

Rozdział 1

Od mechaniki do termodynamiki

1.1 Specyficzność termodynamiki

Termodynamika powstała w pierwszej połowie XIX wieku jako teoria mająca przede wszystkim za zadanie wyjaśnienie działania budowanych już wtedy od blisko wieku maszyn cieplnych przetwarzających ciepło (*thermo*) na pracę (*dynamika*). Jedno z najważniejszych twierdzeń nowej nauki głosiło, że nie jest możliwe stałe przekształcanie ciepła w pracę w wypadku, gdy maszyna sprzężona jest tylko ze źródłem ciepła o ustalonej temperaturze. Konieczne jest jeszcze sprzężenie z drugim układem o niższej temperaturze (chłodnicą). Tradycyjnie, termodynamika zajmowała się więc procesami zachodzącymi w warunkach sprzężenia z otoczeniem o zmiennej temperaturze i to procesami przebiegającymi odwracalnie, bez zaburzania równowagi termodynamicznej. W bilansie energetycznym rozpatrywanych układów wystarczało wtedy uwzględniać tylko ciepło i pracę, ale już nie dysypację (nieodwracalne rozpraszanie) energii.

Ten sposób wykładania termodynamiki pokutuje w wielu podręcznikach praktycznie do dziś. Główną ambicją typowego wstępnego kursu tego przedmiotu jest przedstawienie teoretycznego modelu maszyny cieplnej Sadi Carnota sprzed prawie 200 lat. Z biegiem czasu przedmiot termodynamiki ulegał jednak uogólnieniu obejmując coraz szerszą klasę przemian energetycznych. Procesy badane we współczesnych laboratoriach w większości wypadków zachodzą w warunkach stałej temperatury, stabilizowanej przez wysokiej klasy termostaty. Poza tym, zjawisko dysypacji wcale nie jest zaniedbywane, gdyż dostarcza ogromnej ilości informacji o procesach w skali mikroskopowej. Interesują więc nas głównie *nieodwracalne* procesy *izotermiczne*. W takich zresztą warunkach pracuje wiele maszyn, w tym molekularne maszyny biologiczne, przetwarzające nie ciepło w pracę, lecz jeden rodzaj pracy w drugi.

W XX wieku termodynamika procesów nieodwracalnych poczyniła znaczne postępy. Wspomnijmy teorię reakcji liniowej Onsagera z pierwszej połowy wieku i teorię struktur dysypatywnych Prigogine'a z drugiej. Najważniejsze jednak odkrycia dokonane zostały nie w termodynamice fenomenologicznej, opisującej zja-

wiska, lecz w fizyce statystycznej, wyjaśniającej termodynamikę na gruncie mechaniki mikroobektów leżącej u jej podstaw. Zrozumieliśmy związek pomiędzy makroskopową nieodwracalnością a mikroskopową niestabilnością ruchu.

Termodynamika jest teorią fizyczną bardzo ogólną. Właściwie zakres jej zastosowań nie jest ustalony do dzisiaj. Wydaje się, że przy jej pomocy można wyjaśniać procesy zachodzące zarówno w czarnych dziurach jak i żywych organizmach, ale nie jest to do końca całkiem pewne. Brak dobrze określonego przedmiotu badań jest powodem, że struktura matematyczna termodynamiki jest dużo słabiej określona niż struktura matematyczna innych teorii fizycznych: mechaniki klasycznej i kwantowej, elektrodynamiki, czy kwantowej teorii pola.

Głównym problemem nowoczesnego wykładu termodynamiki jest więc w miarę precyzyjne określenie przedmiotu jej zastosowań. Związki pomiędzy pracą, ciepłem i dysypacją energii można wyjaśnić tylko odwołując się do podstawowego, ale bynajmniej nie prostego dla początkujących adeptów fizyki, pojęcia *entropii*. Rozsądne wydaje się określenie termodynamiki jako teorii opisującej zachowanie się tych i tylko tych stanów układów fizycznych, dla których entropia jest dobrze określoną wielkością fizyczną o możliwej do zmierzenia wartości. Głęboki sens pojęcia entropii można zrozumieć dopiero na gruncie teorii informacji i rachunku prawdopodobieństwa. Naszego wykładu nie chcemy więc zaczynać w tradycyjny sposób od termodynamiki czysto zjawiskowej, fenomenologicznej, w szczególności od uogólniających doświadczenie fenomenologicznych “zasad termodynamiki”. Zamiast tego spróbujemy odpowiedzieć na bardziej fundamentalne pytanie: co to właściwie jest entropia i jaki jest fizyczny mechanizm jej wzrostu?

1.2 Zasada determinizmu mechanicznego

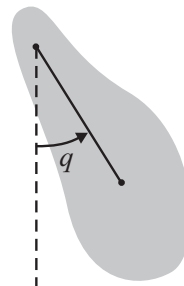
U podstaw całej fizyki, warunkując jej racjonalizację, leży sformułowana ponad trzysta lat temu przez Izaaka Newtona *zasada determinizmu mechanicznego* (Prigogine i Stengers, 1990; Newton, 1996). We współczesnym języku można wyrazić ją zdaniem:

Prawo ruchu i stan układu fizycznego w danej chwili jednoznacznie określają stany układu we wszystkich innych chwilach zarówno późniejszych, jak i wcześniejszych.

Zasada determinizmu nadaje sens pojęciu *czasu*, jako parametrowi liniowo porządkującemu stany układu wzdłuż *trajektorii* — krzywych w wielowymiarowej na ogół *przestrzeni stanów*.

Aby nie czynić wykładu zbyt abstrakcyjnym, rozważmy prosty przykład. Stan wahadła (rys. 1.1) poruszającego się bez tarcia w ustalonej płaszczyźnie jednoznacznie scharakteryzowany jest przez kąt wychylenia q i prędkość kątową lub, równoważnie, proporcjonalny do niej moment pędu p . Przestrzeń stanów tworzy w tym wypadku dwuwymiarową płaszczyznę, a dokładniej, nieskończenie długi pasek, którego brzegi $q = -\pi$ i $q = \pi$ są tożsame ze sobą (taki zbiór można

Rys. 1.1. Wahadło — bryła sztywna poruszająca się w ustalonej płaszczyźnie pod wpływem siły ciężkości. Odległość od osi obrotu do środka masy pozostaje stała



wyobrazić sobie jako powierzchnię walca utworzonego poprzez sklejenie obu brzo-
gów). Kilka trajektorii w tak określonej przestrzeni stanów przedstawia rys. 1.2.a.

Jednoznaczność ewolucji czasowej oznacza, że przez każdy punkt w przestrzeni stanów przechodzi jedna i tylko jedna trajektoria. Lokalnie, w każdym punkcie, prawo ruchu określa wektor styczny do trajektorii, o długości równej prędkości (pochodnej czasowej) zmiany stanu wzdłuż tej trajektorii (por. rys. 1.2.b). Pełne pole wektorów stycznych na całej przestrzeni stanów zadane jest przez różniczkowe *równanie ruchu*. W przypadku wahadła stanowi je układ dwóch równań różniczkowych postaci

$$\dot{q} = \omega p, \quad \dot{p} = -\omega \sin q, \quad (1.1)$$

gdzie kropka oznacza różniczkowanie po czasie (\dot{q} przedstawia składową poziomą a \dot{p} , pionową wektora stycznego, por. rys. 1.2.b) a ω jest odwrotnością pewnej charakterystycznej dla układu jednostki czasu. Przebieg trajektorii w czasie można znaleźć rozwiązując ten układ równań.

Dwa równania różniczkowe pierwszego rzędu (1.1) są równoważne jednemu równaniu różniczkowemu rzędu drugiego, z drugą pochodną względem czasu. Różniczkując pierwsze równanie i podstawiając drugie otrzymujemy równanie

$$\ddot{q} = -\omega^2 \sin q. \quad (1.2)$$

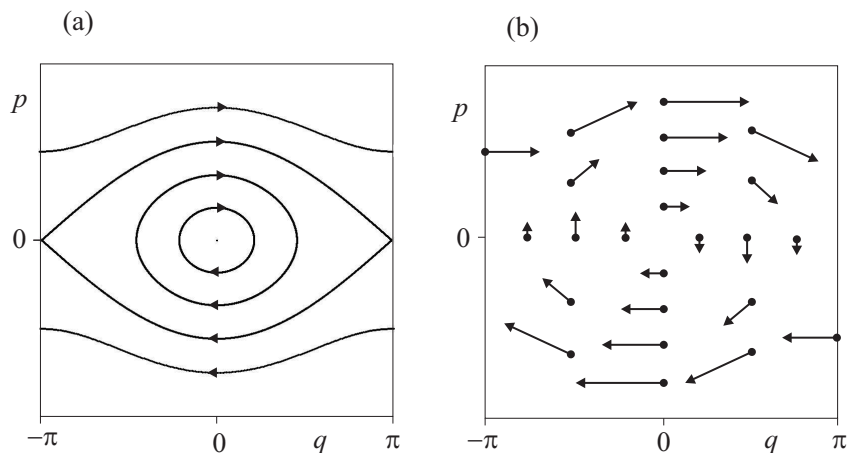
Ma ono postać *równania Newtona: przyspieszenie* (druga pochodna wychylenia po czasie) jest wprost proporcjonalne do *siły* działającej w kierunku ruchu środka masy wahadła, zmieniającej się, jak można zauważyć z rys. 1.1, tak jak sinus kąta wychylenia.

Dla małych wychyleń funkcję sinus można przybliżyć przez wartość jej argumentu, $\sin q \approx q$, i równania (1.1) upraszczają się do

$$\dot{q} = \omega p, \quad \dot{p} = -\omega q. \quad (1.3)$$

Łatwo sprawdzić wykonując proste różniczkowanie, że ich rozwiązanie rozpoczynające się w stanie o współrzędnych $q(0) = A$ i $p(0) = 0$ (nieruchome wahadło wychylone o kąt A) stanowią funkcje

$$q(t) = A \cos \omega t, \quad p(t) = -A \sin \omega t, \quad (1.4)$$



Rys. 1.2. (a) Trajektorie ruchu wahadła. Przy małych wartościach energii całkowitej wahadło wykonuje oscylacje dookoła stanu równowagi $(q, p) = (0, 0)$ (trajektorie zbliżone do elipsy). Przy dużych wartościach energii wahadło porusza się natomiast ruchem rotacyjnym z prędkością kątową nie spadającą nigdy do zera (trajektorie zamknięte przechodzące przez linię $q = \pm\pi$). Granicę pomiędzy obu rodzajami ruchu stanowią dwie trajektorie, wzdłuż których wahadło nieskończenie wolno osiąga maksymalny kąt wychylenia π lub $-\pi$. Stany równowagi stabilnej $(q, p) = (0, 0)$ i niestabilnej $(q, p) = (\pm\pi, 0)$ tworzą trajektorie jednoelementowe. (b) Pole wektorów stycznych na przestrzeni stanów z rysunku (a). Składowa wektora stycznego w kierunku q jest zawsze równa lokalnej wartości prędkości kątowej p (porównaj pierwsze z równań (1.1))

przedstawiające ruch oscylujący z amplitudą A i okresem $2\pi/\omega$. Stały parametr ω nazywamy *częstotliwością kątową* oscylacji.

Równania (1.3) można zapisać w postaci

$$\dot{q} = \frac{\partial \mathcal{H}}{\partial p}, \quad \dot{p} = -\frac{\partial \mathcal{H}}{\partial q}, \quad (1.5)$$

w której prawe strony stanowią pochodne cząstkowe pewnej funkcji wychylenia i prędkości kątowej (momentu pędu)

$$\mathcal{H}(q, p) = \frac{\omega p^2}{2} + \frac{\omega q^2}{2}. \quad (1.6)$$

Ma ona sens całkowitej *energii* układu. Pierwszy składnik, zależny od prędkości (pędu, czy momentu pędu), nosi nazwę energii *kinetycznej* a drugi, zależny od wychylenia, energii *potencjalnej*. Zgodnie z równaniami (1.4) w trakcie ruchu wahadła w zakresie małych kątów wychyleń następuje cykliczna zamiana energii potencjalnej w kinetyczną.

Równania ruchu każdego układu fizycznego opisywanego przez mechanikę klasyczną mają postać (1.5) z tym, że zmienna q jest w ogólności zastąpiona przez wektor, czyli ciąg zmiennych $q = (q_1, q_2, \dots, q_n)$, ogólnie nazywanych *położeniami* a zmienna p przez ciąg zmiennych $p = (p_1, p_2, \dots, p_n)$, ogólnie nazywanych *pędami*. Liczba n określa liczbę *stopni swobody* układu. Stan klasycznego

układu mechanicznego opisany jest więc razem przez wektor o $2n$ współrzędnych $(q, p) = (q_1, \dots, q_n, p_1, \dots, p_n)$. Przestrzeń stanów mechaniki klasycznej określonych przez położenia i pędy nazywamy *przestrzenią fazową*, a trajektorie w tej przestrzeni, *trajektoriami fazowymi*. Rzeczywista funkcja określona na przestrzeni fazowej $\mathcal{H} = \mathcal{H}(q_1, \dots, q_n, p_1, \dots, p_n)$ nazywa się *hamiltonianem* (Newton, 1996).

Pola fizyczne mają continuum (nieskończoną i nieprzeliczalną liczbę) stopni swobody. Stan pola elektromagnetycznego, jak sama nazwa wskazuje, scharakteryzowany jest przez pole zarówno elektryczne jak i magnetyczne, oba będące funkcjami wektorowymi określonymi na trójwymiarowej przestrzeni fizycznej. Zmianę w czasie tych pól opisują różniczkowe cząstkowe równania *Maxwella*. W mechanice kwantowej przestrzenią stanów jest nieskończenie wymiarowa przestrzeń zespolonych *funkcji falowych* a równaniem ruchu jest różniczkowe cząstkowe równanie *Schrödingera*. W wypadku układu wielu nieodróżnialnych cząstek ze spinem właściwszy jest język operatorowych *pól kwantowych* (Newton, 1996). Podkreślmy wyraźnie, że ewolucja funkcji falowej, czy ogólniej, pola kwantowego odbywa się w sposób całkowicie zdeterminowany; tylko proces pomiaru, sprzęgający mikroskopowy obiekt z makroskopowym obserwatorem, zawiera elementy indeterminizmu.

Podstawowe równania fizyki przedstawia rysunek 1.3. Potraktujmy go jako grafikę nie wchodząc w szczegóły, których głębsze zrozumienie wymaga kilku lat solidnych studiów. Wszystkie równania są zapisem matematycznym praw fizyki określających zdeterminowaną w czasie ewolucję stanu układów do których opisu się stosują. Zauważmy, że wszystkie zawierają, podobnie jak równanie (1.1), pierwszą pochodną względem czasu t .

$$\begin{aligned} \dot{q}_i &= \frac{\partial \mathcal{H}}{\partial p_i}, & \dot{p}_i &= -\frac{\partial \mathcal{H}}{\partial q_i} \\ \frac{\partial}{\partial t} \mathbf{E} &= c^2 \frac{\partial}{\partial \mathbf{r}} \times \mathbf{B}, & \frac{\partial}{\partial t} \mathbf{B} &= -\frac{\partial}{\partial \mathbf{r}} \times \mathbf{E} \\ i\hbar \frac{\partial}{\partial t} \psi &= \mathcal{H} \psi \\ i\hbar \frac{\partial}{\partial t} \Psi &= \left(-i\boldsymbol{\alpha} \cdot \frac{\partial}{\partial \mathbf{r}} + m\beta + e\phi - \boldsymbol{\alpha} \cdot \mathbf{A} \right) \Psi \end{aligned}$$

Rys. 1.3. Równania mechaniki klasycznej, elektrodynamiki, mechaniki kwantowej i elektrodynamiki kwantowej (kwantowej teorii pola)

1.3 Nieodwracalność procesów makroskopowych

Z codziennego doświadczenia dobrze wiadomo, że równania (1.3) i ich rozwiązanie (1.4) niezupełnie dokładnie opisują ruch wahadła i to bynajmniej nie dlatego, że przyjęliśmy założenie małych wychyleń. Ruch każdego rzeczywistego wahadła makroskopowego jest tłumiony, co wymaga wprowadzenia do drugiego z równań (1.3) dodatkowej siły *tarcia*, proporcjonalnej do prędkości, a więc pędu:

$$\dot{q} = \omega p, \quad \dot{p} = -\omega q - 2\kappa p. \quad (1.7)$$

Rozwiązania układu (1.7) dla różnych wartości parametru κ przedstawia rys. 1.4.

Charakter rozwiązań zależy od stosunku parametru tłumienia κ do częstości drgań ω . Dla $0 < \kappa/\omega < 1$ wahadło wykonuje oscylacje tłumione (rys. 4.b), natomiast dla κ/ω większego od wartości krytycznej 1, wahadło zmierza do stanu równowagi $(q, p) = (0, 0)$ w sposób aperiodyczny (rys. 4.c). Dla bardzo dużych wartości stosunku κ/ω w przebiegu trajektorii ruchu wahadła wyraźnie wyróżnić można dwa etapy (rys. 4.d). W pierwszym etapie, prędkość początkowa bardzo szybko spada do niewielkiej wartości określonej przez chwilową wartość wychylenia i potem, w drugim etapie, zmienia się zgodnie ze zmianą wartości tej ostatniej zmiennej. Wszystkie trajektorie ruchu w drugim etapie praktycznie pokrywają się, choć niezupełnie dokładnie, bo przeczyłoby to zasadzie determinizmu, zgodnie z którą przez każdy punkt określający stan układu przechodzić może tylko jedna trajektoria.

Jeśli nie interesują nas zmiany stanu układu w przedziałach czasu krótszych od czasu trwania pierwszego, *przejściowego* etapu ruchu, możemy określić zmiany w samym tylko drugim etapie przyjmując warunek *stacjonarności*:

$$\dot{p} = -\omega q - 2\kappa p = 0. \quad (1.8)$$

Oznacza on zanedbywalnie małą zmianę prędkości w etapie drugim w porównaniu ze zmianą tej prędkości w etapie pierwszym i pozwala określić zależność prędkości (pędu) od wychylenia (położenia):

$$p = -(\omega/2\kappa)q. \quad (1.9)$$

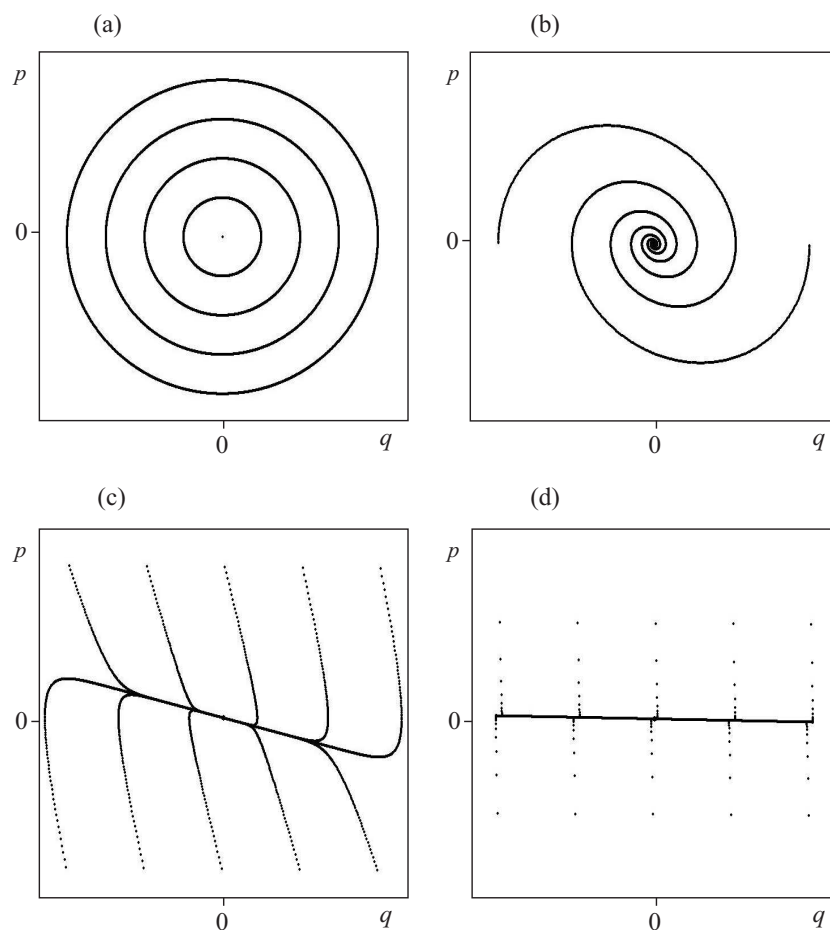
Podstawiając tę zależność do pierwszego z równań (1.7) otrzymujemy jedno tylko równanie dla jednej zmiennej q :

$$\dot{q} = -kq, \quad (1.10)$$

gdzie $k \equiv \omega^2/2\kappa$. Równanie to opisuje zmianę w czasie wychylenia wahadła w drugim, *stacjonarnym* etapie ewolucji i posiada, jak łatwo sprawdzić, rozwiązanie

$$q(t) = q(0)e^{-kt}, \quad (1.11)$$

przedstawiające eksponencjalny w czasie zanik wychylenia do zera. Przedstawione rozumowanie pokazuje jak można uprościć czasowy opis ewolucji układu w



Rys. 1.4. Przykłady rozwiązań układu równań (1.7). (a) $\kappa = 0$: oscylacje niegasnące, porównaj rys. 2.a dla małych wychyleń. (b) $\kappa = 0,2\omega$: oscylacje gasnące. (c) $\kappa = 2\omega$: aperiodyczny zanik wychylenia i prędkości kątowej. (d) $\kappa = 20\omega$: aperiodyczny zanik wychylenia z prędkością praktycznie jednoznacznie określoną przez to wychylenie (początkowy okres szybkich zmian prędkości jest zanedbywalnie krótki, co widać z dużych odległości pomiędzy kropkami odpowiadającym krokom czasowym $0,01/\omega$). Rysunki wykonano za pomocą programu DiGraph Tomasza Jarusa

przypadku występowania dwóch skal czasowych: zaniedbywalnie *szybkiej*, charakteryzującej w naszym wypadku zachowanie zmiennej p , i *wolnej*, charakteryzującej zachowanie ważniejszej zmiennej q .

W porównaniu z determinizmem czystej mechaniki determinizm fizyki układów z tarciem, opisywany przez równania typu (1.7), jest trochę ułomny. Wartości początkowe składowych wektora stanu $q(0)$ i $p(0)$ rzeczywiście jednoznacznie determinują wartości tych składowych we wszystkich chwilach późniejszych. Próbując jednak określić q i p nie w przyszłości, a w przeszłości, może się zdarzyć, że otrzymamy wartości niefizyczne: nieskończenie wielkiego wychylenia (a więc energii potencjalnej) czy nieskończenie wielkiej prędkości kątowej (a więc energii kinetycznej). Trajektorie przedłużane wstecz wychodzą poza obszar fizycznie sensownych stanów. Równania (1.7), w przeciwieństwie do równań (1.1), są *nieodwracalne* w czasie.

W ogólności prawa ruchu na poziomie makroskopowym są nieodwracalne, a na poziomie mikroskopowym, odwracalne w czasie. Jednym z zadań fizyki statystycznej stanowiącej przedmiot tego rozdziału jest wyjaśnienie w jaki sposób makroskopowa nieodwracalność wynika z mikroskopowych, odwracalnych praw ruchu atomów i cząsteczek.

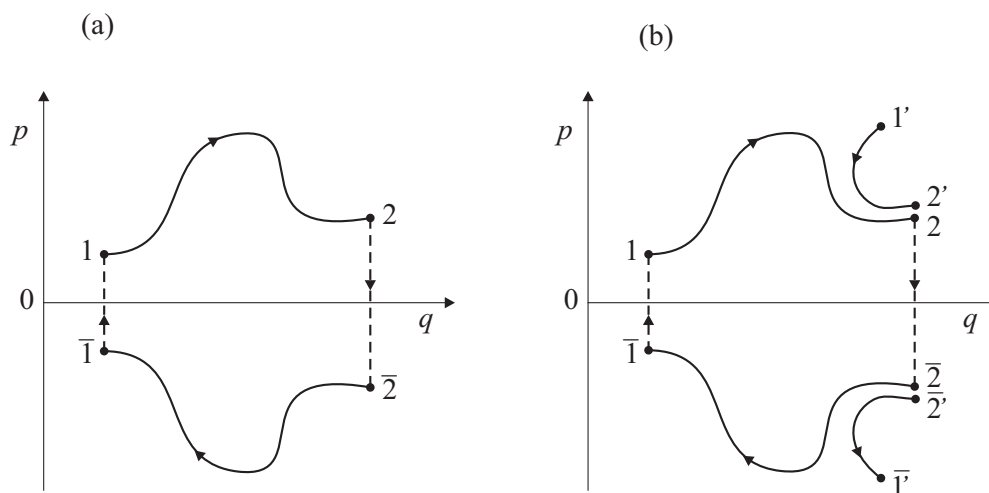
1.4 Niestabilność ruchu jako przyczyna nieodwracalności

W przeciwieństwie do równań (1.7), równania mechaniki (1.5) nie tylko że posiadają wydawałoby się fizycznie sensowne rozwiązania zarówno dla $t \geq 0$ jak i $t < 0$ lecz, co więcej, są niezmiennicze względem transformacji (Penrose, 1979; Prigogine i Stengers, 1990; Newton 1996)

$$t, p_i \mapsto -t, -p_i \tag{1.12}$$

(zamiana kierunku płynięcia czasu połączona z równoczesnym odwróceniem zwrotu wszystkich pędów). Symetria (1.12) oznacza, że jeśli $(q(t), p(t))$ jest rozwiązaniem układu równań (1.5), to również $(q(-t), -p(-t))$ musi być rozwiązaniem tego układu. Schematycznie przedstawiono to na rys. 1.5.a, gdzie q i p reprezentują w ogólności n -wymiarowe wektory $q = (q_1, q_2, \dots, q_n)$, $p = (p_1, p_2, \dots, p_n)$. Rzeczywiście, nie można odróżnić, czy film, przedstawiający ruch wahadła bez tarcia, wyświetlany jest od początku, czy od końca. W układzie mechanicznym wydaje się możliwa nie tylko rekonstrukcja, ale i realny powrót do stanu z przeszłości. Rysunek 1.5.a sugeruje “podróż w czasie” polegającą na dwukrotnym zastosowaniu operacji odwrócenia pędów.

Własność odwracalności ruchu nie zależy od liczby wymiarów przestrzeni stanów, ale w wypadku układów makroskopowych, składających się z ogromnej liczby cząsteczek, wydaje się nie być spełniona. Niech płaszczyzna na rysunku 1.5 reprezentuje bardzo wysoko wymiarową przestrzeń stanów pewnej objętości wody



Rys. 1.5. (a) Z symetrii równań mechaniki wynika, że wraz z odcinkiem trajektorii 12 przestrzeń stanów musi zawierać odcinek trajektorii $\bar{2}\bar{1}$, będący symetrycznym odbiciem odcinka 12 w płaszczyźnie $p = 0$. Odcinek $\bar{2}\bar{1}$ trajektorii opisuje ruch *odwrotny* do ruchu opisanego odcinkiem 12. Układ, który, startując ze stanu 1, osiągnął po czasie t stan 2, powróci do stanu 1, jeśli odwróci pędy (przejdź z 2 do $\bar{2}$), pozwolić na swobodną ewolucję do $\bar{1}$ przez następny okres czasu t , i ponownie odwrócić pędy. (b) W wypadku, gdy ruch jest niestabilny, nieskończenie mała niedokładność w realizacji odwrócenia pędów może wprowadzić układ na zupełnie nieoczekiwaną trajektorię $\bar{2}'\bar{1}'$. Symetryczną do trajektorii $\bar{2}'\bar{1}'$ jest trajektoria $1'2'$. Można ją uważać za alternatywę do trajektorii 12 przy próbie rekonstrukcji przeszłości układu, o którym ze skończoną tylko dokładnością wiadomo, że znajduje się w stanie 2.

w sadzawce. W tym układzie łatwo można sobie wyobrazić na przykład proces zachodzący od momentu wrzucenia do wody kamienia (stan 1) do momentu dojścia do brzegu wywołanych tym zdarzeniem koncentrycznie rozchodzących się fal (stan 2). Dopuszczalny przez mechanikę proces odwrotny, rozpoczynający się falami cofającymi się od brzegu (stan $\bar{2}$) i kończący wyrzuceniem kamienia ze środka sadzawki (stan $\bar{1}$), nie jest jednak nigdy obserwowany. Jeśli zobaczymy coś takiego na filmie, to od razu domyślimy się, że jest on wyświetlany od końca. Sprzeczność pomiędzy wyraźną nieodwracalnością zjawisk makroskopowych a spodziewaną odwracalnością zjawisk mechanicznych, sformułowana w latach 70-ych XIX wieku przez Loschmidta jako tzw. “paradoks odwracalności”, wydawała się świadczyć przeciwko rodzącej się wówczas fizyce statystycznej, usiłującej sprowadzić makroskopowe prawa termodynamiki do mikroskopowych praw mechaniki.

Pozorną sprzeczność pomiędzy nieodwracalnością zjawisk makroskopowych a odwracalnością praw mechaniki tłumaczy asymetryczna *niestabilność* trajektorii (Penrose, 1979; Prigogine i Stengers, 1990), patrz rys. 1.5.b. Kamień można wrzucić do wody (stan 1) na wiele różnych sposobów otrzymując za każdym razem podobny rezultat (stan 2). Kamień wylatujący z wody (stan $\bar{1}$) można jednak zaobserwować tylko pod warunkiem praktycznie niewykonalnej, nieskończenie do-

kładnej preparacji stanu początkowego $\bar{2}$. Już bardzo mała niedokładność sprowadza układ na całkiem inną trajektorię $\bar{2}'1'$. Również niekontrolowalnie mała niedokładność w odwróceniu pędów może prowadzić do nieobliczalnych skutków, co stawia podróżnika wsiadającego do opisanego wyżej “wehikułu czasu” w niezbyt bezpiecznej sytuacji.

Z symetrii (1.12) wynika, że praktyczna niemożliwość powrotu do przeszłości jest równoważna praktycznej niemożliwości określenia stanu tej przeszłości: dwa nieodróżnialne stany końcowe 2 i 2' mogą wynikać z całkiem różnych stanów początkowych 1 lub 1' (patrz rys. 1.5.b).

Niestabilne układy mechaniczne okazują się być praktycznie nieodwracalnymi. Nie ma więc istotnej różnicy pomiędzy równaniami (1.7) i równaniami (1.5). Podobnie jak nie wszystkie wartości wychylenia i prędkości kątowej wahadła mają sens fizyczny, nie wszystkie stany (q, p) w przestrzeni fazowej muszą być praktycznie (z dostateczną dokładnością) możliwe do realizacji (Brillouin, 1969). Poza stanem “dokładnie do góry nogami” wszystkie pozostałe trajektorie płaskiego wahadła są stabilne tylko dlatego, że jest ono układem o jednym stopniu swobody, jego przestrzeń fazowa jest dwuwymiarowa, a ciągłe trajektorie w takiej przestrzeni muszą być stabilne. Stabilne układy o większej liczbie stopni swobody stanowią przypadek niezwykle rzadki. Badania ostatnich lat dostarczają coraz to nowych dowodów na to, że układy mechaniczne o kilku tylko stopniach swobody (w tym także nasz Układ Planetarny, tradycyjny obiekt badań mechaniki od ponad 300 lat) zawierają w przestrzeni fazowej tylko niewielkie obszary wypełnione trajektoriami w pełni stabilnymi. Znacznie bardziej typowy jest ruch częściowo stabilny (tylko dla wybranych stopni swobody) lub wręcz niestabilny (Berry, 1979).

1.5 Atomowo-cząsteczkowa struktura materii

Zjawisko tarcia, czy ogólniej, *dysypacji* (nieodwracalnej utraty energii w wyróżnionym stopniu swobody) związane jest z faktem, że w rozważane procesy zaangażowane są ogromne liczby mikroskopowych cząsteczek. Termodynamika jest teorią zjawisk fizycznych w skali makroskopowej, ale własności obiektów, którymi się zajmuje, są konsekwencją ich struktury mikroskopowej. Blisko dziesięć pokoleń uczonych pracowało nad sformulowaniem i uzasadnieniem dwóch podstawowych tez, których dzisiaj nikt już nie podaje w wątpliwość. Wymieniając je może warto przypomnieć najważniejsze kroki (Ochoa i Corey, 1999).

(1) Materia w skali makroskopowej zbudowana jest z ogromnej liczby cząsteczek

- D. Bernoulli, rok 1738 — połączenie mechaniki newtonowskiej z rachunkiem prawdopodobieństwa, początki kinetycznej teorii gazów,
- A. Avogadro, rok 1811 — sformułowanie hipotezy, że jednakowe objętości

różnych chemicznie gazów zawierają taką samą liczbę cząsteczek,

- A. K. Krönig, rok 1856, R. E. Clausius, rok 1857 — pełna statystyczna interpretacja praw gazowych,
- J. Loschmidt, rok 1865 — oszacowanie rzędu liczby Avogadro na 10^{23} cząsteczek w jednym molu,
- J. C. Maxwell, rok 1866, L. Boltzmann, rok 1872 — statystyczna interpretacja zjawisk transportu,
- A. Einstein, rok 1905, M. Smoluchowski, rok 1906 — fluktuacyjna teoria ruchów Browna,
- T. Svedberg, J. B. Perrin, rok 1909 — potwierdzenie teorii fluktuacji termodynamicznych obserwacjami dyfuzji w roztworach koloidalnych i emulsjach,
- G. Binning, H. Rohrer i współpracownicy, rok 1982 i 1986 — skonstruowanie skaningowego mikroskopu tunelowego, a potem mikroskopu sił atomowych, pozwalających na oglądanie i manipulowanie pojedynczymi cząsteczkami.

(2) Cząsteczki składają się z atomów, a atomy z ujemnie naładowanych elektronów i dodatnio naładowanych jąder

- J. Dalton, rok 1805 — wyjaśnienie chemicznych praw stosunków stałych i wielokrotnych,
- M. Faraday, rok 1833 — prawa elektrolizy,
- D. Mendelejew, L. Meyer, rok 1869 — układ okresowy pierwiastków,
- J. H. van't Hoff, rok 1874 — zasady stereochemii,
- S. Arrhenius, rok 1884 — teoria jonowa elektrolitów,
- J. J. Thomson i E. Wiechert, rok 1897, R. A. Millikan, rok 1911 — wyznaczenie ładunku i masy elektronu,
- E. Rutherford, rok 1911 — odkrycie jądra atomowego,
- M. von Laue, rok 1912, W. H i W. L. Braggowie, rok 1913 — dyfrakcja promieni rentgenowskich na kryształach,
- N. Bohr, rok 1913, W. Heisenberg i E. Schrödinger, rok 1926 — elektrony i jądra oddziałują siłami elektromagnetycznymi zgodnie z prawami mechaniki kwantowej,
- F. Hund, W. Heitler i F. London, rok 1927 — podstawy kwantowo-mechanicznej teorii wiązania chemicznego,

- M. Born, J. R. Oppenheimer i następcy (po roku 1927) — w ustalonym kwantowym stanie elektronowym ruch jąder określony jest przez efektywny potencjał “adiabatyczny” opisujący zarówno silne wiązania chemiczne jak i słabsze, niechemiczne oddziaływania wewnątrz- i międzycząsteczkowe,
- kilka niezależnych zespołów, rok 1986 — obserwacja procesów kwantowych zachodzących w pojedynczych atomach uwięzionych w pułapkach elektromagnetycznych lub optycznych.

1.6 Zespół statystyczny. Własność mieszania i dążenie do równowagi termodynamicznej

Wielka metodologiczna rola zasady determinizmu mechanicznego polega na odróżnieniu ogólnych *praw* ruchu od jednostkowych *faktów* — danych charakteryzujących stan konkretnego układu w wybranym momencie czasu. O ile prawa fizyki znamy już zupełnie nieźle (rys. 1.3), dostatecznie dokładne określenie stanu badanych układów sprawia nam, w ogólności, trudności nie do pokonania. Wartość *metodologiczna* determinizmu mechanicznego nie musi oznaczać więc automatycznie jego wartości *prognostycznej* — praktycznego przewidywania stanów w dowolnej przyszłości lub rekonstruowania ich w dowolnej przeszłości.

Na kłopoty z określeniem stanu początkowego składa się kilka przyczyn (Brillouin, 1969; Prigogine i Stengers, 1990; Kurzyński, 1993). Po pierwsze, *wielkość* układu powodująca, że pełna charakterystyka stanu wymaga podania zbyt dużej ilości liczb. Mikrometr sześcienny ciekłej wody, obiekt o rozmiarach zbliżonych do komórki bakteryjnej, zawiera 3×10^{10} cząsteczek, których same tylko translacyjne stopnie swobodny charakteryzuje wektor stanu o 2×10^{11} składowych. Po drugie, dyskutowana już *niestabilność* ruchu. Stan niestabilnego układu, nawet o kilku tylko stopniach swobody, musi być określony z ogromną precyzją by można było przewidzieć jego ewolucję w rozsądnym przedziale czasu. Z punktu widzenia przetwarzania informacji, do opisu stanu zarówno stabilnego układu o wielkiej ilości stopni swobody jak i układu o małej ilości stopni swobody, lecz niestabilnego, potrzeba praktycznie nieskończonej ilości cyfr (bitów).

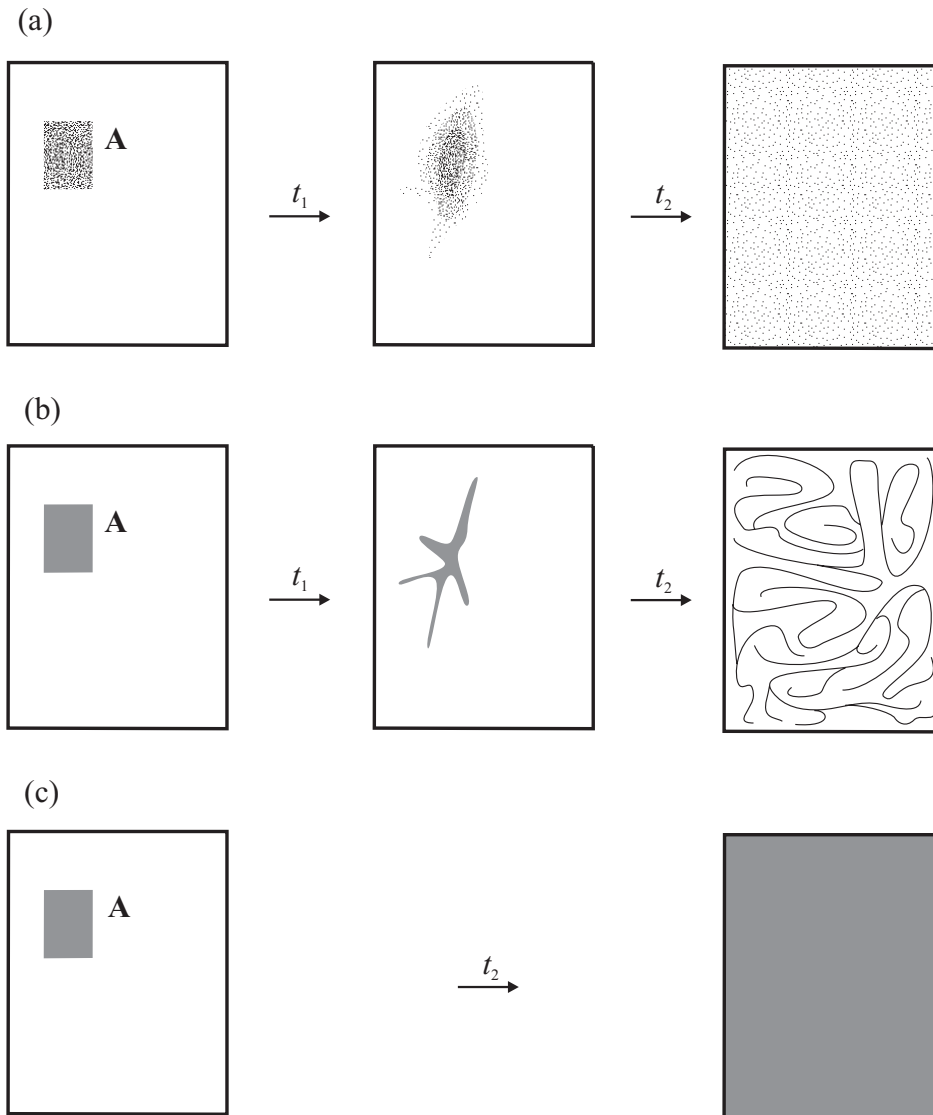
Trzecia przyczyna kłopotów jest natury bardziej fundamentalnej. Aby określić stan układu, trzeba dokonać *pomiaru*, podczas którego można ten stan zaburzyć. Zaburzenie, nieuniknione podczas sprzężenia makroskopowego (klasycznego) przyrządu pomiarowego z mikroskopowym obiektem kwantowym, uniemożliwia pełne określenie jego funkcji falowej, a tym samym wykorzystanie determinizmu zawartego w równaniu Schrödingera. Zaburzenie stanu układu przez obserwację nie jest bynajmniej domeną tylko fizyki kwantowej. Oczywiście jest na przykład rzeczą, że im dokładniej badamy stan organizmu biologicznego, tym mniej jest to dla niego obojętne. Zbyt szczegółowe dane mogą okazać się dla prognozowania wręcz nieużyteczne ze względu na śmierć obiektu w trakcie ich zbierania.

Podobnej natury jest przyczyna czwarta: niemożliwość całkowitego wyeliminowania wpływu *otoczenia* na badany układ. W układach dynamicznie niestabilnych nieprzewidywalne skutki powoduje nie tylko nieskończenie mała niedokładność w określeniu stanu, lecz także nieskończenie mały wpływ otoczenia. Bardzo obrazowy jest często przytaczany przykład pochodzący od matematyka Borela z początku XX wieku, że przesunięcie na Syriuszu masy jednego grama o jeden centymetr powoduje zmianę grawitacji wystarczającą na to, by cząsteczka gazu w atmosferze odległej Ziemi już po kilkudziesięciu zderzeniach zmieniła partnera kolejnego zderzenia! Mechanika kwantowa już z samej natury jest *holistyczna* — z faktu, że całość znajduje się w określonym stanie opisywanym funkcją falową bynajmniej nie wynika, że w takim stanie znajduje się jakaś część tej całości. W fizyce kwantowej na układ oddziałuje nawet otoczenie, którego formalnie nie ma: drgania tak zwanej próżni pola elektromagnetycznego, nie zawierającej ani jednego fotonu, powodujące spontaniczną emisję promieniowania.

Fizyka statystyczna znalazła prosty sposób obejścia problemu nieznaności stanu początkowego. Zamiast pojedynczego układu rozważa zespół wielu identycznych kopii tego układu różniących się tylko stanem początkowym. Pojedyncza trajektoria fazowa zostaje w ten sposób zastąpiona *strumieniem fazowym* (Penrose, 1979). Trajektorie wchodzące w skład strumienia mogą zachowywać się stabilnie, gdy niewielka zmiana stanu początkowego nie prowadzi do istotnych różnic w ich przebiegu, lub niestabilnie, gdy skutki nieskończenie małej zmiany stanu początkowego okazują się ogromne.

Dla fizyki statystycznej interesujące są ekstremalnie niestabilne układy mechaniczne posiadające własność *mieszania* strumienia w postępie wykładniczym (Penrose, 1979; Prigogine i Stengers, 1990). Po pewnym czasie, zwanym *czasem stochastyzacji*, stany zespołu układów o tej własności rozkładają się praktycznie jednorodnie na całym dostępnym dla ruchu obszarze przestrzeni fazowej (rys. 1.6.a). Jednorodny rozkład stanów końcowych jest zupełnie niezależny od rozkładu stanów początkowych, których znajomość okazuje się w tym wypadku zbędna. Skuteczność metod fizyki statystycznej w różnorodnych zastosowaniach jest silnym argumentem wskazującym na to, że niestabilne zachowanie z mieszaniem w postępie wykładniczym jest typowe dla układów składających się z wielu identycznych cząsteczek.

Przechodząc w granicy do continuum (nieskończonej i nieprzeliczalnej liczby) układów, których stany wypełniają początkowo pewien regularny obszar o niezerowej objętości, otrzymujemy obraz strumienia fazowego schematycznie przedstawiony na rys. 1.6.b. Ciągłość ewolucji wymaga, aby obszar stanów zespołu nie tracił swojej spójności, natomiast znane z mechaniki klasycznej twierdzenie Liouville'a żąda zachowania początkowej wartości objętości. Wynikiem ewolucji z mieszaniem w postępie wykładniczym musi być obszar o fantastycznym, rozczłonkowanym kształcie w sposób gęsty rozprzestrzeniający się na cały dostępny podzbiór przestrzeni fazowej, ale go nie pokrywający (Penrose, 1979; Prigogine i Stengers, 1990).



Rys. 1.6. Schematyczny obraz ewolucji czasowej zbioru stanów początkowych układu mechanicznego posiadającego własność mieszania w postępie wykładniczym. Po czasie stochastyzacji zbiór A rozprzestrzenia się jednorodnie na cały dostępny dla ruchu obszar przestrzeni fazowej. Czas t_1 jest krótszy, a t_2 , dłuższy od czasu stochastyzacji. (a) Przypadek, gdy zbiór A jest przeliczalny. (b) Przypadek, gdy zbiór A jest nieprzeliczalny i ma różną od zera objętość. Zgodnie z twierdzeniem Liouville'a objętość zbioru stanów nie może w czasie ewolucji ulegać zmianie, stąd rozprzestrzenienie się wyjściowego zbioru nie oznacza, że *pokrywa* on cały dostępny obszar. (c) Pogrubienie opisu ewolucji przeliczalnego zespołu statystycznego dokonywane nie tylko w chwili początkowej, ale i końcowej.

Przyjmując początkowy rozkład stanów w zespole statystycznym taki jak na rysunku 1.6.b (continuum stanów wypełniających obszar o niezerowej objętości), dokonujemy operacji *pogrubienia opisu* rzeczywistego zespołu, który w najlepszym wypadku może składać się z bardzo dużej, ale zawsze przeliczalnej liczby elementów. Oczywiście wynik ewolucji — jednorodne rozprzestrzenienie się stanów na cały dostępny podzbiór przestrzeni fazowej — jest, w wypadku układów z mieszaniem, niezależny od rozkładu stanów początkowych i od tego, czy jest ich przeliczalna liczba, czy continuum. Operacja pogrubiania opisu ma jednak głęboki sens fizyczny, gdyż uwzględnia niepewność pomiarową przy określaniu stanu (zastąpienie każdego punktu przez wielowymiarowy sześcian lub sferę — “piksel” — o rozmiarach określonych przez błąd pomiaru). Operacja pogrubienia powinna być dokonywana bezpośrednio przed każdym fizycznym opisem rozkładu stanów w zespole. W tym sensie opis rozkładu końcowego stanów przedstawiony na rysunku 1.6.b z pewnością nie jest fizyczny. W wypadku układów mechanicznych z mieszaniem operacja pogrubienia nie jest przemiana z operacją ewolucji. Pogrubienie opisu *po* zakończeniu ewolucji prowadzi do jednorodnego rozkładu stanów zespołu *pokrywającego* w sposób ciągły cały dostępny dla ruchu podzbiór przestrzeni stanów (rys. 1.6.c).

Stan zespołu z jednorodnym rozkładem stanów składających się na niego układów na całym dostępnym podzbiórze przestrzeni stanów nazywamy *stanem równowagi termodynamicznej* (Penrose, 1979; Prigogine i Stengers, 1990).

1.7 Prawdopodobieństwo i entropia. Prawo wzrostu entropii

W sposób formalny pogrubiony opis ewolucji zespołu statystycznego, schematycznie przedstawiony na rys. 1.6, wyrazić można w języku rachunku prawdopodobieństwa (van Kampen, 1990, rozdz. 1). W dalszym ciągu zajmować będziemy się tylko zespołami statystycznymi układów ewoluujących zgodnie z prawami mechaniki klasycznej. Stany układów w zespole (punkty w przestrzeni fazowej) oznaczamy będziemy przez s , a całą przestrzeń stanów przez \mathbf{S} . W ogólności, każdy pogrubiony rozkład stanów w zespole opisuje pewna funkcja *gęstości rozkładu prawdopodobieństwa* $\rho(s)$ zadana na \mathbf{S} . Prawdopodobieństwo tego, że stan układu należy do określonego podzbioru \mathbf{A} przestrzeni fazowej \mathbf{S} , wyraża całka objętościowa (na ogół bardzo wysoko wymiarowa)

$$P(\mathbf{A}) = \int_{\mathbf{A}} ds \rho(s). \quad (1.13)$$

Gęstość prawdopodobieństwa musi być unormowana do jedności, czyli całka po całej przestrzeni stanów

$$\int_{\mathbf{S}} ds \rho(s) = 1, \quad (1.14)$$

co oznacza po prostu, że każdy układ znajduje się na pewno w jakimś stanie.

Dla zadanej gęstości rozkładu prawdopodobieństwa ρ i wybranej *zmiennej dynamicznej*, czyli funkcji rzeczywistej na zbiorze stanów

$$\mathcal{X}(s) = x, \quad (1.15)$$

np. energii (1.6), można, traktując funkcję \mathcal{X} jako *zmienną losową* rachunku prawdopodobieństwa, określić jej *wartość oczekiwaną*

$$X = \langle \mathcal{X} \rangle \equiv \int_{\mathbf{S}} ds \rho(s) \mathcal{X}(s). \quad (1.16)$$

Operacja przypisania zmiennym dynamicznym ich wartości oczekiwanych jest, jako całka, liniowa:

$$\langle a\mathcal{X} + b\mathcal{Y} \rangle = a\langle \mathcal{X} \rangle + b\langle \mathcal{Y} \rangle \quad (1.17)$$

dla dowolnych dwóch zmiennych dynamicznych \mathcal{X} i \mathcal{Y} oraz dowolnych dwóch liczb rzeczywistych a i b . Różnica zmiennej dynamicznej i jej wartości oczekiwanej

$$\delta\mathcal{X} \equiv \mathcal{X} - \langle \mathcal{X} \rangle \quad (1.18)$$

nazywa się *fluktuacją*. Z własności (1.17) wynika, że wartość oczekiwana fluktuacji równa jest zeru. Fluktuacje z równym prawdopodobieństwem mogą przyjmować wartości dodatnie i ujemne. Niezerową miarą fluktuacji jest *odchylenie standardowe*, czyli pierwiastek z wartości oczekiwanej kwadratu fluktuacji:

$$\sigma = \sqrt{\langle (\delta\mathcal{X})^2 \rangle}. \quad (1.19)$$

Załóżmy, że znane jest prawdopodobieństwo $P(\mathbf{A})$ tego, że stan układu w zespole należy do pewnego zbioru \mathbf{A} przestrzeni \mathbf{S} i poddamy układ obserwacji. W wyniku stwierdzenia (zmierzenia), że stan układu fizycznego rzeczywiście należy do zbioru \mathbf{A} (w rachunku prawdopodobieństwa zbiór \mathbf{A} ma sens *zdarzenia*) obserwator uzyskuje pewną *ilość informacji* o badanym układzie i pierwotne prawdopodobieństwo $P(\mathbf{A})$ zamienia się w pewność. Ilość informacji jest tym większa, im mniejsze było prawdopodobieństwo $P(\mathbf{A})$. Odwrotnie, jeśli $P(\mathbf{A})$ było duże, uzyskana ilość informacji jest niewielka; innymi słowy, ilość informacji I jest malejącą funkcją P . Ilość informacji o zajściu dwóch zdarzeń niezależnych (którego prawdopodobieństwo jest równe iloczynowi prawdopodobieństw) powinna być sumą ilości informacji o zajściu każdego zdarzenia z osobna. Jedyna funkcja prawdopodobieństwa posiadająca obie te własności ma postać

$$I(\mathbf{A}) = -k \log_b P(\mathbf{A}), \quad (1.20)$$

gdzie k i b są pewnymi stałymi. Logarytm jedności jest równy zeru: jeśli zajście zdarzenia \mathbf{A} było już przed obserwacją pewne, $P(\mathbf{A}) = 1$, nie uzyskuje się żadnej nowej informacji, $I(\mathbf{A}) = 0$. Formuła (1.20) została wprowadzona w głośnej pracy Shanona z roku 1948 o teorii informacji (Brillouin, 1969).

Wybór stałych k i b określa jednostkę ilości informacji. Przyjmując $k = 1$ i $b = 2$ otrzymujemy jednostkę zwaną *bitem*. Jeden bit jest to ilość informacji uzyskana przy realizacji jednego z dwóch możliwych równoprawdopodobnych ($P = \frac{1}{2}$) zdarzeń reprezentowanych przez cyfrę w układzie dwójkowym (ang. *binary digit*) 0 lub 1:

$$I = -\log_2 \frac{1}{2} = 1. \quad (1.21)$$

Ilość informacji zawarta w trzycyfrowej liczbie zapisanej w układzie dziesiętnym wynosi

$$I = -\log_2 P_1 P_2 P_3 = -\log_2 \frac{1}{9} - \log_2 \frac{1}{10} - \log_2 \frac{1}{10} = 9,28$$

bitów (prawdopodobieństwa wyboru wartości każdej z cyfr są od siebie niezależne, ilość informacji można więc dodawać; pierwsza cyfra nie może być zerem i przyjmuje tylko 9 wartości i dopiero następne mogą przyjmować 10 wartości). Nieco mniejsza jest ilość informacji zawarta w dowolnym ciągu 8 cyfr w układzie dwójkowym

$$I = -\log_2 \left(\frac{1}{2}\right)^8 = \log_2 2^8 = 8$$

bitów, czyli 1 *bajt*.

Jeżeli znane jest nie tylko prawdopodobieństwo $P(\mathbf{A})$ należenia stanu układu fizycznego do pewnego podzbioru \mathbf{A} przestrzeni stanów \mathbf{S} , ale cała gęstość rozkładu prawdopodobieństwa $\rho(s)$ na \mathbf{S} , można, zgodnie z (1.16), określić *oczekiwaną ilość informacji*, jaką zyska obserwator po ustaleniu stanu s układu:

$$S = -k_B \int ds \rho(s) \ln \rho(s). \quad (1.22)$$

Równość (1.22) jest identyczna z wyrażeniem zaproponowanym najpierw przez Boltzmann'a w roku 1872, a później, w ogólnym przypadku dowolnego zespołu statystycznego, przez Gibbs'a w roku 1902, w celu statystycznej interpretacji pojęcia *entropii*. Nazwą tą Clausius w roku 1865 obdarzył iloraz energii "związanej" (tej części energii wewnętrznej, której nie można użyć do wykonania pracy) i temperatury. Tutaj \ln oznacza logarytm naturalny o podstawie $e = 2,718\dots$, natomiast $k_B = 1,38 \times 10^{-23}$ J/K jest wartością stałej występującej w równaniu stanu gazu doskonałego podzielonej przez liczbę cząsteczek w jednym molu i nosi nazwę *stałej Boltzmann'a*. Jednostkami entropii są jednostki energii podzielone przez jednostki temperatury.

Równanie (1.22) wymaga pewnego komentarza. Sens matematyczny może mieć oczywiście tylko logarytm wielkości bezwymiarowej, natomiast gęstość prawdopodobieństwa ma wymiar odwrotności objętości w przestrzeni fazowej. Z postaci równań Hamiltona (1.5) wynika, że iloczyn dowolnej współrzędnej uogólnionej i sprzężonego z nią pędu ma zawsze wymiar energii pomnożonej przez czas, czyli działania. Jednostka objętości przestrzeni fazowej ma więc wymiar działania w potęgze n , gdzie n jest liczbą stopni swobody układu. Ustalając jednostkę

działania przechodzimy do bezwymiarowej funkcji gęstości rozkładu prawdopodobieństwa i nadajemy równaniu (1.22) sens matematyczny. Ustalenie jednostki działania związane jest z dokładnością, z jaką obserwator określa mikroskopowy stan s układu. Jest to oczywiście sprawa takiej czy innej umowy, lecz istnieje naturalna, fizyczna granica dokładności określenia stanu klasycznego wyznaczona przez kwant działania (stałą Plancka) $h = 6,6 \times 10^{-34}$ J·s.

Niech \mathbf{A} będzie wybranym obszarem przestrzeni stanów. Policzmy entropię (1.22) dla gęstości rozkładu prawdopodobieństwa stałej na zbiorze \mathbf{A} , lecz poza nim znikającej (porównaj rys. 1.6.b):

$$\rho(s) = \begin{cases} \Omega^{-1} & \text{jeśli } s \text{ należy do } \mathbf{A} \\ 0 & \text{jeśli } s \text{ nie należy do } \mathbf{A}. \end{cases} \quad (1.23)$$

Aby prawdopodobieństwo było unormowane do jedności, równanie (1.14), stała Ω musi mieć sens objętości obszaru \mathbf{A} :

$$\Omega = \int_{\mathbf{A}} ds. \quad (1.24)$$

Mamy

$$S = -k_B \int_{\mathbf{S}} ds \rho(s) \ln \rho(s) = k_B \int_{\mathbf{A}} ds \Omega^{-1} \ln \Omega = k_B \ln \Omega, \quad (1.25)$$

czyli entropia jest proporcjonalna do logarytmu objętości obszaru \mathbf{A} , a dokładniej, liczby zawartych w \mathbf{A} jednostkowych elementów objętości określających dokładność wyznaczenia stanu. W oryginalnym sformułowaniu Boltzmanna entropia jest proporcjonalna do logarytmu liczby możliwych realizacji danego makrostanu (czyli stanu zespołu statystycznego) przez mikrostan (stany układów tworzących ten zespół).

Ponieważ objętość w przestrzeni fazowej mechaniki klasycznej jest zachowana w czasie ewolucji (twierdzenie Liouville'a), również entropia zespołu statystycznego, policzona dla rozkładu prawdopodobieństwa $\rho(s)$ ewoluującego ściśle według równań ruchu, pozostaje niezmienna (rysunek 1.6.b). Ale, jak stwierdziliśmy, sens fizyczny ma rozkład prawdopodobieństwa pogrubiany za każdym razem gdy chcemy opisać zespół. W układach z mieszaniem pogrubiany rozkład prawdopodobieństwa rozszerza się w określonej skali czasowej na cały dostępny podzbiór stanów (rysunek 1.6.c). W konsekwencji, entropia pogrubianego rozkładu w stanie równowagi termodynamicznej jest *większa* niż w stanie początkowym. Niestabilność ruchu, czyli zjawisko mieszania, prowadzi do prawa wzrostu entropii w czasie.

1.8 Prawo wielkich liczb

Wydawałoby się, że prawdopodobieństwo, a co za tym idzie i entropia, jako miara informacji posiadanej przez obserwatora przed pomiarem, charakteryzuje raczej

obserwatora niż układ. Prawdopodobieństwo można jednak wyznaczyć eksperymentalnie (Jaynes, 1978). Jako przykład rozważmy doświadczenie z rzucaniem monetą. Zasada racji niedostatecznej (sformułowana przez Daniela Bernoulliego, którego ojciec i stryj wiele godzin spędzili na tej zabawie) każe przypuszczać, że prawdopodobieństwa wyrzucenia orła lub reszki są takie same, czyli każde jest równe $P = 0,5$. Po serii rzutów może okazać się jednak, że orzeł pojawiał się średnio 0,6 razy na rzut. Jeśli liczba rzutów była dostatecznie duża wyciągamy stąd wniosek, że moneta nie jest symetryczna. W dalszych doświadczeniach spodziewamy się wyrzucenia orła z prawdopodobieństwem $P = 0,6$. Wartość prawdopodobieństwa charakteryzuje więc nie obserwatora a badany układ, w tym wypadku jego asymetrię.

Wartość prawdopodobieństwa wyznacza się w *doświadczeniu statystycznym* wykonanym na dostatecznie dużej *próbie statystycznej*. O elementach składowych próby zakładamy, że są *statystycznie niezależne*, pod czym rozumiemy brak korelacji pomiędzy zmiennymi losowymi określającymi poszczególne elementy próby. Dwie zmienne losowe (zmienne dynamiczne) \mathcal{X} i \mathcal{Y} nazywamy *nieskorelowanymi*, jeśli wartość oczekiwana ich iloczynu jest równa iloczynowi wartości oczekiwanych:

$$\langle \mathcal{X}\mathcal{Y} \rangle = \langle \mathcal{X} \rangle \langle \mathcal{Y} \rangle. \quad (1.26)$$

Ponieważ z definicji fluktuacji (1.18) i własności liniowości (1.17) wynika, że

$$\langle \mathcal{X}\mathcal{Y} \rangle - \langle \mathcal{X} \rangle \langle \mathcal{Y} \rangle = \langle \delta\mathcal{X}\delta\mathcal{Y} \rangle, \quad (1.27)$$

brak korelacji zmiennych \mathcal{X} i \mathcal{Y} oznacza znikanie odpowiedniej *funkcji korelacji fluktuacji*:

$$\langle \delta\mathcal{X}\delta\mathcal{Y} \rangle = 0. \quad (1.28)$$

Konsekwencją warunku statystycznej niezależności jest niezwykle ważne twierdzenie zwane *prawem wielkich liczb*. Rozważmy próbę statystyczną N identycznych zmiennych losowych (dynamicznych) $\mathcal{X}^{(l)}$ o tej samej wartości oczekiwanej X i odchyleniu standardowym σ :

$$\langle \mathcal{X}^{(l)} \rangle = X, \quad \langle (\delta\mathcal{X}^{(l)})^2 \rangle = \sigma^2. \quad (1.29)$$

O zmiennych $\mathcal{X}^{(l)}$ zakładamy, że są nieskorelowane:

$$\langle \delta\mathcal{X}^{(l)}\delta\mathcal{X}^{(l')} \rangle = 0 \quad (1.30)$$

dla $l \neq l'$. Przez *średnią statystyczną* ciągu zmiennych ($\mathcal{X}^{(l)}$) rozumiemy zmienną losową będącą ich średnią arytmetyczną:

$$\bar{\mathcal{X}} \equiv \frac{1}{N} \sum_{l=1}^N \mathcal{X}^{(l)}. \quad (1.31)$$

Z własności liniowości (1.17) wynika, że wartość oczekiwana średniej statystycznej $\bar{\mathcal{X}}$ pokrywa się z wartością oczekiwaną X :

$$\langle \bar{\mathcal{X}} \rangle = \left\langle \frac{1}{N} \sum_{l=1}^N \mathcal{X}^{(l)} \right\rangle = \frac{1}{N} \sum_{l=1}^N \langle \mathcal{X}^{(l)} \rangle = X. \quad (1.32)$$

Korzystając natomiast z niezależności statystycznej (1.30) zmiennych $\mathcal{X}^{(l)}$, otrzymujemy

$$\begin{aligned} \langle (\delta \bar{\mathcal{X}})^2 \rangle &= \langle (\bar{\mathcal{X}} - X)^2 \rangle = \left\langle \left[\frac{1}{N} \sum_l (\mathcal{X}^{(l)} - X) \right]^2 \right\rangle \\ &= \frac{1}{N^2} \left\langle \left[\sum_l \delta \mathcal{X}^{(l)} \right]^2 \right\rangle = \frac{1}{N^2} \sum_{l,l'} \langle \delta \mathcal{X}^{(l)} \delta \mathcal{X}^{(l')} \rangle = \frac{\sigma^2}{N} \end{aligned} \quad (1.33)$$

(z sumy N^2 składników w przedostatnim wyrażeniu tylko N różnych jest od zera), czyli odchylenie standardowe średniej statystycznej $\bar{\mathcal{X}}$ od wartości oczekiwanej X ,

$$\sqrt{\langle (\delta \bar{\mathcal{X}})^2 \rangle} = \frac{\sigma}{\sqrt{N}}. \quad (1.34)$$

W granicy $N \rightarrow \infty$ to odchylenie dąży do zera, a więc wartość średniej statystycznej $\bar{\mathcal{X}}$ pokrywa się z wartością oczekiwaną X , czyli

$$\bar{\mathcal{X}} = X. \quad (1.35)$$

Jest to teza prawa wielkich liczb. Średnia statystyczna $\bar{\mathcal{X}}$ z dostatecznie dużej próby jest zmienną losową która, w przeciwieństwie do zmiennych $\mathcal{X}^{(l)}$, nie wykazuje *żadnych* fluktuacji.

Prawdopodobieństwo P wyrzucenia orła w ciągu doświadczeń Bernoulliego można uważać za wartość oczekiwaną dychotomicznej (przyjmującej tylko dwie wartości) zmiennej losowej \mathcal{P} określonej formułą

$$\mathcal{P} = \begin{cases} 1 & \text{jeśli wynikiem rzutu jest orzeł} \\ 0 & \text{jeśli wynikiem rzutu jest reszka.} \end{cases} \quad (1.36)$$

Rzeczywiście, z równań (1.16) i (1.13) wynika, że $\langle \mathcal{P} \rangle$ określa prawdopodobieństwo zbioru \mathbf{A} stanów prowadzących do upadku monety orłem do góry. Na mocy prawa wielkich liczb P jest równe wartości średniej $\bar{\mathcal{P}}$, czyli ułamkowi określającemu stosunek ilości rzutów pomyslnych (liczonych ze współczynnikiem 1) do ilości wszystkich rzutów w dostatecznie dużej próbie. Możliwa różnica pomiędzy wartością średnią uzyskaną dla niedostatecznie dużej próby a rzeczywistą wartością P jest interpretowana jako błąd pomiaru w doświadczeniu statystycznym (Jaynes, 1978).

1.9 Fizyczne realizacje próby statystycznej

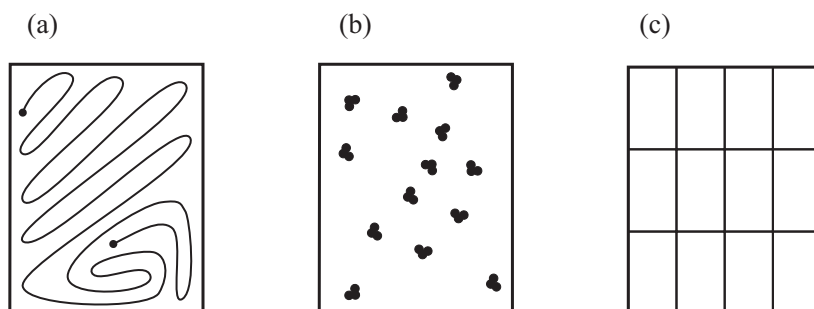
Prawo wielkich liczb pozwala utożsamiać wartości oczekiwane rachunku prawdopodobieństwa z wartościami średnimi z dostatecznie dużej próby statystycznej. Fizyka statystyczna nazywa zresztą wprost wielkość $X = \langle \mathcal{X} \rangle$ *średnią zmienną dynamiczną* \mathcal{X} i nie używa nazwy *wartość oczekiwana* (van Kampen, 1990, rozdz. 1). Próbę statystyczną może tworzyć zbiór *wielu* identycznych doświadczeń wykonywanych w różnych chwilach czasu nad tym samym *pojedynczym* układem lub *jedno* doświadczenie (pomiar) wykonane równocześnie nad *wieloma* identycznymi układami tworzącymi *zespół* statystyczny. O ile jednak nie ma w zasadzie problemu z realizacją obu rodzajów prób w przypadku wyznaczania prawdopodobieństwa wyrzucenia orła (zamiast podrzucić wiele razy tę samą monetę możemy rozrzucić od razu cały worek monet, jeżeli tylko taki posiadamy), to zastosowanie zarówno jednej jak i drugiej metody w wypadku statystycznego opisu makroskopowych układów fizycznych wymaga pewnej ostrożności.

Ludwig Boltzmann, jeden z wielkich twórców fizyki statystycznej, podkreślał, że podstawową cechą pomiaru makroskopowego jest uśrednianie *po czasie* wartości obserwowanej zmiennej dynamicznej w następujących po sobie różnych stanach mikroskopowych układu. Wynika ono ze skończoności czasu trwania pomiaru oraz z bezwładności używanego makroskopowego przyrządu. Pomiar takim przyrządem automatycznie realizuje doświadczenie statystyczne na próbie pierwszego rodzaju. Kłopot polega na tym, że nie można oczekiwać by zmienne dynamiczne $\mathcal{X}(t)$, opisujące daną własność ciągle tego samego układu w różnych chwilach czasowych,¹ były w ogólności statystycznie niezależne. Drugi wielki twórca fizyki statystycznej, Josiah Willibard Gibbs, autor pojęcia zespołu statystycznego które już wprowadziliśmy w paragrafie 1.6, traktował obserwowane wielkości makroskopowe jako średnie z próby statystycznej stanowiącej ten właśnie zespół. Tutaj kłopot polega na tym, że zespół statystyczny wielu identycznych kopii tego samego układu był w zamysle swego autora tworem raczej *konceptyjnym* niż rzeczywistym. Zawsze mamy do czynienia z jedną tylko kopią rozważanego układu makroskopowego, stąd trudno mówić o wykonywaniu na nim doświadczeń statystycznych.

Jeśli fizyka statystyczna ma prawo wykorzystywać rzeczywiste realizacje obu rodzajów prób statystycznych, to tylko dlatego, że zajmuje się układami makroskopowymi oraz przypisuje tym układom dyskutowaną w paragrafie 1.6 własność eksponencjalnie szybkiego mieszania (podkreślmy, że mieszanie jest indywidualną własnością układu, a nie zespołu). Własność mieszania jest matematycznie niezwykle silna. Po pierwsze, wynika z niej *ergodyczność* (Penrose, 1979), równość *średniej czasowej* zmiennej dynamicznej \mathcal{X} ,

$$\overline{\mathcal{X}} \equiv \frac{1}{\tau} \int_0^\tau dt \mathcal{X}(t) \quad (1.37)$$

¹Zamiast o zmianie stanu układu w czasie można mówić, równoważnie, o zmianie w czasie zmiennej dynamicznej — funkcji rzeczywistej określonej na tym stanie.



Rys. 1.7. Trzy fizyczne realizacje próby statystycznej. (a) Pojedyncza kopia układu ergodycznego, który w czasie pomiaru odwiedza dostatecznie jednorodnie rozłożony zbiór stanów w całym dostępnym dla ruchu obszarze przestrzeni fazowej. (b) Gaz doskonały wielu identycznie statystycznie nieskorelowanych cząsteczek. (c) Układ silnie oddziałujących cząsteczek, który w myśli można podzielić na wiele w przybliżeniu identycznych i praktycznie niezależnych podukładów wielocząsteczkowych.

dla czasu τ rzędu czasu stochastyzacji, z wartością oczekiwaną $\langle \mathcal{X} \rangle$ określoną przez formułę (1.16):

$$\bar{\mathcal{X}} = \langle \mathcal{X} \rangle. \quad (1.38)$$

W ten sposób słuszna staje się teza prawa wielkich liczb (1.35) bez konieczności przyjęcia założenia braku korelacji pomiędzy poszczególnymi elementami próby $\mathcal{X}(t)$. Ergodyczność oznacza, że w czasie uśredniania układ odwiedził dostatecznie jednorodnie rozłożony zbiór stanów w całym dostępnym dla ruchu obszarze przestrzeni fazowej (Kurzyński, 1993), patrz rys. 1.7.a.

Warunek ergodyczności przez ponad pół wieku wydawał się kluczowy dla statystycznego uzasadnienia praw termodynamiki, ale dzisiaj jest oczywiste, że ma znaczenie tylko w kontekście silniejszego warunku mieszania, gdyż tylko ten warunek zapewnia *dochodzenie* do równowagi termodynamicznej. Sama ergodyczność wystarcza, gdy rozpatruje się własności układu, który już jest w równowadze termodynamicznej, ale procesy dochodzenia do równowagi muszą być opisane z wykorzystaniem innej własności niestabilnych układów z mieszaniem. Tą własnością jest *czasowy zanik korelacji*: dla dowolnych dwóch zmiennych dynamicznych \mathcal{X} i \mathcal{Y}

$$\langle \mathcal{X}(\tau)\mathcal{Y} \rangle \rightarrow \langle \mathcal{X} \rangle \langle \mathcal{Y} \rangle \quad (1.39)$$

dla czasu τ rzędu czasu stochastyzacji (w stanie równowagi termodynamicznej $\langle \mathcal{X}(t) \rangle = \langle \mathcal{X} \rangle$). Zanik korelacji jest formalnym przejawem procesu stochastyzacji (Penrose, 1979).

Podkreślmy wyraźnie, że własność ergodyczności czy mieszania mogą posiadać już bardzo małe układy mechaniczne, na przykład dwie sztywne kule na stole bilardowym, co formalnie udowodnił Sinai w roku 1966. Sama własność mieszania nie wystarcza więc jeszcze do uzasadnienia podstaw fizyki statystycznej. Dopiero w wypadku układu *makroskopowego*, składającego się z wielu identycznych cząsteczek jednego lub kilku rodzajów, proces stochastyzacji zakończyć się może

podziałem układu na wiele *identycznych* nieskorelowanych podukładów (Penrose, 1979). Pojedynczy układ makroskopowy w takim stanie stanowi *rzeczywisty* zespół statystyczny i można sensownie mówić o tym stanie jako o stanie równowagi termodynamicznej (pamiętajmy, że pojęcie to odnosi się tylko do zespołu, czy ogólniej, próby statystycznej).

Rodzaj podukładów, na jakie rozpada się konkretny makroskopowy układ fizyczny zależy od charakteru tego układu, jego struktury wewnętrznej i siły wewnętrznych oddziaływań. Najprostszy zespół statystycznie niezależnych podukładów tworzy *gaz doskonały*, zbiór wielu identycznych cząsteczek z zanedbywalnie słabymi oddziaływaniami (rys. 1.7.b).¹ Pomimo prostoty, fizyka statystyczna gazów doskonałych, zarówno klasycznych, jak i kwantowych, wyjaśnia wiele nietrywialnych zjawisk. Ale model ten nie jest już w stanie opisać niektórych własności fazy skondensowanej np. krótkofalowych wzbudzeń kolektywnych czy zjawiska przemian fazowych, związanych ze spontanicznym porządkowaniem się układu, w których oddziaływania pomiędzy cząsteczkami i ich wzajemne korelacje są istotne (Chandler, 1987). W takim wypadku powszechnie stosowanym rozwiązaniem jest podział układu (ze względu na jego dowolność raczej w myśli niż w rzeczywistości) na większe, w przybliżeniu identyczne części (rys. 1.7.c). Jeśli części te są dostatecznie duże, wzajemne korelacje i oddziaływania na granicach są do zanedbania, tak jak w gazie doskonałym. Wyjątek stanowi *stan krytyczny* układu w warunkach bliskich *ciągłej* przemiany fazowej, gdy korelacje stają się nieskończenie daleko zasięgowe, a odchylenia standardowe nieskończenie wielkie (patrz rozdz. 4).

¹ Jeśli cząsteczki są nieodróżnialne, elementarna objętość przestrzeni fazowej określająca dokładność, z jaką obserwator wyznacza stan układu makroskopowego (patrz komentarz do definicji (1.22)), musi zawierać wszystkie $N!$ permutacji elementarnych objętości z ustaloną numeracją cząsteczek (Chandler, 1987).