(U) \cap T^l(U) \neq \emptyset. \quad (7.5)$$

Zastosujmy teraz do obu stron odwrotne odwzorowanie T^{-l} . Dostajemy

$$T^{k-l}(U) \cap U \neq \emptyset. \quad (7.6)$$

Niech y będzie dowolnym punktem tej części wspólnej. Wtedy $y \in U$, ale zarazem $y = T^{k-l}(x)$ dla pewnego $x \in U$. Innymi słowy, znaleźliśmy punkt x , który po $n = k - l$ krokach wraca do obszaru U .

Teraz drugi krok. Niech

$$V = \{x \in U : T^n(x) \notin U \text{ dla każdego } n \geq 1\} \quad (7.7)$$

oznacza zbiór punktów, które po wyjściu z U już nigdy do niego nie wracają. Rozważmy jego obrazy

$$V, T(V), T^2(V), \dots \quad (7.8)$$

Twierdzimy, że są one parami rozłączne. Gdyby bowiem dla pewnych $k > l \geq 0$ zachodziło

$$T^k(V) \cap T^l(V) \neq \emptyset, \quad (7.9)$$

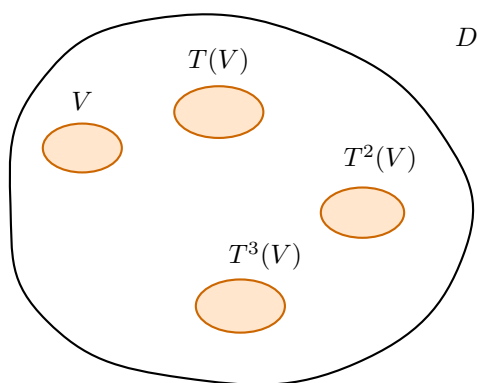
to po zastosowaniu T^{-l} otrzymalibyśmy

$$T^{k-l}(V) \cap V \neq \emptyset. \quad (7.10)$$

Istniałby więc punkt z V , który po $k - l$ krokach znów trafił do $V \subset U$, a więc wróciłby do U . To przeczy definicji zbioru V .

A zatem zbiory $V, T(V), T^2(V), \dots$ są rozłączne. Każdy z nich ma jednak tę samą objętość co V . Gdyby objętość V była dodatnia, to znowu otrzymalibyśmy nieskończenie wiele rozłącznych kopii tego samego obszaru upakowanych w skończonym D . To niemożliwe. Wynika stąd, że objętość V musi być równa zeru.

Warto podkreślić, że twierdzenie o powracaniu jest własnością znacznie słabszą niż ergodyczność. Mówi ono tylko, że trajektoria po bardzo długim czasie wraca do wcześniej odwiedzanych rejonów przestrzeni fazowej. Nie mówi jeszcze, jak często odwiedza różne obszary ani czy robi to w proporcjach zgodnych z opisem mikrokanonicznym.



Rysunek 8: Drugi krok dowodu. Gdyby zbiór punktów, które po wyjściu z U już nigdy do niego nie wracają, miał dodatnią objętość, to jego obrazy $V, T(V), T^2(V), \dots$ byłyby parami rozłączne i wszystkie miałyby tę samą objętość. Nie mogłyby się więc zmieścić w skończonym obszarze fazowym D .

Argument Zermela. Z twierdzenia o powracaniu wynika od razu klasyczny zarzut wobec dosłownie rozumianego wzrostu entropii. Załóżmy, że układ wystartował z makrostanu o małej entropii, a następnie przeszedł do makrostanu o entropii większej. Skoro jednak trajektoria ma po dostatecznie długim czasie wrócić dowolnie blisko stanu początkowego, to po drodze wielkość utożsamiana z entropią musiałaby również *zmaleć*. Nie może więc istnieć żadna funkcja dokładnego mikrostanu, która dla odwracalnej dynamiki Hamiltona rosłaby ściśle monotonicznie dla wszystkich czasów.

To właśnie był sens argumentu Zermela przeciw naiwnej, zbyt dosłownej interpretacji prawa wzrostu entropii. Nie obala on jednak fizyki statystycznej. Pokazuje tylko, że wzrost entropii nie może być rozumiany jako *bezwzględne twierdzenie o każdej trajektorii i każdym czasie*. Jest to twierdzenie statystyczne: dla typowych warunków początkowych i na skalach czasowych dostępnych makroskopowo układ niemal zawsze przechodzi ku makrostanom o większej liczbie mikrostanów.

Warto zestawić ten argument z omówionym wcześniej paradoksem Loschmidta (podrozdz. 5.4). Oba zarzuty uderzają w inną lukę naiwnego obrazu nieodwracalności. **Loschmidt** przypomina, że równania ruchu są odwracalne: po odwróceniu wszystkich prędkości układ powinien wracać ku stanowi bardziej uporządkowanemu. **Zermelo** dodaje, że nawet bez takiej sztucznej operacji sama dynamika w skończonym obszarze fazowym prowadzi po bardzo długim czasie do nawrotów. Innymi słowy, Loschmidt mówi: *można odwrócić bieg filmu*, a Zermelo: *film i tak kiedyś prawie wróci do pierwszych klatek*. Odpowiadając na te zarzuty, Boltzmann wskazał na ogromne skale czasowe, po których mogą zdarzyć się spadki entropii, odpowiadające rzadkim fluktuacjom, mówiąc *tak, powroty istnieją, ale trzeba na nie czekać absurdalnie długo*.

Loschmidt, Zermelo i Boltzmann – w skrócie

Loschmidt: skoro równania ruchu są odwracalne w czasie, to po odwróceniu wszystkich prędkości układ powinien wracać ku stanowi bardziej uporządkowanemu. Nie może więc istnieć ściśle mechaniczne prawo mówiące, że entropia zawsze rośnie dla każdej trajektorii.

Zermelo: nawet bez odwracania prędkości układ w skończonym obszarze przestrzeni fazowej musi po bardzo długim czasie wracać dowolnie blisko stanów początkowych. Dosłownie monotoniczny wzrost entropii dla wszystkich czasów jest więc niemożliwy.

Boltzmann: druga zasada ma charakter statystyczny, a nie absolutnie dynamiczny. Dla typowych warunków początkowych i na czasach makroskopowych entropia gruboziarnista rośnie niemal zawsze, choć wyjątkowo rzadkie fluktuacje i nawroty nie są matematycznie

wykluczone.

7.1 Boltzmann i zagadkaiskoentropijnej przeszłości

Po odpowiedziach na zarzuty Loschmidta i Zermela pozostaje jeszcze jeden problem, trudniejszy od poprzednich. Nawet jeśli zgodzimy się, że druga zasada ma charakter statystyczny i że entropia gruboziarnista rośnie dla typowych warunków początkowych, to natychmiast rodzi się nowe pytanie: **dłaczego warunki początkowe były nietypowe?** Dlaczego przeszłość była tak szczególna? Dlaczego Wszechświat zaczął w stanie o tak małej entropii?

Sam Boltzmann zmagał się z tym problemem przez wiele lat i stopniowo przesunął akcenty swojej odpowiedzi. We wczesnym okresie wierzył, że prawo wzrostu entropii da się bezpośrednio wyprowadzić z mechaniki. Po krytyce Loschmidta coraz wyraźniej przechodził jednak do języka prawdopodobieństwa: spadki entropii nie są niemożliwe, lecz skrajnie nieprawdopodobne. To jednak nie usuwało trudności, tylko przesunęło ją o jeden krok głębiej: Skoro wszystko opiera się na typowości, to dlaczego nasze dzieje zaczynają się od stanu tak wybitnie nietypowego?

Współcześnie myśl tę ujmuje się często w postaci **hipotezy przeszłości** (*Past Hypothesis*): Wszechświat rozpoczął ewolucję w makroście o wyjątkowo małej entropii. Nie jest to jednak wyjaśnienie w ścisłym sensie, lecz raczej dodatkowe założenie o warunku początkowym, bez którego trudno zrozumieć obserwowaną asymetrię między przeszłością a przyszłością.

Świat jako fluktuacja Boltzmann próbował także innej drogi. W latach 90. XIX wieku, w odpowiedzi na trudności związane z kosmologiczną interpretacją drugiej zasady, rozważał możliwość, że globalnie Wszechświat przez większość czasu pozostaje bardzo blisko równowagi, a tylko lokalnie i bardzo rzadko pojawiają się obszary dalekie od równowagi. Nasz świat mógłby być właśnie jednym z takich wyjątkowych obszarów: wielką fluktuacją na tle niemal martwego termicznie kosmosu.

Pomysł ten ma pewien urok: nie wymaga wyróżniania szczególnego początku całego Wszechświata. Ma jednak swoją ciemną stronę. Jeżeli stanem typowym jest równowaga, to bardziej prawdopodobna od olbrzymiej fluktuacji odtwarzającej cały obserwowany Wszechświat jest fluktuacja minimalna: taka, która wytwarza jedynie mały fragment uporządkowanego świata. W wersji skrajnej najtańsza okaże się fluktuacja tworząca tylko **mózg** z przypadkowo uformowanymi wspomnieniami, wrażeniami i przekonaniem, że istnieje zewnętrzny świat oraz długa uporządkowana przeszłość. Taki obiekt nazywa się **mózgiem Boltzmann**.

Mózg Boltzmann

Mózg Boltzmann to nie tylko filozoficzna osobliwość. Mózgi takie mogą rozplenić się we Wszechświatach, w których modele kosmologiczne przewidują bardzo długą, niemal pustą przyszłość de Sittera, czyli fazę z dodatnią stałą kosmologiczną. W takim świecie zwykli obserwatorzy, tacy jak my, powstają tylko w skończonym przedziale czasu: zanim zgaśnie formowanie gwiazd, planet i złożonych struktur. Ich liczba pozostaje więc skończona i jest ograniczona przez ciepłą śmierć Wszechświata.

Jeżeli jednak późny, prawie pusty etap ewolucji Wszechświata trwa bez końca, to zmienia to statystykę możliwych obserwatorów. Wszechświat de Sittera, ze względu na swoją przyspieszającą ekspansję, posiada własny horyzont kosmologiczny. Z tym horyzontem wiąże się subtelny efekt kwantowy (tzw. promieniowanie Gibbonsa-Hawkinga), który sprawia, że próżnia de Sittera nie ma temperatury zera bezwzględnego. Oznacza to, że w absurdalnie długich skalach czasu nie można wykluczyć skrajnie rzadkich fluktuacji tworzących lokalnie uporządkowane struktury. Wśród tych przypadkowych struktur mogą znaleźć się także obiekty zdolne do chwilowego podtrzymania świadomości. Według jednego oszacowania mózg Boltzmann jako fluktuacja w termicznej próżni de Sittera pojawia się raz na $10^{10^{69}}$ lat. Nie są to liczby do zapamiętania, lecz znak, jak niewiarygodnie rzadkie są takie zdarzenia. Tyle że w wiecz-

nym Wszechświecie mają one nieskończenie wiele okazji, by zajść. Wtedy liczba zwykłych obserwatorów pozostaje skończona, a liczba mózgów Boltzmann'a może rosnąć bez ograniczeń. W takim obrazie staje się nieskończenie mało prawdopodobne, aby Czytelnicy tego skryptu byli "normalnymi" umysłami, a nie właśnie taką fluktuacją. Rozpoznając absurd tej sytuacji, w nowoczesnej kosmologii traktuje się pojawianie się mózgów Boltzmann'a jako test teorii: dobry model nie powinien prowadzić do świata, w którym typowym obserwatorem jest efemeryczna fluktuacja w niemal pustej przyszłości de Sittera. Powinna raczej przewidywać, że typowy obserwator jest produktem zwykłej historii kosmicznej: gorącego początku, formowania struktur, gwiazd, planet i długiej ewolucji materii. Jednym z możliwych sposobów uniknięcia tej trudności jest to, by próżnia de Sittera rozpadała się szybciej, niż zdąży wyprodukować mózgi Boltzmann'a. Niektórzy uważają jednak takie argumenty za pozanaukowe, a mózgi Boltzmann'a za realną możliwość.

Argument Gotta

Amerykański kosmolog J. Richard Gott III zaproponował prosty test: jeśli jesteśmy mózgami Boltzmann'a, wystarczy poczekać 10 sekund i sprawdzić, czy jeszcze istniejemy. Większość takich mózgów trwa tylko chwilę, więc powinniśmy niemal natychmiast zniknąć...

Jest z tym jednak kłopot: trudno stwierdzić, że już się nie istnieje (znacie kogoś, kto to stwierdził?). Z drugiej strony w wiecznym Wszechświecie długowiecznych mózgów Boltzmann'a jest i tak nieskończenie wiele...

Część IV

Fizyka statystyczna stanów równowagi

*“Time isn’t holding up,
Time isn’t after us
Same as it ever was,
Same as it ever was.”*

— Talking Heads, *Once in a Lifetime*

1 Zespół mikrokanoniczny

Wiemy już, że w stanie równowagi **układu izolowanego** wszystkie dopuszczalne mikrostanu są jednakowo prawdopodobne. Taki rozkład prawdopodobieństwa nazywamy **zespołem mikrokanonicznym**. Jest to więc statystyczny opis układu o ustalonych wartościach energii E , objętości V i liczby cząstek N .

W dyskretnym obrazie można było po prostu powiedzieć, że wszystkim dopuszczalnym mikrostanom przypisujemy jednakowe prawdopodobieństwo. Dla układu klasycznego wygodniej jest jednak przejść do opisu ciągłego. Cząstki klasyczne mają ciągłe położenia (x, y, z) i pędy (p_x, p_y, p_z) , a stan układu N cząstek reprezentowany jest przez punkt Γ_N w $6N$ -wymiarowej **przestrzeni fazowej**.

Ponieważ układ jest izolowany, jego całkowita energia jest stała:

$$H(\Gamma_N) = E = \text{const.} \quad (1.1)$$

Równanie to wyznacza $(6N - 1)$ -wymiarową hiperpowierzchnię w przestrzeni fazowej. Zespół mikrokanoniczny jest więc rozkładem prawdopodobieństwa skoncentrowanym właśnie na tej powierzchni energii. W opisie ciągłym mówimy już nie o prawdopodobieństwie pojedynczego mikrostanu, lecz o **gęstości prawdopodobieństwa**:

$$\rho(\Gamma_N) = \frac{1}{\omega(E, V, N)} \delta(E - H(\Gamma_N, V)). \quad (1.2)$$

Wielkość $\omega(E, V, N)$ jest stałą normalizacyjną, wyznaczoną warunkiem $\int \rho(\Gamma_N) d\Gamma_N = 1$. Ma ona postać

$$\omega(E, V, N) = \int \delta(E - H(\Gamma_N, V)) d\Gamma_N. \quad (1.3)$$

Jest to gęstość stanów, czyli miara powierzchni fazowej dostępnej układowi przy zadanej energii E .

Bezwymiarowa miara całkowania, uwzględniająca nierozróżnialność cząstek ($N!$) oraz stałą Plancka, ma postać

$$d\Gamma_N = \frac{d^{3N}q d^{3N}p}{N! h^{3N}}. \quad (1.4)$$

Nieoczekiwane h

Pojawienie się stałej Plancka w mierze całkowania wynika z kwantowej natury świata, a dokładniej z zasady nieoznaczoności Heisenberga ($\Delta x \Delta p \geq \hbar/2$). Z jej powodu nie możemy określić położenia i pędu z dowolną dokładnością, co oznacza, że przestrzeń fazowa nie jest ciągła w sensie fizycznym. Minimalna objętość komórki fazowej dla pojedynczej cząstki w jednym wymiarze jest rzędu h . Dla cząstki w trzech wymiarach daje to objętość h^3 , a dla układu N cząstek – całkowitą elementarną objętość równą h^{3N} . Dzielenie przez ten czynnik sprowadza całkę po ciągłej przestrzeni fazowej do poprawnego zliczania dyskretnych stanów kwantowych. Wróćmy do tego zagadnienia w przyszłości, pokazując, że w przejściu od sum

kwantowych do całek taki człon pojawia się naturalnie.

Często wygodniej jest całkować nie po samej powierzchni energii, lecz po całej objętości fazowej obejmującej wszystkie stany o energii mniejszej niż E . Objętość tę oznaczamy przez wielkie Ω :

$$\Omega(E, V, N) = \int_{H(\Gamma_N) < E} d\Gamma_N. \quad (1.5)$$

Wtedy gęstość stanów otrzymujemy przez różniczkowanie:

$$\omega(E, V, N) = \frac{\partial \Omega(E, V, N)}{\partial E}. \quad (1.6)$$

1.1 Gaz doskonały i wzór Sackura-Tetrode

Zastosujmy teraz formalizm zespołu mikrokanonicznego do gazu doskonałego. Dla gazu złożonego z N nieoddziałujących cząstek energia układu jest wyłącznie energią kinetyczną:

$$H(\Gamma_N) = \frac{1}{2m} \sum_{i=1}^{3N} p_i^2.$$

Chcemy obliczyć objętość przestrzeni fazowej dostępnej dla układu o energii nie większej niż E :

$$\Omega(E, V, N) = \int_{H(\Gamma_N) < E} d\Gamma_N.$$

Całka po przestrzeni fazowej rozdziela się na część położeniową i pędową. Ponieważ hamiltonian gazu doskonałego nie zależy od położenia cząstek, całkowanie po współrzędnych daje po prostu czynnik V^N . Pozostaje więc całka po pędach, ograniczona warunkiem

$$\sum_{i=1}^{3N} p_i^2 < 2mE.$$

Warunek ten wyznacza wnętrze kuli w przestrzeni $3N$ -wymiarowej o promieniu

$$R = \sqrt{2mE}.$$

Objętość kuli jednostkowej w d wymiarach wynosi

$$A_d = \frac{\pi^{d/2}}{\Gamma\left(\frac{d}{2} + 1\right)},$$

a więc dla $d = 3N$ otrzymujemy

$$\Omega(E, V, N) = \frac{V^N}{N! h^{3N}} \frac{\pi^{3N/2}}{\Gamma\left(\frac{3N}{2} + 1\right)} (2mE)^{3N/2}.$$

Po uporządkowaniu:

$$\Omega(E, V, N) = \frac{V^N}{N! h^{3N}} \frac{(2\pi mE)^{3N/2}}{\Gamma\left(\frac{3N}{2} + 1\right)}.$$

Gęstość stanów otrzymujemy przez różniczkowanie po energii:

$$\omega(E, V, N) = \frac{\partial \Omega(E, V, N)}{\partial E}.$$

Daje to

$$\omega(E, V, N) = \frac{V^N}{N! h^{3N}} \frac{(2\pi m)^{3N/2}}{\Gamma\left(\frac{3N}{2}\right)} E^{\frac{3N}{2}-1}.$$

Aby wyznaczyć entropię, bierzemy logarytm i przechodzimy do granicy termodynamicznej, korzystając ze wzoru Stirlinga. W tej granicy człony rzędu $\log N$ są pomijalne wobec członów proporcjonalnych do N , tak że

$$S(E, V, N) = k \log \omega(E, V, N)$$

przyjmuje postać

$$S(E, V, N) = Nk \left[\log\left(\frac{V}{N}\right) + \frac{3}{2} \log\left(\frac{E}{N}\right) + \frac{5}{2} + \frac{3}{2} \log\left(\frac{4\pi m}{3h^2}\right) \right].$$

Jest to **wzór Sackura–Tetrode**. Jeśli wprowadzimy objętość właściwą

$$v = \frac{V}{N}$$

oraz energię właściwą

$$u = \frac{E}{N},$$

to entropia na jedną cząsteczkę ma postać

$$s = \frac{S}{N} = k \left[\log v + \frac{3}{2} \log u + \frac{5}{2} + \frac{3}{2} \log\left(\frac{4\pi m}{3h^2}\right) \right].$$

Wynik ten pokazuje kilka ważnych własności:

- Entropia ma postać

$$S(E, V, N) = N f\left(\frac{E}{N}, \frac{V}{N}\right),$$

a więc jest wielkością ekstensywną.

- Ekstensywność ta wynika z braku oddziaływań długiego zasięgu. Gdy dzielimy układ na dwie duże części, energia oddziaływania między nimi rośnie jak pole powierzchni styku, a nie jak objętość. W granicy termodynamicznej wkład ten staje się pomijalny, co prowadzi do addytywności entropii:

$$S = S_1 + S_2.$$

- Zależność

$$S = S(E, V, N)$$

nazywamy **związkiem podstawowym**.

- Liczba dostępnych mikrostanów rośnie z energią bardzo szybko, jak potęga $E^{3N/2}$, a więc dla układu makroskopowego jest wielkością astronomiczną.

2 Statystyczna definicja temperatury

*„Baby’s on fire
And all the instruments agree that
Her temperature’s rising
But any idiot would know that”*

— Brian Eno, *Baby’s on Fire*

Opierając się na postulatcie o maksimum entropii w stanie równowagi, wprowadzimy teraz podstawowe wielkości termodynamiczne. W każdym z poniższych przypadków postępujemy tak samo: stykamy dwa układy, pozwalamy im wymieniać konkretną wielkość (energię, objętość lub cząstki) i sprawdzamy warunek na maksimum całkowitej entropii w nowym stanie równowagi.

Poprzednio rozważaliśmy dwa jednakowe pojemniki z gazem, którym pozwalaliśmy wymieniać cząstki. W stanie równowagi całkowita entropia była największa, co prowadziło do wyrównania liczby cząstek po obu stronach. Teraz wykonamy analogiczne doświadczenie, ale tym razem dopuścimy wymianę energii, zachowując stałe objętości i liczby cząstek obu podukładów.

Rozważmy więc dwa układy stykające się przez ścianę przewodzącą ciepło. Niech pierwszy z nich ma objętość V_1 , liczbę cząstek N_1 i energię E_1 , a drugi odpowiednio V_2 , N_2 i E_2 . Zakładamy, że w trakcie kontaktu mogą zmieniać się tylko energie E_1 i E_2 , natomiast

$$V_1, V_2, N_1, N_2 = \text{const.} \quad (2.1)$$

Ponieważ cały układ jest izolowany, całkowita energia pozostaje stała:

$$E_1 + E_2 = E = \text{const.} \quad (2.2)$$

To znaczy, że tylko jedna z energii jest zmienną niezależną. Wygodnie jest wziąć za nią E_1 , a wtedy

$$E_2 = E - E_1. \quad (2.3)$$

Całkowita entropia układu złożonego ma więc postać

$$S(E_1) = S_1(E_1, V_1, N_1) + S_2(E - E_1, V_2, N_2). \quad (2.4)$$

W stanie równowagi entropia musi być maksymalna względem dopuszczalnego przepływu energii, czyli względem zmian E_1 . Różniczkując powyższy wzór po E_1 , przy ustalonych V_1, V_2, N_1, N_2 oraz ustalonej całkowitej energii E , dostajemy

$$\frac{dS}{dE_1} = \left(\frac{\partial S_1}{\partial E_1} \right)_{V_1, N_1} + \left(\frac{\partial S_2}{\partial E_2} \right)_{V_2, N_2} \frac{dE_2}{dE_1}. \quad (2.5)$$

Ponieważ $E_2 = E - E_1$, mamy

$$\frac{dE_2}{dE_1} = -1, \quad (2.6)$$

a więc warunek maksimum entropii przyjmuje postać

$$\frac{dS}{dE_1} = \left(\frac{\partial S_1}{\partial E_1} \right)_{V_1, N_1} - \left(\frac{\partial S_2}{\partial E_2} \right)_{V_2, N_2} = 0. \quad (2.7)$$

Stąd wynika warunek równowagi cieplnej:

$$\left(\frac{\partial S_1}{\partial E_1} \right)_{V_1, N_1} = \left(\frac{\partial S_2}{\partial E_2} \right)_{V_2, N_2}. \quad (2.8)$$

Widzimy więc, że po zetknięciu dwóch układów wyrównuje się nie sama energia, lecz pewna pochodna entropii po energii, liczona przy stałych V i N . Z klasycznej termodynamiki wiemy, że wielkością wyrównującą się w takim procesie jest temperatura. To prowadzi do statystycznej definicji temperatury:

$$\frac{1}{T} \equiv \left(\frac{\partial S}{\partial E} \right)_{V, N}. \quad (2.9)$$

Zobaczmy teraz, jaki kierunek ma samorzutny przepływ energii. Załóżmy, że układy nie są jeszcze w równowadze i następuje mały transfer energii:

$$E_1 \rightarrow E_1 + \Delta E_1, \quad E_2 \rightarrow E_2 + \Delta E_2, \quad (2.10)$$

przy czym z zachowania całkowitej energii wynika

$$\Delta E_1 + \Delta E_2 = 0. \quad (2.11)$$

Dla małych zmian całkowita zmiana entropii wynosi

$$\Delta S = \left(\frac{\partial S_1}{\partial E_1} \right)_{V_1, N_1} \Delta E_1 + \left(\frac{\partial S_2}{\partial E_2} \right)_{V_2, N_2} \Delta E_2. \quad (2.12)$$

Po wstawieniu $\Delta E_2 = -\Delta E_1$ dostajemy

$$\Delta S = \left[\left(\frac{\partial S_1}{\partial E_1} \right)_{V_1, N_1} - \left(\frac{\partial S_2}{\partial E_2} \right)_{V_2, N_2} \right] \Delta E_1 = \left(\frac{1}{T_1} - \frac{1}{T_2} \right) \Delta E_1. \quad (2.13)$$

W procesie samorzutnym całkowita entropia układu izolowanego rośnie, więc

$$\Delta S > 0. \quad (2.14)$$

Jeżeli $T_1 > T_2$, to

$$\frac{1}{T_1} < \frac{1}{T_2}, \quad (2.15)$$

a zatem nawias w ostatnim wzorze jest ujemny. Aby wtedy ΔS było dodatnie, musi zachodzić

$$\Delta E_1 < 0. \quad (2.16)$$

Oznacza to, że pierwszy układ traci energię, a drugi ją zyskuje. Innymi słowy, energia samorzutnie przepływa od ciała o wyższej temperaturze do ciała o niższej temperaturze. Właśnie dlatego definiujemy temperaturę przez odwrotność pochodnej $\partial S/\partial E$, a nie odwrotnie.

Zastosujmy naszą nową definicję do jednoatomowego gazu doskonałego. Korzystamy z gęstości entropii ze wzoru Sackura-Tetrode'a, różniczkując ją po energii właściwej $e = E/N$ przy stałej objętości właściwej $v = V/N$:

$$\frac{1}{T} = \left(\frac{\partial s}{\partial e} \right)_v = \frac{\partial}{\partial e} \left(\frac{3}{2} k_B \log e + \dots \right) = \frac{3k_B}{2e}$$

Przekształcając to równanie względem energii e , odzyskujemy klasyczny wynik wynikający z zasady ekwipartycji energii:

$$e = \frac{3}{2} k_B T$$

Dlaczego definiujemy odwrotność temperatury?

Pojawia się w tym miejscu naturalne pytanie: dlaczego wielkość $\left(\frac{\partial S}{\partial E} \right)_{V, N}$ definiujemy jako odwrotność temperatury ($1/T$), a nie wprost jako temperaturę (T)?

Powód jest w zasadzie wyłącznie historyczny. Temperatura w skali absolutnej została zdefiniowana na długo przed powstaniem fizyki statystycznej, przy założeniu całkowicie empirycznym: ciepło samorzutnie przepływa z ciała gorętszego (o wyższym T) do chłodniejszego (o niższym T).

Tymczasem z wyprowadzonego przed chwilą warunku statystycznego $\Delta S > 0$ wynika, że energia samorzutnie przepływa z podukładu o mniejszej wartości pochodnej $\partial S/\partial E$ do podukładu o większej jej wartości. Aby połączyć nasz mikroskopowy model z klasyczną konwencją, musimy utożsamić tę pochodną z odwrotnością temperatury. Wówczas mniejsza wartość $1/T$ oznacza wyższe T , co wymusza poprawny, znany z doświadczenia kierunek przepływu ciepła.

3 Ciśnienie

„Pressure pushing down on me
Pressing down on you, no man ask for
Under pressure that burns a building down
Splits a family in two, puts people on streets.”
— Queen & David Bowie, *Under Pressure*

Rozważmy pionowy cylinder o przekroju A wypełniony gazem, zamknięty od góry ruchomym tłokiem o masie m . Objętość gazu wynosi $V = Ah$, gdzie h to wysokość położenia tłoka. Układ jako całość jest izolowany, zatem całkowita energia (energia wewnętrzna gazu E_g oraz energia potencjalna tłoka mgh) jest stała:

$$E_{tot} = E_g + mgh = \text{const.} \quad (3.1)$$

Entropię gazu możemy zapisać jako funkcję położenia tłoka h , korzystając z pełnej statystycznej postaci entropii $S_g(E_g, V, N)$:

$$S(h) = S_g(E_{tot} - mgh, Ah, N) \quad (3.2)$$

W stanie równowagi mechanicznej tłok spoczywa w położeniu, dla którego entropia układu osiąga maksimum. Obliczmy pochodną S względem h :

$$\frac{dS}{dh} = \left(\frac{\partial S_g}{\partial E_g} \right)_{V,N} \frac{dE_g}{dh} + \left(\frac{\partial S_g}{\partial V} \right)_{E_g,N} \frac{dV}{dh} = 0 \quad (3.3)$$

Podstawiając $dE_g/dh = -mg$ oraz $dV/dh = A$, a także wykorzystując wyprowadzony w poprzednim podrozdziale warunek $\left(\frac{\partial S_g}{\partial E_g} \right)_{V,N} = \frac{1}{T}$, otrzymujemy:

$$-\frac{mg}{T} + A \left(\frac{\partial S_g}{\partial V} \right)_{E_g,N} = 0 \quad (3.4)$$

Wielkość mg/A to po prostu zdefiniowane makroskopowo ciśnienie p , jakie tłok wywiera na gaz (i jakie gaz, będąc w równowadze, wywiera na tłok). Przekształcając to równanie, uzyskujemy statystyczną definicję ciśnienia:

$$\left(\frac{\partial S}{\partial V} \right)_{E,N} = \frac{p}{T} \quad (3.5)$$

Zbadajmy teraz zachowanie układu poza stanem równowagi mechanicznej. Wyobraźmy sobie izolowany układ o stałej objętości całkowitej ($V_1 + V_2 = V = \text{const.}$), przedzielony ruchomą membraną. Załóżmy dla uproszczenia, że membrana swobodnie przewodzi ciepło, co gwarantuje stałą, wspólną temperaturę T obu podukładów. Przesunięcie membrany oznacza zmianę $\Delta V_2 = -\Delta V_1$. Całkowita zmiana entropii układu wynosi wówczas:

$$\Delta S = \left(\frac{\partial S_1}{\partial V_1} \right)_{E_1,N_1} \Delta V_1 + \left(\frac{\partial S_2}{\partial V_2} \right)_{E_2,N_2} \Delta V_2 = \left(\frac{p_1}{T} - \frac{p_2}{T} \right) \Delta V_1 \quad (3.6)$$

Zgodnie z postulatem dążenia do stanu równowagi, entropia musi rosnać ($\Delta S > 0$). Jeśli ciśnienie z lewej strony jest większe ($p_1 > p_2$), wyrażenie w nawiasie jest dodatnie. Aby zachować dodatni przyrost entropii, musi zachodzić $\Delta V_1 > 0$. Oznacza to, że membrana przesuwa się, zwiększając objętość układu o wyższym ciśnieniu kosztem układu o niższym. Ruch ten ustanie dopiero wtedy, gdy ciśnienia po obu stronach się zrównają.

Zastosujmy nową definicję do jednoatomowego gazu doskonałego. Korzystamy z wyprowadzonej wcześniej gęstości entropii $s(e, v)$, różniczkując ją po objętości właściwej $v = V/N$ przy zachowaniu stałej energii właściwej $e = E/N$:

$$\frac{p}{T} = \left(\frac{\partial S}{\partial V} \right)_{E,N} = \left(\frac{\partial s}{\partial v} \right)_e = \frac{\partial}{\partial v} (k_B \log v + \dots) = \frac{k_B}{v} \quad (3.7)$$

Przekształcając ten wynik, otrzymujemy zależność:

$$p = \frac{k_B T}{v} \quad (3.8)$$

co jest równoważne klasycznemu równaniu stanu gazu doskonałego ($pV = Nk_B T$).

Korzystając dodatkowo z udowodnionej wcześniej relacji wynikającej z ekwipartycji energii ($k_B T = \frac{2}{3}e$), możemy zapisać ciśnienie w użytecznej postaci niezależnej od temperatury:

$$p = \frac{2e}{3v} \implies pv = \frac{2}{3}e \quad (3.9)$$

W ten sposób wszystkie własności makroskopowe gazu doskonałego zostały wyprowadzone ze statystycznego zachowania jego mikrostanów.

4 Wymiana cząstek i potencjał chemiczny

„Should I stay or should I go now?

If I go, there will be trouble

And if I stay it will be double

So come on and let me know.”

— The Clash, *Should I Stay or Should I Go*

Rozważmy na koniec warunek równowagi związanej z wymianą cząstek. Dwa podukłady są przedzielone ścianką, która utrzymuje stałe objętości V_1 i V_2 , ale pozwala na wymianę zarówno energii, jak i cząstek. Całkowita energia oraz całkowita liczba cząstek układu złożonego są zachowane:

$$E_1 + E_2 = E = \text{const.}, \quad N_1 + N_2 = N = \text{const.} \quad (4.1)$$

Całkowita entropia układu złożonego jest więc funkcją tych dwóch zmiennych:

$$S_{\text{tot}}(E_1, N_1) = S_1(E_1, V_1, N_1) + S_2(E - E_1, V_2, N - N_1). \quad (4.2)$$

Stan równowagi odpowiada maksimum tej funkcji. Oznacza to, że

$$\frac{\partial S_{\text{tot}}}{\partial E_1} = \left(\frac{\partial S_1}{\partial E_1} \right)_{V_1, N_1} - \left(\frac{\partial S_2}{\partial E_2} \right)_{V_2, N_2} = 0, \quad (4.3)$$

$$\frac{\partial S_{\text{tot}}}{\partial N_1} = \left(\frac{\partial S_1}{\partial N_1} \right)_{E_1, V_1} - \left(\frac{\partial S_2}{\partial N_2} \right)_{E_2, V_2} = 0. \quad (4.4)$$

Pierwszy z tych warunków oznacza równość temperatur ($T_1 = T_2$), a drugi warunek mówi, że w stanie równowagi wyrównuje się pochodna entropii po liczbie cząstek:

$$\left(\frac{\partial S_1}{\partial N_1} \right)_{E_1, V_1} = \left(\frac{\partial S_2}{\partial N_2} \right)_{E_2, V_2}. \quad (4.5)$$

Aby powiązać ten wynik z klasyczną termodynamiką, definiujemy potencjał chemiczny:

$$-\frac{\mu}{T} \equiv \left(\frac{\partial S}{\partial N} \right)_{E, V}. \quad (4.6)$$

a równość (4.5) interpretujemy wtedy jako wyrównywanie się potencjałów, $\mu_1 = \mu_2$.

Rozpatrzmy teraz wariację entropii poza stanem równowagi, spowodowaną niewielkim przepływem materii ΔN_1 :

$$\Delta S = \left(\frac{\partial S_1}{\partial N_1} \right)_{E_1, V_1} \Delta N_1 + \left(\frac{\partial S_2}{\partial N_2} \right)_{E_2, V_2} \Delta N_2 = -\frac{\mu_1 - \mu_2}{T} \Delta N_1 \geq 0 \quad (4.7)$$

Z postulatu wzrostu entropii wynika, że jeśli $\mu_1 > \mu_2$, ułamek jest dodatni, więc dla zachowania nierówności $\Delta S \geq 0$ musi zachodzić $\Delta N_1 \leq 0$. Oznacza to, że pierwszy podukład traci materię na rzecz drugiego. Samorzutny przepływ cząstek odbywa się zawsze z rejonu o wyższym potencjale μ w stronę niższego. Z tego powodu parametr ten nazywany jest często w chemii fizycznej miarą skłonności ucieczki (*escaping tendency*) – określa on dążenie cząsteczek do opuszczenia swojej fazy.

Potencjał chemiczny dla gazu doskonałego można wyliczyć startując ze wzoru Sackura-Tetrode’a

$$S = Nk_B \left(\log \frac{V}{N} + \frac{3}{2} \log \frac{E}{N} + \frac{5}{2} + \frac{3}{2} \log \frac{4\pi m}{3h^2} \right) \quad (4.8)$$

i różniczkując to wyrażenie po liczbie cząstek N . To prowadzi do

$$\mu = -T \frac{\partial S}{\partial N} = -k_B T \left(\log v + \frac{3}{2} \log e + \frac{5}{2} + \frac{3}{2} \log \frac{4\pi m}{3h^2} \right) - k_B T N \left(-\frac{1}{N} - \frac{3}{2N} \right) \quad (4.9)$$

co można uprościć do

$$\mu = -k_B T \log \left[v \cdot e^{3/2} \left(\frac{4\pi m}{3h^2} \right)^{3/2} \right] \quad (4.10)$$

Wyrażenie to można przepisać w bardziej użytecznej formie, korzystając z tego, iż (dla gazu jednoatomowego) $e = \frac{3}{2} k_B T$ oraz z równania stanu $v = kT/p$:

$$\mu = -k_B T \log \left[\frac{k_B T}{p} \cdot (k_B T)^{3/2} \left(\frac{2\pi m}{h^2} \right)^{3/2} \right] = k_B T \log \left(\frac{p}{k_B T} \lambda^3 \right) \quad (4.11)$$

gdzie termiczna długość fali de Broglie’a λ to

$$\lambda = \frac{h}{\sqrt{2\pi m k_B T}} \quad (4.12)$$

Wielkość ta wyznacza skalę kwantowego rozmycia cząstek w temperaturze T . Jeśli średnia odległość między cząstkami jest dużo większa od λ , układ można opisywać klasycznie, natomiast gdy staje się z nią porównywalna, istotne stają się efekty statystyki kwantowej.

5 Pierwsza zasada termodynamiki

„Reformator może zmieniać
Rozkład łapek po kieszeniach
I więcej nic
I więcej nic, już nic.”

— Lech Janerka, *Reformator*

Rozważmy układ, w którym zmianie mogą ulegać energia, objętość oraz liczba cząstek. Entropia jest w takiej sytuacji funkcją trzech zmiennych: $S = S(E, V, N)$. (Uwaga: w wielu podręcznikach energia wewnętrzna oznaczana jest literą U , my jednak pozostaniemy przy stosowanym dotychczas E).

Zmiana entropii wyraża się różniczką zupełną:

$$dS = \left(\frac{\partial S}{\partial E} \right)_{V,N} dE + \left(\frac{\partial S}{\partial V} \right)_{E,N} dV + \left(\frac{\partial S}{\partial N} \right)_{E,V} dN$$

Zgodnie z wprowadzonymi wcześniej definicjami, pochodne te wynoszą odpowiednio $1/T$, p/T oraz $-\mu/T$. Podstawiając je, otrzymujemy fundamentalne równanie termodynamiki:

$$dS = \frac{1}{T} dE + \frac{p}{T} dV - \frac{\mu}{T} dN$$

Odwracając to równanie względem zmiany energii dE , dostajemy:

$$dE = TdS - pdV + \mu dN$$

Równanie to niesie za sobą jasną interpretację fizyczną: zmiana energii wewnętrznej układu może nastąpić w wyniku przepływu ciepła, wykonania pracy mechanicznej lub wymiany materii z otoczeniem. Definiujemy zatem:

$$\delta W = -pdV$$

jako pracę sił zewnętrznych wykonaną *nad* układem (znak minus oznacza, że w przypadku sprężania układu $dV < 0$, więc jego energia rośnie),

$$\delta Q = TdS$$

jako ciepło dostarczone *do* układu, natomiast wyraz μdN określamy często mianem pracy chemicznej, związanej z dodaniem lub ujęciem cząstek z układu.

Otrzymujemy w ten sposób pełny matematyczny zapis **pierwszej zasady termodynamiki**:

$$dE = \delta Q + \delta W + \mu dN$$

Należy pamiętać, że w przeciwieństwie do E , wielkości takie jak „praca” δW czy „ciepło” δQ nie są różniczkami funkcji stanu (nie istnieje wielkość taka jak Całkowita Praca czy Całkowite Ciepło układu). Podział zmiany energii ΔE na pracę i ciepło zależy od konkretnego procesu termodynamicznego. Sama znajomość punktu początkowego i końcowego nie wystarczy do obliczenia, jaka praca została wykonana przy przejściu między nimi.

O różniczkach

Dla ułatwienia rozważmy przypadek dwóch zmiennych. W ogólności forma różniczkowa ma postać:

$$\omega = F_x(x, y)dx + F_y(x, y)dy$$

Mówimy o różniczce zupełnej i zapisujemy ją jako df , jeśli formę tę daje się zapisać jako różniczkę pewnej funkcji $f(x, y)$:

$$df = \left(\frac{\partial f}{\partial x}\right)_y dx + \left(\frac{\partial f}{\partial y}\right)_x dy$$

Z twierdzenia o pochodnych mieszanych wynika wówczas warunek:

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$$

Można pokazać, że jeśli przestrzeń, na której określona jest forma, jest pozbawiona „dziur” (jest jednospójna), to warunkiem koniecznym i wystarczającym na to, by różniczka była zupełna, jest równość krzyżowych pochodnych cząstkowych:

$$\left(\frac{\partial F_x(x, y)}{\partial y}\right)_x = \left(\frac{\partial F_y(x, y)}{\partial x}\right)_y$$

Wróćmy do naszego przykładu z termodynamiki. Mieliliśmy tam formę ciepła:

$$\delta Q = 1 \cdot dE + p \cdot dV$$

Sprawdźmy równość pochodnych krzyżowych:

$$\left(\frac{\partial(1)}{\partial V}\right)_E \stackrel{?}{=} \left(\frac{\partial(p)}{\partial E}\right)_V$$

Lewa strona wynosi zero. Aby obliczyć prawą dla gazu doskonałego, korzystamy z relacji $pV = \frac{2}{3}E$, skąd $p = \frac{2E}{3V}$. Obliczamy pochodną:

$$\left(\frac{\partial(-p)}{\partial E}\right)_V = -\frac{2}{3V} \neq 0$$

Ponieważ pochodne cząstkowe nie są równe, wykazaliśmy, że elementarne ciepło δQ nie jest różniczką zupełną.

5.1 Procesy termodynamiczne

Zdefiniowane przez nas parametry stanu, takie jak ciśnienie p czy temperatura T , mają precyzyjny sens fizyczny tylko w stanie równowagi termodynamicznej. Aby poprawnie opisywać ewolucję układu za pomocą tych wielkości, musimy na samym początku ostro rozróżnić trzy kluczowe pojęcia:

- **Proces termodynamiczny** – dowolne przejście układu między dwoma stanami równowagi (początkowym i końcowym). Stany pośrednie w trakcie trwania procesu nie muszą być stanami równowagi.
- **Proces kwazistatyczny** – proces zachodzący na tyle wolno, że układ przechodzi przez ciągłą sekwencję (kontinuum) stanów równowagi. W każdej chwili trwania procesu możemy mu przypisać dobrze określone parametry makroskopowe (np. p i T).
- **Proces odwracalny** – szczególny przypadek procesu kwazistatycznego, w którym dodatkowo **nie występują żadne efekty dyssypacyjne** (takie jak tarcie mechaniczne, lepkość, czy opór elektryczny). Proces ten można w dowolnym momencie odwrócić poprzez nieskończenie małą zmianę warunków zewnętrznych, przywracając zarówno układ, jak i jego otoczenie do stanu początkowego bez pozostawiania jakichkolwiek zmian we Wszechświecie.

Warto wyraźnie podkreślić tę różnicę: **każdy proces odwracalny jest kwazistatyczny, ale nie każdy proces kwazistatyczny jest odwracalny**. Klasycznym kontrprzykładem jest nieskończenie powolne przesuwanie tłoka, który trze o ścianki cylindra. Proces jest kwazistatyczny (gaz w środku ma ustalone p i T), ale tarcie bezpowrotnie rozprasza część pracy w postaci ciepła, co nieodwracalnie zwiększa entropię otoczenia.

W ogólności praca mechaniczna wykonywana nad układem zależy od ciśnienia zewnętrznego p_{zew} :

$$\delta W = -p_{zew} dV.$$

Jeśli jednak proces jest mechanicznie odwracalny, tłok porusza się bez tarcia, a różnica między siłą działającą z zewnątrz a siłą od strony gazu jest nieskończenie mała ($p_{zew} = p \pm dp$). W granicy tej powolnej przemiany ciśnienie zewnętrzne zrównuje się z ciśnieniem gazu, co pozwala nam napisać:

$$\delta W = -p dV, \quad W = -\int_{V_A}^{V_B} p(V) dV.$$

Jeżeli przeprowadzamy układ ze stanu A do stanu B w sposób gwałtowny (niekwazistatyczny), w ogólności nie możemy skorzystać z powyższej całki, ani napisać $Q = \int T dS$. Na drodze między A i B układ nie przechodzi wtedy przez stany równowagi, więc wielkości takie jak T i p po prostu nie istnieją. Ekstremalnym przykładem jest nagłe usunięcie przegrody w naczyniu z gazem (rozprężanie próżniowe): stan równowagi istnieje przed rozpoczęciem procesu i po jego zakończeniu, ale w trakcie samego rozprężania w naczyniu panuje chaos hydrodynamiczny.

Co w praktyce oznacza postulat, by proces zachodził „dostatecznie wolno”? Zmiana parametrów zewnętrznych musi następować wolniej niż najdłuższy czas charakterystyczny (czas

relaksacji) odpowiedzialny za wyrównywanie zaburzeń w układzie. Przy rozprężaniu gazu tłokiem należy więc porównać czas ruchu tłoka z czasem, jakiego fala dźwiękowa potrzebuje na przebycie długości naczynia. Dla $L = 3\text{ m}$ i prędkości dźwięku $v = 332\text{ m/s}$ dostajemy

$$\tau \approx \frac{L}{v} \approx 0.01\text{ s}.$$

Oznacza to, że tłok należy przesuwac z prędkością znacznie mniejszą od prędkości dźwięku, aby zaburzenia ciśnienia zdążyły wyrównać się w całej objętości.

5.2 Procesy adiabatyczne i czynnik całkujący

Jeśli proces zachodzi w całkowitej izolacji cieplnej od otoczenia, czyli

$$\delta Q = 0,$$

nazywamy go procesem **adiabatycznym**. Przy ustalonej liczbie cząstek pierwsza zasada termodynamiki przyjmuje wtedy postać

$$dE = \delta W.$$

W procesie adiabatycznym cała zmiana energii wewnętrznej pochodzi wyłącznie z pracy mechanicznej wykonanej nad układem. Oznacza to, że dla układu izolowanego termicznie wykonana praca zależy tylko od stanu początkowego i końcowego, stając się funkcją stanu.

Zauważmy na koniec, że choć ciepło nie jest różniczką zupełną, to po podzieleniu przez T staje się różniczką entropii

$$dS = \frac{\delta Q_{\text{odwr}}}{T}.$$

W formalizmie równań różniczkowych czynnik $1/T$ nazywamy **czynnikiem całkującym**. Oznacza to, że choć całka

$$\int_A^B \delta Q$$

zależy od wybranej drogi, to całka

$$\int_A^B \frac{\delta Q_{\text{odwr}}}{T} = \int_A^B dS = S_B - S_A$$

zależy już wyłącznie od stanów brzegowych układu.

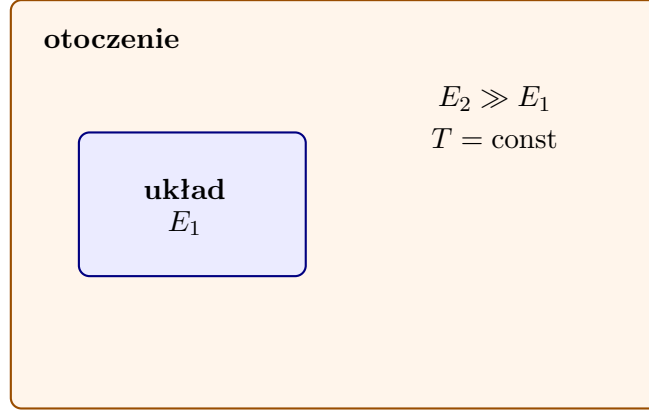
Jeszcze raz podkreślmy kluczowy warunek: zależność $dS = \delta Q/T$ jest ściśle zdefiniowana tylko dla procesów odwracalnych (δQ_{odwr}), w których temperatura układu T jest dobrze określona i jednorodna w każdym punkcie na drodze pomiędzy stanem A i B .

6 Zespół kanoniczny i rozkład Boltzmanna

*„I want to live like common people
I want to do whatever common people do
I want to sleep with common people
I want to sleep with common people like you.”*
— Pulp, *Common People*

Rozważmy teraz niewielki układ 1, który może wymieniać energię z bardzo dużym układem 2. Całość $(1 + 2)$ traktujemy jako układ izolowany o ustalonej energii całkowitej E , objętości V i liczbie cząstek N , a więc opisujemy ją zespołem mikrokanonicznym:

$$\rho(\Gamma^{(1)}, \Gamma^{(2)}) = \frac{1}{\omega(E, V, N)} \delta(E - H_1(\Gamma^{(1)}) - H_2(\Gamma^{(2)})). \quad (6.1)$$



Rysunek 9: Mały układ zanurzony w dużym otoczeniu, które pełni rolę termostatu.

Interesuje nas jednak nie pełny mikrostan układu złożonego, lecz prawdopodobieństwo danego mikrostanu małego układu 1, niezależnie od tego, w jakim stanie znajduje się termostat. Otrzymujemy je przez scałkowanie po wszystkich stopniach swobody układu 2:

$$\rho_1(\Gamma^{(1)}) = \int d\Gamma^{(2)} \rho(\Gamma^{(1)}, \Gamma^{(2)}). \quad (6.2)$$

Po wstawieniu rozkładu mikrokanonicznego dostajemy

$$\rho_1(\Gamma^{(1)}) = \frac{1}{\omega(E, V, N)} \int d\Gamma^{(2)} \delta(E - H_1(\Gamma^{(1)}) - H_2(\Gamma^{(2)})). \quad (6.3)$$

Całka po zmiennych układu 2 jest po prostu jego gęstością stanów przy energii równej temu, co zostało po odjęciu energii układu 1:

$$\rho_1(\Gamma^{(1)}) = \frac{1}{\omega(E, V, N)} \omega_2(E - H_1(\Gamma^{(1)}), V_2, N_2). \quad (6.4)$$

Korzystając z relacji między gęstością stanów a entropią,

$$\omega_2(E_2, V_2, N_2) \propto e^{S_2(E_2, V_2, N_2)/k}, \quad (6.5)$$

możemy zapisać

$$\rho_1(\Gamma^{(1)}) \propto \exp\left[\frac{1}{k} S_2(E - H_1(\Gamma^{(1)}), V_2, N_2)\right]. \quad (6.6)$$

Ponieważ układ 1 jest mały w porównaniu z termostatem, jego energia jest małą poprawką do całkowitej energii rezerwuaru. Możemy więc rozwinąć entropię S_2 w szereg Taylora wokół E :

$$S_2(E - H_1(\Gamma^{(1)}), V_2, N_2) \approx S_2(E, V_2, N_2) - \left(\frac{\partial S_2}{\partial E}\right)_{V_2, N_2} H_1(\Gamma^{(1)}). \quad (6.7)$$

Korzystając z definicji temperatury,

$$\left(\frac{\partial S_2}{\partial E}\right)_{V_2, N_2} = \frac{1}{T}, \quad (6.8)$$

notujemy więc

$$\rho_1(\Gamma^{(1)}) \propto \exp\left(\frac{S_2(E, V_2, N_2)}{k}\right) \exp\left(-\frac{H_1(\Gamma^{(1)})}{kT}\right). \quad (6.9)$$

Pierwszy czynnik nie zależy od mikrostanu układu 1, więc może zostać wchłonięty do stałej normalizacyjnej. Ostatecznie otrzymujemy **rozkład kanoniczny**:

$$\rho_1(\Gamma_N) = \frac{1}{Z(T, V, N)} e^{-\beta H(\Gamma_N, V)}, \quad \beta = \frac{1}{kT}. \quad (6.10)$$

Wielkość

$$Z(T, V, N) = \int e^{-\beta H(\Gamma_N, V)} d\Gamma_N \quad (6.11)$$

nazywamy **sumą statystyczną** (lub klasyczną całką statystyczną). Zawiera ona całą informację termodynamiczną układu w kontakcie z termostatem. W szczególności średnie wartości obserwabli liczymy teraz z użyciem rozkładu Boltzmanna, a nie rozkładu mikrokanonicznego. W zespole kanonicznym energia układu nie jest więc ustalona dokładnie, lecz fluktuuje wokół wartości średniej wyznaczonej przez temperaturę termostatu.

6.1 Energia, jej fluktuacje i energia swobodna Helmholtza

W zespole kanonicznym średnia energia układu wynosi z definicji

$$U = \langle E \rangle = \int E(\Gamma_N) \rho_1(\Gamma_N) d\Gamma_N = \frac{1}{Z(T, V, N)} \int E(\Gamma_N) e^{-\beta E(\Gamma_N)} d\Gamma_N,$$

gdzie

$$\rho_1(\Gamma_N) = \frac{1}{Z(T, V, N)} e^{-\beta E(\Gamma_N)}, \quad Z(T, V, N) = \int e^{-\beta E(\Gamma_N)} d\Gamma_N, \quad \beta = \frac{1}{kT}.$$

Ponieważ

$$\frac{\partial Z}{\partial \beta} = - \int E(\Gamma_N) e^{-\beta E(\Gamma_N)} d\Gamma_N,$$

to

$$U = - \frac{1}{Z} \frac{\partial Z}{\partial \beta} = - \frac{\partial \log Z}{\partial \beta}.$$

Policzmy teraz fluktuacje energii:

$$\langle (\Delta E)^2 \rangle = \langle (E - \langle E \rangle)^2 \rangle = \langle E^2 \rangle - \langle E \rangle^2.$$

Ponieważ

$$\langle E \rangle = \frac{\int E e^{-\beta E(\Gamma_N)} d\Gamma_N}{\int e^{-\beta E(\Gamma_N)} d\Gamma_N},$$

to różniczkując po β , dostajemy

$$\frac{\partial \langle E \rangle}{\partial \beta} = - \frac{\int E^2 e^{-\beta E(\Gamma_N)} d\Gamma_N}{\int e^{-\beta E(\Gamma_N)} d\Gamma_N} + \left(\frac{\int E e^{-\beta E(\Gamma_N)} d\Gamma_N}{\int e^{-\beta E(\Gamma_N)} d\Gamma_N} \right)^2.$$

A więc

$$\frac{\partial U}{\partial \beta} = - \langle (\Delta E)^2 \rangle.$$

Stąd

$$\langle (\Delta E)^2 \rangle = - \left(\frac{\partial U}{\partial \beta} \right)_{V, N} = kT^2 \left(\frac{\partial U}{\partial T} \right)_{V, N} \sim N$$

gdyż dla układów makroskopowych energia jest ekstensywna. Wobec tego względne fluktuacje energii zachowują się jak

$$\frac{\sqrt{\langle (\Delta E)^2 \rangle}}{U} \sim \frac{\sqrt{N}}{N} = \frac{1}{\sqrt{N}}.$$

W granicy termodynamicznej fluktuacje energii stają się więc zaniedbywalne: rozkład energii w zespole kanonicznym jest bardzo silnie skupiony wokół wartości średniej U . Dlatego w tej granicy zespół mikrokanoniczny i kanoniczny dają takie same wyniki dla wielkości makroskopowych.

Skoro w granicy termodynamicznej zespół mikrokanoniczny i kanoniczny dają te same wyniki dla wielkości makroskopowych, porównajmy teraz oba opisy bardziej bezpośrednio.

W zespole mikrokanonicznym wszystkie dopuszczalne mikrostanu realizujące makrostan równowagowy o energii U są jednakowo prawdopodobne. Zgodnie z postulatem równych prawdopodobieństw *a priori* prawdopodobieństwo pojedynczego takiego mikrostanu wynosi

$$P_{\text{mik}} = \frac{1}{\Sigma(U, N, V)},$$

gdzie $\Sigma(U, N, V)$ oznacza liczbę wszystkich dopuszczalnych mikrostanów o energii U .

Z kolei w zespole kanonicznym prawdopodobieństwo mikrostanu o energii E_i dane jest przez rozkład Boltzmanna:

$$P_i = \frac{e^{-\beta E_i}}{Z(T, V, N)}.$$

Dla mikrostanu realizującego energię U dostajemy więc

$$P_{\text{kan}} = \frac{e^{-\beta U}}{Z(T, V, N)}.$$

W granicy termodynamicznej rozkład kanoniczny jest bardzo silnie skupiony wokół energii średniej U , a więc oba opisy powinny przypisywać tę samą wagę mikrostanom realizującym tę energię. Otrzymujemy zatem

$$\frac{e^{-\beta U}}{Z(T, V, N)} = \frac{1}{\Sigma(U, N, V)}.$$

Korzystając z definicji entropii Boltzmanna

$$S(U, V, N) = k \log \Sigma(U, N, V),$$

możemy zapisać

$$\Sigma(U, N, V) = e^{S(U, V, N)/k},$$

a więc

$$\frac{e^{-\beta U}}{Z(T, V, N)} = e^{-S(U, V, N)/k}.$$

Po przekształceniu dostajemy

$$\log Z(T, V, N) = \frac{S(U, V, N)}{k} - \beta U,$$

czyli

$$-kT \log Z = U - TS.$$

Widzimy więc, że w naturalny sposób pojawia się kombinacja

$$U - TS.$$

Oznaczamy ją przez F i nazywamy **energiją swobodną Helmholtza**:

$$F = U - TS.$$

Ostatecznie otrzymujemy

$$F = -kT \log Z.$$

Jest to równanie podstawowe w reprezentacji energii swobodnej:

$$F = F(T, V, N).$$

Policzmy teraz różniczkę F . Mamy

$$F = U - TS,$$

a więc

$$dF = dU - T dS - S dT.$$

Korzystając z pierwszej zasady termodynamiki w postaci

$$dU = T dS - p dV + \mu dN,$$

otrzymujemy

$$dF = -S dT - p dV + \mu dN.$$

Stąd odczytujemy:

$$S = -\left(\frac{\partial F}{\partial T}\right)_{V,N}, \quad p = -\left(\frac{\partial F}{\partial V}\right)_{T,N}, \quad \mu = \left(\frac{\partial F}{\partial N}\right)_{T,V}.$$

6.2 Gaz doskonały w zespole kanonicznym

Policzmy energię swobodną gazu doskonałego, korzystając z rozkładu kanonicznego. Funkcja partycji ma postać

$$Z(T, V, N) = \frac{1}{N!h^{3N}} \int \exp\left(-\beta \sum_{i=1}^N \frac{|\mathbf{p}_i|^2}{2m}\right) d^{3N}p d^{3N}r. \quad (6.12)$$

Ponieważ energia nie zależy od położenia, całkowanie po współrzędnych daje po prostu czynnik V^N , więc

$$Z(T, V, N) = \frac{V^N}{N!h^{3N}} \left(\int \exp\left(-\beta \frac{|\mathbf{p}|^2}{2m}\right) d^3p \right)^N. \quad (6.13)$$

Całka w nawiasie rozkłada się na iloczyn trzech jednowymiarowych całek Gaussowskich, z których każda jest równa

$$\int_{-\infty}^{\infty} \exp\left(-\beta \frac{p_x^2}{2m}\right) dp_x = \sqrt{\frac{2\pi m}{\beta}}. \quad (6.14)$$

Stąd

$$\int \exp\left(-\beta \frac{|\mathbf{p}|^2}{2m}\right) d^3p = \left(\frac{2\pi m}{\beta}\right)^{3/2}, \quad (6.15)$$

a więc

$$Z(T, V, N) = \frac{V^N}{N!h^{3N}} \left(\frac{2\pi m}{\beta}\right)^{3N/2}. \quad (6.16)$$

Energia swobodna wynosi zatem

$$F(T, V, N) = -kT \log Z. \quad (6.17)$$

Korzystając ze wzoru Stirlinga,

$$\log N! \simeq N \log N - N, \quad (6.18)$$

dostajemy

$$F(T, V, N) = -kT \left[N \log V - N \log N + N - 3N \log h + \frac{3N}{2} \log(2\pi m kT) \right]. \quad (6.19)$$

W granicy termodynamicznej wygodnie jest przejść do energii swobodnej na cząstkę,

$$f(T, v) = \lim_{N \rightarrow \infty} \frac{F(T, V, N)}{N}, \quad v = \frac{V}{N}. \quad (6.20)$$

Otrzymujemy wtedy

$$f(T, v) = -kT \left[\log v + 1 - 3 \log h + \frac{3}{2} \log(2\pi mk) + \frac{3}{2} \log T \right]. \quad (6.21)$$

Możemy teraz łatwo wyznaczyć funkcje stanu. Na przykład

$$p = - \left(\frac{\partial f}{\partial v} \right)_T = \frac{kT}{v}, \quad (6.22)$$

czyli ponownie dostajemy równanie stanu gazu doskonałego:

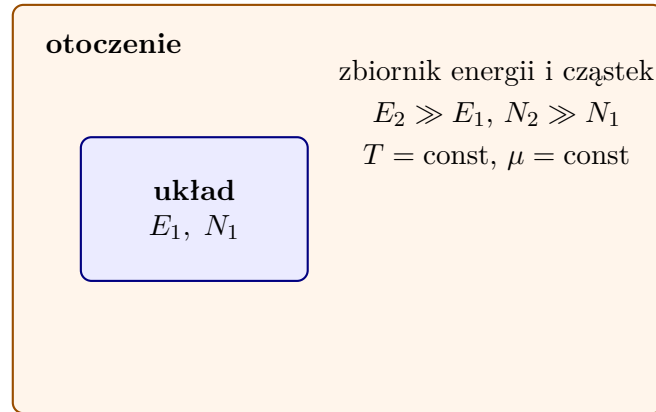
$$pV = NkT. \quad (6.23)$$

7 Wielki zespół kanoniczny

*„They’re piling in the back seat
They’re generating steam heat
Pulsating to the back beat
The blitzkrieg bop.”*

— Ramones, *Blitzkrieg Bop*

W poprzednim rozdziale rozważaliśmy układ zanurzony w **zbiorniku energii**, czyli w termostacie. Teraz zrobmy krok dalej i założymy, że nasz układ może wymieniać z otoczeniem nie tylko energię, ale również **cząstki**. Takie otoczenie nazywamy czasem **zbiornikiem energii i cząstek**. Jest to naturalny model na przykład dla gazu znajdującego się za przepuszczalną przegrodą, przez którą mogą przechodzić zarówno cząstki, jak i ciepło.



Rysunek 10: Mały układ zanurzony w dużym otoczeniu, które pełni rolę zbiornika energii i cząstek. Układ może wymieniać z otoczeniem zarówno energię, jak i cząstki, natomiast temperatura i potencjał chemiczny otoczenia pozostają stałe.

Rozważmy zatem mały układ 1, który pozostaje w kontakcie z bardzo dużym układem 2. Całość (1 + 2) traktujemy jako układ izolowany o ustalonych wartościach E , V i N , a więc opisujemy ją zespołem mikrokanonicznym:

$$\rho(\Gamma^{(1)}, \Gamma^{(2)}) = \frac{1}{\omega(E, V, N)} \delta(E - H_1(\Gamma^{(1)}) - H_2(\Gamma^{(2)})). \quad (7.1)$$

Interesuje nas jednak nie pełny stan układu złożonego, lecz prawdopodobieństwo danego mikrostanu małego układu 1, niezależnie od tego, co dzieje się w rezerwuarze. Trzeba więc scałkować po wszystkich stopniach swobody układu 2 oraz dopuścić możliwość, że układ 1 zawiera różne liczby cząstek N_1 :

$$\rho_1(\Gamma_{N_1}^{(1)}) = \int d\Gamma^{(2)} \rho(\Gamma^{(1)}, \Gamma^{(2)}) = \frac{1}{\omega(E, V, N)} \omega_2(E - H_1(\Gamma_{N_1}^{(1)}), V_2, N - N_1). \quad (7.2)$$

Korzystając ponownie z relacji między gęstością stanów a entropią,

$$\omega_2(E_2, V_2, N_2) \propto e^{S_2(E_2, V_2, N_2)/k}, \quad (7.3)$$

dostajemy

$$\rho_1(\Gamma_{N_1}^{(1)}) \propto \exp\left[\frac{1}{k} S_2(E - H_1(\Gamma_{N_1}^{(1)}), V_2, N - N_1)\right]. \quad (7.4)$$

Ponieważ układ 1 jest mały, zarówno jego energia, jak i liczba cząstek stanowią tylko niewielką poprawkę do odpowiednich wielkości rezerwuaru. Możemy więc rozwinąć entropię S_2 w szereg Taylora:

$$S_2(E - H_1, V_2, N - N_1) \approx S_2(E, V_2, N) - \left(\frac{\partial S_2}{\partial E}\right)_{V_2, N} H_1 - \left(\frac{\partial S_2}{\partial N}\right)_{E, V_2} N_1. \quad (7.5)$$

Teraz korzystamy z definicji temperatury i potencjału chemicznego:

$$\left(\frac{\partial S}{\partial E}\right)_{V, N} = \frac{1}{T}, \quad \left(\frac{\partial S}{\partial N}\right)_{E, V} = -\frac{\mu}{T}. \quad (7.6)$$

Po wstawieniu tych zależności otrzymujemy

$$\rho_1(\Gamma_{N_1}) \propto \exp[-\beta(H(\Gamma_{N_1}, V) - \mu N_1)], \quad \beta = \frac{1}{kT}. \quad (7.7)$$

Po normalizacji dostajemy **wielki rozkład kanoniczny**:

$$\rho_1(\Gamma_N) = \frac{1}{\Xi(T, V, \mu)} \exp[-\beta(H(\Gamma_N, V) - \mu N)]. \quad (7.8)$$

Wielkość

$$\Xi(T, V, \mu) = \sum_{N=0}^{\infty} \int e^{-\beta(H(\Gamma_N, V) - \mu N)} d\Gamma_N \quad (7.9)$$

nazywamy **wielką sumą statystyczną**. Jest ona odpowiednikiem sumy statystycznej $Z(T, V, N)$ z zespołu kanonicznego, tyle że teraz dodatkowo sumujemy po wszystkich możliwych liczbach cząstek.

Często wygodnie jest wprowadzić **aktywność**

$$z = e^{\beta\mu}. \quad (7.10)$$

Wtedy rozkład wielkokanoniczny można zapisać jako

$$\rho_1(\Gamma_N) = \frac{1}{\Xi(T, V, \mu)} z^N e^{-\beta H(\Gamma_N, V)}, \quad (7.11)$$

a sama wielka suma statystyczna przyjmuje postać

$$\Xi(T, V, \mu) = \sum_{N=0}^{\infty} z^N Z(T, V, N), \quad (7.12)$$

gdzie

$$Z(T, V, N) = \int e^{-\beta H(\Gamma_N, V)} d\Gamma_N. \quad (7.13)$$

Widzimy więc, że wielki zespół kanoniczny można traktować jako ważoną sumę zespołów kanonicznych o różnych liczbach cząstek.

7.1 Wielki potencjał termodynamiczny

Analogicznie jak poprzednio można pokazać, że fluktuacje liczby cząstek rosną jak $\sqrt{N}$, a więc względne fluktuacje maleją jak

$$\frac{\sqrt{\langle(\Delta N)^2\rangle}}{\langle N\rangle} \sim \frac{1}{\sqrt{N}}. \quad (7.14)$$

Oznacza to, że w granicy termodynamicznej wszystkie trzy rozkłady, mikrokanoniczny, kanoniczny i wielki kanoniczny, prowadzą do tych samych wyników dla wielkości makroskopowych.

Porównując kanoniczny i wielki kanoniczny dla typowych stanów realizujących wartości średnie, dostajemy

$$\frac{e^{-\beta\langle E\rangle}}{Z} = \frac{e^{-\beta(\langle E\rangle - \mu\langle N\rangle)}}{\Xi}. \quad (7.15)$$

Po skróceniu przez czynnik Boltzmannowski mamy

$$\Xi = Z e^{\beta\mu\langle N\rangle}. \quad (7.16)$$

Logarytmując:

$$\log Z + \beta\mu\langle N\rangle = \log \Xi. \quad (7.17)$$

Po pomnożeniu stronami przez $-kT$ dostajemy

$$F - \mu\langle N\rangle = -kT \log \Xi. \quad (7.18)$$

W granicy termodynamicznej możemy utożsamiać N z jego wartością średnią, a więc

$$\Omega(T, V, \mu) = F(T, V, N) - \mu N = -kT \log \Xi. \quad (7.19)$$

Ten nowy obiekt nazywać będziemy **wielkim potencjałem termodynamicznym**.

Policzmy jego przyrost, żeby zobaczyć, jak wyznaczać z niego wielkości termodynamiczne:

$$\begin{aligned} d\Omega &= dF - \mu dN - N d\mu \\ &= -p dV - S dT + \mu dN - \mu dN - N d\mu \\ &= -p dV - S dT - N d\mu. \end{aligned} \quad (7.20)$$

Stąd odpowiednie funkcje termodynamiczne można otrzymać przez różniczkowanie Ω :

$$p = - \left(\frac{\partial \Omega}{\partial V} \right)_{T, \mu}, \quad (7.21)$$

$$S = - \left(\frac{\partial \Omega}{\partial T} \right)_{V, \mu}, \quad (7.22)$$

oraz

$$\langle N \rangle = - \left(\frac{\partial \Omega}{\partial \mu} \right)_{T, V}. \quad (7.23)$$

Zaś średnią energię można policzyć, analogicznie do zespołu kanonicznego, różniczkując wielką sumę statystyczną po β :

$$\left(\frac{\partial \log \Xi}{\partial \beta} \right)_{V, \mu} = -\langle E - \mu N \rangle = -(U - \mu\langle N \rangle). \quad (7.24)$$

Zatem

$$U = - \left(\frac{\partial \log \Xi}{\partial \beta} \right)_{V, \mu} + \mu\langle N \rangle. \quad (7.25)$$

7.2 Gaz doskonały w wielkim zespole kanonicznym

Zastosujmy teraz ten formalizm do klasycznego gazu doskonałego. Dla ustalonej liczby cząstek N energia ma postać

$$H(\Gamma_N) = \sum_{i=1}^N \frac{p_i^2}{2m}, \quad (7.26)$$

a więc kanoniczna suma statystyczna wynosi

$$Z(T, V, N) = \frac{1}{N! h^{3N}} \int e^{-\beta \sum_{i=1}^N p_i^2 / (2m)} d^{3N} p d^{3N} r. \quad (7.27)$$

Całkę po położeniach liczymy natychmiast i dostajemy czynnik V^N , zaś całka po pędach daje standardowo

$$Z(T, V, N) = \frac{1}{N!} \left(\frac{V}{\lambda^3} \right)^N, \quad (7.28)$$

gdzie

$$\lambda = \frac{h}{\sqrt{2\pi m k T}} = \sqrt{\frac{h^2}{2\pi m k T}} \quad (7.29)$$

jest termiczną długością de Broglie'a.

Wobec tego wielka suma statystyczna przyjmuje postać

$$\Xi(T, V, \mu) = \sum_{N=0}^{\infty} \frac{1}{N!} \left(\frac{zV}{\lambda^3} \right)^N = \exp \left(\frac{zV}{\lambda^3} \right). \quad (7.30)$$

Stąd natychmiast dostajemy wielki potencjał

$$\Omega(T, V, \mu) = -kT \log \Xi = -kT \frac{zV}{\lambda^3}. \quad (7.31)$$

Ponieważ ogólnie zachodzi

$$p = - \left(\frac{\partial \Omega}{\partial V} \right)_{T, \mu}, \quad (7.32)$$

a w tym przypadku Ω jest proporcjonalne do V , otrzymujemy po prostu

$$\Omega = -pV. \quad (7.33)$$

A zatem

$$pV = kT \frac{zV}{\lambda^3}. \quad (7.34)$$

Z drugiej strony średnia liczba cząstek wynosi

$$\langle N \rangle = \frac{1}{\beta} \left(\frac{\partial \log \Xi}{\partial \mu} \right)_{T, V} = \frac{zV}{\lambda^3}. \quad (7.35)$$

Po połączeniu obu wyników dostajemy zwykle równanie stanu gazu doskonałego:

$$pV = \langle N \rangle kT. \quad (7.36)$$

Widzimy więc, że wielki zespół kanoniczny nie zmienia samej fizyki gazu doskonałego, ale daje bardzo wygodny formalizm wszędzie tam, gdzie liczba cząstek nie jest z góry ustalona.

7.3 Twierdzenie Eulera o funkcjach jednorodnych i co z niego wynika

Przyda nam się teraz proste twierdzenie matematyczne dotyczące tzw. funkcji jednorodnych.

Mówimy, że funkcja $f(x_1, \dots, x_N)$ jest **jednorodna stopnia n -tego**, jeśli dla dowolnego λ zachodzi

$$f(\lambda x_1, \dots, \lambda x_N) = \lambda^n f(x_1, \dots, x_N). \quad (7.37)$$

Na przykład funkcja $3x^2$ jest jednorodna stopnia drugiego, funkcja $xy^2 + z^3$ stopnia trzeciego, natomiast $e^{xy} - xy$ nie jest jednorodna w ogóle.

Twierdzenie Eulera mówi, że dla funkcji jednorodnej stopnia n zachodzi

$$nf(x_1, \dots, x_N) = \sum_{i=1}^N x_i \frac{\partial f}{\partial x_i}. \quad (7.38)$$

Dowód jest prosty. Różniczkujemy definicję funkcji jednorodnej stronami po λ :

$$\frac{d}{d\lambda} f(\lambda x_1, \dots, \lambda x_N) = \frac{d}{d\lambda} [\lambda^n f(x_1, \dots, x_N)]. \quad (7.39)$$

Lewa strona daje, na mocy reguły łańcuchowej,

$$\sum_{i=1}^N x_i \frac{\partial f}{\partial (\lambda x_i)}, \quad (7.40)$$

zaś prawa

$$n\lambda^{n-1} f(x_1, \dots, x_N). \quad (7.41)$$

Po podstawieniu $\lambda = 1$ dostajemy tezę.

Skorzystajmy teraz z tego twierdzenia, rozpatrując entropię $S(U, V, N)$. Entropia jest wielkością ekstensywną, a więc po zwiększeniu rozmiarów układu λ razy mamy

$$S(\lambda U, \lambda V, \lambda N) = \lambda S(U, V, N). \quad (7.42)$$

Oznacza to, że S jest funkcją jednorodną stopnia pierwszego. Z twierdzenia Eulera dostajemy więc

$$S = U \left(\frac{\partial S}{\partial U} \right)_{V,N} + V \left(\frac{\partial S}{\partial V} \right)_{U,N} + N \left(\frac{\partial S}{\partial N} \right)_{U,V}. \quad (7.43)$$

Korzystając z definicji temperatury, ciśnienia i potencjału chemicznego

$$\left(\frac{\partial S}{\partial U} \right)_{V,N} = \frac{1}{T}, \quad (7.44)$$

$$\left(\frac{\partial S}{\partial V} \right)_{U,N} = \frac{p}{T}, \quad (7.45)$$

$$\left(\frac{\partial S}{\partial N} \right)_{U,V} = -\frac{\mu}{T}, \quad (7.46)$$

możemy powyższy wzór przepisać w postaci

$$S = \frac{1}{T}(U + pV - \mu N). \quad (7.47)$$

Teraz to wyrażenie na entropię możemy wykorzystać do obliczenia innych potencjałów termodynamicznych. Dla energii swobodnej Helmholtza dostajemy

$$F = U - TS = U - (U + pV - \mu N) = -pV + \mu N. \quad (7.48)$$

Natomiast dla wielkiego potencjału termodynamicznego

$$\Omega = F - \mu N = -pV. \quad (7.49)$$

Widzimy więc, że wielki potencjał termodynamiczny ma wyjątkowo prostą postać.

8 O kurach i grzędach: interpretacja pracy i ciepła w języku stanów energetycznych

„Ponieważ kury lubią nocować wysoko, przybija się więc wewnątrz kurnika w równej wysokości ponad podłogę grzędy czyli banty; grzędy powinny być zrobione z łat 6-7 cm. szerokich, o brzegach górnych zaokrąglonych i znajdować się od siebie w odległości 50 cm., umieszczone na stałych słupkach 50, najwyżej 65 cm. wysokich, tak, aby łatwo dały się z nich zdjąć.

— Józef Zagaja, *Hodowla Kur* (1909)

Rozważmy układ o poziomach energii ε_i oraz prawdopodobieństwach obsadzenia p_i . Energia wewnętrzna układu ma postać

$$U = \sum_i \varepsilon_i p_i. \quad (8.1)$$

Różniczkując, dostajemy

$$dU = \sum_i p_i d\varepsilon_i + \sum_i \varepsilon_i dp_i. \quad (8.2)$$

Drugi człon w tej sumie możemy powiązać ze zmianą entropii

$$S = -k \sum_i p_i \ln p_i. \quad (8.3)$$

Jej różniczka wynosi

$$dS = -k \sum_i dp_i \ln p_i - k \sum_i p_i d(\ln p_i). \quad (8.4)$$

Ponieważ

$$\sum_i p_i = 1, \quad \sum_i dp_i = 0, \quad (8.5)$$

drugi składnik znika i zostaje

$$dS = -k \sum_i dp_i \ln p_i. \quad (8.6)$$

Założmy teraz, że układ ma rozkład kanoniczny

$$p_i = \frac{e^{-\beta \varepsilon_i}}{Z}, \quad Z = \sum_i e^{-\beta \varepsilon_i}, \quad \beta = \frac{1}{kT}. \quad (8.7)$$

Wówczas

$$\ln p_i = -\beta \varepsilon_i - \ln Z. \quad (8.8)$$

Po podstawieniu do (8.6) otrzymujemy

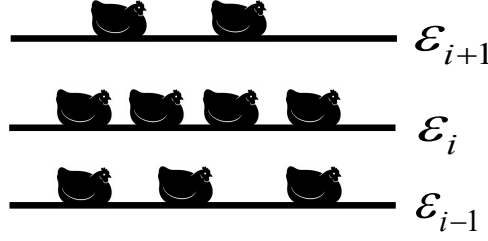
$$dS = -k \sum_i dp_i (-\beta \varepsilon_i - \ln Z) = \beta \sum_i \varepsilon_i dp_i + k(\ln Z) \sum_i dp_i. \quad (8.9)$$

Ponieważ $\sum_i dp_i = 0$, drugi składnik znika, więc

$$dS = \frac{1}{T} \sum_i \varepsilon_i dp_i. \quad (8.10)$$

Widzimy więc, że drugi człon w rozkładzie (8.2) jest związany z przyrostem entropii.

Rozważmy teraz dwa szczególne przypadki



Rysunek 11: Obsadzenia poziomów energetycznych w układzie przypominają rozkład kur na grzędach w kurniku.

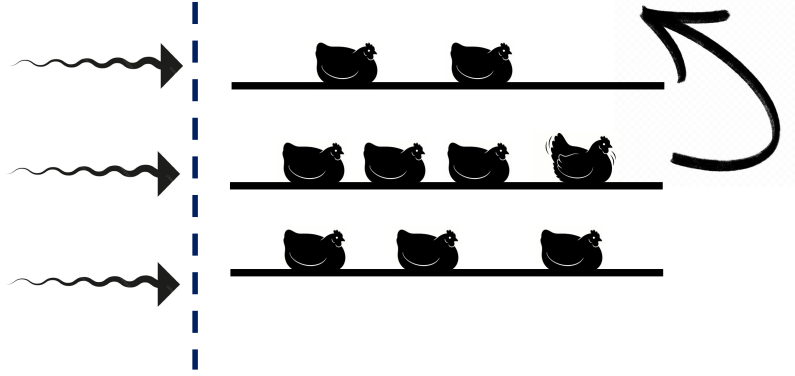
1. Przekaz ciepła Najpierw wyobraźmy sobie, że układ znajduje się w kontakcie cieplnym z drugim ciałem o innej temperaturze. Wtedy energie poziomów się nie zmieniają: $d\varepsilon_i = 0$, natomiast zmieniają się prawdopodobieństwa ich obsadzenia. Z rozkładu (8.2) dostajemy wtedy

$$dQ = dU = \sum_i \varepsilon_i dp_i, \quad (8.11)$$

z czym, na podstawie (8.10) wiąże się przyrost entropii

$$d_e S = \frac{dQ}{T} \quad (8.12)$$

gdzie indeks e oznacza, że zmiana entropii jest w tym przypadku związana z przekazem energii z zewnątrz.



Rysunek 12: Przekaz ciepła - poziomy nie zmieniają się, a obsadzenia - tak

2. Powolna zmiana parametru zewnętrznego Teraz wyobraźmy sobie sytuację, w której układ jest izolowany cieplnie, ale - przy pomocy sił zewnętrznych - bardzo powoli zmieniamy w nim poziomy energetyczne. Jako przykład może tu posłużyć cząstka w nieskończonej studni kwantowej, w której poziomy energetyczne wynoszą

$$\varepsilon_n = \frac{\hbar^2 \pi^2 n^2}{2mL^2} \quad (8.13)$$

Rozciągając studnię, zwiększamy L , co wpłynie na zmniejszenie ε_n .

Jeżeli będziemy to robić bardzo wolno, to będzie można zastosować tzw. *Twierdzenie adiabacyjne* z mechaniki kwantowej:

Twierdzenie adiabatyczne. Jeżeli hamiltonian $H(\lambda(t))$ zależy od parametru $\lambda(t)$ zmieniającego dostatecznie wolno, a układ początkowo znajduje się w nienaznaczonym degeneracją stanie własnym odseparowanym przerwą energetyczną od reszty widma, to podczas ewolucji układ pozostaje w odpowiadającym mu chwilowym stanie własnym, z dokładnością do fazy dynamicznej i geometrycznej.

Oznacza to, że przy dostatecznie powolnej zmianie parametru zewnętrznego nie dochodzi do przejść między poziomami, a więc

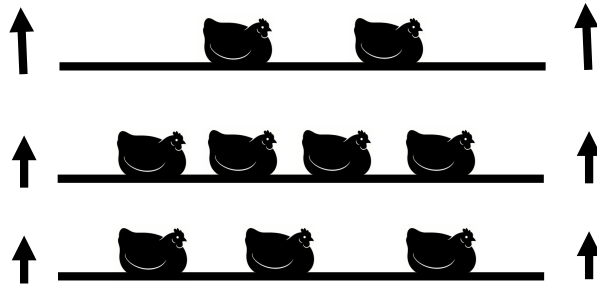
$$dp_i = 0. \quad (8.14)$$

Energia takiego układu zmienia się, ale nie na skutek przepływu ciepła, ale poprzez wykonanie pracy

$$dU = \delta W = \sum_i p_i d\varepsilon_i. \quad (8.15)$$

bez zmiany entropii

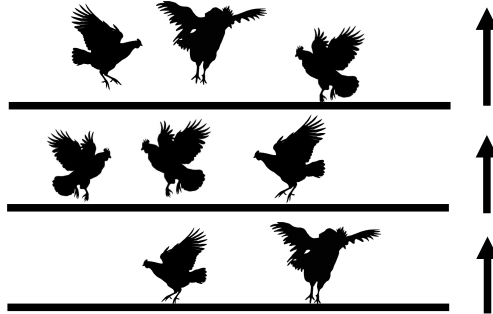
$$dS = 0 \quad (8.16)$$



Rysunek 13: Każdy hodowca wie: przy powolnym podnoszeniu grzęd, kury na nich zostaną.

8.1 Nagła zmiana parametru zewnętrznego

Rozważmy teraz sytuację przeciwną: parametr zewnętrzny zmieniamy nagle. Nasze kury mogą



Rysunek 14: Zbyt szybko podniesione grzędy skutkują spadaniem znajdujących się na nich kur.

Wtedy twierdzenie adiabatyczne nie obowiązuje i pojawiają się przejścia między nowymi chwilowymi stanami własnymi. Oznacza to, że po nagłej zmianie

$$dp_i \neq 0. \quad (8.17)$$

W konsekwencji oprócz członu pracowego

$$\sum_i p_i d\varepsilon_i \quad (8.18)$$

pojawi się także dodatkowy wkład

$$\sum_i \varepsilon_i dp_i, \quad (8.19)$$

nawet jeśli układ jest izolowany.

Pokażmy to na przykładzie jednowymiarowego pudła o długości L . Przy małej zmianie długości

$$d\varepsilon_n = \frac{\partial \varepsilon_n}{\partial L} dL < 0 \quad (\text{dla } dL > 0) = -2 \frac{\varepsilon_n}{L} dL. \quad (8.20)$$

Niech teraz długość pudła zwiększy się nagle. W chwili skoku stan układu nie zdąży się zmienić, zmienia się tylko hamiltonian. Wartość średnia energii pozostaje taka sama jak przed rozszerzeniem, tak że $dU = 0$. Wstawiając to do rozkładu (8.2), dostajemy

$$\sum_n \varepsilon_n dp_n = - \sum_n p_n d\varepsilon_n = 2 \frac{dL}{L} \sum_n p_n \varepsilon_n = 2U \frac{dL}{L} > 0 \quad (8.21)$$

Jest to dodatni wkład “ciepły” wywołany nieadiabatycznością procesu. Nie wiąże się on jednak z transportem ciepła z otoczeniem. Rudolf Clausius nazwał ten człon *ciepłem nieskompensowanym* (uncompensated heat) - nie jest ono związane z przepływem energii przez granicę układu, lecz z nieodwracalnym procesem zachodzącym w samym układzie. Jeżeli temperatura układu jest dobrze określona, to - zgodnie z relacją (8.10) - wielkość ta wiąże się z wewnętrzną produkcją entropii

$$dS_i = \sum_n \varepsilon_n dp_n = \frac{2U dL}{TL} > 0. \quad (8.22)$$

W przykładzie jednowymiarowego pudła widzimy to bezpośrednio. Za chwilę pokażemy, że ten sam wniosek wynika także ogólnie z rachunku zaburzeń dla małego nagłego skoku parametru.

8.2 Mały nagły skok parametru: ujęcie perturbacyjne

Rozważmy teraz ogólny hamiltonian zależny od parametru zewnętrznego λ . Niech

$$H(\lambda)|n\rangle = \varepsilon_n|n\rangle. \quad (8.23)$$

Zakładamy, że widmo jest niezdegenerowane oraz że parametr ulega małej nagłej zmianie

$$\lambda \rightarrow \lambda + d\lambda, \quad (8.24)$$

więc

$$H(\lambda + d\lambda) = H(\lambda) + d\lambda V + O(d\lambda^2), \quad V = \frac{\partial H}{\partial \lambda}. \quad (8.25)$$

Stan układu w chwili skoku nie zdąży się zmienić, zmienia się natomiast baza własna hamiltonianu. Nowe stany własne oznaczamy przez $|n'\rangle$, a odpowiadające im energie przez ε'_n .

Z rachunku zaburzeń dla stanów własnych otrzymujemy

$$|n'\rangle = |n\rangle + d\lambda \sum_{k \neq n} \frac{\langle k|V|n\rangle}{\varepsilon_n - \varepsilon_k} |k\rangle + O(d\lambda^2). \quad (8.26)$$

Mnożąc teraz (8.26) z lewej przez $\langle m|$, dostajemy dla $m \neq n$

$$\langle m|n'\rangle = d\lambda \frac{\langle m|V|n\rangle}{\varepsilon_n - \varepsilon_m} + O(d\lambda^2). \quad (8.27)$$

Z drugiej strony, z ortogonalności stanów $\langle m'|n'\rangle = 0$ dla $m \neq n$ wynika

$$0 = \langle m'|n'\rangle = \langle m + \delta m|n + \delta n\rangle = \langle m'|n\rangle + \langle m|n'\rangle \quad (8.28)$$

co daje

$$\langle m'|n\rangle = -\langle m|n'\rangle + O(d\lambda^2). \quad (8.29)$$

Prawdopodobieństwo przejścia ze starego poziomu n na nowy poziom m wynosi zatem

$$P_{mn} = |\langle m'|n\rangle|^2 = d\lambda^2 \frac{|\langle m|V|n\rangle|^2}{(\varepsilon_n - \varepsilon_m)^2} = P_{nm}, \quad m \neq n. \quad (8.30)$$

oraz

$$\tilde{P}_{nn} = 1 - \sum_{m \neq n} P_{mn}. \quad (8.31)$$

Nowe obsadzenia poziomów są równe

$$p'_m = \langle m'|\rho|m'\rangle = \sum_n P_{mn} p_n, \quad (8.32)$$

a więc

$$dp_m \equiv p'_m - p_m = \sum_n P_{mn} (p_n - p_m), \quad (8.33)$$

gdzie skorzystaliśmy z warunku

$$\sum_n P_{mn} = \sum_n |\langle m'|n\rangle|^2 = \sum_n \langle m'|n\rangle \langle n|m'\rangle = \langle m'|m'\rangle = 1 \quad (8.34)$$

co wynika z zupełności bazy stanów własnych.

Człon ciepły przepisać zatem możemy jako

$$\begin{aligned} \sum_m \varepsilon_m dp_m &= \sum_{mn} \tilde{P}_{mn} \varepsilon_m (p_n - p_m) = \frac{1}{2} \sum_{mn} (P_{mn} \varepsilon_m (p_n - p_m) + P_{nm} \varepsilon_n (p_m - p_n)) \\ &= \frac{1}{2} \sum_{mn} P_{mn} (\varepsilon_m - \varepsilon_n) (p_n - p_m) \end{aligned} \quad (8.35)$$

gdzie skorzystaliśmy z symetrii $P_{nm} = P_{mn}$. Ponieważ prawdopodobieństwa obsadzeń maleją z energią, dla każdych dwóch poziomów zachodzi $(\varepsilon_m - \varepsilon_n)(p_n - p_m) \geq 0$. Każdy składnik sumy (8.35) jest więc nieujemny, a zatem

$$\delta Q_{\text{niesk}} = \sum_m \varepsilon_m dp_m \geq 0. \quad (8.36)$$

Nierówność jest ostra, jeżeli rzeczywiście występują niezerowe przejścia między poziomami o różnych obsadzeniach. Widzimy więc, że dla małego nagłego skoku parametru drugi człon w rozkładzie energii znów dostaliśmy dodatnie ciepło nieskompensowane. Jeżeli temperatura układu jest dobrze określona, to dodatniemu ciepłu nieskompensowanemu odpowiada dodatnia produkcja entropii wewnętrznej:

$$dS_i = \frac{1}{T} \sum_m \varepsilon_m dp_m > 0. \quad (8.37)$$

Dodatniość wewnętrznej produkcji entropii jest manifestacją drugiej zasady termodynamiki

9 Druga zasada termodynamiki

„Kipple is useless objects, like junk mail or match folders after you use the last match or gum wrappers or yesterday’s homeopape. When nobody’s around, kipple reproduces itself. For instance, if you go to bed leaving any kipple around your apartment, when you wake up the next morning there’s twice as much of it. It always gets more and more

— Philip K. Dick, *Do Androids Dream of Electric Sheep?*

Pierwsza zasada termodynamiki mówi, że energia jest zachowana, ale sama nie rozstrzyga, w którą stronę mogą zachodzić procesy. Nie zabrania ona na przykład samorzutnego przepływu ciepła od ciała zimniejszego do cieplejszego ani samoczynnego sprężenia gazu. Dopiero druga zasada wprowadza rozróżnienie między przemianami możliwymi samorzutnie a takimi, które mogą zajść jedynie pod wpływem działania zewnętrznego.

Najzwyczajniej można ją sformułować dla układu izolowanego:

Entropia układu izolowanego nigdy nie maleje, $dS \geq 0$. Równość zachodzi w procesach odwracalnych, natomiast dla procesów spontanicznych $dS > 0$.

Jest to najbardziej podstawowa postać drugiej zasady. W bardziej ogólnej sytuacji, gdy układ może wymieniać ciepło z otoczeniem, wygodnie jest rozdzielić całkowitą zmianę entropii na część związaną z wymianą zewnętrzną oraz część produkowaną wewnątrz:

$$dS = dS_e + dS_i. \quad (9.1)$$

Wówczas

$$dS_e = \frac{dQ_{\text{ext}}}{T}, \quad dS_i \geq 0, \quad (9.2)$$

a więc druga zasada przyjmuje postać

$$dS \geq \frac{dQ_{\text{ext}}}{T}. \quad (9.3)$$

Należy jednak pamiętać, że entropia samego układu nie musi zawsze rosnąć. W procesach wymuszanych zewnątrz może ona maleć, ale tylko kosztem zmian w otoczeniu: wykonania pracy, odpływu ciepła itd.

10 Termodynamiczny taniec

*„See how eagerly the lobsters and the turtles all advance!
They are waiting on the shingle—will you come and join the dance?
Will you, won’t you, will you, won’t you, will you join the dance?
Will you, won’t you, will you, won’t you, won’t you join the dance?*

— Lewis Carroll, *Lobster Quadrille*

Zauważyliśmy zapewne, że w termodynamice pojawia się cała rodzina potencjałów: energia wewnętrzna, entalpia, energia swobodna, funkcja Gibbsa, wielki potencjał termodynamiczny. Skąd bierze się ten pozorny nadmiar?

Odpowiedź jest prosta: różne potencjały są wygodne w różnych warunkach zewnętrznych. Każdy z nich ma swoje *naturalne zmienne*, czyli takie, względem których jego różniczka ma najprostszą postać.

Zacznijmy od energii wewnętrznej:

$$dU = TdS - pdV + \mu dN. \quad (10.1)$$

Widzimy więc, że naturalnymi zmiennymi funkcji U są S , V i N :

$$U = U(S, V, N). \quad (10.2)$$

Załóżmy teraz, że chcemy przejść do opisu, w którym zamiast entropii S wygodniej jest posługiwać się temperaturą T . Tak dzieje się na przykład przy kontakcie z termostatem, gdzie temperatura jest wielkością ustaloną z zewnątrz. Definiujemy wtedy energię swobodną Helmholtza

$$F = U - TS. \quad (10.3)$$

Jej różniczka wynosi

$$dF = dU - T dS - S dT = T dS - p dV + \mu dN - T dS - S dT = -S dT - p dV + \mu dN. \quad (10.4)$$

A zatem

$$F = F(T, V, N). \quad (10.5)$$

Widzimy wyraźnie, że w przejściu od U do F zmienna S została zastąpiona przez T .

Podobnie jest z entalpią, która pojawia się naturalnie wtedy, gdy wygodniej jest pracować przy stałym ciśnieniu. Definiujemy

$$H = U + pV, \quad (10.6)$$

skąd

$$dH = dU + p dV + V dp = T dS - p dV + \mu dN + p dV + V dp = T dS + V dp + \mu dN. \quad (10.7)$$

Zatem

$$H = H(S, p, N). \quad (10.8)$$

Tym razem zamieniliśmy zmienną V na jej sprzężone ciśnienie p .

W termodynamice pojawiają się więc pary wielkości sprzężonych:

$$(T, S), \quad (-p, V), \quad (\mu, N), \quad (10.9)$$

a ogólnie można pisać symbolicznie

$$(X, x), \quad (10.10)$$

gdzie x jest zmienną ekstensywną, a X odpowiadającą jej zmienną intensywną.

Jeśli chcemy przejść od opisu w zmiennej x do opisu w zmiennej X , wykonujemy *transformację Legendre'a*. Jeżeli dla pewnego potencjału P_1 mamy

$$dP_1 = \dots + X dx, \quad (10.11)$$

to definiujemy nowy potencjał

$$P'_1 = P_1 - Xx. \quad (10.12)$$

Wtedy

$$dP'_1 = dP_1 - X dx - x dX = \dots - x dX, \quad (10.13)$$

czyli w nowym potencjale zmienna x została zastąpiona przez X . W szczególności

$$x = - \left(\frac{\partial P'_1}{\partial X} \right). \quad (10.14)$$

Pobawmy się teraz tym mechanizmem dla standardowych potencjałów termodynamicznych. Zaczynamy od energii wewnętrznej:

$$U = U(S, V, N), \quad dU = T dS - p dV + \mu dN. \quad (10.15)$$

Po transformacji względem V otrzymujemy entalpię:

$$H = U[-p] = U + pV, \quad (10.16)$$

$$H = H(S, p, N), \quad dH = TdS + Vdp + \mu dN. \quad (10.17)$$

Po transformacji względem S otrzymujemy energię swobodną Helmholtza:

$$F = U[T] = U - TS, \quad (10.18)$$

$$F = F(T, V, N), \quad dF = -SdT - pdV + \mu dN. \quad (10.19)$$

Po transformacji względem S i V dostajemy funkcję Gibbsa:

$$G = U[T, -p] = U - TS + pV, \quad (10.20)$$

$$G = G(T, p, N), \quad dG = -SdT + Vdp + \mu dN. \quad (10.21)$$

Wreszcie, transformując względem S i N , otrzymujemy wielki potencjał termodynamiczny:

$$\Omega = U[T, \mu] = U - TS - \mu N. \quad (10.22)$$

Dla układu jednorodnego mamy dodatkowo

$$\Omega = -pV, \quad (10.23)$$

a jego różniczka wynosi

$$\Omega = \Omega(T, V, \mu), \quad d\Omega = -SdT - pdV - Nd\mu. \quad (10.24)$$

Możemy teraz zadać pytanie na czujność: co się stanie, jeśli wykonamy transformację Legendre'a względem wszystkich zmiennych ekstensywnych i zdefiniujemy

$$SUP = U[T, -p, \mu] = U - TS + pV - \mu N? \quad (10.25)$$

Dla układu jednorodnego odpowiedź brzmi:

$$SUP = 0. \quad (10.26)$$

Wynika to bezpośrednio z relacji Eulera,

$$U = TS - pV + \mu N, \quad (10.27)$$

lub równoważnie

$$S = \frac{1}{T}(U + pV - \mu N). \quad (10.28)$$

Skoro $SUP = 0$, to również

$$dSUP = 0, \quad (10.29)$$

a więc

$$0 = -SdT + Vdp - Nd\mu. \quad (10.30)$$

Ten bardzo ważny związek nazywamy **relacją Gibbsa-Duhema**.

Kontynuujmy nasz taniec. Wszystkie wypisane wyżej różniczki są formami zupełnymi. Dla każdej formy

$$df = A_1 da_1 + A_2 da_2 + \dots \quad (10.31)$$

mamy

$$A_i = \frac{\partial f}{\partial a_i}. \quad (10.32)$$

Z równości pochodnych mieszanych wynika więc

$$\frac{\partial A_i}{\partial a_j} = \frac{\partial A_j}{\partial a_i}. \quad (10.33)$$

Relacje otrzymywane w ten sposób nazywamy **tożsamościami Maxwella**.

Na przykład z

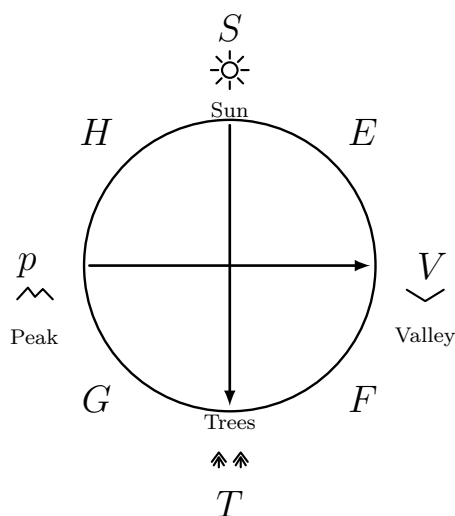
$$dF = -S dT - p dV \quad (10.34)$$

otrzymujemy

$$\left(\frac{\partial S}{\partial V}\right)_T = \left(\frac{\partial p}{\partial T}\right)_V. \quad (10.35)$$

Jest to przykład relacji wiążącej pochodne dotyczące różnych par wielkości sprzężonych.

W zapamiętywaniu tych zależności bardzo pomocny bywa tzw. kółko Borna.



Kółko Borna to patent mnemotechniczny, wymyślony przez Borna, która pozwala orientować się w gęstwie potencjałów termodynamicznych, szybko odtworzyć ich naturalne zmienne oraz znaki w ich różniczkach.

Na obwodzie koła umieszczamy cztery zmienne termodynamiczne: u góry S (jak słońce), na dole T (trees), po lewej p (peak) i po prawej V (valley). Te pary połączone są strzałkami - słońce świeci na drzewa, więc strzałka biegnie od S do T . Woda płynie ze szczytu do doliny, więc strzałka biegnie od p do V .

Na obwodzie wpisujemy następnie, zgodnie z ruchem wskazówek zegara, alfabetycznie potencjały termodynamiczne

$$E, F, G, H. \quad (10.36)$$

Każdy z tych potencjałów zależy od dwóch zmiennych leżących po jego obu stronach, a więc

$$E = E(S, V), \quad F = F(T, V), \quad G = G(T, p), \quad H = H(S, p). \quad (10.37)$$

Rysunek pozwala też natychmiast odczytać różniczki potencjałów. Reguła jest następująca: przy danej zmiennej stoi wielkość leżąca naprzeciw niej na tej samej średnicy, a znak zależy od kierunku strzałki. Przejście zgodnie ze strzałką daje znak plus, a przejście przeciwnie do strzałki znak minus: przy dS stoi $+T$, przy dT stoi $-S$, przy dp stoi $+V$, a przy dV stoi $-p$.

$$dS \leftrightarrow +T, \quad dT \leftrightarrow -S, \quad dp \leftrightarrow +V, \quad dV \leftrightarrow -p. \quad (10.38)$$

W rezultacie dostajemy od razu cztery podstawowe różniczki:

$$dE = T dS - p dV, \quad (10.39)$$

$$dF = -S dT - p dV, \quad (10.40)$$

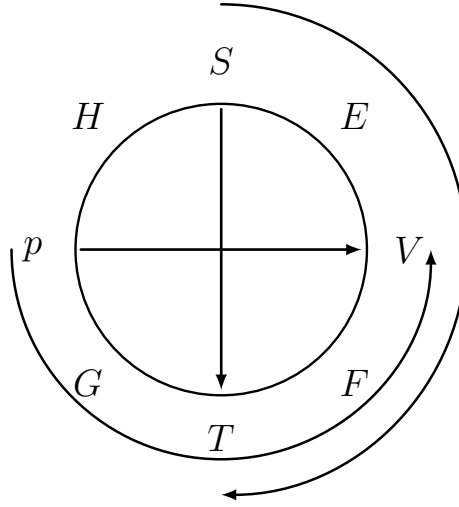
$$dG = -S dT + V dp, \quad (10.41)$$

$$dH = T dS + V dp. \quad (10.42)$$

To samo kółko można również traktować jako skrótowy zapis transformacji Legendre'a:

$$F = E - TS, \quad H = E + pV, \quad G = E - TS + pV. \quad (10.43)$$

Z tego samego rysunku wygodnie odczytuje się także tożsamości Maxwella.



Dla wybranego potencjału rysujemy dwa łuki kończące się na jego ćwiartce. Każdy taki łuk odczytujemy według reguły

$$A \rightarrow B \rightarrow C \quad \Rightarrow \quad \left(\frac{\partial A}{\partial B} \right)_C. \quad (10.44)$$

Innymi słowy: zmienna, od której startujemy, trafia do licznika; zmienna mijana po drodze trafia do mianownika; a zmienna, przy której kończymy, jest utrzymywana na stałym poziomie.

Pozostaje jeszcze znak. Odczytujemy go patrząc na strzałki przy końcach łuku. Jeśli na obu końcach strzałki mają ten sam zwrot, to znak jest dodatni, a dla odwrotnie skierowanych strzałek - ujemny.

Na przykład dla funkcji Helmholtza F łuki narysowane na rysunku są następujące:

$$S \rightarrow V \rightarrow T \quad \Rightarrow \quad \left(\frac{\partial S}{\partial V} \right)_T, \quad (10.45)$$

oraz

$$p \rightarrow T \rightarrow V \quad \Rightarrow \quad \left(\frac{\partial p}{\partial T} \right)_V. \quad (10.46)$$

W tym przypadku oba końce ćwiartki związanej z F mają ten sam zwrot, stąd znak dodatni:

$$\left(\frac{\partial S}{\partial V} \right)_T = \left(\frac{\partial p}{\partial T} \right)_V. \quad (10.47)$$

Dla funkcji Gibbsa G analogiczne dostajemy

$$\left(\frac{\partial S}{\partial p} \right)_T = - \left(\frac{\partial V}{\partial T} \right)_p, \quad (10.48)$$

bo tym razem odpowiednie strzałki (przy G) mają przeciwne znaki

W ten sam sposób odczytujemy cały komplet relacji Maxwella:

$$\left(\frac{\partial T}{\partial V}\right)_S = -\left(\frac{\partial p}{\partial S}\right)_V, \quad (10.49)$$

$$\left(\frac{\partial S}{\partial V}\right)_T = \left(\frac{\partial p}{\partial T}\right)_V, \quad (10.50)$$

$$\left(\frac{\partial S}{\partial p}\right)_T = -\left(\frac{\partial V}{\partial T}\right)_p, \quad (10.51)$$

$$\left(\frac{\partial T}{\partial p}\right)_S = \left(\frac{\partial V}{\partial S}\right)_p. \quad (10.52)$$

11 Prawa ekstremalne, czyli do czego przydają się potencjały

„Porównaj to z resztą,
wybierz wartość największą”
— Kult, *Czarne Słońce*

Prawo wzrostu entropii sformułowaliśmy dla układu izolowanego, a więc przy ustalonych wartościach energii, objętości i liczby cząstek. W praktyce często interesują nas jednak układy pozostające w kontakcie z otoczeniem: z termostatem, barostatem albo rezerwuarem cząstek. Wtedy wygodniej jest posługiwać się nie samą entropią, lecz odpowiednim potencjałem termodynamicznym.

Kontakt z termostatem: minimum energii swobodnej Helmholtza

Rozważmy układ znajdujący się początkowo w stanie równowagi, ale utrzymywany przez pewne więzy. Następnie więzy te usuwamy i pytamy, w jakim kierunku układ będzie dalej ewoluował, jeśli pozostaje on w kontakcie cieplnym z termostatem o ustalonej temperaturze T . Zakładamy przy tym, że objętość i liczba cząstek układu nie zmieniają się. Po usunięciu więzów zachodzi więc *proces spontaniczny*, a to właśnie dla takich procesów druga zasada termodynamiki w przypadku układów izolowanych mówi o wzroście entropii.

Układ wraz z termostatem tworzą duży układ izolowany, zatem dla tego spontanicznego procesu mamy

$$\Delta S_{\text{ot}} + \Delta S \geq 0. \quad (11.1)$$

Ponieważ termostat przez cały czas pozostaje w stanie równowagi o temperaturze T , mamy

$$T \Delta S_{\text{ot}} = \Delta U_{\text{ot}}. \quad (11.2)$$

Z zachowania całkowitej energii wynika zaś

$$\Delta U_{\text{ot}} + \Delta U = 0. \quad (11.3)$$

Po podstawieniu dostajemy

$$\Delta S_{\text{ot}} + \Delta S = -\frac{\Delta U}{T} + \Delta S \geq 0, \quad (11.4)$$

co po wymnożeniu przez $-T$ daje

$$\Delta U - T \Delta S \leq 0. \quad (11.5)$$

A więc

$$\Delta F \leq 0. \quad (11.6)$$

Oznacza to, że w procesach spontanicznych zachodzących przy ustalonych T , V i N energia swobodna Helmholtza nie rośnie. W stanie równowagi osiąga ona minimum.

Kontakt z termostatem i barostatem: minimum funkcji Gibbsa

Rozważmy teraz proces spontaniczny zachodzący przy ustalonych T , p i N . Dla otoczenia mamy wtedy

$$\Delta U_{\text{ot}} = T \Delta S_{\text{ot}} - p \Delta V_{\text{ot}}, \quad (11.7)$$

więc

$$\Delta S_{\text{ot}} = \frac{1}{T} (\Delta U_{\text{ot}} + p \Delta V_{\text{ot}}). \quad (11.8)$$

Ponieważ całkowita energia i całkowita objętość układu wraz otoczeniem są stałe,

$$\Delta U_{\text{ot}} + \Delta U = 0, \quad \Delta V_{\text{ot}} + \Delta V = 0, \quad (11.9)$$

to

$$\Delta S_{\text{ot}} = \frac{1}{T} (-\Delta U - p \Delta V). \quad (11.10)$$

Wstawiając to prawa wzorstu entropii otrzymujemy

$$\Delta G = \Delta U - T \Delta S + p \Delta V \leq 0. \quad (11.11)$$

Oznacza to, że przy ustalonych T , p i N funkcja Gibbsa nie rośnie i w stanie równowagi osiąga minimum.

Kontakt z termostatem i rezerwuarem cząstek: minimum wielkiego potencjału

Analogicznie można potraktować układ pozostający w kontakcie cieplnym i materiałowym z otoczeniem przy ustalonych T , V i μ . Dla rezerwuaru, którego objętość się nie zmienia, mamy

$$\Delta U_{\text{ot}} = T \Delta S_{\text{ot}} + \mu \Delta N_{\text{ot}}. \quad (11.12)$$

Korzystając ze stałości całkowitej energia i liczby cząstek dostaniemy

$$\Delta S_{\text{ot}} = -\frac{\Delta U - \mu \Delta N}{T}. \quad (11.13)$$

co po skorzystaniu z prawa wzrostu entropii da nam

$$\Delta \Omega = \Delta U - T \Delta S - \mu \Delta N \leq 0. \quad (11.14)$$

A zatem przy ustalonych T , V i μ wielki potencjał termodynamiczny również przyjmuje minimum w stanie równowagi.

Zasady ekstremum

Widzimy ogólną regułę: potencjał termodynamiczny osiąga minimum w stanie równowagi wtedy, gdy utrzymujemy stałe jego naturalne zmienne. Dla najważniejszych potencjałów otrzymujemy:

- entropia S osiąga maksimum przy zadanych U , V i N ,
- entalpia H osiąga minimum przy zadanych S , p i N ,
- energia swobodna Helmholtza F osiąga minimum przy zadanych T , V i N ,
- funkcja Gibbsa G osiąga minimum przy zadanych T , p i N ,
- wielki potencjał termodynamiczny Ω osiąga minimum przy zadanych T , V i μ .

To właśnie dlatego potencjały są tak użyteczne: nie są barokowymi ozdobnikami termodynamiki, lecz narzędziem do rozpoznawania, w jaką stronę spontanicznie ewoluuje układ i jaki warunek musi spełniać stan równowagi. Jak niebawem zobaczymy, przydają się też do określenia, jaką pracę można otrzymać z układu w różnych warunkach fizycznych.

Maksimum entropii a minimum energii wewnętrznej

Uważny czytelnik zauważył być może, że w gronie potencjałów wymienionych w powyższej ramce zabrakło energii wewnętrznej. Dowodzi się jej nieco inaczej, niż w przypadku innych potencjałów, wyprowadzając ją bezpośrednio z zasady maksimum dla entropii układu izolowanego. Załóżmy, że w układzie izolowanym stan równowagi opisuje pewien parametr X . Może to być położenie ruchomej przegrody, stopień wymieszania albo inna zmienna wewnętrzna opisująca dopuszczalne stany układu. Skoro w równowadze entropia ma maksimum przy ustalonych U , V i N , to

$$\left(\frac{\partial S}{\partial X}\right)_U = 0, \quad \left(\frac{\partial^2 S}{\partial X^2}\right)_U < 0. \quad (11.15)$$

(dla przejrzystości pomijamy w zapisie V i N). Chcemy teraz opisać ten sam punkt równowagi funkcją odwrotną

$$U = U(S, X), \quad (11.16)$$

czyli zapytać, jak zmienia się energia, gdy poruszamy się po powierzchni stałej entropii.

Tożsamość -1 dla pochodnych cząstkowych

Dla trzech powiązanych zmiennych x, y, z zachodzi tożsamość cykliczna

$$\left(\frac{\partial y}{\partial x}\right)_z \left(\frac{\partial x}{\partial z}\right)_y \left(\frac{\partial z}{\partial y}\right)_x = -1. \quad (11.17)$$

Oto krótki dowód. Niech $z = z(x, y)$. Wtedy

$$dz = \left(\frac{\partial z}{\partial x}\right)_y dx + \left(\frac{\partial z}{\partial y}\right)_x dy. \quad (11.18)$$

Na krzywej $z = \text{const}$ mamy $dz = 0$, zatem

$$\left(\frac{\partial y}{\partial x}\right)_z = -\frac{\left(\frac{\partial z}{\partial x}\right)_y}{\left(\frac{\partial z}{\partial y}\right)_x}, \quad (11.19)$$

co daje tezę.

Znak minus w (11.17) jest ważny: pojawia się dlatego, że poruszając się po krzywej $z = \text{const}$ zwiększenie jednej zmiennej musi być kompensowane zmianą drugiej.

Zastosujmy teraz tę tożsamość do trójki zmiennych X, U, S :

$$\left(\frac{\partial U}{\partial X}\right)_S = -\frac{\left(\frac{\partial S}{\partial X}\right)_U}{\left(\frac{\partial S}{\partial U}\right)_X}. \quad (11.20)$$

W punkcie równowagi licznik znika, bo entropia ma tam ekstremum przy stałym U :

$$\left(\frac{\partial S}{\partial X}\right)_U = 0. \quad (11.21)$$

Stąd natychmiast

$$\left(\frac{\partial U}{\partial X}\right)_S = 0. \quad (11.22)$$

Energia wewnętrzna ma więc w tym samym punkcie ekstremum przy stałym S . Pozostaje ustalić, czy jest to maksimum, czy minimum.

Najprościej to zobaczyć analizując równanie definiujące powierzchnię stałej entropii:

$$S(U(S, X), X) = S. \quad (11.23)$$

Różniczkując je po X przy stałym S , dostajemy

$$0 = \left(\frac{\partial S}{\partial U} \right)_X \left(\frac{\partial U}{\partial X} \right)_S + \left(\frac{\partial S}{\partial X} \right)_U. \quad (11.24)$$

Różniczkujemy jeszcze raz po X przy stałym S . W punkcie równowagi $(\partial U / \partial X)_S = 0$, więc wszystkie wyrazy zawierające ten czynnik znikają i zostaje

$$0 = \left(\frac{\partial S}{\partial U} \right)_X \left(\frac{\partial^2 U}{\partial X^2} \right)_S + \left(\frac{\partial^2 S}{\partial X^2} \right)_U. \quad (11.25)$$

A zatem, w punkcie równowagi,

$$\left(\frac{\partial^2 U}{\partial X^2} \right)_S = - \frac{\left(\frac{\partial^2 S}{\partial X^2} \right)_U}{\left(\frac{\partial S}{\partial U} \right)_X}. \quad (11.26)$$

Teraz używamy definicji temperatury:

$$\left(\frac{\partial S}{\partial U} \right)_X = \frac{1}{T} > 0. \quad (11.27)$$

Ponieważ entropia ma w stanie równowagi maksimum,

$$\left(\frac{\partial^2 S}{\partial X^2} \right)_U < 0, \quad (11.28)$$

dostajemy

$$\left(\frac{\partial^2 U}{\partial X^2} \right)_S > 0. \quad (11.29)$$

Widzimy zatem, że warunek maksimum entropii przy ustalonych U , V i N jest równoważny warunkowi minimum energii wewnętrznej przy ustalonych S , V i N .

Najkrócej

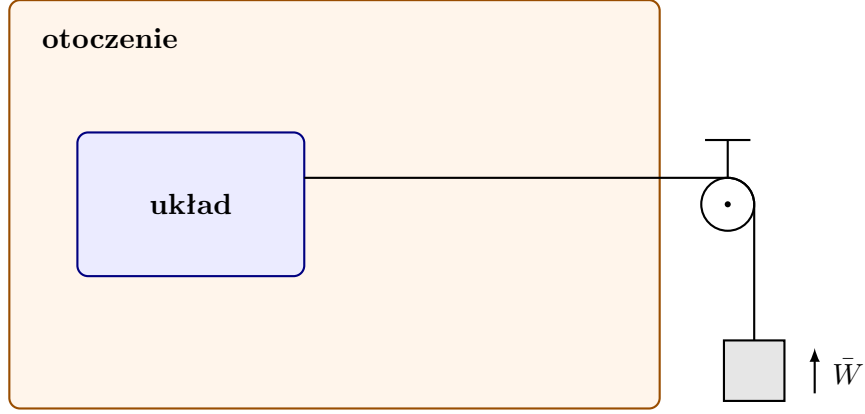
Dla układu izolowanego maksymalizowana jest entropia. Gdy część zmiennych kontroluje otoczenie, wygodniej przejść do odpowiedniego potencjału termodynamicznego. W stanie równowagi potencjał ten przyjmuje minimum przy ustalonych swoich naturalnych zmiennych.

12 Praca minimalna i maksymalna

„Kapitał to praca umarła, która jak wampir ożywa tylko wtedy, gdy wysysa żywą pracę, i tym więcej ma życia, im więcej jej wyssie.”

— Karol Marks, *Kapitał*

Termodynamika wyrosła ze stawiania praktycznych pytań: ile pracy można wydobyć z układu fizycznego? Pierwotnie chodziło o silniki cieplne i maszyny parowe, ale samo pytanie jest znacznie ogólniejsze. Jeżeli układ przechodzi ze stanu początkowego do końcowego, to chcielibyśmy wiedzieć, jaka jest największa praca użyteczna, którą można z niego uzyskać. Równoważnie, jeśli to my chcemy wymusić daną przemianę, możemy zapytać, jaka jest najmniejsza praca, którą musimy w tym celu wykonać.



Rysunek 15: Układ zanurzony w otoczeniu może wykonywać pracę użyteczną, na przykład podnosząc ciężarek.

Rozważmy układ w kontakcie z otoczeniem, scharakteryzowanym przez stałe wielkości intensywne $T_{\text{ot}}, p_{\text{ot}}, \mu_{\text{ot}}$. Całość, to znaczy badany układ oraz otoczenie traktujemy jako układ adiabatycznie izolowany. Z tej całości można jednak wyprowadzić pracę użyteczną $\bar{W}$, na przykład podnosząc ciężarek albo ładując akumulator. Nie wliczamy so niej pracy objętościowej przeciw ciśnieniu otoczenia, ponieważ ta jest już uwzględniona przez kontakt z rezerwuarem o ciśnieniu p_{ot} . Zauważmy, że zgodnie z przyjętą przez nas wcześniej konwencją, W (czyli praca wykonana nad układem) jest dana przez $W = -\bar{W}$. Bilans energii dla układu i otoczenia ma postać

$$\Delta U_{\text{ot}} + \Delta U = -\bar{W}. \quad (12.1)$$

Jeżeli $\bar{W} > 0$, energia układu i otoczenia maleje, bo część energii została wyprowadzona na zewnątrz jako praca użyteczna. Dla zmiennych ekstensywnych mamy

$$\Delta V_{\text{ot}} + \Delta V = 0, \quad \Delta N_{\text{ot}} + \Delta N = 0. \quad (12.2)$$

Całość jest izolowana cieplnie, więc z drugiej zasady termodynamiki wynika

$$\Delta S_{\text{ot}} + \Delta S \geq 0. \quad (12.3)$$

Ponieważ otoczenie pozostaje cały czas w stanie równowagi, jego energia wewnętrzna zmienia się zgodnie z równaniem fundamentalnym

$$\Delta U_{\text{ot}} = T_{\text{ot}} \Delta S_{\text{ot}} - p_{\text{ot}} \Delta V_{\text{ot}} + \mu_{\text{ot}} \Delta N_{\text{ot}}. \quad (12.4)$$

Po podstawieniu (12.4) do (12.1) oraz skorzystaniu z (12.2) oraz drugiej zasady dostajemy

$$\bar{W} = -\Delta U - T_{\text{ot}} \Delta S_{\text{ot}} - p_{\text{ot}} \Delta V + \mu_{\text{ot}} \Delta N \leq -\Delta U + T_{\text{ot}} \Delta S - p_{\text{ot}} \Delta V + \mu_{\text{ot}} \Delta N. \quad (12.5)$$

Jest to ograniczenie z góry na pracę użyteczną, którą można wydobyć z układu. Maksymalna praca jest osiągana wtedy i tylko wtedy, gdy całkowita produkcja entropii znika, czyli gdy $\Delta S_{\text{ot}} + \Delta S = 0$, co odpowiada procesowi odwracalnemu

Ten sam wynik można odczytać od drugiej strony. Jeśli pytamy nie o to, ile pracy da się uzyskać, lecz ile trzeba włożyć, aby wymusić pewną przemianę (np. sprężanie gazu), mamy problem pracy minimalnej, czyli: Jak robić, żeby się nie narobić? Ponieważ $W = -\bar{W}$, z (12.5) wynika

$$W \geq \Delta U - T_{\text{ot}} \Delta S + p_{\text{ot}} \Delta V - \mu_{\text{ot}} \Delta N. \quad (12.6)$$

Minimalna praca także odpowiada procesowi odwracalnemu.

W procesie odwracalnym osiągamy zarówno maksimum pracy wydobytej z układu, jak i minimum pracy nad nim wykonanej

Ogólne nierówności (12.5) i (12.6) zawierają w sobie najważniejsze przypadki spotykane w termodynamice. W zależności od tego, jakie parametry są ustalone przez otoczenie, prawa strona nierówności przyjmuje postać zmiany jednego ze znanych potencjałów termodynamicznych. Dla stanów początkowego i końcowego pozostających w równowadze z odpowiednim rezerwuarem otrzymujemy:

warunki	potencjał	$W_{\min}$	$\bar{W}_{\max}$
T, V, N	$F = U - TS$	ΔF	$-\Delta F$
T, p, N	$G = U - TS + pV$	ΔG	$-\Delta G$
T, V, μ	$\Omega = U - TS - \mu N$	$\Delta \Omega$	$-\Delta \Omega$

W szczególnym przypadku procesu adiabatycznego odwracalnego przy stałych V i N mamy $\Delta S = \Delta V = \Delta N = 0$, więc z (12.6)

$$W = \Delta U. \quad (12.7)$$

13 Wypukłość i wklęsłość potencjałów termodynamicznych

*„There ain't no mountain high enough
Ain't no valley low enough
Ain't no river wide enough
To keep me from getting to you.”*

— Marvin Gaye & Tammi Terrell,

Ain't No Mountain High Enough

Warunki stabilności równowagi przekładają się bezpośrednio na geometrię funkcji stanu: entropia $S(U, V, N)$ jest wklęsła względem swoich naturalnych zmiennych, natomiast energia $U(S, V, N)$ jest względem nich wypukła. Potencjały otrzymane z U przez transformacje Legendre'a są, w obszarze stabilności, wypukłe względem swoich naturalnych zmiennych ekstensywnych i wklęsłe względem zmiennych intensywnych. Te własności geometryczne narzucają znaki pojemności cieplnych, ściśliwości oraz innych podatności.

Zacznijmy od energii wewnętrznej. Rozważmy dwa jednorodne stany

$$X_1 = (S_1, V_1, N_1), \quad X_2 = (S_2, V_2, N_2)$$

oraz parameter $\lambda \in [0, 1]$. Zbudujmy układ z dwóch makroskopowych części: jedna znajduje się w stanie X_1 i zajmuje ułamek λ całości, druga w stanie X_2 i zajmuje ułamek $1 - \lambda$. Całkowite zmienne ekstensywne takiego układu wynoszą wtedy

$$S = \lambda S_1 + (1 - \lambda) S_2, \quad V = \lambda V_1 + (1 - \lambda) V_2, \quad N = \lambda N_1 + (1 - \lambda) N_2, \quad (13.1)$$

a jego energia jest równa

$$U_{\text{złożony}} = \lambda U(X_1) + (1 - \lambda) U(X_2). \quad (13.2)$$

Po usunięciu wewnętrznych więzów układ może zrelaksować do jednorodnego stanu równowagi o tych samych całkowitych S, V i N . Ponieważ przy ustalonych S, V, N stan równowagi minimalizuje energię wewnętrzną, dostajemy nierówność

$$U(\lambda X_1 + (1 - \lambda) X_2) \leq \lambda U(X_1) + (1 - \lambda) U(X_2). \quad (13.3)$$

A zatem $U(S, V, N)$ jest funkcją wypukłą swoich zmiennych ekstensywnych. Równoważnie, entropia $S(U, V, N)$ jest funkcją wklęsłą.

Dla funkcji gładkich wypukłość oznacza, że odpowiednie drugie pochodne są nieujemne. Dla układu jednoskładnikowego mamy

$$U = U(S, V, N), \quad dU = T dS - p dV + \mu dN. \quad (13.4)$$

Stąd wynika w szczególności

$$\left(\frac{\partial^2 U}{\partial S^2} \right)_{V, N} = \left(\frac{\partial T}{\partial S} \right)_{V, N} \geq 0 \quad (13.5)$$

oraz

$$\left(\frac{\partial^2 U}{\partial V^2} \right)_{S, N} = - \left(\frac{\partial p}{\partial V} \right)_{S, N} \geq 0. \quad (13.6)$$

Po odwróceniu odpowiednich pochodnych dostajemy więc

$$C_V = T \left(\frac{\partial S}{\partial T} \right)_{V, N} \geq 0, \quad \kappa_S = -\frac{1}{V} \left(\frac{\partial V}{\partial p} \right)_{S, N} \geq 0. \quad (13.7)$$

W stabilnym obszarze jednofazowym nierówności te są ostre.

Przejdźmy teraz do entalpii

$$H = U + pV, \quad dH = T dS + V dp + \mu dN. \quad (13.8)$$

Przy ustalonym ciśnieniu p ten sam argument z dzieleniem na podukłady pokazuje, że $H(S, p, N)$ jest funkcją wypukłą względem swoich naturalnych zmiennych ekstensywnych, w szczególności względem S . Otrzymujemy więc

$$\left(\frac{\partial^2 H}{\partial S^2} \right)_{p, N} = \left(\frac{\partial T}{\partial S} \right)_{p, N} \geq 0, \quad (13.9)$$

a zatem

$$C_p = T \left(\frac{\partial S}{\partial T} \right)_{p, N} \geq 0. \quad (13.10)$$

Z drugiej strony, ponieważ z wypukłości U wynika

$$\left(\frac{\partial p}{\partial V} \right)_{S, N} \leq 0, \quad (13.11)$$

to w stabilnym obszarze funkcja $V(p)$ jest nierosnąca, a z różniczki entalpii dostajemy

$$\left(\frac{\partial^2 H}{\partial p^2} \right)_{S, N} = \left(\frac{\partial V}{\partial p} \right)_{S, N} \leq 0. \quad (13.12)$$

Entalpia jest więc wypukła względem S , ale wklęsła względem p . Jest to typowy skutek transformacji Legendre'a: wypukłość względem zmiennej ekstensywnej przechodzi we wklęsłość względem sprzężonej z nią zmiennej intensywnej.

Pozostałe potencjały

Dokładnie tę samą logikę stosujemy do innych potencjałów. Dla energii swobodnej Helmholtza

$$F = U - TS, \quad dF = -S dT - p dV + \mu dN \quad (13.13)$$

otrzymujemy

$$\left(\frac{\partial^2 F}{\partial T^2}\right)_{V,N} = -\left(\frac{\partial S}{\partial T}\right)_{V,N} = -\frac{C_V}{T} \leq 0, \quad \left(\frac{\partial^2 F}{\partial V^2}\right)_{T,N} = -\left(\frac{\partial p}{\partial V}\right)_{T,N} \geq 0. \quad (13.14)$$

A więc $F(T, V, N)$ jest wklęsła względem T , ale wypukła względem V .

Dla funkcji Gibbsa

$$G = U - TS + pV, \quad dG = -S dT + V dp + \mu dN \quad (13.15)$$

mamy natomiast

$$\left(\frac{\partial^2 G}{\partial T^2}\right)_{p,N} = -\left(\frac{\partial S}{\partial T}\right)_{p,N} = -\frac{C_p}{T} \leq 0, \quad \left(\frac{\partial^2 G}{\partial p^2}\right)_{T,N} = \left(\frac{\partial V}{\partial p}\right)_{T,N} \leq 0. \quad (13.16)$$

Zatem $G(T, p, N)$ jest wklęsła zarówno względem T , jak i względem p .

Podobnie dla wielkiego potencjału termodynamicznego

$$\Omega = U - TS - \mu N, \quad d\Omega = -S dT - p dV - N d\mu \quad (13.17)$$

odpowiednie znaki drugich pochodnych są następujące:

$$\left(\frac{\partial^2 \Omega}{\partial T^2}\right)_{V,\mu} \leq 0, \quad \left(\frac{\partial^2 \Omega}{\partial V^2}\right)_{T,\mu} \geq 0, \quad \left(\frac{\partial^2 \Omega}{\partial \mu^2}\right)_{T,V} = -\left(\frac{\partial N}{\partial \mu}\right)_{T,V} \leq 0. \quad (13.18)$$

Krótko mówiąc: potencjały termodynamiczne są wypukłe względem swoich naturalnych zmiennych ekstensywnych i wklęsłe względem naturalnych zmiennych intensywnych.

14 Podatności termodynamiczne

„*She is susceptible
He is impossible
They have their cross to share
Three of a perfect pair*”
— King Crimson, *Three of a perfect pair*

Wielkości takie jak pojemność cieplna, ściśliwość czy współczynnik rozszerzalności cieplnej opisują odpowiedź układu pozostającego w stanie równowagi na niewielką zmianę warunków zewnętrznych. Nazywamy je *podatnościami termodynamicznymi*. Są one szczególnie ważne, bo z jednej strony można je bezpośrednio mierzyć doświadczalnie, a z drugiej strony ich znaki i wzajemne relacje są silnie ograniczone przez warunki stabilności równowagi.

Jeżeli do układu dostarczamy niewielką ilość ciepła przy ustalonym parametrze zewnętrznym A , to pojemność cieplna jest zdefiniowana jako

$$C_A = \left(\frac{dQ}{dT}\right)_A = T \left(\frac{\partial S}{\partial T}\right)_A. \quad (14.1)$$

Na jedną cząstkę można analogicznie zdefiniować ciepło właściwe

$$c_A = T \left(\frac{\partial s}{\partial T}\right)_A. \quad (14.2)$$

Najczęściej spotykamy pojemności przy stałej objętości i przy stałym ciśnieniu:

$$C_V = T \left(\frac{\partial S}{\partial T}\right)_V, \quad C_p = T \left(\frac{\partial S}{\partial T}\right)_p. \quad (14.3)$$

Dla pary sprzężonej $(-p, V)$ naturalną podatnością jest ściśliwość:

$$\kappa_A = -\frac{1}{V} \left(\frac{\partial V}{\partial p} \right)_A. \quad (14.4)$$

W szczególności

$$\kappa_T = -\frac{1}{V} \left(\frac{\partial V}{\partial p} \right)_T, \quad \kappa_S = -\frac{1}{V} \left(\frac{\partial V}{\partial p} \right)_S. \quad (14.5)$$

Wielkości takie jak C_A czy κ_A , w których pojawiają się elementy tej samej pary sprzężonych zmiennych, można traktować jako przykłady *efektów prostych*. Oprócz nich istnieją także *efekty krzyżowe*, w których mieszają się różne pary zmiennych termodynamicznych. Najprostszym przykładem jest współczynnik rozszerzalności cieplnej:

$$\alpha_p = \frac{1}{V} \left(\frac{\partial V}{\partial T} \right)_p. \quad (14.6)$$

Efektowi krzyżowemu zawsze towarzyszy sprzężony efekt krzyżowy, wynikający z odpowiedniej tożsamości Maxwella. Na przykład z różniczki funkcji Gibbsa

$$dG = -S dT + V dp \quad (14.7)$$

wynika

$$\left(\frac{\partial V}{\partial T} \right)_p = - \left(\frac{\partial S}{\partial p} \right)_T. \quad (14.8)$$

Oznacza to, że jeśli objętość silnie reaguje na zmianę temperatury przy stałym ciśnieniu, to entropia równie silnie reaguje na zmianę ciśnienia przy stałej temperaturze.

Znaki podstawowych podatności wynikają bezpośrednio z rozważań z poprzedniego rozdziału:

$$C_V \geq 0, \quad C_p \geq 0, \quad \kappa_S \geq 0, \quad \kappa_T \geq 0. \quad (14.9)$$

W stabilnym obszarze jednofazowym nierówności te są ostre. Dodatnia pojemność cieplna oznacza, że dostarczenie ciepła nie obniża temperatury układu. Dodatnia ściśliwość oznacza, że wzrost ciśnienia zmniejsza objętość. Przeciwnie znaki prowadziłyby do niestabilności.

Stabilność mówi jednak więcej niż tylko to, po której stronie zera leżą podatności. Pozwala również porównywać odpowiedzi mierzone przy różnych więzach. Dlatego pojemność cieplna przy stałym ciśnieniu i przy stałej objętości, podobnie jak ściśliwość izotermiczna i adiabatyczna, nie są od siebie niezależne.

Dlaczego $C_p \geq C_V$?

Wygodnie zacząć od prostego rachunku ogólnego. Niech

$$df(x, y) = M(x, y) dx + N(x, y) dy. \quad (14.10)$$

Wykonujemy transformację Legendre'a po zmiennej y :

$$g = f - Ny, \quad dg = M dx - y dN. \quad (14.11)$$

Z drugiej z tych równości wynika tożsamość Maxwella

$$\left(\frac{\partial M}{\partial N} \right)_x = - \left(\frac{\partial y}{\partial x} \right)_N. \quad (14.12)$$

Z reguły łańcuchowej dostajemy następnie

$$\left(\frac{\partial M}{\partial x} \right)_y = \left(\frac{\partial M}{\partial x} \right)_N + \left(\frac{\partial M}{\partial N} \right)_x \left(\frac{\partial N}{\partial x} \right)_y. \quad (14.13)$$

Korzystając z tożsamości “minus jedynki” dla funkcji $N(x, y)$,

$$\left(\frac{\partial y}{\partial x}\right)_N = -\frac{\left(\frac{\partial N}{\partial x}\right)_y}{\left(\frac{\partial N}{\partial y}\right)_x}, \quad (14.14)$$

otrzymujemy po krótkim rachunku

$$\left(\frac{\partial M}{\partial x}\right)_y - \left(\frac{\partial M}{\partial x}\right)_N = \left(\frac{\partial M}{\partial N}\right)_x \left(\frac{\partial N}{\partial y}\right)_x. \quad (14.15)$$

Prawa strona jest nieujemna, jeśli tylko

$$\left(\frac{\partial N}{\partial y}\right)_x \geq 0. \quad (14.16)$$

Wtedy mamy nierówność

$$\left(\frac{\partial M}{\partial x}\right)_y \geq \left(\frac{\partial M}{\partial x}\right)_N. \quad (14.17)$$

Zastosujmy teraz ten wynik do energii wewnętrznej

$$dU = T dS - p dV. \quad (14.18)$$

Bierzemy zatem

$$x = S, \quad y = V, \quad M = T, \quad N = -p. \quad (14.19)$$

Ponieważ

$$\left(\frac{\partial N}{\partial y}\right)_x = \left(\frac{\partial(-p)}{\partial V}\right)_S = -\left(\frac{\partial p}{\partial V}\right)_S \geq 0, \quad (14.20)$$

z nierówności (14.17) dostajemy

$$\left(\frac{\partial T}{\partial S}\right)_{V,N} \geq \left(\frac{\partial T}{\partial S}\right)_{p,N} \geq 0. \quad (14.21)$$

A ponieważ

$$\left(\frac{\partial T}{\partial S}\right)_{V,N} = \frac{T}{C_V}, \quad \left(\frac{\partial T}{\partial S}\right)_{p,N} = \frac{T}{C_p}, \quad (14.22)$$

przy dodatniej temperaturze otrzymujemy

$$\frac{1}{C_V} \geq \frac{1}{C_p} \geq 0, \quad (14.23)$$

czyli

$$C_p \geq C_V \geq 0. \quad (14.24)$$

Dlaczego $\kappa_T \geq \kappa_S$?

Stosujemy ten sam rachunek jeszcze raz, ale tym razem zamieniamy role zmiennych:

$$x = V, \quad y = S, \quad M = -p, \quad N = T. \quad (14.25)$$

Warunek nieujemności prawej strony wzoru (14.15) ma teraz postać

$$\left(\frac{\partial N}{\partial y}\right)_x = \left(\frac{\partial T}{\partial S}\right)_{V,N} \geq 0, \quad (14.26)$$

co właśnie wynika z dodatniości C_V . Otrzymujemy więc

$$-\left(\frac{\partial p}{\partial V}\right)_{S,N} \geq -\left(\frac{\partial p}{\partial V}\right)_{T,N} \geq 0. \quad (14.27)$$

Po odwróceniu pochodnych dostajemy

$$-\left(\frac{\partial V}{\partial p}\right)_{T,N} \geq -\left(\frac{\partial V}{\partial p}\right)_{S,N} \geq 0, \quad (14.28)$$

a więc

$$\kappa_T \geq \kappa_S \geq 0. \quad (14.29)$$

Nierówności

$$C_p \geq C_V, \quad \kappa_T \geq \kappa_S$$

można rozumieć jako dwa przykłady tej samej zasady Le Chateliera-Brauna. Wygodnie sformułować ją w języku par zmiennych sprzężonych.

Niech małe zaburzenie zmiennej ekstensywnej X_1 wywołuje zmianę sprzężonej z nią zmiennej intensywnej P_1 . Zasada Le Chateliera mówi, że bezpośrednia odpowiedź układu działa tak, aby osłabić tę zmianę P_1 . Zasada Le Chateliera-Brauna dodaje, że jeśli zaburzenie uruchamia także procesy wtórne, związane z innymi parami zmiennych sprzężonych, to one również osłabiają początkową zmianę P_1 .

W przypadku pojemności cieplnych zaburzeniem jest dopływ ciepła, czyli wzrost entropii. Zmienną sprzężoną z entropią jest temperatura. Przy ustalonej objętości wzrost entropii powoduje wzrost temperatury opisany pochodną

$$\left(\frac{\partial T}{\partial S}\right)_V.$$

Jeśli jednak układ pozostaje pod stałym ciśnieniem, może uruchomić dodatkowy kanał odpowiedzi: zmianę objętości. Ten proces wtórny osłabia początkowy wzrost temperatury, dlatego

$$\left(\frac{\partial T}{\partial S}\right)_p \leq \left(\frac{\partial T}{\partial S}\right)_V.$$

Ponieważ

$$\left(\frac{\partial T}{\partial S}\right)_V = \frac{T}{C_V}, \quad \left(\frac{\partial T}{\partial S}\right)_p = \frac{T}{C_p},$$

otrzymujemy

$$C_p \geq C_V.$$

W przypadku ściśliwości zaburzeniem jest zmiana objętości, a zmienną sprzężoną z objętością jest $-p$. Przy ustalonej entropii zmiana objętości wywołuje zmianę ciśnienia opisaną pochodną

$$-\left(\frac{\partial p}{\partial V}\right)_S.$$

Jeśli układ pozostaje w kontakcie cieplnym z otoczeniem, pojawia się dodatkowy kanał odpowiedzi: wymiana ciepła. Ten proces wtórny osłabia zmianę ciśnienia wywołaną tą samą zmianą objętości, dlatego

$$-\left(\frac{\partial p}{\partial V}\right)_T \leq -\left(\frac{\partial p}{\partial V}\right)_S.$$

Po odwróceniu pochodnych dostajemy równoważnie

$$\kappa_T \geq \kappa_S.$$

W obu przypadkach schemat jest więc ten sam. Zaburzamy zmienną ekstensywną i patrzymy na zmianę zmiennej intensywnej z nią sprzężonej. Gdy dopuszczony jest dodatkowy kanał odpowiedzi, zmiana tej zmiennej intensywnej jest mniejsza. Dlatego przy stałym ciśnieniu wzrost temperatury jest mniejszy niż przy stałej objętości, a przy stałej temperaturze zmiana ciśnienia jest mniejsza niż przy stałej entropii.

15 Przejścia fazowe

*„All of the handsome fiction will melt away
And when the flame burns brighter, evaporate”
— Dance Gavin Dance, Evaporate*

Do tej pory rozważaliśmy głównie układy jednorodne, występujące w jednej fazie. Tymczasem ta sama substancja może występować w kilku różnych fazach, różniących się własnościami fizycznymi, mimo że jej skład chemiczny pozostaje ten sam. Najbardziej oczywiste przykłady to gaz, ciecz i ciało stałe, ale fazami są także ferromagnetyk i paramagnetyk, nadprzewodnik i zwykły metal, czy nadciekła i zwykła faza helu.

15.1 Układy wieloskładnikowe i wielofazowe

Dotychczas w naszych rozważaniach zakładaliśmy najczęściej, że mamy do czynienia z czystą substancją. Jeśli jednak układ zawiera kilka składników, to wzór na różniczkę entropii uogólnia się do postaci

$$dS = \frac{1}{T} \left(dU + p dV - \sum_i \mu_i dN_i \right). \quad (15.1)$$

Widzimy stąd, że potencjał chemiczny i -tego składnika można zdefiniować jako

$$\mu_i = -T \left(\frac{\partial S}{\partial N_i} \right)_{U, V, N_{j \neq i}}. \quad (15.2)$$

Przechodząc do energii wewnętrznej, dostajemy równoważnie

$$dU = T dS - p dV + \sum_i \mu_i dN_i. \quad (15.3)$$

Człon $\mu_i dN_i$ można więc interpretować jako wkład do energii związany z dopływem substancji i .

O wyznaczaniu potencjału chemicznego

Dla układu jednoskładnikowego zaczniemy od wyznaczenia równania podstawowego

$$S = S(U, V, N). \quad (15.4)$$

Można to zrobić następująco. Najpierw rozważamy jeden mol substancji i doświadczalnie wyznaczamy zależności

$$\frac{1}{T} = \frac{1}{T}(U, V, 1), \quad (15.5)$$

oraz

$$\frac{p}{T} = \frac{p}{T}(U, V, 1). \quad (15.6)$$

Korzystamy teraz z relacji

$$\left(\frac{\partial S}{\partial U} \right)_{V, N} = \frac{1}{T}, \quad \left(\frac{\partial S}{\partial V} \right)_{U, N} = \frac{p}{T}. \quad (15.7)$$

Znajomość funkcji $\frac{1}{T}(U, V, 1)$ oraz $\frac{p}{T}(U, V, 1)$ pozwala więc, przynajmniej w zasadzie, odtworzyć entropię jednego mola,

$$S = S(U, V, 1), \quad (15.8)$$

z dokładnością do stałej addytywnej.

Dopiero teraz przechodzimy do układu zawierającego N moli. Korzystamy z ekstensywności entropii:

$$S(\lambda U, \lambda V, \lambda N) = \lambda S(U, V, N). \quad (15.9)$$

Podstawiając $\lambda = \frac{1}{N}$, dostajemy

$$S(U, V, N) = N S\left(\frac{U}{N}, \frac{V}{N}, 1\right). \quad (15.10)$$

A zatem znajomość równania podstawowego dla jednego mola wystarcza do wyznaczenia go dla dowolnej liczby moli.

Mając już funkcję $S(U, V, N)$, możemy obliczyć potencjał chemiczny:

$$\mu = -T \left(\frac{\partial S}{\partial N} \right)_{U, V}. \quad (15.11)$$

W przypadku mieszaniny sytuacja jest podobna, ale dla każdego składnika mamy osobny potencjał chemiczny. Ogólnie

$$\mu_i = -T \left(\frac{\partial S}{\partial N_i} \right)_{U, V, N_{j \neq i}}. \quad (15.12)$$

Potencjał chemiczny składnika w mieszaninie można wyznaczyć myślowo w następujący sposób. Tworzymy układ złożony z dwóch części: w jednej znajduje się badana mieszanina, w drugiej czysty składnik i . Części te oddzielone są błoną półprzepuszczalną, która przepuszcza tylko ten składnik. Jeżeli układ pozostaje w równowadze, to potencjały chemiczne po obu stronach muszą być równe. Ponieważ potencjał chemiczny czystego składnika potrafimy wyznaczyć niezależnie, dostajemy w ten sposób także potencjał chemiczny tego składnika w mieszaninie.

15.2 Warunki współistnienia faz

Nawet dla czystej substancji możliwe jest współistnienie kilku różnych faz, które nie różnią się składem chemicznym, ale mają inne własności fizyczne. Klasycznym przykładem są lód, woda i para wodna.

Jeśli dwie fazy pozostają w równowadze termodynamicznej, to muszą mieć tę samą temperaturę, to samo ciśnienie oraz ten sam potencjał chemiczny. Dla fazy ciekłej i parowej oznacza to

$$T_c = T_p, \quad p_c = p_p, \quad \mu_c = \mu_p. \quad (15.13)$$

Ogólniej, załóżmy, że mamy f faz: $\alpha, \beta, \dots$, a w każdej z nich występuje r składników chemicznych. Wtedy warunki równowagi mają postać

$$\begin{aligned} T^\alpha &= T^\beta = \dots, & (\text{równowaga termiczna}), \\ p^\alpha &= p^\beta = \dots, & (\text{równowaga mechaniczna}), \\ \mu_1^\alpha &= \mu_1^\beta = \dots, & (\text{równowaga chemiczna}), \\ \mu_2^\alpha &= \mu_2^\beta = \dots, \\ &\vdots \\ \mu_r^\alpha &= \mu_r^\beta = \dots \end{aligned} \quad (15.14)$$

Dla opisu przejść fazowych najwygodniejszym potencjałem termodynamicznym jest energia swobodna Gibbsa,

$$G = U - TS + pV. \quad (15.15)$$

Jej różniczka dla układu wieloskładnikowego ma postać

$$dG = -S dT + V dp + \sum_i \mu_i dN_i. \quad (15.16)$$

Dodatkowo, dla każdej fazy osobno możemy napisać relację Gibbsa-Duhema

$$-S^\alpha dT + V^\alpha dp - \sum_i N_i^\alpha d\mu_i^\alpha = 0. \quad (15.17)$$

Policzmy teraz liczbę niezależnych parametrów intensywnych. Mamy $r + 2$ wielkości intensywnych: T , p oraz $\mu_1, \dots, \mu_r$. Jednak dla każdej z f faz zachodzi relacja Gibbsa-Duhema, a ponadto w stanie współistnienia potencjały chemiczne, temperatura i ciśnienie są równe we wszystkich fazach. Ostatecznie liczba niezależnych parametrów intensywnych wynosi

$$w = r + 2 - f. \quad (15.18)$$

Jest to **reguła faz Gibbsa**.

Dla substancji jednoskładnikowej, czyli dla $r = 1$, dostajemy:

- dla jednej fazy, $f = 1$, mamy $w = 2$, więc można niezależnie wybierać na przykład temperaturę i ciśnienie;
- dla dwóch współistniejących faz, $f = 2$, mamy $w = 1$, więc współistnienie zachodzi na krzywej na płaszczyźnie (p, T) ;
- dla trzech współistniejących faz, $f = 3$, mamy $w = 0$, więc odpowiada temu pojedynczy punkt, czyli **punkt potrójny**.

15.3 Równanie Clapeyrona

Rozważmy teraz dwie współistniejące fazy jednoskładnikowe, na przykład ciecz i parę. Dla każdej z nich relacja Gibbsa-Duhema przyjmuje postać

$$-S_p dT + V_p dp - N_p d\mu = 0, \quad (15.19)$$

$$-S_c dT + V_c dp - N_c d\mu = 0. \quad (15.20)$$

Dzieląc przez liczbę cząstek lub moli i przechodząc do wielkości właściwych, dostajemy

$$-s_p dT + v_p dp - d\mu = 0, \quad (15.21)$$

$$-s_c dT + v_c dp - d\mu = 0. \quad (15.22)$$

Odejmując te dwa równania stronami, otrzymujemy

$$(s_c - s_p) dT + (v_p - v_c) dp = 0. \quad (15.23)$$

Jeśli oznaczymy przez

$$\Delta v = v_p - v_c \quad (15.24)$$

przyrost objętości właściwej przy parowaniu oraz przez

$$\Delta s = s_p - s_c \quad (15.25)$$

przyrost entropii właściwej, to ponieważ w przemianie fazowej w stałej temperaturze

$$\Delta s = \frac{q_p}{T}, \quad (15.26)$$

gdzie $q_p = T\Delta s$ jest ciepłem parowania na mol lub na cząstkę, dostajemy

$$-\frac{q_p}{T} dT + \Delta v dp = 0. \quad (15.27)$$

Stąd

$$\frac{dp}{dT} = \frac{q_p}{T \Delta v}. \quad (15.28)$$

Jest to **równanie Clapeyrona**. Przejście fazowe, w którym $\Delta s \neq 0$ oraz $\Delta v \neq 0$, nazywamy **przejściem pierwszego rodzaju** albo **przejściem nieciągłym**. Wielkość $q_p = T \Delta s$ to **ciepło utajone** przemiany. Oznacza ono, że w czasie przejścia fazowego można dostarczać lub odbierać energię bez zmiany temperatury, zmieniając jedynie udział obu faz.

Dla przemiany ciecż–para zwykle mamy $q_p > 0$ oraz $\Delta v > 0$, a więc

$$\frac{dp}{dT} > 0. \quad (15.29)$$

Dla topnienia większości substancji znak jest taki sam, ale dla wody jest odwrotnie, bo lód ma większą objętość właściwą niż ciecz. Dlatego linia topnienia wody ma na diagramie fazowym nachylenie ujemne.

Jeśli dodatkowo przybliżymy parę gazem doskonałym, tak że

$$v_p = \frac{RT}{p}, \quad (15.30)$$

oraz założymy, że $v_p \gg v_c$, czyli $\Delta v \simeq v_p$, to dostajemy

$$\frac{dp}{dT} = \frac{q_p p}{RT^2}. \quad (15.31)$$

Stąd

$$\frac{d \ln p}{dT} = \frac{q_p}{RT^2}. \quad (15.32)$$

Jeśli jeszcze przyjmiemy, że q_p słabo zależy od temperatury, to po scałkowaniu mamy

$$p(T) = C \exp\left(-\frac{q_p}{RT}\right), \quad (15.33)$$

czyli przybliżoną postać równania Clausiusa-Clapeyrona.

Linia współistnienia ciecż–para kończy się w punkcie krytycznym. W miarę zbliżania się do tego punktu gęstość pary rośnie, a gęstość cieczy maleje. W punkcie krytycznym różnica między cieczą i gazem zanika. Powyżej temperatury krytycznej nie można już przeprowadzić ostrego rozróżnienia między tymi fazami; mówimy wtedy o płynie nadkrytycznym.

15.4 Przejścia pierwszego rodzaju i przejścia ciągłe

Nie wszystkie przejścia fazowe mają postać linii współistnienia dwóch faz rozdzielonych granicą fazową. Istnieją też przejścia, w których układ zmienia swój stan w sposób ciągły, bez ciepła utajonego i bez skokowej zmiany objętości. Takie przejścia nazywamy **przejściami ciągłymi** albo **przejściami krytycznymi**.

Klasycznymi przykładami są przejście paramagnetyk–ferromagnetyk, przejście do stanu nadprzewodzącego oraz przejście do stanu nadciekłego. W takich układach nie chodzi zwykle o zmianę stanu skupienia, lecz o pojawienie się nowego rodzaju uporządkowania, np. spontanicznego namagnesowania dla ferromagnetyka.

Do opisu przejść ciągłych wprowadza się **parametr porządku**, czyli wielkość, która po jednej stronie przejścia jest równa zero, a po drugiej stronie przyjmuje wartość niezerową. Dla ferromagnetyka parametrem porządku jest namagnesowanie:

$$M = 0 \quad \text{dla } T > T_c, \quad M \neq 0 \quad \text{dla } T < T_c. \quad (15.34)$$

Powyżej temperatury Curie układ nie ma wyróżnionego kierunku namagnesowania, natomiast poniżej tej temperatury taki kierunek pojawia się spontanicznie.

Termodynamicznie przejście ciągłe oznacza, że pierwsze pochodne energii Gibbsa pozostają ciągłe, natomiast osobliwości pojawiają się dopiero w funkcjach odpowiedzi. Do takich funkcji odpowiedzi należą na przykład molowa pojemność cieplna, ściśliwość izotermiczna oraz współczynnik rozszerzalności cieplnej:

$$c_p = -T \left(\frac{\partial^2 g}{\partial T^2} \right)_p, \quad (15.35)$$

$$\kappa_T = -\frac{1}{v} \left(\frac{\partial^2 g}{\partial p^2} \right)_T, \quad (15.36)$$

$$\alpha_p = \frac{1}{v} \left(\frac{\partial^2 g}{\partial T \partial p} \right). \quad (15.37)$$

W pobliżu przejścia ciągłego wielkości te mogą mieć skoki, ostre maksima albo rozbieżności. Na przykład w pobliżu temperatury krytycznej często zapisuje się zależności typu

$$c_p \sim |T - T_c|^{-\alpha_{cr}}, \quad \chi \sim |T - T_c|^{-\gamma}, \quad (15.38)$$

gdzie χ oznacza podatność, na przykład podatność magnetyczną. Wykładniki α_{cr} i γ nazywamy **wykładnikami krytycznymi**.

Szczególnie ważną wielkością jest **długość korelacji** ξ . Mówi ona, na jak dużych odległościach fluktuacje układu są ze sobą powiązane. W pobliżu punktu krytycznego często

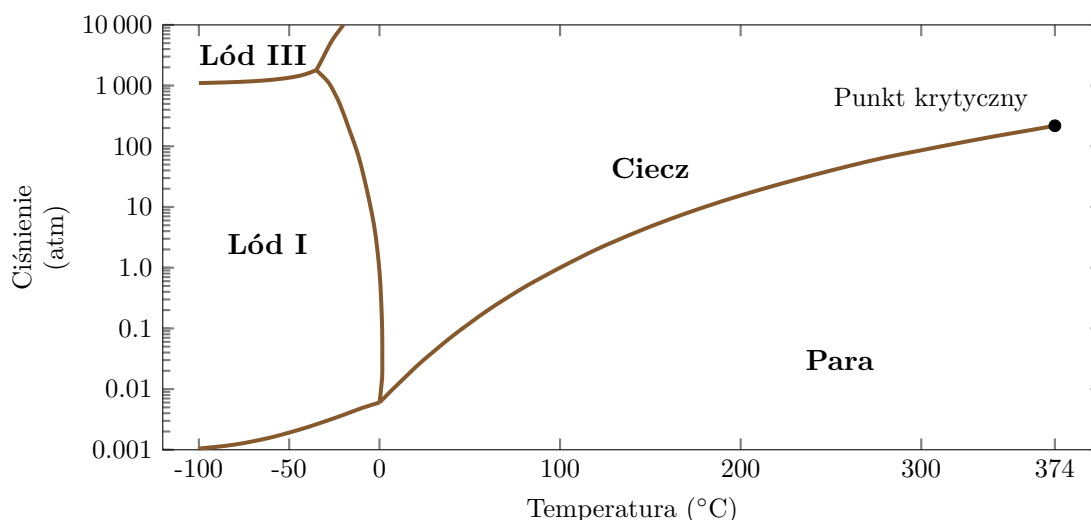
$$\xi \sim |T - T_c|^{-\nu}. \quad (15.39)$$

Oznacza to, że fluktuacje obejmują coraz większe obszary układu. W samym punkcie krytycznym układ traci charakterystyczną skalę długości.

15.5 Diagram fazowy

Diagram fazowy pokazuje, która faza substancji jest stabilna przy zadanych warunkach zewnętrznych. W przypadku substancji jednoskładnikowej najczęściej rysujemy go w płaszczyźnie p - T . Obszary na takim diagramie odpowiadają pojedynczym fazom, na przykład fazie stałej, ciekłej lub gazowej. Linie oddzielające te obszary są liniami współistnienia dwóch faz.

Na początek rzućmy okiem na diagram fazowy wody:

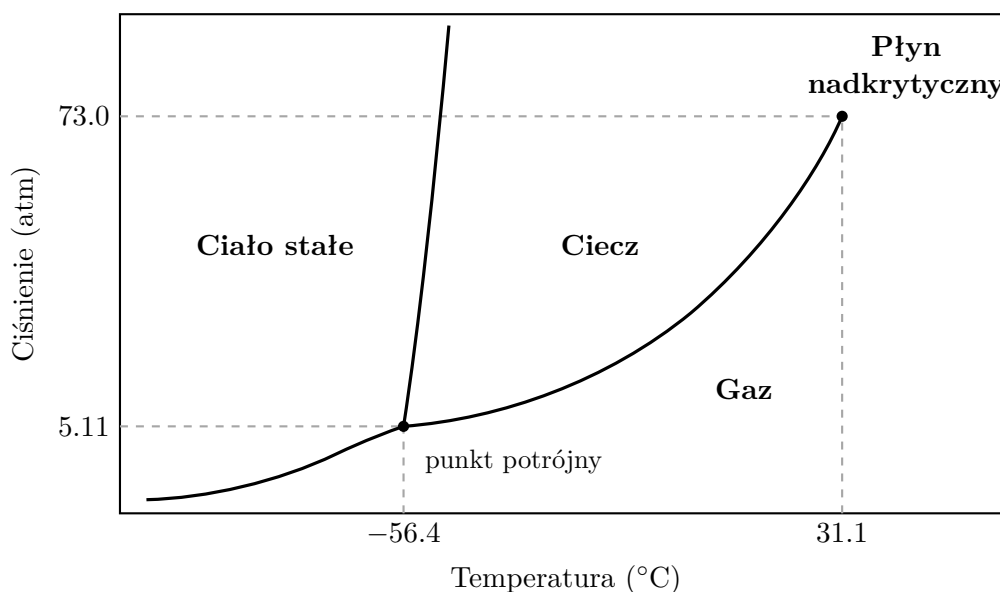


Punkt potrójny leży przy temperaturze 0.01°C i ciśnieniu około 0.006 atm. W tym punkcie mogą współistnieć lód, ciekła woda i para wodna.

Najbardziej charakterystyczną cechą diagramu fazowego wody jest ujemne nachylenie linii równowagi lód I–ciecz. Wynika ono z tego, że lód I ma mniejszą gęstość niż woda w temperaturze topnienia, zatem $\Delta v = v_{\text{ciecz}} - v_{\text{lód}} < 0$, co prowadzi do $\left(\frac{dp}{dT}\right)_{\text{topnienia}} < 0$. Zwiększenie ciśnienia obniża więc temperaturę topnienia lodu I. Dla większości substancji jest odwrotnie: faza stała jest gęstsza od ciekłej, a linia topnienia ma nachylenie dodatnie.

Warto też pamiętać, że lód występujący w warunkach codziennych to tylko jedna z odmian krystalicznych wody, oznaczana jako lód I. Przy wyższych ciśnieniach stabilne stają się inne, zwykle gęstsze fazy lodu, na przykład lód III, V, VI czy VII. Pełny diagram fazowy wody jest więc znacznie bogatszy niż prosty schemat z trzema fazami. Na rysunku zaznaczamy tylko wybrane obszary, aby podkreślić najważniejsze własności: punkt potrójny, punkt krytyczny oraz nietypowe nachylenie linii lód I–ciecz.

Prostszy i bardziej typowy od diagramu wody jest diagram fazowy dwutlenku węgla:

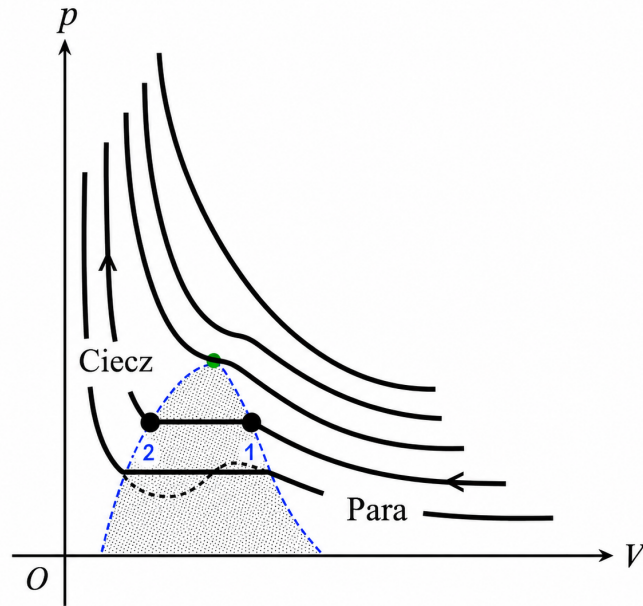


W przypadku CO₂ punkt potrójny leży przy temperaturze około 217 K, czyli -56.4°C, oraz ciśnieniu 5.11 atm. Punkt krytyczny znajduje się przy temperaturze około 304 K, czyli 31.1°C, oraz ciśnieniu około 72.8 atm.

Ponieważ ciśnienie punktu potrójnego CO₂ jest większe od ciśnienia atmosferycznego, dwutlenek węgla przy 1 atm nie może występować jako ciecz. Podczas ogrzewania przechodzi bezpośrednio z fazy stałej do gazowej, czyli sublimuje. Stąd bierze się nazwa *suchy lód*: ciało stałe znika, nie pozostawiając cieczy.

Warto pamiętać, że diagram (p, T) nie zawiera informacji o wielkościach ekstensywnych, a więc nie rozróżnia stanów różniących się ilościami współlistniejących faz. Pojedynczemu punktowi na linii współlistnienia w diagramie (p, T) odpowiada na diagramie (p, V) cały odcinek.

Najlepiej widać to dla izoterm poniżej temperatury krytycznej. Podczas sprężania gazu, po osiągnięciu ciśnienia współlistnienia, gaz zaczyna się skraplać. Dalsze zmniejszanie objętości nie zwiększa ciśnienia, lecz zmienia proporcje między parą i cieczą.



Jeżeli całkowita liczba moli wynosi N , to na początku odcinka współistnienia mamy

$$V_1 = Nv_g, \quad (15.40)$$

a na jego końcu

$$V_2 = Nv_c. \quad (15.41)$$

Wewnątrz obszaru współistnienia objętość ma postać

$$V = N_1v_g + N_2v_c, \quad N_1 + N_2 = N. \quad (15.42)$$

Na diagramie (p, V) punkt współistnienia z diagramu (p, T) odpowiada więc poziomemu odcinkowi.

W pobliżu punktu krytycznego (oznaczonego zieloną kropką na powyższym wykresie) ściśliwość izotermiczna rośnie, a w idealizacji może stać się nieskończona:

$$\kappa_T = -\frac{1}{V} \left(\frac{\partial V}{\partial p} \right)_T. \quad (15.43)$$

Związek z fluktuacjami widać szczególnie wyraźnie w wielkim zespole kanonicznym, gdzie pokazywaliśmy, że

$$\sigma_N^2 = \frac{kT}{V} \langle N \rangle^2 \kappa_T. \quad (15.44)$$

A zatem wielka ściśliwość oznacza ogromne fluktuacje gęstości. To właśnie one odpowiadają za **opalescencję krytyczną**, czyli silne rozpraszanie światła przez układ blisko punktu krytycznego.

15.6 Model van der Waalsa

Modele oparte na gazie doskonałym nie są w stanie opisać przejścia ciecz–gaz, ponieważ nie ma w nim mechanizmu kondensacji. Aby otrzymać choćby jakościowy model tego przejścia, trzeba uwzględnić dwa efekty nieobecne w równaniu gazu doskonałego:

1. skończony rozmiar cząsteczek,
2. przyciąganie między cząsteczkami.

Poprawka objętościowa. W gazie doskonałym cząsteczki traktowaliśmy jak punkty materialne. Rzeczywiste cząsteczki mają jednak skończony rozmiar, dlatego część objętości naczynia jest niedostępna dla ruchu pozostałych cząsteczek. Jeśli wyobrazimy je sobie jako twarde kule o promieniu r_0 , to środek jednej cząsteczki nie może znaleźć się w odległości mniejszej niż $2r_0$ od środka drugiej. Oznacza to pojawienie się objętości wykluczonej.

W najprostszym przybliżeniu efekt ten uwzględniamy, zastępując całkowitą objętość V objętością dostępną dla ruchu cząsteczek:

$$V_{\text{eff}} = V - nb, \quad (15.45)$$

gdzie n jest liczbą moli, a b stałą charakteryzującą objętość wykluczoną przez jeden mol cząsteczek.

Poprawka ciśnieniowa. Druga poprawka wynika z przyciągania międzycząsteczkowego. Cząsteczka uderzająca w ściankę jest przyciągana przez cząsteczki znajdujące się głębiej w naczyniu, dlatego przekazuje ściance nieco mniejszy pęd niż w gazie doskonałym. Mierzone ciśnienie jest więc mniejsze od ciśnienia, jakie gaz wywierałby bez oddziaływań przyciągających.

Wielkość tej poprawki powinna być proporcjonalna do liczby oddziałujących par cząsteczek, a więc do kwadratu gęstości molowej:

$$\Delta p \sim \left(\frac{n}{V}\right)^2. \quad (15.46)$$

Dlatego w równaniu stanu należy użyć ciśnienia skorygowanego

$$p_{\text{eff}} = p + a \left(\frac{n}{V}\right)^2, \quad (15.47)$$

gdzie a jest stałą mierzącą siłę przyciągania międzycząsteczkowego.

Wstawiając obie poprawki do równania gazu doskonałego, otrzymujemy **równanie van der Waalsa**

$$\left(p + \frac{a}{v^2}\right)(v - b) = RT, \quad (15.48)$$

gdzie $v = V/n$ jest objętością molową. Przy wysokich temperaturach dominuje wkład kinetyczny, natomiast przy niższych temperaturach i większych gęstościach coraz ważniejsze staje się przyciąganie międzycząsteczkowe. To właśnie ono pozwala temu prostemu modelowi opisać jakościowo kondensację i przejście ciec–gaz.

Dla wysokich temperatur izotermy $p(v)$ są monotoniczne, podobnie jak w gazie doskonałym. Dla temperatur dostatecznie niskich pojawia się natomiast charakterystyczna pętla van der Waalsa: krzywa ma lokalne minimum i maksimum (linia przerywana na rys. X). Jest to pierwszy sygnał, że model zawiera informację o przejściu fazowym. W pewnym zakresie objętości molowych otrzymujemy bowiem

$$\left(\frac{\partial p}{\partial v}\right)_T > 0. \quad (15.49)$$

co odpowiada ujemnej ściśliwości izotermicznej. Taki stan jest niestabilny: niewielka fluktuacja gęstości nie byłaby tłumiona, lecz wzmacniana. Środkowa część pętli van der Waalsa nie może więc opisywać rzeczywistego stanu równowagi. Jest to sygnał, że poniżej temperatury krytycznej jednorodny opis układu musi zostać zastąpiony opisem współistnienia cieczy i gazu.

15.7 Konstrukcja Maxwella

Rozważmy izotermę dla $T < T_c$. Dla pewnego zakresu ciśnień równanie van der Waalsa ma trzy rozwiązania względem objętości molowej. Najmniejsza objętość odpowiada fazie ciekłej, największa fazie gazowej, natomiast gałąź środkowa przechodzi przez obszar niestabilny, w którym $\left(\frac{\partial p}{\partial v}\right)_T > 0$. W rzeczywistym układzie jednorodny stan o takiej objętości nie jest stanem równowagi. Zamiast niego pojawia się separacja faz: część substancji tworzy ciecz, a część gaz.

Na wykresie (p, v) oznacza to, że rzeczywista izoterma nie biegnie po falującym fragmencie krzywej van der Waalsa. Zostaje on zastąpiony poziomym odcinkiem przy pewnym ciśnieniu $p_0(T)$. Podczas sprężania w tej temperaturze objętość zmienia się od v_g do v_l , ale ciśnienie pozostaje stałe. Wykonywana praca nie ściska już jednorodnego gazu, lecz zamienia kolejne porcje gazu w ciecz.

Można to ująć jeszcze inaczej: poniżej temperatury krytycznej model van der Waalsa sam sygnalizuje, że jednorodne stany o pośrednich objętościach nie są stabilne. Układ obniża energię swobodną, rozdzielając się na dwie fazy o objętościach molowych v_l i v_g . Dla objętości pośrednich zmieniają się jedynie proporcje cieczy i gazu, a nie objętości molowe samych faz.

Tak czy inaczej pętla van der Waalsa nie opisuje poprawnie stanu równowagi. Trzeba ją zastąpić poziomym odcinkiem odpowiadającym współistnieniu cieczy i gazu. Powstaje pytanie: jak wyznaczyć ciśnienie $p_0(T)$ tego poziomego odcinka? Warunkiem równowagi faz przy ustalonej temperaturze jest równość potencjałów chemicznych, czyli molowych energii Gibbsa:

$$g_l(T, p_0) = g_g(T, p_0). \quad (15.50)$$

Aby przełożyć ten warunek na zależność geometryczną na wykresie (p, v) , przejdźmy do energii swobodnej, której różniczka, przy stałej temperaturze to

$$dF = -p dV, \quad (15.51)$$

Dla wielkości molowej $f = F/n$ daje to

$$\left(\frac{\partial f}{\partial v}\right)_T = -p. \quad (15.52)$$

Stąd

$$f(v_g, T) - f(v_l, T) = - \int_{v_l}^{v_g} p(v, T) dv. \quad (15.53)$$

Z drugiej strony, warunek współistnienia faz

$$g_l = g_g, \quad (15.54)$$

możemy zapisać w języku energii swobodnej, korzystając z tego, iż $g = f + pv$. Wtedy przy ciśnieniu p_0 , dostaniemy

$$f(v_g, T) - f(v_l, T) = -p_0(v_g - v_l). \quad (15.55)$$

Porównanie obu wyrażeń prowadzi do warunku

$$- \int_{v_l}^{v_g} p(v, T) dv = -p_0(v_g - v_l), \quad (15.56)$$

czyli

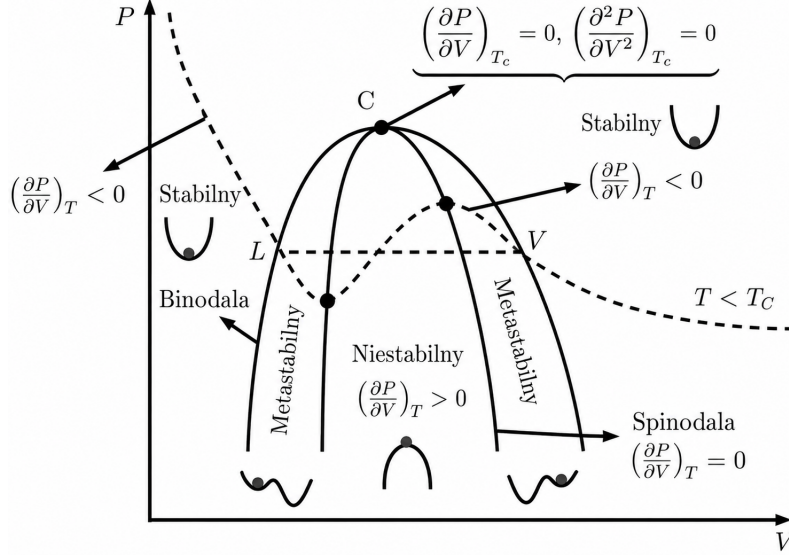
$$\int_{v_l}^{v_g} [p(v, T) - p_0] dv = 0. \quad (15.57)$$

Jest to **konstrukcja Maxwella**, zwana też **regułą równych pól**.

Jej sens geometryczny jest prosty. Pozioma prosta $p = p_0$ musi być dobrana tak, aby pole między izotermą van der Waalsa a tą prostą powyżej prostej było równe polu poniżej prostej. Wtedy falujący fragment izotermy zostaje zastąpiony odcinkiem współistnienia faz.

15.8 Binodała, spinodała i stany metastabilne

Konstrukcja Maxwella wyznacza dla każdej temperatury $T < T_c$ dwie objętości molowe, v_l i v_g , odpowiadające współistniejącym fazom ciekłej i gazowej. Zbiór takich punktów, wyznaczony dla różnych temperatur, nazywamy **binodałą** (patrz Rys. 16).



Rysunek 16: Binodała, spinodała i obszar metastabilny

Binodała jest granicą obszaru współistnienia faz. Na binodali spełniony jest warunek

$$\mu_l(T, p) = \mu_g(T, p).$$

Wewnątrz binodali stan równowagi nie jest jednorodny: układ rozdziela się na ciecz i parę. Nie cały ten obszar ma jednak ten sam charakter. W jego środku leży obszar bezwzględnej niestabilności, w którym

$$\left(\frac{\partial p}{\partial v}\right)_T > 0.$$

Jednorodna faza jest tam niestabilna już wobec nieskończenie małych fluktuacji gęstości.

Granice tego obszaru wyznaczają punkty

$$\left(\frac{\partial p}{\partial v}\right)_T = 0.$$

Ich zbiór nazywamy **spinodałą**. Między binodałą a spinodałą znajdują się stany metastabilne: przegrzana ciecz po stronie ciekłej i przesycona para po stronie gazowej. Są one stabilne wobec małych fluktuacji, ale mogą przejść do stanu dwufazowego po pojawieniu się dostatecznie dużego zarodka nowej fazy.

W punkcie krytycznym binodała i spinodała spotykają się. Izoterma krytyczna ma tam poziomą styczną i zerową krzywiznę:

$$\left(\frac{\partial p}{\partial v}\right)_{T_c} = 0, \quad \left(\frac{\partial^2 p}{\partial v^2}\right)_{T_c} = 0.$$

Dla równania van der Waals,

$$p(v, T) = \frac{RT}{v - b} - \frac{a}{v^2}, \quad (15.58)$$

warunki te dają

$$-\frac{RT}{(v - b)^2} + \frac{2a}{v^3} = 0, \quad (15.59)$$

$$\frac{2RT}{(v-b)^3} - \frac{6a}{v^4} = 0. \quad (15.60)$$

Rozwiązując ten układ, otrzymujemy

$$v_c = 3b, \quad p_c = \frac{a}{27b^2}, \quad T_c = \frac{8a}{27Rb}. \quad (15.61)$$

W szczególności czynnik ściśliwości w punkcie krytycznym,

$$Z_c = \frac{p_c v_c}{RT_c}, \quad (15.62)$$

ma w modelu van der Waalsa wartość uniwersalną:

$$Z_c = \frac{3}{8}. \quad (15.63)$$

Nie zależy ona od parametrów a i b , choć dla rzeczywistych substancji wartość Z_c zwykle różni się od przewidywania van der Waalsa.

15.9 Uwagi końcowe

Model van der Waalsa nie opisuje dokładnie rzeczywistych substancji, zwłaszcza bardzo blisko punktu krytycznego, ale ma dużą wartość dydaktyczną. Pokazuje bowiem w prosty sposób:

- jak skończony rozmiar cząsteczek i przyciąganie międzycząsteczkowe prowadzą do bardziej realistycznego równania stanu,
- skąd bierze się punkt krytyczny,
- dlaczego pewien fragment izotermy odpowiada stanom niestabilnym,
- jak z warunku równości potencjałów chemicznych wynika konstrukcja Maxwella,
- jak obszar współistnienia faz zastępuje niefizyczny, falujący fragment izotermy.

W tym sensie równanie van der Waalsa jest najprostszym modelem, w którym widać całą logikę przejścia ciecz–gaz: od jednorodnej fazy, przez niestabilność i separację faz, aż po zanik różnicy między cieczą i gazem w punkcie krytycznym.

Część V

Kwantowa fizyka statystyczna

*„Elektron niańczy w magnetycznym polu
Ślepy wzrok widma wstrzymany w swej drodze
Względność przestrzeni wszystkie linje zgina,
A ślepiec liże bezbarwne przedmioty
Duch, jak łza czysty, bejakościowym
rzyga wciąż fluidem...”*

— Witkacy, *Tumor Mózgowicz*

1 Macierz gęstości

W klasycznej fizyce statystycznej stan mikroskopowy układu opisywaliśmy przez punkty w przestrzeni fazowej, a zespoły statystyczne przez rozkład prawdopodobieństwa na tej przestrzeni. W mechanice kwantowej rolę stanu mikroskopowego odgrywa wektor stanu, czyli funkcja falowa $|\Psi\rangle$. Jeżeli jednak nie wiemy dokładnie, w którym stanie znajduje się układ, musimy opisywać nie pojedynczy stan, lecz mieszaninę stanów kwantowych.

Założmy więc, że układ znajduje się w stanach $|\Psi_i\rangle$ z prawdopodobieństwami p_i , gdzie

$$p_i \geq 0, \quad \sum_i p_i = 1. \quad (1.1)$$

Podkreślmy, że nie chodzi tu o superpozycję stanów, lecz o mieszaninę statystyczną: układ jest w jednym ze stanów $|\Psi_i\rangle$, ale nie wiemy, w którym.

Policzmy teraz wartość średnią obserwabli opisanej operatorem A . Dla takiej mieszaniny statystycznej mamy

$$\langle A \rangle = \sum_i p_i \langle \Psi_i | A | \Psi_i \rangle. \quad (1.2)$$

Rozwijając to wyrażenie w dowolnej bazie ortonormalnej $\{|\phi_j\rangle\}$ i korzystając z relacji zupełności $\sum_j |\phi_j\rangle\langle\phi_j| = 1$, dostajemy

$$\langle A \rangle = \sum_i p_i \sum_j \langle \Psi_i | \phi_j \rangle \langle \phi_j | A | \Psi_i \rangle = \sum_j \left\langle \phi_j \left| A \left(\sum_i p_i |\Psi_i\rangle\langle\Psi_i| \right) \right| \phi_j \right\rangle = \sum_j \langle \phi_j | A \rho | \phi_j \rangle \quad (1.3)$$

gdzie operator

$$\rho = \sum_i p_i |\Psi_i\rangle\langle\Psi_i|. \quad (1.4)$$

nazywamy **macierzą gęstości**. Relację (1.3) można zapisać jako

$$\langle A \rangle = \text{Tr}(A\rho) \quad (1.5)$$

co pokazuje, że operator gęstości pełni w mechanice kwantowej rolę analogiczną do klasycznej gęstości prawdopodobieństwa.

Z normalizacji prawdopodobieństw wynika normalizacja operatora gęstości:

$$\text{Tr } \rho = \sum_j \langle \phi_j | \rho | \phi_j \rangle = \sum_{i,j} p_i \langle \phi_j | \Psi_i \rangle \langle \Psi_i | \phi_j \rangle = \sum_i p_i \sum_j |\langle \phi_j | \Psi_i \rangle|^2 = 1, \quad (1.6)$$

gdzie skorzystaliśmy z zupełności bazy $|\phi_j\rangle$.

Założmy teraz, że stany $|\Psi_i\rangle$ zmieniają się w czasie zgodnie z równaniem Schrödingera:

$$\frac{\partial}{\partial t} |\Psi_i(t)\rangle = -\frac{i}{\hbar} H |\Psi_i(t)\rangle. \quad (1.7)$$

Z hermitowskości hamiltonianu, równanie sprzężone na $\langle \Psi |$ ma postać

$$\frac{\partial}{\partial t} \langle \Psi_i(t) | = \frac{i}{\hbar} \langle \Psi_i(t) | H, \quad (1.8)$$

Powyższe równania posłużą nam do znalezienia prawa ewolucji dla macierzy gęstości $\rho(t) = \sum_i p_i |\Psi_i(t)\rangle \langle \Psi_i(t)|$, przy czym zakładamy, że prawdopodobieństwa p_i nie zależą jawnie od czasu; zmieniają się tylko same stany kwantowe. Różniczkując definicję $\rho(t)$, dostajemy

$$\frac{\partial \rho}{\partial t} = \sum_i p_i \left(\frac{\partial |\Psi_i\rangle}{\partial t} \langle \Psi_i| + |\Psi_i\rangle \frac{\partial \langle \Psi_i|}{\partial t} \right) = \sum_i p_i \left(-\frac{i}{\hbar} H |\Psi_i\rangle \langle \Psi_i| + \frac{i}{\hbar} |\Psi_i\rangle \langle \Psi_i| H \right). \quad (1.9)$$

Wyciągając operatory H przed sumy, otrzymujemy

$$i\hbar \frac{\partial \rho}{\partial t} = [H, \rho]. \quad (1.10)$$

Jest to **równanie von Neumanna**. Jest ono kwantowym odpowiednikiem równania Liouville'a dla klasycznej gęstości prawdopodobieństwa w przestrzeni fazowej.

W stanie równowagi operator gęstości nie zależy od czasu, więc $\frac{\partial \rho}{\partial t} = 0$. Z równania von Neumanna wynika wtedy

$$[H, \rho] = 0. \quad (1.11)$$

A zatem w równowadze operator gęstości komutuje z hamiltonianem i może być zdiagonalizowany w bazie stanów własnych energii:

$$H|\alpha_k\rangle = E_k|\alpha_k\rangle. \quad (1.12)$$

W tej bazie

$$\rho = \sum_k p_k |\alpha_k\rangle \langle \alpha_k|. \quad (1.13)$$

2 Kwantowe zespoły statystyczne

*„And if you complain once more
You'll meet an army of me”
— Björk, Army of Me*

Zespół mikrokanoniczny. W zespole mikrokanonicznym wszystkie stany o energiach należących do wąskiego przedziału

$$E \leq E_k \leq E + \Delta E$$

są jednakowo prawdopodobne, a pozostałe mają prawdopodobieństwo zerowe:

$$p_k = \begin{cases} \text{const.}, & E \leq E_k \leq E + \Delta E, \\ 0, & \text{w przeciwnym razie.} \end{cases} \quad (2.1)$$

Macierz gęstości ma więc postać

$$\rho = \frac{1}{\Gamma(E, V, N, \Delta E)} \sum_k' |\alpha_k\rangle \langle \alpha_k|, \quad (2.2)$$

gdzie prim przy sumie oznacza sumowanie tylko po stanach spełniających warunek energetyczny, a $\Gamma(E, V, N, \Delta E)$ jest liczbą takich stanów. Entropię definiujemy analogicznie jak w przypadku klasycznym:

$$S(E, V, N) = k_B \ln \Gamma(E, V, N, \Delta E). \quad (2.3)$$

Zespół kanoniczny. W zespole kanonicznym prawdopodobieństwo stanu własnego energii jest proporcjonalne do czynnika Boltzmanna:

$$p_k = \frac{e^{-\beta E_k}}{Z}, \quad \rho = \frac{1}{Z} \sum_k e^{-\beta E_k} |\alpha_k\rangle \langle \alpha_k|. \quad (2.4)$$

Ponieważ jednak

$$e^{-\beta H} |\alpha_k\rangle = e^{-\beta E_k} |\alpha_k\rangle,$$

to wynik ten możemy zapisać w postaci operatorowej:

$$\boxed{\rho = \frac{e^{-\beta H}}{Z}} \quad (2.5)$$

gdzie suma statystyczna jest dana przez

$$Z = \text{Tr } e^{-\beta H} = \sum_k e^{-\beta E_k}. \quad (2.6)$$

Jeżeli poziomy energii mają degeneracje $g(E_n)$, to tę samą sumę można zapisać jako

$$Z = \sum_n g(E_n) e^{-\beta E_n}. \quad (2.7)$$

gdzie suma jest tym razem nie po stanach kwantowych ale po poziomach energetycznych, a $g(E_n)$ jest degeneracją poziomu E_n .

Zespół wielki kanoniczny. W zespole wielkim kanonicznym układ może wymieniać z otoczeniem zarówno energię, jak i cząstki. Wprowadzamy operator liczby cząstek $\hat{N}$. Wspólne stany własne hamiltonianu i operatora liczby cząstek oznaczmy przez $|\alpha_{N,k}\rangle$:

$$H |\alpha_{N,k}\rangle = E_{N,k} |\alpha_{N,k}\rangle, \quad \hat{N} |\alpha_{N,k}\rangle = N |\alpha_{N,k}\rangle. \quad (2.8)$$

Prawdopodobieństwo takiego stanu jest proporcjonalne do

$$e^{-\beta(E_{N,k} - \mu N)}.$$

Dlatego operator gęstości ma postać

$$\rho = \frac{1}{\Xi} \sum_{N,k} e^{-\beta(E_{N,k} - \mu N)} |\alpha_{N,k}\rangle \langle \alpha_{N,k}|. \quad (2.9)$$

W zapisie operatorowym

$$\boxed{\rho = \frac{e^{-\beta(H - \mu \hat{N})}}{\Xi}} \quad (2.10)$$

gdzie wielka suma statystyczna wynosi

$$\Xi = \text{Tr } e^{-\beta(H - \mu \hat{N})} = \sum_{N,k} e^{-\beta(E_{N,k} - \mu N)}. \quad (2.11)$$

3 Kwantowe gazy doskonałe. Nierozróżnialność.

„I am he as you are he as you are me and we are all together”

— The Beatles, *I Am the Walrus*

3.1 Kwantowe gazy doskonałe

Kwantowym gazem doskonałym nazywamy układ nieoddziałujących cząstek. Oznacza to, że hamiltonian układu N cząstek można zapisać jako sumę hamiltonianów jednocząstkowych:

$$H(1, \dots, N) = \sum_{i=1}^N H_0(i), \quad (3.1)$$

gdzie $H_0(i)$ działa na zmienne i -tej cząstki. Hamiltoniany jednocząstkowe komutują ze sobą, a więc problem wielocząstkowy można zbudować z rozwiązań problemu jednej cząstki:

$$H_0 \Psi_k = \epsilon_k \Psi_k. \quad (3.2)$$

Indeks k oznacza tutaj zestaw liczb kwantowych opisujących stan jednocząstkowy.

Ponieważ cząstki kwantowe są nierozróżnialne, nie pytamy, która cząstka znajduje się w którym stanie. Sens fizyczny ma jedynie pytanie, ile cząstek obsadza dany stan jednocząstkowy. Stan wielocząstkowy opisujemy więc przez **liczby obsadzeń**

$$n_k. \quad (3.3)$$

Dla danego zestawu liczb obsadzeń całkowita liczba cząstek i energia układu wynoszą

$$N = \sum_k n_k, \quad E = \sum_k \epsilon_k n_k. \quad (3.4)$$

Dopuszczalne wartości n_k zależą od rodzaju cząstek. Dla fermionów, czyli cząstek o spinie połówkowym, obowiązuje zakaz Pauliego:

$$n_k = 0, 1. \quad (3.5)$$

Dla bozonów, czyli cząstek o spinie całkowitym, jeden stan jednocząstkowy może być obsadzony dowolną liczbą cząstek:

$$n_k = 0, 1, 2, \dots \quad (3.6)$$

Opis przez liczby obsadzeń jest szczególnie prosty dla gazu doskonałego, ponieważ energie stanów jednocząstkowych są określone raz na zawsze przez objętość i warunki brzegowe. Nie zależą one od tego, ile cząstek znajduje się w innych stanach. Sumując po liczbach obsadzeń, automatycznie uwzględniamy nierozróżnialność cząstek. Dla cząstek rozróżnialnych ta sama konfiguracja obsadzeń odpowiadałaby wielu różnym przypisaniom cząstek do stanów, a więc pojawiłyby się dodatkowe czynniki kombinatoryczne.

Przykład: dwie cząstki i trzy stany.

Załóżmy, że mamy dwie cząstki oraz trzy dostępne stany jednocząstkowe, oznaczone przez 1, 2, 3.

Dla cząstek rozróżnialnych, które roboczo oznaczmy przez A i B , możliwe przypisania są

następujące:

1	2	3
AB	–	–
–	AB	–
–	–	AB
A	B	–
A	–	B
–	A	B
B	A	–
B	–	A
–	B	A

Mamy trzy przypadki, w których obie cząstki znajdują się w tym samym stanie, oraz sześć przypadków, w których znajdują się w stanach różnych.

Dla bozonów przypisania A, B oraz B, A nie są odrębnymi stanami fizycznymi. Liczymy tylko konfiguracje obsadzeń:

1	2	3
2	0	0
0	2	0
0	0	2
1	1	0
1	0	1
0	1	1

Są trzy konfiguracje z podwójnym obsadzeniem jednego stanu i trzy konfiguracje z cząstkami w różnych stanach.

Dla fermionów podwójne obsadzenie jednego stanu jest zabronione:

1	2	3
1	1	0
1	0	1
0	1	1

Nie ma więc żadnej konfiguracji, w której dwie cząstki zajmują ten sam stan jednocząstkowy. Porównajmy teraz wielkość

$$v = \frac{\text{liczba przypadków, w których dwie cząstki są w tym samym stanie}}{\text{liczba przypadków, w których dwie cząstki są w różnych stanach}}. \quad (3.7)$$

Dla trzech rodzajów statystyki dostajemy

$$v_{\text{klas}} = \frac{3}{6} = \frac{1}{2}, \quad v_{\text{BE}} = \frac{3}{3} = 1, \quad v_{\text{FD}} = \frac{0}{3} = 0. \quad (3.8)$$

W tym prostym przykładzie widać już charakter statystyk kwantowych: bozony łatwiej obsadzają wspólny stan niż cząstki klasyczne, natomiast fermiony nie mogą tego zrobić w ogóle.

3.2 Wielka suma statystyczna w języku liczb obsadzeń

Dla kwantowych gazów doskonałych najwygodniejszy jest wielki zespół kanoniczny, bo sumy po liczbach obsadzeń łatwiej się liczy, gdy nie ma górnego ograniczenia na N . Wielka suma statystyczna ma postać

$$\Xi(T, V, \mu) = \sum_{N=0}^{\infty} \sum_i \exp[-\beta(E_{N,i} - \mu N)], \quad (3.9)$$

gdzie $E_{N,i}$ jest energią całego układu złożonego z N cząstek w i -tym stanie wielocząstkowym.

Przejdźmy teraz do liczb obsadzeń. Dla ustalonej liczby cząstek N sumę po stanach wielocząstkowych można zapisać jako sumę po wszystkich konfiguracjach obsadzeń, z warunkiem narzuconym przez deltę:

$$\sum_i^{(N)} \rightarrow \sum_{n_1} \sum_{n_2} \cdots \delta\left(\sum_k n_k - N\right) \equiv \sum_{\{n_k\}} \delta\left(\sum_k n_k - N\right). \quad (3.10)$$

Symbol $\delta(\sum_k n_k - N)$ to delta Kroneckera wycinająca ze wszystkich kombinacji $\{n_k\}$ te, w których całkowita liczba cząstek jest równa N . Wielka suma statystyczna przyjmuje więc postać

$$\Xi = \sum_{N=0}^{\infty} \sum_{\{n_k\}} \delta\left(\sum_k n_k - \sum_k n_k\right) \exp\left[-\beta\left(\sum_k n_k \epsilon_k - \mu N\right)\right]. \quad (3.11)$$

Zamieniamy teraz kolejność sumowania. Delta wymusza $N = \sum_k n_k$, ale wcześniej sumujemy po wszystkich N więc w efekcie nie ma żadnych ograniczeń na liczby obsadzeń. Dostajemy zatem:

$$\Xi = \sum_{\{n_k\}} \exp\left[-\beta \sum_k n_k (\epsilon_k - \mu)\right]. \quad (3.12)$$

Teraz korzystamy z tego, że wykładnik jest sumą wkładów od poszczególnych stanów jednocząstkowych. EkspONENTA sumy zamienia się więc na iloczyn:

$$\Xi = \sum_{\{n_k\}} \prod_k \exp[-\beta n_k (\epsilon_k - \mu)] = \sum_{n_1} e^{-\beta n_1 (\epsilon_1 - \mu)} \sum_{n_2} e^{-\beta n_2 (\epsilon_2 - \mu)} \dots \quad (3.13)$$

Wprowadzając oznaczenie

$$\Xi_k = \sum_{n_k} \exp[-\beta n_k (\epsilon_k - \mu)], \quad (3.14)$$

dostajemy

$$\Xi = \prod_k \Xi_k. \quad (3.15)$$

Dalszy rachunek zależy tylko od tego, jakie wartości n_k są dozwolone.

Fermiony. Dla fermionów $n_k = 0, 1$, więc

$$\Xi_k^f = \sum_{n_k=0}^1 \left[e^{-\beta(\epsilon_k - \mu)} \right]^{n_k} = 1 + e^{-\beta(\epsilon_k - \mu)}. \quad (3.16)$$

Wprowadzając aktywność

$$z = e^{\beta\mu}, \quad (3.17)$$

możemy zapisać to jako

$$\Xi_k^f = 1 + z e^{-\beta\epsilon_k}. \quad (3.18)$$

Bozony. Dla bozonów $n_k = 0, 1, 2, \dots$, więc otrzymujemy szereg geometryczny:

$$\Xi_k^b = \sum_{n_k=0}^{\infty} \left[e^{-\beta(\epsilon_k - \mu)} \right]^{n_k} = \frac{1}{1 - e^{-\beta(\epsilon_k - \mu)}}. \quad (3.19)$$

Czyli

$$\Xi_k^b = \frac{1}{1 - z e^{-\beta\epsilon_k}}. \quad (3.20)$$

Szereg geometryczny jest zbieżny tylko wtedy, gdy

$$e^{-\beta(\epsilon_k - \mu)} < 1, \quad (3.21)$$

a więc

$$\mu < \epsilon_k. \quad (3.22)$$

Warunek ten musi zachodzić dla każdego stanu jednocząstkowego. Jeśli energię stanu podstawowego wybierzemy jako $\epsilon_0 = 0$, to dla bozonów

$$\mu < 0. \quad (3.23)$$

Oba wyniki można zapisać w zwarty sposób jako

$$\Xi_k^{\text{FD}} = 1 + ze^{-\beta\epsilon_k}, \quad \Xi_k^{\text{BE}} = \left(1 - ze^{-\beta\epsilon_k}\right)^{-1}. \quad (3.24)$$

Albo, używając górnego znaku dla fermionów i dolnego dla bozonów:

$$\Xi_k^{f,b} = \left(1 \pm ze^{-\beta\epsilon_k}\right)^{\pm 1}. \quad (3.25)$$

Stąd

$$\Xi_{f,b} = \prod_k \left(1 \pm ze^{-\beta\epsilon_k}\right)^{\pm 1}. \quad (3.26)$$

3.3 Ciśnienie i średnie liczby obsadzeń

Z wielkiego zespołu kanonicznego wiemy, że

$$\Omega = -k_B T \log \Xi. \quad (3.27)$$

Dla jednorodnego układu

$$\Omega = -pV, \quad (3.28)$$

a więc

$$pV = k_B T \log \Xi. \quad (3.29)$$

Korzystając z faktoryzacji wielkiej sumy statystycznej, otrzymujemy dla fermionów

$$pV = k_B T \sum_k \log \left(1 + e^{-\beta(\epsilon_k - \mu)}\right), \quad (3.30)$$

a dla bozonów

$$pV = -k_B T \sum_k \log \left(1 - e^{-\beta(\epsilon_k - \mu)}\right). \quad (3.31)$$

Średnia liczba cząstek wynika z pochodnej po aktywności:

$$\langle N \rangle = z \frac{\partial}{\partial z} \log \Xi. \quad (3.32)$$

Ponieważ

$$\log \Xi_{f,b} = \pm \sum_k \log \left(1 \pm ze^{-\beta\epsilon_k}\right), \quad (3.33)$$

dostajemy

$$\langle N \rangle = \pm z \frac{\partial}{\partial z} \sum_k \log \left(1 \pm ze^{-\beta\epsilon_k}\right). \quad (3.34)$$

Po zróżniczkowaniu:

$$\langle N \rangle = \sum_k \frac{ze^{-\beta\epsilon_k}}{1 \pm ze^{-\beta\epsilon_k}}. \quad (3.35)$$

Ponieważ $N = \sum_k n_k$, to również

$$\langle N \rangle = \sum_k \langle n_k \rangle. \quad (3.36)$$

Wyrażenie pod sumą powinno być zatem średnią liczbą obsadzeń danego stanu jednocząstkowego:

$$\boxed{\langle n_k \rangle = \frac{1}{e^{\beta(\epsilon_k - \mu)} \pm 1}}. \quad (3.37)$$

Górny znak odpowiada fermionom:

$$\langle n_k \rangle^f = \frac{1}{e^{\beta(\epsilon_k - \mu)} + 1}, \quad (3.38)$$

a dolny bozonom:

$$\langle n_k \rangle^b = \frac{1}{e^{\beta(\epsilon_k - \mu)} - 1}. \quad (3.39)$$

Są to odpowiednio rozkład Fermiego-Diraca i rozkład Bosego-Einsteina. Dla fermionów natychmiast widać, że

$$0 \leq \langle n_k \rangle^f \leq 1, \quad (3.40)$$

zgodnie z zakazem Pauliego. Dla bozonów średnia liczba obsadzeń może być dowolnie duża, gdy μ zbliża się do energii stanu podstawowego.