Chapter 1

Introdução

1.1 Métodos

A física de equilíbrio é responsável por uma enorme quantidade de resultados importantes, tanto do ponto de vista teórico quanto do ponto de vista aplicado. O comportamento de gases, sistemas termodinâmicos, reações químicas, transições de fase, entre outros, pode ser bastante bem entendido ao usarmos o ferramental desenvolvido para o estudo da física de equilíbrio. O teorema fundamental da mecânica estatística é o principal instrumento a partir do qual a mecânica estatística e a termodinâmica derivam.

A física de equilíbrio é apenas uma aproximação. Essa aproximação é bastante boa para sistemas grandes ($N \to \infty$) submetidos a condições externas estáveis. Contudo, classes muito importantes de sistemas não se enquadram nestas categorias. Dentre os mais importantes estão os sistemas biológicos vivos, que são sistemas abertos e totalmente fora do equilíbrio.

Infelizmente não existem ferramentas semelhantes desenvolvidas para o tratamento de sistemas fora de equilíbrio. Existem coleções de métodos apropriados para cada situação especıfica dado que a solução exata das equações de movimento para um sistema macro-(ou meso-)scópico é totalmente impraticável.

Nos últimos anos, dois desenvolvimentos gerais levaram a um aumento do interesse geral por sistemas fora do equilíbrio: o gigantesco progresso de técnicas computacionais, que em muito ampliaram a quantidade de sistemas accessíveis a estudo direto; e o desenvolvimento de técnicas experimentais que permitem a manipulação de sistemas pequenos (átomos individuais, transistores a um elétron, experimentos com moléculas únicas, nanotubos,

etc), para os quais a termodinâmica tradicional não se aplica.

Neste curso tentaremos dar ao estudante um painel introdutório sobre diversos aspectos de sistemas fora do equilíbrio, seguindo uma linha que começa nas equações de Hamilton e termina em aspectos atuais, tais como Teoremas de Flutuação.

1.2 Probabilidades

1.2.1 Visão empírica

A necessidade do uso de probabilidades em física ocorre por duas razões principais: a impossibilidade experimental de pordermos levar em conta, com precisão arbitrária, a imensa quantidade de informação que pode ser obtida em um sistema; e a impossibilidade que temos de computar com precisão arbitrária a evolução do sistema devido às dificuldades numéricas envolvidas.

Uma distribuição de probabilidades tenta refletir, da maneira mais fiel possível, a frequência de resultados de uma medida, quando esta é realizada um número $N \to \infty$ de vezes. Um problema imediato é que nunca podemos ter **certeza** se a frequência **observada** corresponde à correta. Ex.: Podemos jogar um dado e obter o resultado 1 por 10 vezes seguidas. Será que temos $p(1) = 1$ e $p(2) = p(3) = p(4) = p(5) = p(6) = 0$? Provavelmente não, mas poderia se tratar de um dado viciado ou então a medida que o número de experimentos aumentasse veríamos que $p(1) \to 1/6$. Mas para um número finito de experimentos, nunca poderemos afirmar que sabemos a distribuição correta.

O objetivo da Mecânica Estatística é justamente desenvolver uma metodologia para o cálculo *a priori* de uma distribuição de probabilidades para as variáveis microscópicas de um sistema. A partir daí podemos derivar as leis da termodinâmica e hidrodinâmica. Na sequência ofereceremos argumentos convincentes de que o Postulado Fundamental da Mecânica Estatística (PFME) funciona, tanto do ponto de vista teórico como experimental. Mas é crucial observar que o PFME ainda é um postulado! Sua validade é extremamente difícil de provar para sistemas de muitas partículas.

Voltando às probabilidades, para sistemas que possuam algum tipo de simetria entre seus p estados, digamos, e quando não houver nenhuma razão física para que um estado seja o escolhido em algum experimento, em detrimento dos outros, então a escolha a priori equiprovável de $1/p$ como probabilidade pode ser justificada. Este é o caso de um dado "honesto".

1.2.2 Definição mais formal: espaço de probabilidades

O conceito fundamental em teoria de probabilidades é o espaço de probabilidades. Este consiste de três ingredientes básicos:

- Um espaço amostral de eventos elementares Ω (todos os outros eventos podem ser expressos como combinação desses eventos elementares). Estes são resultados de algum tipo de medida;
- Uma σ-álgebra de eventos;
- Uma medida de probabilidade sobre a σ-álgebra.

σ-álgebra de eventos

Um espaço amostral pode possuir muitos subconjuntos que não nos interessam (como quando Ω é infinito não-contável). Podemos especificar apenas aqueles subconjuntos que farão parte de nossa teoria. Para isso é importante que esses subconjuntos formem uma σ-álgebra, que é um sistema $\mathcal{A}$ de subconjuntos de Ω com as seguintes três propriedades:

1. $\Omega \in \mathcal{A}$ e $\emptyset \in \mathcal{A}$;
2. Se $A_1 \in \mathcal{A}$ e $A_2 \in \mathcal{A}$, então a união $A_1 \bigcup A_2 \in \mathcal{A}$, a intersecção $A_1 \bigcap A_2 \in \mathcal{A}$ e a diferença $A_1 \setminus A_2 \in \mathcal{A}$[1];
3. Dada uma coleção contável de eventos $A_1, A_2, \ldots, A_n, \ldots \in \mathcal{A} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$.

A terceira condição, que a união contável de um número infinito de subconjuntos (eventos) é um evento, é o que faz $\mathcal{A}$ ser uma σ-álgebra.

Medidas de probabilidade e axiomas de Kolmogorov

A medida de probabilidade é definida sobre $\mathcal{A}$ como o mapa

$$\mu : \mathcal{A} \to R,$$

que associa a cada $A \in \mathcal{A}$ um número real $\mu(A)$.

A medida de probabilidades deve obedecer aos axiomas de Kolmogorov

[1] A diferença pode ser definida como conjunto $D \equiv A_1 \setminus A_2 \iff \forall x \in D, x \in A_1 \wedge x \notin A_2$. Algumas vezes a diferença é notada como $A_1 - A_2$. Podemos definir que $A > B$ se $A \setminus B \neq \emptyset$ e $B \setminus A = \emptyset$.

1. $\forall A \in \mathcal{A} : 0 \leq \mu(A) \leq 1$;
2. A probabilidade é normalizada: $\mu(\Omega) = 1$;
3. Seja uma coleção contável de eventos disjuntos $A_1, A_2, \ldots, A_n, \ldots \in \mathcal{A}$ onde $A_i \bigcap A_j = \emptyset$ para todo $i \neq j$, então a probabilidade de sua união é a soma de suas probabilidades

$$\Rightarrow \mu(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n).$$

Exemplo:

Seja $\Omega = R_+$. Uma σ-álgebra pode ser definida como o conjunto de todos os segmentos $[a, b) \in \mathcal{A}$, onde $(a, b) \in R_+^2$ com $a < b$. A medida de probabilidades sobre $\mathcal{A}$ é definida como

$$\mu\left([a, b)\right) = e^{-a} - e^{-b}.$$

1.2.3 Algumas propriedades de distribuições

De uma maneira um pouco menos formal, probabilidades são medidas $p_X(x) = \mu(x)$ definidas sobre um conjunto Φ_X definido como o espaço amostral de uma variável aleatória X. Os pontos $x \in \Phi_X$ são os eventos possíveis. Ex: para o dado honesto, $\Phi_x = \{1, 2, 3, 4, 5, 6\}$ e $p_X(i = 1, \ldots, 6) = 1/6$. Uma apresentação muito boa pode ser encontrada no livro do Feller, em especial no volume I [1].

- Probabilidades nunca são negativas:

$$\forall x \in \Phi_X : p_X(x) \geq 0.$$

- Para $A, B \subset \Phi_X$, probabilidade de A ou B acontecerem:

$$p_X(A), p_X(B) \geq 0 \rightarrow p_X(A \bigcup B) = p_X(A) + p_X(B) - p_X(A \bigcap B).$$

- Para $A, B \subset \Phi_X$, probabilidade de A e B acontecerem:

$$p_X(A), p_X(B) \geq 0 \rightarrow p_X(A, B) = p_X(A \bigcap B).$$

- Probabilidades são normalizadas:

$$p_X(\Phi_X) = 1.$$

- Probabilidades condicionais $p(B|A)$ se aplicam quando sabemos que o evento $A \subset \Phi_X$ acontece equeremos saber então qual é a probabilidade de B acontecer. Podemos escrever (Regra de Bayes):

$$p(B|A) = \frac{\mu(A \bigcap B)}{\mu(A)} = \frac{p(A,B)}{p(A)}.$$

- Eventos independentes $p(B|A) = p(B)$

$$p(B|A) = \frac{p(A,B)}{p(A)} \rightarrow p(A,B) = p(A)\, p(B|A) = p(A)\, p(B).$$

1.2.4 Médias e momentos

Por simplicidade suporemos que as distribuições são contínuas.

- Médias ou momentos de X de ordem n:

$$\langle X^n \rangle = \int dx\, p(x)\, x^n;$$

- Função geratriz de médias:

$$G(k) = \left\langle e^{-ikX} \right\rangle = \sum_{n=0}^{\infty} \frac{(-ik)^n}{n!} \, \langle X^n \rangle \, ;$$

$$\langle X^n \rangle = \frac{\partial^n}{\partial(-ik)^n} G(k)\Big|_{k=0} .$$

- Função geratriz de cumulantes:

$$\ln G(k) = \ln \left\langle e^{-ikX} \right\rangle = \sum_{n=1}^{\infty} \frac{(-ik)^n}{n!} \, \langle X^n \rangle_c \, ;$$

$$\langle X^n \rangle_c = \frac{\partial^n}{\partial(-ik)^n} \ln G(k)\Big|_{k=0} .$$

  Ex:

$$\langle X \rangle_c = \langle X \rangle .$$

$$\left\langle X^2 \right\rangle_c = \left\langle X^2 \right\rangle - \langle X \rangle^2 .$$

A próxima etapa é entender a evolução no tempo das distribuições de probabilidade.

1.3 O começo de tudo: a Equação de Boltzmann

Por volta de 1870, Ludwig Boltzmann publicou vários trabalhos [2, 3, 4, 5, 6, 7, 8, 9] onde pela primeira vez se fazia, de maneira consistente e profunda, um tratamento microscópico da evolução dos gases utilizando uma teoria cinética.

Teorias cinéticas são formalmente definidas a partir das densidades de distribuição multipartículas de um modelo. Seja um modelo a N partículas; sua distribuição de probabilidades completa, ou densidade a N-partículas é

$$\rho_N(\mathbf{q}^N;\mathbf{p}^N;t) \to \int \prod_{i=1}^{N} dq_i\, dp_i\, \rho_N(\mathbf{q}^N;\mathbf{p}^N;t) = 1.$$

A distribuição a $s < N$-partículas é obtida integrando sobre as $N - s$:

$$\rho_s(q_1,\ldots,q_s;p_1,\ldots,p_s;t) = \frac{N!}{(N-s)!}\int \prod_{i=s+1}^{N} dq_i\, dp_i\, \rho_N(\mathbf{q}^N;\mathbf{p}^N;t).$$

As densidades a s-partículas estão normalizadas também, como se pode ver facilmente. A densidade a 1-partícula pode ser usada para a obtenção da maioria das grandezas termodinâmicas do sistema.

Para sistemas Hamiltonianos, podemos mostrar que existe uma hierarquia [15] de equações para a evolução de ρ_s de modo que $\dot{\rho}_s = F[\rho_1,\ldots,\rho_s,\rho_{s+1}]$. Esta é a chamada hierarquia BBGKY. [2] Assim, temos

$$\dot{\rho}_1 = F_{12}[\rho_1,\rho_2].$$

Se conseguimos escrever ρ_2 como um funcional de ρ_1, podemos então mostrar que a definição de uma **Teoria Cinética**, que é dada por

$$\dot{\rho}_1 = f[\rho_1], \tag{1.1}$$

é respeitada por ρ_1.

Boltzmann conseguiu derivar uma equação cinética para ρ_1 utilizando a hipótese do **caos molecular**: as velocidades das moléculas do gás estão descorrelacionadas antes de cada colisão. Isto permite escrever $\rho_2 \approx \rho_1\rho_1$ e obter uma equação para ρ_1. Sem entrar em muitos detalhes [15], podemos escrever a equação de Boltzmann como:

$$\frac{\partial\rho_1}{\partial t} + \frac{\mathbf{p_1}}{m}\cdot\frac{\partial\rho_1}{\partial\mathbf{q_1}} + \mathbf{F}\cdot\frac{\partial\rho_1}{\partial\mathbf{p_1}} = \frac{\partial\rho_1}{\partial t}_{coll}, \tag{1.2}$$

[2] Bogoliubov, Born, Green, Kirkwood, Yvon.

onde o termo de colisão é [15]

$$\frac{\partial \rho_1}{\partial t}_{coll} = \int d\Omega\, dp_1\, |v_2 - v_1|\, I(\Omega, |v_2 - v_1|)\, [\rho_1(p_1', t)\rho_1(p_2', t) - \rho_1(p_1, t)\rho_1(p_2, t)]\,. \tag{1.3}$$

Esta forma representa o efeito de dois termos: um termo de destruição de probabilidades

$$-\int d\Omega\, dp_1\, |v_2 - v_1|\, I(\Omega, |v_2 - v_1|)\rho_1(p_1, t)\rho_1(p_2, t),$$

que mostra que, ao colidir, uma partícula de momento p_1 tem esse momento alterado; o termo de criação

$$\int d\Omega\, dp_1\, |v_2 - v_1|\, I(\Omega, |v_2 - v_1|)\rho_1(p_1', t)\rho_1(p_2', t),$$

representa o ganho de probabilidade, ou seja, quando as partículas de momento p_1' ep_2' colidem, elas geram uma partícula de momento p_1.

A equação de Boltzmann determina a evolução no tempo da densidade a 1-partícula. Inicialmente em um estado de não-equilíbrio, o sistema se dirige ao equilíbrio. Isto é acompanhado de um aumento de entropia do sistema.

1.4 Teorema-H e Entropia

Boltzmann criou uma quantidade, H, relacionada à entropia como segue:

$$H = \int dp_1\ \ \rho_1\, \ln\, \rho_1, \tag{1.4}$$

da qual podemos calcular a evolução no tempo. Quando $\rho = \rho_{eq}$ vemos que $H = -S + c_0$. Assim, devemos ter H decrescente no tempo.

Derivando H:

$$\frac{d}{dt}H \;=\; -\int dp_1\, \frac{\partial \rho_1}{\partial t}_{coll}$$

$$= -\int d\Omega\, dp_1\, |v_2 - v_1|\, I(\Omega, |v_2 - v_1|)\ [\rho_1(p_1', t)\rho_1(p_2', t) - \rho_1(p_1, t)\rho_1(p_2, t)] \times$$

$$\times \ln\left[\frac{\rho_1(p_1', t)\rho_1(p_2', t)}{\rho_1(p_1, t)\rho_1(p_2, t)}\right] \leq 0,$$

como esperado. No equilíbrio a função H chega a um valor constante e podemos mostrar que isto acontece quando

$$\rho_1(p_1')\rho_1(p_2') = \rho_1(p_1)\rho_1(p_2) \Rightarrow \rho_1(p_1) \propto e^{-\frac{p_1^2}{2mk_BT}},$$

que é a distribuição correta de equilíbrio.

A partir da equação de Boltzmann podemos derivar equações para as médias das variáveis dinâmicas do sistema e a partir daí obter as equações hidrodinâmicas para o mesmo, no limite de longos comprimentos de onda ($k \to 0$)

1.5 Equações para médias

Uma equação de evolução para as probabilidades do tipo $\dot{\rho}(x) = F[\rho(x)]$ pode ser usada para calcular a evolução de $\langle f(x) \rangle$:

$$\frac{d}{dt} \langle f(x) \rangle = \frac{d}{dt} \int dx\, \rho(x)\, f(x) = \int dx\, \dot{\rho}(x)\, f(x) = \int dx\, F[\rho(x)]\, f(x),$$

onde supumos que ρ era normalizada.

Existem vários tipos diferentes de equações para médias. Essas são chamadas de "equações macroscópicas" por van Kampen [17]. Equações de decaimento radiativo, evolução de populações, hidrodinâmica, mode-mode-coupling são exemplos destes tipos de tratamento.

Como exemplo, para um fluido, a conservação de massa leva à equação da continuidade:

$$\frac{\partial}{\partial t} n(q,t) = \nabla \cdot (n\, v)$$

onde $n = \int dp\, \rho(p,q,t)$ e $n\, v = \int dp\, \rho(p,q,t)\, \dot{q}$.

Dados os tipos básicos de equações que encontraremos na sequência, nossos objetivos serão entender o comportamento temporal destes sistemas quando os mesmos não estão em equilíbrio.

Chapter 2

Mecânica estatística e Termodinâmica

2.1 Médias Temporais e Equilíbrio

Um sistema está em equilíbrio termodinâmico é quando, durante o tempo típico do ex[erimento Δt_{exp}, as grandezas termodinâmicas $A(t)$ medidas sobre o sistema possuem uma variação desprezível, $\frac{\partial A}{\partial t} \approx 0$. No que segue suporemos um sistema clássico com N partículas, cujas posições e momentos são dados por $(\mathbf{q}^N, \mathbf{p}^N)$.

A razão microscópica para isso é que, dado o microestado

$$\Gamma(t) \equiv (\mathbf{q}^N(t), \mathbf{p}^N(t)),$$

a distribuição de probabilidade de um sistema é definida através da **média temporal** dada por (no intervalo $(0, T_T)$):

$$p^{T_T}(\mathbf{q}^N, \mathbf{p}^N) \;=\; \frac{\text{tempo gasto em } d\mathbf{q}^N\, d\mathbf{p}^N \;\equiv\; \delta t(\Gamma)}{\text{tempo total} \equiv T_T}. \tag{2.1}$$

A média de uma variável termodinâmica $A(\Gamma(t))$, no intervalo $(0, T_T)$, é dada então por:

$$\begin{aligned} \overline{A} &= \sum_{\Gamma} A(\Gamma)\, p^{T_T}(\Gamma) \\ &= \sum_{\Gamma} A(\Gamma)\, \frac{\delta t(\Gamma)}{T_T} \\ &= \frac{\int_0^{T_T} A(t)\, dt}{T_T}. \end{aligned} \tag{2.2}$$

Quando tomamos o limite $T_T \to \infty$ em geral o sistema tende ao equilíbrio, ou para algum estado estacionário, caso as condições externas se mantenham constantes. Assim, a média de equilíbrio é definida por

$$\overline{A}_{eq} = \lim_{T_T \to \infty} \frac{\int_0^{T_T} A(t)\, dt}{T_T}. \tag{2.3}$$

Quando a função converge para um valor de equilíbrio, podemos observar que a média acima pode ser substituída por outra equivalente. Seja $A(t) = \overline{A}_{eq} + a(t)$, onde $a(t)$ é uma função de média temporal nula. A integral $\lim_{T_T \to \infty} \int_0^{T_T} a(t)\, dt \to a_0 < \infty$, pois caso contrário sua média não seria nula. Assim, podemos definir a média temporal de equilíbrio de Laplace via

$$\overline{A}_L = \lim_{z \to 0} \frac{\int_0^\infty e^{-zt} A(t)\, dt}{\int_0^\infty e^{-zt}\, dt} = \lim_{z \to 0} z \int_0^\infty e^{-zt} A(t)\, dt = \lim_{z \to 0} z\, \tilde{A}(z). \tag{2.4}$$

Podemos ver que são equivalentes as definições acima, quando acontece a convergência: $\overline{A}_L = \overline{A}_{eq}$. Podemos ver:

$$\begin{aligned} \overline{A}_L &= \lim_{z \to 0} \frac{\int_0^\infty e^{-zt} \left(\overline{A}_{eq} + a(t)\right) dt}{\int_0^\infty e^{-zt}\, dt} \\ &= \overline{A}_{eq} + \lim_{z \to 0} \frac{\int_0^\infty e^{-zt}\, a(t)\, dt}{\int_0^\infty e^{-zt}\, dt} \\ &= \overline{A}_{eq} + \lim_{z \to 0} z \int_0^\infty e^{-zt}\, a(t)\, dt. \end{aligned}$$

Como $a(t)$ tem média temporal nula, devemos ter

$$\int_0^T a(t)\, dt \sim T^\alpha \Rightarrow a(t) \sim t^{\alpha - 1}$$

onde $\alpha < 1$. Assim

$$\int_0^\infty e^{-zt}\, a(t)\, dt \sim \int_0^\infty e^{-zt}\, t^{\alpha - 1}\, dt \sim z^{-\alpha}.$$

Portanto

$$\lim_{z \to 0} z \int_0^\infty e^{-zt}\, a(t)\, dt \sim \lim_{z \to 0} z\, z^{-\alpha} = 0.$$

Isto justifica que se a função $f(t)$ possui uma média temporal convergente, essa média pode ser obtida por $\lim_{z \to 0} z\, \tilde{f}(z)$.

2.2 Ensemble Microcanônico

Mas o que leva o sistema a um estado de equilíbrio? Claramente é a dinâmica que o mesmo obedece. O teorema de Liouville nos dá a pista de que precisamos: os estados do sistema que são funções do Hamiltoniano do próprio sistema são invariantes no tempo.

Para um sistema isolado, a energia é fixa, $E = H(\Gamma)$, e dado que a distribuição estacionária deve ser função de $H(\Gamma)$, esta pode assumir a forma **microcanônica** coerente com a estacionaridade no tempo:

$$p^{eq}(\Gamma) = \frac{1}{\Omega(E)}\,\delta(E - H(\Gamma)),$$

onde $\Omega(E) \equiv \int d\Gamma\,\delta(E - H(\Gamma))$ é um fator de normalização que mede o volume accessível do espaço de fases.

Da equação acima vemos que todos os estados tem o mesmo peso estatístico. Portanto, para um sistema clássico isolado, todos os estados são equiprováveis no equilíbrio. Daí derivamos o **Teorema Fundamental da Mecaânica Estatística**:
Para um sistema clássico isolado, todo (micro)estado accessível no espaço de fases terá a mesma probabilidade se o sistema está em equilíbrio.

O contato com a termodinâmica se dá através da Entropia $S(\Omega)$. Esta função é uma medida interessante da desordem de um sistema. A entropia deve ser extensiva para sistemas com interfaces desprezíveis com relação a seus interiores (bulk). Dados dois sistemas independentes, a energia total, o números total de estados e a entropia total variam como:

$$\begin{aligned} E_T &= E_1 + E_2 \\ \Omega_T(E_T) &= \Omega_1(E_1) \times \Omega_3(E_2) \\ S_T(E_T, \Omega_T) &= S_1(E_1, \Omega_1) + S_2(E_2, \Omega_2). \end{aligned}$$

Portanto, a forma da entropia consistente com a extensividade e a independência é:

$$S_(E, \Omega) = k_B \ln \Omega(E),$$

onde k_B é a constante de Boltzmann. Esta última muitas vezes é tomada por $k_B = 1$, de modo que a temperatura passa a ter dimensões de energia.

A partir da entropia podemos construir a termodinâmica.

2.3 Ensemble Canônico

A forma microcanônica para a distribuição de probabilidades é muitas vezes complicada de ser usada do ponto de vista prático. para sistemas grandes, uma maneira mais accessível é a **distribuição canônica**, obtida para sistemas em contatos com reservatórios térmicos e que podem trocar calor com os mesmos.

Seja um sistema completo T constituído de um sistema pequeno S ($N_S \gg 1$) e um reservatório térmico R ($N_R \gg N_S$). A energia e número total de partículas obedecem

$$\begin{aligned} E_T &= E_S + E_R \\ N_T &= N_S + N_R, \end{aligned}$$

onde desprezamos a energia de interação entre R e S.

Dado que a entropia mede o volume de fases explorado pelo sistema quando as condições externas, ou seja, o **macroestado**, estão fixas, ao variarmos este macroestado, a entropia procurará seu máximo espontaneamente. Assim, devemos maximizar a entropia do sistema formado por R +S. Como $\Omega = e^{\frac{S}{k_B}} \rightarrow p^{eq} \propto e^{\frac{S}{k_B}}$, podemos escrever

$$\Omega_{R+S}(E_T) = \sum_{E_S} \Omega_R(E_T - E_S) \times \Omega_S(E_S).$$

Como Ω cresce exponencialmente como função da entropia (que é extensiva $\propto N$), praticamente toda a contribuição para o volume total de fases vem do termo $\mathbf{max}[\Omega_R(E_T - E_S) \times \Omega_S(E_S)]_{E_S}$. Isto se deve ao fato de que se trata de uma integração tipo ponto de sela [18].

Sabendo que $\frac{\partial S}{\partial E} = \frac{1}{T}$, obtemos que

$$\begin{aligned} 0 &= \frac{\partial S_R(E_T - E_S)}{\partial E_S} + \frac{\partial S_S(E_S)}{\partial E_S} \\ 0 &= -\frac{1}{T_R} + \frac{1}{T_S} \\ \Rightarrow T_S &= T_R. \end{aligned}$$

A probabilidade canônica de que um estado Γ_S seja escolhido é

$$\begin{aligned} p_S(\Gamma_S) &= \frac{\Omega_R(E_T - E_S(\Gamma_S))}{\Omega_{R+S}(E_T)} \\ \Rightarrow p_S(E_S) &= \frac{\Omega_R(E_T - E_S) \times \Omega_S(E_S)}{\Omega_{R+S}(E_T)}. \end{aligned}$$

Portanto

$$\begin{aligned} p_S(\Gamma_S) &= \frac{\Omega_R(E_T - E_S(\Gamma_S))}{\Omega_{R+S}(E_T)} \\ &= \frac{e^{S_R(E_T - E_S(\Gamma_S))/k_B}}{\Omega_{R+S}(E_T)} \\ &= \frac{e^{S_R(E_T)/k_B - \frac{\partial S_R}{\partial E_S} E_S(\Gamma_S))/k_B}}{\Omega_{R+S}(E_T)} \\ &\propto e^{-\frac{E_S(\Gamma_S)}{k_B T_R}}. \end{aligned}$$

Assim, a distribuição canônica é dada por

$$\Rightarrow p_S(\Gamma_S) = \frac{1}{Z} e^{-\frac{H_S(\Gamma_S)}{k_B T_R}},$$

onde $Z = \sum_{\Gamma_S} e^{-\frac{H_S(\Gamma_S)}{k_B T_R}}$ é a **função de partição.**

A interpretação da função de partição como fator de proporcionalidade é simples de entender ($E_S = H_S(\Gamma_S)$):

$$\begin{aligned} Z &= \sum_{\Gamma_S} e^{-\frac{H_S(\Gamma_S)}{k_B T_R}} \\ &= \sum_{E_S} \Omega(E_S)\, e^{-\frac{E_S}{k_B T_R}} \\ &= \sum_{E_S} e^{\frac{T_R S_S - E_S}{k_B T_R}}, \end{aligned}$$

cujo valor aproximado é dado por

$$-k_B T_R \ln Z = -\mathbf{max}[T_R S_S - E_S]_{E_S^*} = \mathbf{min}[E_S - T_R S_S]_{E_S^*} = F(T_R),$$

a energia-livre de Helmholtz do sistema.

2.4 Ensemble grand-canônico (de partículas)

Um sistema aberto pode trocar desde energia até partículas com sua vizinhança. Quando isto acontece, podemos supor que o sistema está em equilíbrio com um reservatório (de partículas, de energia, etc) pois em caso

contrário o sistema S não poderia estar em equilíbrio. Assim, podemos determinar as condições para o equilíbrio, que são a maximização do volume de fases accessível:

$$\begin{aligned}
\Omega_{R+S}(E_T, N_T) &= \sum_{E_S, N_S} \Omega_R(E_T - E_S, N_T - N_S) \times \Omega_S(E_S, N_S) \\
S_{R+S}(E_T, N_T) &= S_R(E_T - E_S, N_T - N_S) + S_S(E_S, N_S) \\
E_T &= E_S + E_R \\
N_T &= N_S + N_R,
\end{aligned}$$

Maximizando a entropia total com respeito a N_S e E_S obtemos ($\partial S/\partial N = -\mu/T$)

$$\begin{aligned}
T_S &= T_R \\
\mu_S &= \mu_R.
\end{aligned}$$

Similarmente ao caso canônico, a probabilidade de equilíbrio agora é dada por:

$$\begin{aligned}
p_S(\Gamma_S) &= \frac{\Omega_R(E_T - E_S(\Gamma_S), N_T - N_S(\Gamma_S))}{\Omega_{R+S}(E_T, N_T)} \\
&\propto e^{S_R(E_T - E_S(\Gamma_S), N_T - N_S(\Gamma_S))/k_B} \\
&\propto e^{-\frac{\partial S_R}{\partial E_S} E_S(\Gamma_S))/k_B - \frac{\partial S_R}{\partial N_S} N_S(\Gamma_S))/k_B} \\
&\propto e^{\frac{\mu_R N_S(\Gamma_S) - E_S(\Gamma_S)}{k_B T_R}}.
\end{aligned}$$

A grand-função de partição se obtêm (onde "dropamos" os índices do reservatório):

$$\begin{aligned}
\Xi &= \sum_{N_S=0}^{\infty} \sum_{\Gamma_S(N_S)} e^{\frac{\mu N_S(\Gamma_S) - E_S(\Gamma_S)}{k_B T}} \\
&= \sum_{N_S=0}^{\infty} e^{\frac{\mu N_S}{k_B T}} \sum_{\Gamma_S(N_S)} e^{-\frac{E_S(\Gamma_S)}{k_B T}} \\
&= \sum_{N_S=0}^{\infty} e^{\frac{\mu N_S}{k_B T}} e^{-\frac{F_S(\Gamma_S)}{k_B T}},
\end{aligned}$$

de onde obtemos que (ao escolhermos o valor máximo da integração de ponto de sela)

$$-k_B T_R \ln \Xi = \mathbf{max} \left[\frac{\mu N_S - F_S(\Gamma_S)}{k_B T} \right]_{N_S^*} = \mathbf{min} [E_S - T S_S - \mu N_S]_{E_S^*} = \Phi(T_R),$$

o grão-potencial do sistema.

2.5 Termodinâmica

No LT a distribuição de probabilidades de equilíbrio para um sistema se torna exata, ou seja, como funções do tipo delta de Dirac, exceto para situações tais como a que acontece na vizinhança de um ponto crítico, onde as flutuações são dominantes e se tornam infinitas em extensão (comprimento de correlação $\xi \to \infty$).

Leis da Termodinâmica, um pequeno resumo:

- Lei 0: define a existência de um termostato.
- 1^a Lei: expressa a conservação da energia.
- 2^a Lei: leva à "seta do tempo".
- 3^a Lei: uma consequência da existência de um estado fundamental de degenerecência finita.

2.5.1 Lei 0 da Termodinâmica

Se um sistema **C** estiver simultaneamente em equilíbrio térmico com dois sistemas **A** e **B**, então **A** e **B** estão em equilíbrio térmico entre si.

Deduzimos que

- $T(\mathbf{A}) = T(\mathbf{C}) \,\&\, T(\mathbf{B}) = T(\mathbf{C}) \;\Rightarrow\; T(\mathbf{A}) = T(\mathbf{B})$.
- Alguma variável do corpo **C** (termostato), tal qual o seu volume V, pode ser usada para criarmos um termômetro.

2.5.2 1^a Lei da Termodinâmica

A 1^a Lei reflete a conservação de energia térmica total E quando um sistema está em contato com fontes de calor Q (trabalho microscópico transmitido de forma desorganizada) e fontes de trabalho W (trabalho microscópico transmitido de forma ordenada). Trabalho e calor dependem do processo termodinâmico enquanto que a energia é função de estado.

Forma infinitesimal (gases em processos reversíveis)

- $dE = \cancel{d}Q + \cancel{d}W = T\,dS - p\,dV + \mu\,dN$

- $\not{d}Q$ é o calor recebido pelo sistema
- $\not{d}W$ é o trabalho recebido pelo sistema

2.5.3 Processos Reversíveis

São processos termodinâmicos realizados de modo extremamente lento e de maneira que o sistema esteja em cada instante em equilíbrio. São chamados de quase-estáticos.

Um motor realizando um processo reversível proverá potência nula! Não são muito eficientes!

Para estes processos podemos escrever $dE = T\,dS + J\,dX$, onde S é a entropia e (J, X) são a força e o deslocamento generalizados, respectivamente.

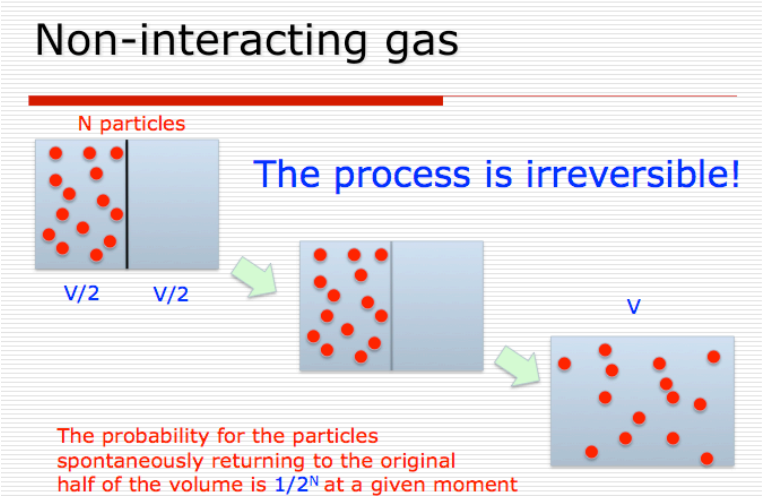


Figure 2.1: Processo irreversível leva a aumento espontâneo de entropia.

2.5.4 2^a Lei da Termodinâmica

A 2^a Lei reflete o fato que, durante um processo termodinâmico espontâneo, vale para os volumes de fase $\Omega_{final} \gg \Omega_{inicial}$, dada uma mudança nos vínculos externos sobre o sistema.

A entropia é então definida como $S = k_B \ln \Omega$

Assim temos a forma para a 2^a Lei: $S_{final} - S_{inicial} = \Delta S \geq 0$

2.5.5 3^a Lei da Termodinâmica

No limite $T \to 0$, o sistema tende a ocupar o seu estado fundamental (quântico) ψ_0.

Supondo que a degenerecência de ψ_0 seja $g_0 = 1$. Assim $S = k_B \ln \Omega = k_B \ln 1 = 0$.

Isto leva o calor específico também a se tornar zero nesse limite dado que $\lim_{T\to 0} \Delta S = \lim_{T\to 0} \int_0^T dT \, \frac{C(T)}{T} = 0$.

Uma consequência da 3^a Lei é que não é possível atingir $T = 0$, em um número finito de processos, pois a condutividade térmica (proporcional ao calor específico) se torna nula também.

2.6 Informação e Temodinâmica

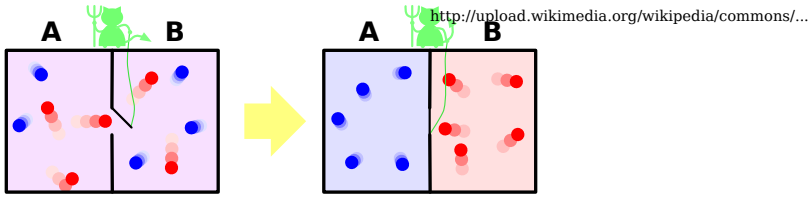
2.6.1 Demônio de Maxwell

O "Demônio de Maxwell" (DM) é um medidor-operador que age sobre um sistema e permite a violação da 2^a Lei da termodinâmica.

O DM deixa passar as moléculas quentes (rápidas) para a direita e as frias (lentas) para a esquerda, mas nunca ao contrário. Isto é realizado sem a necessidade de realizar nenhum trabalho. Ele parece violar a 2^{nd} Lei (Clausius).

Após transformar as temperaturas em cada metade, podemos fazer uma máquina operar entre essas duas metades extraindo trabalho.

A ideia é tentar realizar o DM por algum meio. R. Feynmann sugeriu a criação de um ratchet, como na figura 2.2. Como poderíamos "exorcizá-lo"? O próximo passo é tentar usar o DM para algo útil, tal como extração de trabalho.



2.6.2 Máquina de Szilard

O DM é na verdade um aparato baseado em medida e feedback, a partir do resultado da medida. A segunda lei da termodinâmica é muito limitante, e pareceria ser possível evitá-la, o que em princípio parece simples. Apenas seria necessário ter informação sobre o sistema.

L. Szilard [10] propôs uma máquina capaz de extrair trabalho apenas usando uma única fonte de calor. Em um processo isotérmico em que um gás expande seu volume de $V/2$ até V, o trabalho total é

$$W = \int_{V/2}^{V} p\, dV = N\, k_B\, T \int_{V/2}^{V} \frac{dV}{V} = N\, k_B\, T\, \ln 2.$$

Quando apenas uma partícula realiza o trabalho temos $W = k_B\, T\, \ln 2$. A entropia do gás aumenta, ao se expandir isotermicamente, em $N\, k_B\, \ln 2$.

A Máquina de Szilard (MS) funciona da seguinte maneira: um gás de 1 partícula ocupa uma caixa de volume V, e temperatura T. Sem realização de trabalho, uma partição é rapidamente colocada no meio da caixa, deixando a

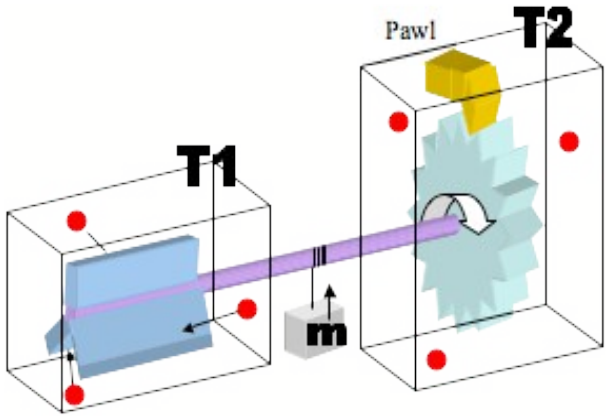


Figure 2.2: Demônio Browniano de Feynmann: Gedankenexperiment com $T_1 = T_2 = T$ (Wikipedia)

partícula aleatoriamente do lado esquerdo ou direito da partição. Realizando uma medida sem (quase) nenhum custo energético colocamos o peso do lado apropriado e deixamos a partícula empurrar a partição até que ela ocupe o volume total da caixa. Assim extraímos um trabalho $W_G = k_B\,T\,\ln 2$ apenas de um único reservatório térmico. O exemplo acima é ilustrativo de um caso de medida e feed-back. Mas a MS viola mesmo a 2^a Lei?

A resposta é não. Para podermos usar o detetor, o lado medido necessita ser registrado em apenas 1 bit (1 é E e 0 é D). Para o sistema voltar ao estado inicial e completar o ciclo, devemos apagar esse bit de memória, o que custa dissipar um trabalho dissipado $W_D \leq -k_B\,T\,\ln 2$, que é o Princípio de Landauer [12]. Assim, o trabalho total gerado produzido pela máquina corresponde a $W = W_G + W_D \leq k_B\,T\,\ln 2 - k_B\,T\,\ln 2 = 0$. E a 2^a Lei não é violada.

2.6.3 Informação e trabalho

Podemos estudar o problema da informação x trabalho x calor a partir do trabalho de Sagawa e Ueda **??**. Eles mostram que vale

$$\left\langle e^{-\beta\,W - \Delta I} \right\rangle = e^{-\beta\,\Delta F},$$

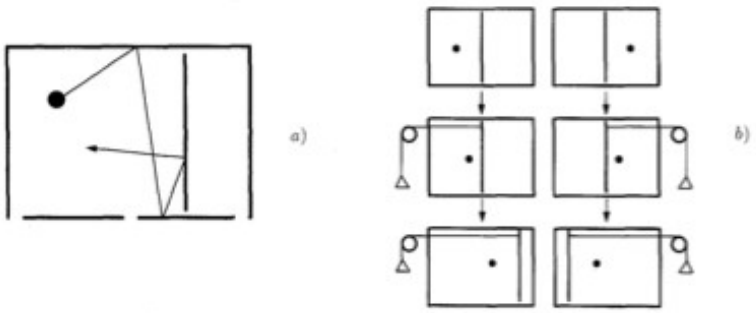


Fig. 1. – The operation of the Szilard heat engine. *a*) The one-molecule gas and the partition Collisions are not ballistic: the particle exchanges energy with the cylinder and the partition. *b*) If we know on which side the molecule is, we know how to attach a weight.

© Les Editions de Physique

Figure 2.3: Máquina de Szilard: ao se medir, se a partícula está na direita ou na esquerda, podemos determinar em que lado instalar o puxador, e assim extrair trabalho apenas de uma fonte térmica de maneira cíclica. Ver M. Magnasco [11]

onde ΔI é a variação da correlação mútua

$$I_{xy} = \ln \frac{p(x,y)}{p(x)\,p(y)},$$

onde a média é a informação mútua

$$\langle I_{xy} \rangle = \sum_{xy} p(x,y)\, \ln \frac{p(x,y)}{p(x)\,p(y)} \geq 0.$$

As variáveis x correspondem ao sistema e as variáveis y ao DM.

A desigualdade vem da entropia relativa $D(p||q)$

$$D(p||q) = \sum_x p \ln \frac{p}{q} = -\sum_x p \ln \frac{q}{p} \geq -p(\frac{q}{p} - 1) = 0,$$

onde usamos

$$\ln w \leq w - 1 \Leftrightarrow w \leq e^{w-1}.$$

Desta última obtemos ($w = e^u$)

$$u \leq e^u - 1 \Leftrightarrow \langle u \rangle < 1 + \langle u \rangle \leq \langle e^u \rangle \, .$$

Desigualdade de Jensen

Para uma função convexa $f(x)$, temos que

$$f(x) \leq \frac{b-x}{b-a} f(a) + \frac{x-a}{b-a} f(b) \Rightarrow f\left(x = t\, a + (1-t)\, b\right) \leq t\, f(a) + (1-t)\, f(b),$$

que pode ser generalizado para $(t_1 + t_2 + \ldots + t_N = 1)$ para a Desigualdade de Jansen:

$$f\left(x = t_1\, a_1 + t_2\, a_2 + \ldots + t_N\, a_N\right) \leq t_1\, f(a_1) + t_2\, f(a_2) + \ldots + t_N\, f(a_N), \;\; a_1 \leq x \leq a_N .$$

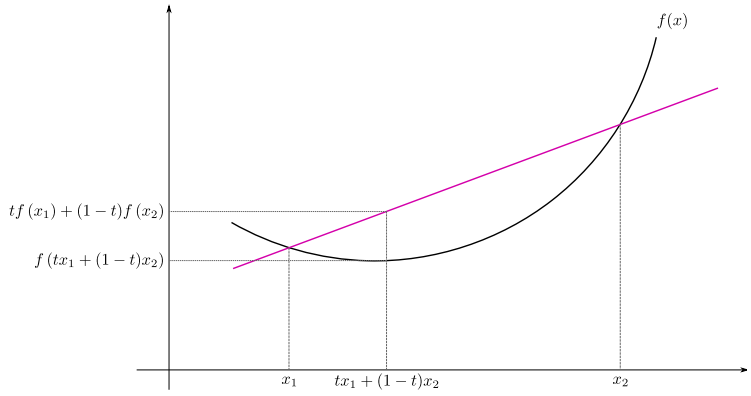


Figure 2.4: Usando $t = (b-x)/(b-a)$ podemos mostrar a propriedade de convexidade do texto. By Eli Osherovich - Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=10764763

O resultado acima pode ser provado por indução dado que vale para $n = 2$, e se vale para algum $n > 2$, podemos escolher um t_1 que seja positivo

não-nulo e obter

$$f\left(\sum_{i=1}^{n+1} t_i\, a_i\right) = f\left(t_1\, a_1 + (1-t_1)\sum_{i=2}^{n+1} \frac{t_i}{1-t_1}\, a_i\right) \leq t_1\, f\,(a_1) + (1-t_1) f\left(\sum_{i=2}^{n+1} \frac{t_i}{1-t_1}\, a_i\right), \text{ etc,}$$

pois vale a desigualdade para o segundo termo, e também

$$\sum_{i=2}^{n+1} \frac{t_i}{1-t_1} = 1.$$

A forma probabilística $\langle f(x)\rangle \leq f(\langle x\rangle)$ pode ser obtida no limite $n \to \infty$ de modo que $(t_i \to dx\, p(x), a_i \to x)$ e assim

$$\begin{aligned} \sum_{i=1}^{n+1} t_i\, a_i &\to \int dx\, p(x)\, x = \langle x\rangle \\ \sum_{i=1}^{n+1} t_i\, f(a_i) &\to \int dx\, p(x)\, f(x) = \langle f(x)\rangle\,, \end{aligned}$$

e portanto, quando $f(x)$ é uma função convexa em x, vale

$$\langle f(x)\rangle \leq f(\langle x\rangle). \tag{2.5}$$

Violando a forma simples da 2^a Lei

Desta forma, a partir da quação de jarzynski generalizada

$$\left\langle e^{-\beta\, W - \Delta I + \beta\, \Delta F}\right\rangle = 1$$

podemos escrever

$$1 = \left\langle e^{-\beta\, W - \Delta I + \beta\, \Delta F}\right\rangle \geq e^{-\beta\, \langle W\rangle - \langle \Delta I\rangle + \beta\, \Delta F} \Rightarrow 0 \geq -\beta\, \langle W\rangle - \langle \Delta I\rangle + \beta\, \Delta F.$$

Assim o trabalho dissipado pode ser negativo!! Pois temos

$$W_D = \langle W\rangle - \Delta F \geq -k_B T\, \langle \Delta I\rangle\,.$$

Podemos mover calor da fonte fria para a fonte quente sem realizar trabalho! Basta utilizar a correlação entre o sistema e o DM (reservatório de informação).

Chapter 3

Evolução temporal de um sistema

Nesta seção estudaremos a evolução de um sistema que parte de uma condição inicial arbitrária e tende ao equilíbrio quando $t \to \infty$. Esta é uma maneira eficiente de estudarmos os estados de não-equilíbrio pelos quais um sistema passa.

3.1 Evolução de um sistema Hamiltoniano clássico

Começaremos utilizando a evolução exata clássica de um sistema, baseada nas equações de Hamilton, e a partir dela determinar a evolução de distribuição de probabilidades do sistema. Esta será dada pelo Teorema de Liouville que desenvolveremos na sequência.

A mecânica clássica determina a evolução no tempo das variáveis dinâmicas de um sistema. Essa evolução é governada pelas equações de Hamilton. Para um sistema isolado de N partículas temos

$$\dot{\mathbf{p}}_i = -\frac{\partial H}{\partial \mathbf{q}_i}, \tag{3.1}$$

$$\dot{\mathbf{q}}_i = \frac{\partial H}{\partial \mathbf{p}_i} \tag{3.2}$$

onde o Hamiltoniano $H(\mathbf{q}^N, \mathbf{p}^N)$, para um sistema isolado, é dado por

$$H = \sum_{i=1}^{N} \frac{\mathbf{p}_i^2}{2m} + V(\mathbf{q}^N). \tag{3.3}$$

Estas equações são equivalentes às Leis de Newton.

Como vimos, a definição $\mathbf{q}^N \equiv (\mathbf{q}_1, \ldots, \mathbf{q}_N)$ e $\mathbf{p}^N \equiv (\mathbf{p}_1, \ldots, \mathbf{p}_N)$ as posições e momentos do sistema. Seja uma função $A(\mathbf{q}^N, \mathbf{p}^N, t)$

$$\frac{dA}{dt} = \sum_{i=1}^{N} \frac{d\mathbf{q}_i}{dt} \frac{\partial A}{\partial \mathbf{q}_i} + \sum_{i=1}^{N} \frac{d\mathbf{p}_i}{dt} \frac{\partial A}{\partial \mathbf{p}_i} + \frac{\partial A}{\partial t},$$

$$\frac{dA}{dt} = \sum_{i=1}^{N} \frac{\partial H}{\partial \mathbf{p}_i} \frac{\partial A}{\partial \mathbf{q}_i} - \sum_{i=1}^{N} \frac{\partial H}{\partial \mathbf{q}_i} \frac{\partial A}{\partial \mathbf{p}_i} + \frac{\partial A}{\partial t} = \{A, H\} + \frac{\partial A}{\partial t},$$

onde definimos o Parêntesis de Poisson:

$$\{A, B\} \equiv \sum_{i=1}^{N} \frac{\partial A}{\partial \mathbf{q}_i} \frac{\partial B}{\partial \mathbf{p}_i} - \sum_{i=1}^{N} \frac{\partial B}{\partial \mathbf{q}_i} \frac{\partial A}{\partial \mathbf{p}_i}. \tag{3.4}$$

Definindo o operador de Liouville

$$i\,\mathcal{L}\, f \equiv \{H, f\} \rightarrow i\,\mathcal{L} \equiv \{H, \}\,, \tag{3.5}$$

$$\Rightarrow \frac{dA}{dt} = -i\,\mathcal{L}\, A + \frac{\partial A}{\partial t}. \tag{3.6}$$

Mostraremos que para a distribuição de probabilidades a N corpos, $\rho(\mathbf{q}, \mathbf{p}, t)$, o Teorema de Liouville nos dá:

$$\frac{d}{dt}\rho(\mathbf{q}^N, \mathbf{p}^N, t) = 0. \tag{3.7}$$

3.2 Exemplos calculáveis usando $\mathcal{L}$

Vejamos dois casos onde uma solução exata, e simples, é possível utilizando-se o formalismo de Liouville.

3.2.1 Oscilador harmônicos

Apesar das dificuldades matemáticas que aparecem para sistemas grandes, podemos resolver alguns sistemas simples utilizando o formalismo de Liouville. Seja por exemplo o sistema 1D dado pelo oscilador harmônico descrito pelo Hamiltoniano independente do tempo

$$H = \frac{p^2}{2m} + \frac{1}{2}kq^2,$$

para o qual temos

$$i\,\mathcal{L} = \frac{\partial\, H}{\partial\, q}\frac{\partial}{\partial\, p} - \frac{\partial\, H}{\partial\, p}\frac{\partial}{\partial\, q} = kq\frac{\partial}{\partial\, p} - \frac{p}{m}\frac{\partial}{\partial\, q}.$$

Podemos estudar a evolução no tempo de $q(t)$ dado que este obdedece a (já que $\frac{\partial\, q}{\partial\, t} = 0$)

$$q(t) = e^{-i\,\mathcal{L}\,t}\, q(0),$$

onde o sinal de $\mathcal{L}$ é trocado em relação àquele do caso da distribuição de probabilidade ρ.

Vejamos algumas propriedades de $\mathcal{L}$:

$$\begin{aligned}
\mathcal{L}q &= i\,\frac{p}{m},\\
\mathcal{L}p &= -i\,kq,\\
\mathcal{L}^2 q &= \omega_0^2\, q,\\
\mathcal{L}^2 p &= \omega_0^2\, p,
\end{aligned}$$

onde definimos $\omega_0^2 = k/m$. os autovalores de $\mathcal{L}$ são $\pm\,\omega_0$, reais.

Assim temos

$$\begin{aligned}
q(t) &= e^{-i\,\mathcal{L}\,t}\, q(0)\\
&= \sum_{n=0}^{\infty} \frac{(-t)^n}{n!}\,(i\,\mathcal{L})^n\, q(0)\\
&= \sum_{n=0}^{\infty}(-1)^n\,\frac{(-t)^{2n}}{(2n)!}\,\mathcal{L}^{2n}\, q(0) + \sum_{n=0}^{\infty}(-1)^n\, i\,\frac{(-t)^{2n+1}}{(2n+1)!}\,\mathcal{L}^{2n}\,\mathcal{L}\, q(0)\\
&= \sum_{n=0}^{\infty}(-1)^n\,\frac{(-t)^{2n}}{(2n)!}\,\mathcal{L}^{2n}\, q(0) - \frac{1}{m}\sum_{n=0}^{\infty}(-1)^n\,\frac{(-t)^{2n+1}}{(2n+1)!}\,\mathcal{L}^{2n}\, p(0).
\end{aligned}$$

Como q e p são autofunções de $\mathcal{L}^2$, com autovalor ω_0^2, podemos escrever

$$\begin{aligned}
q(t) &= \sum_{n=0}^{\infty}\frac{(-t)^{2n}}{(2n)!}\,(-\omega_0^2)^n\, q(0) - \frac{1}{m}\sum_{n=0}^{\infty}\frac{(-t)^{2n+1}}{(2n+1)!}\,(-\omega_0^2)^n\, p(0)\\
&= \sum_{n=0}^{\infty}\frac{(\omega_0\, t)^{2n}}{(2n)!}\, q(0) + \frac{1}{m\omega_0}\sum_{n=0}^{\infty}\frac{(\omega_0\, t)^{2n+1}}{(2n+1)!}\, p(0)\\
&= \cos(\omega_0\, t)\; q(0) + \sin(\omega_0\, t)\;\frac{p(0)}{m\omega_0},
\end{aligned}$$

onde a solução correta aparece diretamente. A extensão para um número arbitrário de osciladores independentes é imediata. Para uma solução mais geral da Equação de Liouville, ver [19].

3.3 Teorema de Liouville

A evolução da distribuição de probabilidades do sistema conserva a probabilidade total, pois não há fontes nem "sumidouros" de probabilidades no sistema. Assim, a variação local de probabilidade em um volume arbitrário corresponde ao fluxo de probabilidades na borda desse volume. Diferencialmente isto corresponde a:

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot J_\rho \tag{3.8}$$

Temos $J_\rho = \rho V_\rho$, onde definimos vetor velocidade generalizada $V_\rho = (\dot{\mathbf{q}}^N, \dot{\mathbf{p}}^N)$. Portanto

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot (\rho V_\rho) = -\rho \nabla \cdot V_\rho - V_\rho \cdot \nabla \rho$$

Temos:

$$\nabla \rho = \left(\frac{\partial\, \rho}{\partial\, \mathbf{q}^N}, \frac{\partial\, \rho}{\partial\, \mathbf{p}^N} \right),$$

e

$$\nabla \cdot V_\rho = \sum_{i=1}^{N} \left(\frac{\partial\, \dot{q}_i}{\partial\, \mathbf{q}_i} + \frac{\partial\, \dot{p}_i}{\partial\, \mathbf{p}_i} \right) = \sum_{i=1}^{N} \left(\frac{\partial^2 H}{\partial\, \mathbf{p}_i\, \partial\, \mathbf{q}_i} - \frac{\partial^2 H}{\partial\, \mathbf{q}_i\, \partial\, \mathbf{p}_i} \right) = 0,$$

que é a característica de um fluido incompressível. Temos então

$$\frac{\partial \rho}{\partial t} = -\sum_{i=1}^{N} \left(\dot{q}_i \frac{\partial\, \rho}{\partial\, \mathbf{q}_i} + \dot{p}_i \frac{\partial\, \rho}{\partial\, \mathbf{p}_i} \right) = -\{\rho, H\} \Rightarrow \frac{d\rho}{dt} = 0.$$

O resultado acima verifica o **Teorema de Liouville**.

3.4 O fluido incompressível de probabilidades

A derivada total no tempo pode ser interpretada como a variação ao longo da trajetória do fluxo, quando tratamos de um fluido. Isto pode ser visto quando escolhemos uma função $\phi(x, y, z, t) \equiv \phi(x(t), y(t), z(t), t)$ e calculamos a variação de ϕ entre os instantes de fluxo t e $t + dt$:

$$d\phi = \phi(x(t + dt), y(t + dt), z(t + dt), t + dt) - \phi(x(t), y(t), z(t), t) =$$

$$dt \left[\frac{\partial \phi}{\partial t} + \frac{\partial \phi}{\partial x}\dot{x} + \frac{\partial \phi}{\partial y}\dot{y} + \frac{\partial \phi}{\partial z}\dot{z} \right] = dt \left[\frac{d\phi}{dt} \right].$$

Vemos então que o significado do teorema de Liouville é que o fluido de probabilidades é incompressível. Uma região ocupada do espaço de fases evolui ao longo do tempo dado que para um sistema Hamiltoniano as trajetórias formam uma variedade onde as linhas de fluxos não tem intersecções, já que intersecções correspondem a duas velocidades distintas para o mesmo corpo no mesmo instante. Um fluido formado por pontos de fase ($\mathcal{N} \gg N$ micro-estados de um sistema, formando um "ensemble" dos mesmos) evolui com sua densidade constante.

Um certo elemento de volume de fases inicial se deforma (por exemplo, se bifurca a cada evento do tipo colisional entre as N moléculas do sistema), mas mantêm seu volume constante. De fato, para sistemas bem comportados, uma distribuição inicial de pontos de fase concentrada se "espaguetificação" ao longo do tempo, de modo que após um certo tempo o "espaguete" extremamente fino ocupa a hiper-superfície de energia constante de modo uniforme, levando à equiprobabilidade que define o ensemble microcanônico.

3.5 Resolvendo a Equação de Liouville

A partir da equação de movimento mais básica para a distribuição de probabilidades, e Equação de Liouville, equivalente às equações de Hamilton:

$$\frac{\partial}{\partial\, t}\rho(t) = i\,\mathcal{L}\,\rho(t) \equiv \{H, \rho(t)\}\,, \tag{3.9}$$

podemos obter a solução formal

$$\rho(t) = e^{i\,\mathcal{L}\,t}\,\rho(0) \equiv e^{i\,\mathcal{L}\,t}\,\rho(\mathbf{q}^N, \mathbf{p}^N)\Big|_{(\mathbf{q}^N,\mathbf{p}^N)=(\mathbf{q}_0^N,\mathbf{p}_0^N)}\,. \tag{3.10}$$

Como calculamos a média temporal de $\rho(t)$? Como vimos em 2.4, esta pode ser definida por

$$\begin{aligned} < \rho >_{ta} &= \lim_{T\to\infty} \frac{1}{T}\int_0^T dt\,\rho(t) \\ &= \lim_{z\to 0^+} z\int_0^\infty dt\, e^{-zt}\,\rho(t) \\ &= \lim_{z\to 0^+} z\int_0^\infty dt\, e^{-zt}\, e^{i\,\mathcal{L}\,t}\,\rho(0) \end{aligned}$$

Podemos usar uma expansão de $\rho(0)$ em autofunções do operador de Liouville ϕ_n [19]:

$$i\,\mathcal{L}\phi_n = \lambda_n\phi_n,$$

onde temos, por razões de convergência, $\Re(\lambda_n) \leq 0$. O caso $\lambda_n = 0$ corresponde às quantidades conservadas do problema. Em geral, os autovalores serão complexos, como vimos na seção 3.2. Assim

$$\rho(0) = \sum_n c_n^{(0)} \phi_n.$$

Podemos escrever:

$$e^{i\,\mathcal{L}\,t}\rho(0) = \sum_n c_n^{(0)} e^{i\,\mathcal{L}\,t}\phi_n = \sum_n c_n^{(0)} e^{\lambda_n\,t}\phi_n,$$

e

$$\begin{aligned} < \rho >_{ta} &= \lim_{z\to 0^+} z \int_0^\infty dt\, e^{-zt}\, e^{L\,t}\,\rho(0) \\ &= \lim_{z\to 0^+} z \int_0^\infty dt\, e^{-zt} \sum_n c_n^{(0)} e^{\lambda_n\,t}\phi_n \\ &= \lim_{z\to 0^+} \sum_n \frac{z}{z-\lambda_n}\, c_n^{(0)}\phi_n. \end{aligned}$$

Apenas as quantidades conservadas contribuirão, pois no limite $z \to 0^+$ o numerador será nulo, com um denominador não-nulo, sempre quando $\lambda_n \neq 0$. Assim

$$\begin{aligned} < \rho >_{ta} &= \lim_{z\to 0^+} \sum_n^{conserved} \frac{z}{z}\, c_n^{(0)}\phi_n \\ &= \sum_n^{conserved} c_n^{(0)}\phi_n. \end{aligned}$$

Da maneira pouco formal anunciada acima podemos ver que o equilíbrio acontece quando o sistema evolui ($z \to 0 \Leftrightarrow T \to \infty$), e assim apenas as quantidades conservadas contribuem. Encontrar as autofunções ϕ_n não é exatamente uma tarefa simples mas em teoria pode ser feito. Existem modos mais eficazes de tratar este problema, como veremos na sequência.

3.6 Médias temporais gerais

O formalismo mencionado acima pode ser aplicado de modo mais geral e abrangente. A idéia é que dada uma função $f(t) \equiv f(\mathbf{q}^N, \mathbf{p}^N, t)$ das variáveis dinâmicas, poderemos obter sua propriedades médias utilizando a distribuição de probabilidades de equilíbrio.

Assumimos que $f(t)$ é suficientemente bem comportada de modo que sua média temporal existe. Como vimos:

$$\begin{aligned} < f >_{ta} &= \lim_{T\to\infty} \frac{1}{T}\int_0^T dt\, f(t) \\ &= \lim_{z\to 0^+} z\int_0^\infty dt\, e^{-zt} f(t) \\ &= \lim_{z\to 0^+} z\,\tilde{f}(z). \end{aligned}$$

No Apêndice D mostramos como calcular a média temporal para a distribuição de probabilidades para um sistema simples Browniano, que pode ser expressa como D.10

$$\begin{aligned} p^{eq}(x,v) &= \lim_{z\to 0}\lim_{\epsilon\to 0^+} \sum_{n=0}^{\infty}\sum_{n=0}^{\infty}\int_{-\infty}^{\infty}\frac{dQ}{2\pi}\frac{dP}{2\pi}\, e^{iQx+iPv}\frac{(-iQ)^n\,(-iP)^r}{n!\,r!}\int_{-\infty}^{\infty}\prod_{a=1}^{n+r}\frac{dq_a}{2\pi}\,\frac{z}{z-\sum_{a=1}^{n+r}(iq_a+\epsilon)}\times \\ &\times \frac{\prod_{a=n+1}^{n+r}(iq_a+\epsilon)}{m^{n+r}\prod_{a=1}^{n+r}R(iq_a+\epsilon)}\left\langle\prod_{a=1}^{n+r}\tilde{\eta}(iq_a+\epsilon)\right\rangle. \end{aligned}$$

A partir desta expressão todas as variáveis termodinâmicas podem ser obtidas.

Para muitos problemas mais interessantes, especialmente quando o número de partículas é grande, soluções simples para o cálculo de médias temporais não são realmente praticáveis. Nas próximas seções veremos como tratar sistemas realistas, grandes e pequenos, de modo a estudar a transição entre não-equilíbrio e equilíbrio.

3.7 Ergodicidade: médias temporais e médias de ensemble

Um sistema é dito ergódico quando, entre outras propriedades, inicialmente a partir de qualquer ponto de fases do sistema Γ_0, o ponto de fases Γ_t passa tão perto quanto se queira de qualquer outro ponto que se determine Γ_D, ou seja, não existem pontos inacessíveis (mesma equisuperfícies de energia) isolados.

Este é o caso de um sistema com uma barreira finita de potencial. Quando a energia total é mais alta que a barreira o sistema pode transicionar entre as configurações de um lado e do outro da barreira (caso ergódico). Se a energia total for mais baixa que a altura da barreira o sistema fica preso

uma das configurações, que depende das condições iniciais) R (direita) ou L (esquerda). Este último caso é não-ergódico.

Neste último caso a média temporal da distribuição de probabilidades do sistema é perfeitamente definida e descreverá o comportamento correto do sistema enquanto que o ensemble canônico assinala probabilidades simultaneamente para estados tipo R e tipo L. Assim, médias temporais e de ensemble só coincidem quando o sistema for ergódico.

3.8 Movimento Browniano e Difusão

Uma referência moderna e interessante é [20], onde as principais definições e aspectos históricos são discutidos.

O Movimento Browniano (MB) é uma consequência da finitude ("lumpiness") da matéria na escala atômica (nm). De fato, nessa escala as flutuações térmicas são fator preponderante na dinâmica das partículas. Partículas da ordem microscópica (μm) são grandes o suficiente para serem observadas e pequenas o suficiente para ainda sofrerem a influência das flutuações térmicas.

A chave para entendermos o MB é a existência de uma separação clara de escalas de tempo entre a partícula Browniana (M) e as partículas (moléculas) do ambiente (m): a razão entre suas massas corresponde ao parâmetro pequeno $\epsilon = \sqrt{\frac{m}{M}}$ que nos permite separar as escalas de tempo da dinâmica. No próximo capítulo trabalharemos os detalhes da abordagem de separação de variáveis, mas a razão é que o operador de liouville do sistema se separa em duas partes, térmica (T) e Browniana (B)

$$\mathcal{L} = \mathcal{L}_T + \epsilon\, \mathcal{L}_B,$$

o que nos permite obter equações efetivas, como na seção 4.1, para o MB. A principal delas é a equação de langevin, que veremos a seguir.

Como para partículas Brownianas típicas $\epsilon \sim 10^{-5} - 10^{-9}$, a separação de escalas acima se torna muito boa. O movimento que resulta é tipicamente difusivo, ou seja uma caminhada aleatória que não tem característica balística ($s \sim vt$) ou acelerada ($s \sim at^2$). A trajetória do MB é difusiva ($s \sim D\sqrt{t}$).

Um exemplo simples da caminhada difusiva é dado pelo "random walk". Sejam N variáveis aleatórias independentes $X_{i=1,\ldots,N}$ (passos) com distribuição

$$p(X = \pm 1) \;\; = \;\; \frac{1}{2}.$$

Portanto temos os momentos

$$\begin{aligned}\langle X_i\rangle &= \frac{1}{2}(1+(-1)) = 0,\\ \langle X_iX_j\rangle &= \delta_{ij}\frac{1}{2}(1^2+(-1)^2) = \delta_{ij}.\end{aligned}$$

Seja então a caminhada

$$S = \sum_{i=1}^{N} X_i.$$

Supondo que o caminhante começa em $R = 0$, após N passos, sua posição será S. Cada caminhada realizada corresponde a um distinto. Se a média $\langle S\rangle \neq 0$, a velocidade típica (balística) será

$$V_B = \frac{\langle S\rangle}{N},$$

onde vemos que N tem o papel de tempo. Quando $\langle S\rangle = 0$, a velocidade típica (difusiva) será definida então como

$$V_D = \frac{\sqrt{\langle S^2\rangle}}{N},$$

que mede o afastamento típico da orígem, por unidade de tempo. Calculemos então as médias de S:

$$\begin{aligned}\langle S\rangle &= \sum_{i=1}^{N}\langle X_i\rangle = 0,\\ \left\langle S^2\right\rangle &= \sum_{i=1}^{N}\sum_{j=1}^{N}\langle X_iX_j\rangle\\ &= \sum_{i=1}^{N}\sum_{j=1}^{N}\delta_{ij}\\ &= \sum_{i=1}^{N}1 = N.\end{aligned}$$

Portanto, a caminhada aleatória descreve um movimento aleatório de velocidade típica

$$V_D = \frac{\sqrt{N}}{N} = \frac{1}{\sqrt{N}} \to 0.$$

Este tipo de movimento é típico de soluções da equação de Langevin.

Podemos também estudar a probabilidade de que o caminhante esteja na posição $S = R$ após N passos. Os passos possíveis são dados para a esquerda N_L e para a direita N_R de modo que $N_L + N_R = N$. Portanto

$$p(S = R = (N_R - N_L) \leq n, N) = \frac{1}{2^N} \frac{N!}{N_R! N_L!}$$

$$\rightarrow \sum_{N_R=0}^{N} p(R) = \sum_{N_R=0}^{N} \frac{1}{2^N} \frac{N!}{N_R!(N-N_R)!} 1^{N_R} 1^{N-N_R} = \frac{(1+1)^N}{2^N} = 1.$$

Podemos utilisar a aproximação de Stirling

$$N! \approx \sqrt{2\pi} e^{-N} N^{\frac{1}{2}+N},$$

onde

$$N = \frac{t}{\tau}, \;\; R = \frac{x}{a},$$

onde τ e a são as escalas de tempo e dislocamento, respectivamente. O limite para o contínuo é dado por $\lim_{\tau,a\rightarrow 0}$.

Vemos que pela expressão de Stirling, a probabilidade será da forma $e^{-N\,\phi(N_R)}$ que pode ser aproximada, after a little algebra, como

$$p(R) = p(N_R) \propto e^{-\frac{(N_R - N_L)^2}{2N}}.$$

Tomando o limite do contínuo, escrevemos:

$$\begin{aligned} p(x,t) &\propto \lim_{\tau,a\rightarrow 0} e^{-\frac{(x/a)^2}{2t/\tau}} \\ &= e^{-\lim_{\tau,a\rightarrow 0} \frac{\tau}{a^2} \frac{(x)^2}{2t}} \\ \Rightarrow p(x,t) &= \frac{1}{\sqrt{2\pi\, D\, t}} e^{-\frac{x^2}{2Dt}}, \end{aligned} \tag{3.11}$$

onde o coeficiente de difusão é

$$D = \lim_{\tau,a\rightarrow 0} \frac{a^2}{\tau}.$$

É facil verificar que $p(x,t)$ satisfaz a equação de difusão

$$\frac{\partial}{\partial t} p(x,t) = \frac{D}{2} \frac{\partial^2}{\partial x^2} p(x,t), \tag{3.12}$$

com a condição inicial

$$p(x, t = 0) = \delta(x).$$

Na sequência trataremos dos tipos de equações de evolução mais comumente encontradas. Elas se dividem basicamente em três tipos: equações de evolução para as variáveis dinâmicas (eq. canônicas, eq. de Langevin); equações de evolução para as distribuições de probabilidade (eq. de Fokker-Planck, eq. mestra, eq. de Boltzmann, eq. de Liouville); eq. de evolução para as médias (eq. hidrodinâmicas, mode-mode coupling).

3.9 Equação de Langevin

3.9.1 Exemplo de osciladores acoplados

Podemos entender o aparecimento da aproximação de Langevin via um exemplo trabalhado por Zwanzig [16]. Vamos supor que um conjunto de osciladores (x_i, p_i, com $i = 1, \ldots, N$) unidimensonais de massa m, e frequências ω_i, está acoplado a uma partícula Browniana de massa M. O Hamiltoniano total é dado por

$$H = \frac{P^2}{2M} + U(X) + \sum_{i=1}^{N} \left(\frac{p_i^2}{2m} + \frac{m\,\omega_i^2}{2} \left\{ x_i - \frac{\gamma_i}{m\,\omega_i^2}\, X \right\}^2 \right), \tag{3.13}$$

onde γ_i são constantes de acoplamento.

As equações canônicas são

$$\frac{dX}{dt} = \frac{P}{M}, \tag{3.14}$$

$$\frac{dP}{dt} = -\frac{\partial U}{\partial X} + \sum_{i=1}^{N} \gamma_i \left(x_i - \frac{\gamma_i}{m\,\omega_i^2}\, X \right), \tag{3.15}$$

$$\frac{dx_i}{dt} = \frac{p_i}{m}, \tag{3.16}$$

$$\frac{dp_i}{dt} = -m\,\omega_i^2\, x_i + \gamma_i\, X. \tag{3.17}$$

Tomando $X \equiv X(t)$ como fonte externa, uma solução particular para as equações 3.16-3.17 pode ser obtida pelo método (matricial neste caso) da função de Green. Podemos escrever

$$\frac{d\psi}{dt} \equiv \frac{d}{dt} \begin{pmatrix} x_i \\ p_i \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{m} \\ -m\,\omega_i^2\, x_i & 0 \end{pmatrix} \begin{pmatrix} x_i \\ p_i \end{pmatrix} + \begin{pmatrix} 0 \\ \gamma_i\, X \end{pmatrix} \equiv \mathrm{M}\,\psi + \eta.$$

A função de Green é

$$G(t-t') = \theta(t-t')\, e^{\mathrm{M}(t-t')},$$

e a solução é dada por

$$\begin{aligned} x_i(t) &= \frac{\gamma_i\, X(t)}{m\,\omega_i^2} - \int_0^t dt'\, \frac{\gamma_i}{m\,\omega_i^2}\, \cos\left[\omega_i\,(t-t')\right]\, \frac{d\,X(t')}{dt'} + \\ &= \left\{ x_i(0) - \frac{\gamma_i\, X(0)}{m\,\omega_i^2} \right\} \cos\left[\omega_i\, t\right] + \dot{x}_i(0)\, \frac{\sin\left[\omega_i\, t\right]}{\omega_i}. \end{aligned} \tag{3.18}$$

A equação 3.18 pode ser insertada na equação 3.15 dando a equação

$$\frac{dP}{dt} = -\frac{\partial U}{\partial X} - \int_0^t dt'\, \zeta(t-t')\, \frac{d\,X(t')}{dt'} + \xi(t), \tag{3.19}$$

onde

$$\zeta(t) = \sum_{i=1}^{N} \frac{\gamma_i^2}{m\,\omega_i^2}\, \cos\left[\omega_i\, t\right],$$

e

$$\xi(t) = \sum_{i=1}^{N} \gamma_i \left[\left\{ x_i(0) - \frac{\gamma_i\, X(0)}{m\,\omega_i^2} \right\} \cos\left[\omega_i\, t\right] + \dot{x}_i(0)\, \frac{\sin\left[\omega_i\, t\right]}{\omega_i} \right].$$

Observe que até este ponto as equações são totalmente sujeitas a reversibilidade temporal, e nenhuma informação sobre a temperatura do meio foi dada.

Se tomarmos as condições iniciais como sendo canonicamente distribuídas por

$$\mathcal{P}(\dot{x}_i(0), x_i(0), X(0)) \propto \exp\left\{ -\beta \sum_{i=1}^{N} \left(\frac{m\dot{x}_i^2(0)}{2} + \frac{m\,\omega_i^2}{2} \left\{ x_i(0) - \frac{\gamma_i}{m\,\omega_i^2}\, X(0) \right\}^2 \right) \right\},$$

obteremos as médias seguintes

$$\langle \xi(t) \rangle = 0 \tag{3.20}$$

$$\left\langle \xi(t)\, \xi(t') \right\rangle = k_B T\, \zeta(t-t'). \tag{3.21}$$

Se as escalas de tempo dissipativo forem pequenas em comparação com a escala do experimento, podemos fazer $\zeta(t-t') \to \gamma\, \delta(t-t')$ e obtemos o resultado correspondente ao ruído Gaussiano branco.

3.9.2 Caso geral

Quando um parâmetro pequeno ϵ existe podemos, em muitos casos, reduzir a equação de Liouville, equivalente às equações canônicas, para equações efetivas para as variáveis lentas do sistema, como na seção 3.7.

A essa equação chamamos em geral de equação de Langevin, ou "Langevin-like" (quando não -Markoviana). Este tipo de equação, que se aplica às variáveis lentas do sistema, tais como a posição e o momento da partícula Browniana (M), tem a forma geral

$$\dot{\mathbf{P}}(t) = -\int_{-\infty}^{t} d\tau\, \Gamma(t-\tau)\, \mathbf{P}(\tau) - \frac{\partial V}{\partial \mathbf{R}} + \eta(t), \tag{3.22}$$

onde

$$\Gamma(t-t') = \frac{1}{M\,T}\, \langle \eta(t)\, \eta(t') \rangle\,.$$

Quando o ruído η é Markoviano temos em geral (quando o ruído é descorrelacionado no tempo)

$$\langle \eta(t)\, \eta(t') \rangle = 2\,\gamma\,T\,\delta(t-t'),$$

o que nos dá a equação usual

$$\dot{\mathbf{P}}(t) = -\frac{\gamma\,\mathbf{P}(t)}{M} - \frac{\partial V}{\partial \mathbf{R}} + \eta(t). \tag{3.23}$$

Esta equação é a base de muitos tratamentos de sistemas que vão de partículas Brownianas a preços em uma economia. Uma aproximação muito usada é dita superamortecida (overdamped) quando $M/\gamma \to 0$. Neste caso o movimento fica escravizado pela força instantânea:

$$\mathbf{V}(t) = -\frac{1}{\gamma}\frac{\partial V}{\partial \mathbf{R}} + \frac{1}{\gamma}\eta(t). \tag{3.24}$$

3.10 Processos Markovianos

Uma importante classe de processos corresponde aos processos em que a memória dos sistema tem curta duração, ou seja, se a escala de tempo de uma medida é longa, comparada à escala de dissipação da memória do processo, podemos prever o próximo estado de um sistema apenas sabendo o seu estado atual. Neste caso, a configuração futura do sistema pode ser prevista como

função apenas da configuração instantânea do sistema, sem necessidade de levar em conta seu passado.

De um ponto de vista mais matemático, dado um processo estocástico definido por $Y(t)$ onde a variável Y assume os valores $y \in \Re$ em cada instante. Assim, para um processo Markoviano, a probabilidade condicional[1] não depende dos instantes anteriores:

$$p_{1|n-1}(y_n, t_n|y_1, t_1; y_2, t_2; \ldots; y_{n-1}, t_{n-1}) = p_{1|1}(y_n, t_n|y_{n-1}, t_{n-1}). \quad (3.25)$$

Assim, podemos escrever a forma de p_n como

$$p_n(y_1, t_1; \ldots; y_n, t_n) = p_{1|n-1}(y_n, t_n|y_1, t_1; \ldots; y_{n-1}, t_{n-1}) p_{n-1}(y_1, t_1; \ldots; y_{n-1}, t_{n-1}), \quad (3.26)$$

o que, continuando para os n instantes, pode ser escrita como

$$p_n(y_1, t_1; \ldots; y_n, t_n) = p_{1|1}(y_n, t_n|y_{n-1}, t_{n-1}) \ \ldots \ p_{1|1}(y_2, t_2|y_1, t_1) p_1(y_1, t_1) \quad (3.27)$$

onde $p_{1|1}(y_i, t_i|y_{i-1}, t_{i-1})$ é a probabilidade condicional de transição de que o sistema esteja em (y_i, t_i), dado que estivesse em (y_{i-1}, t_{i-1}). A probabilidade em multiplos instantes pode ser expressa como função da probabilidade condicional $p_{1|1}$ e da probabilidade inicial.

Duas propriedades importantes são a normalização

$$\int dy_1 \, dy_2 \ \ldots \ dy_n \, p_n(y_1, t_1; y_2, t_2; \ldots; y_n, t_n) = 1, \quad (3.28)$$

e soma marginal

$$\int dy_1 \, dy_2 \ \ldots \ dy_k \, p_n(y_1, t_1; y_2, t_2; \ldots; y_n, t_n) = p_{n-k}(y_{k+1}, t_{k+1}; \ldots; y_n, t_n). \quad (3.29)$$

A probabilidade de transição também é normalizada:

$$\int dy_{i+1} \, p_{1|1}(y_{i+1}, t_{i+1}|y_i, t_i) = \int dy_{i+1} \, \frac{p_2(y_{i+1}, t_{i+1}; y_i, t_i)}{p_1(y_i, t_i)} = \frac{p_1(y_i, t_i)}{p_1(y_i, t_i)} = 1. \quad (3.30)$$

[1] Relembrando que a probabilidade condicional de que o evento A aconteça dado que B aconteceu é dada pela relação de Bayes

$$p_{1|1}(A|B) = \frac{p_2(A, B)}{p_1(B)} \Leftrightarrow p_{1|1}(A|B) \, p_1(B) = p_2(A, B).$$

Para um processo Markoviano, toda probabilidade conjunta p_n pode ser escrita como um produto entre $p_{1|1}$ e p_1 : [2]

$$p_n(y_1,t_1;\ldots;y_n,t_n) = p_{1|1}(y_n,t_n|y_{n-1},t_{n-1})\ldots p_{1|1}(y_2,t_2|y_1,t_1)\,p_1(y_1,t_1). \tag{3.31}$$

3.10.1 Equação de Chapman-Kolmogoroff

Um processo Markoviano obedece a uma propriedade importante, que é a equação de Chapman-Kolmogorov (CK)

$$\begin{aligned} p_{1|1}(y_3,t_3|y_1,t_1) &= \frac{p_2(y_3,t_3;y_1,t_1)}{p_1(y_1,t_1)} \\ &= \int dy_2\,\frac{p_3(y_3,t_3;y_1,t_1;y_2,t_2)}{p_1(y_1,t_1)} \\ &= \int dy_2\,\frac{p_3(y_3,t_3;y_1,t_1;y_2,t_2)}{p_2(y_1,t_1;y_2,t_2)}\,\frac{p_2(y_1,t_1;y_2,t_2)}{p_1(y_1,t_1)}, \end{aligned}$$

onde obtemos a CK:

$$\begin{aligned} p_{1|1}(y_3,t_3|y_1,t_1) &= \int dy_2\,p_{1|2}(y_3,t_3|y_1,t_1;y_2,t_2)\,p_{1|1}(y_2,t_2|y_1,t_1) \\ &= \int dy_2\,p_{1|2}(y_3,t_3|y_2,t_2)\,p_{1|1}(y_2,t_2|y_1,t_1). \end{aligned} \tag{3.32}$$

Outra relação importante:

$$\begin{aligned} p_{1|1}(y_3,t_3|y_1,t_1) &= \frac{p_2(y_3,t_3;y_1,t_1)}{p_1(y_1,t_1)} \\ \Rightarrow p_2(y_3,t_3;y_1,t_1) &= p_{1|1}(y_3,t_3|y_1,t_1)\,p_1(y_1,t_1), \end{aligned}$$

e finalmente

$$p_1(y_3,t_3) = \int dy_1\,p_{1|1}(y_3,t_3|y_1,t_1)\,p_1(y_1,t_1). \tag{3.33}$$

[2] Sejam (A,B,C) eventos tais que $t_A > t_B > t_C$; podemos escrever:

$$p(A,B,C) = \frac{p(A,B,C)}{p(B,C)}\,\frac{p(B,C)}{p(C)}p(C) = p(A|B,C)\,p(B|C)\,p(C) = p(A|B)\,p(B|C)\,p(C).$$

Estas equações são típicas para a evolução de sistemas que perderam sua coerência. Para sistemas quânticos coerentes, A fase da função de onda é essencial na interação (como no caso do experimento de Young). Quando a coerência do feixe é perdida, apenas a probabilidade ($\Psi^* \Psi$) importa, não mais a fase. Isto é o que se passa com um sistema quando os termos não-diagonais da matriz densidade desaparecem e apenas os termos diagonais (probabilidades) sobrevivem.

3.10.2 Equações Mestras

Uma equação mestra é uma equação para a evolução da distribuição de probabilidades (ou mais corretamente, da matriz de transição entre os estados) [17] no tempo, para um processo Markoviano estacionário [3]

Assim, escrevemos para estes processos:

$$T_\tau(y_2|y_1) \equiv p_{1|1}(y_2, t+\tau|y_1, t). \tag{3.34}$$

As matrizes de transição $T_\tau(y_2|y_1)$ pode ser expandida em potências de τ. Para processos em tempo contínuo (em contraste com tempos discretizados) escrevemos [17]

$$T_\tau(y_2|y_1) = [1 - a_0(y_1)\,\tau]\,\delta(y_2 - y_1) + \tau\, W(y_2|y_1) + \mathcal{O}(\tau^2), \tag{3.35}$$

onde $W(y_2|y_1)$ é a probabilidade de transição, por unidade de tempo, de $y_1 \to y_2$. A quantidade a_0 é

$$a_0(y_1) = \int dy_2\, W(y_2|y_1).$$

Observe que $T_\tau(y_2|y_1)$ está perfeitamente normalizada até a ordem $\mathcal{O}(\tau^2)$.

Substituindo 3.35 em 7.5 temos:

$$T_{\tau+\tau'}(y_3|y_1) = [1 - a_0(y_3)\,\tau']\,T_\tau(y_3|y_1) + \tau' \int dy_2\, W(y_3|y_2)\, T_\tau(y_2|y_1).$$

Remanejando termos e dividindo por τ', e em seguida tomando o limite $\tau' \to 0$, temos a Equação Mestra

$$\frac{\partial}{\partial \tau} T_\tau(y_3|y_1) = \int dy_2\, \{W(y_3|y_2)\, T_\tau(y_2|y_1) - W(y_2|y_3)\, T_\tau(y_3|y_1)\}. \tag{3.36}$$

[3] Processos estacionários só dependem de $t_2 - t_1$, mas não de t_1

Observe que esta equação é uma equação para a evolução de uma matriz de transição $T_\tau(y_3|y_1)$, e não para uma probabilidade de ocupação de um estado! Contudo, podemos multiplicar ambos os lados de 3.36 por $p_1(y_1, t_1 = 0)$ e integrar em y_1 obtendo (usando $T_\tau(y, y_1) = T_\tau(y|y_1)p_1(y_1, t_1)$):

$$\frac{\partial}{\partial\tau}p_1(y,\tau) \;=\; \int dy'\,\left\{W(y|y')\,p_1(y',\tau) - W(y'|y)\,p_1(y,\tau)\right\}. \quad (3.37)$$

Quando o conjunto de estados é discreto, podemos escrever ($p_1(y, \tau) \to p_n(\tau)$):

$$\frac{\partial}{\partial\tau}p_n(\tau) \;=\; \sum_{n'}\left\{W_{nn'}\,p_{n'}(\tau) - W_{n'n}\,p_n(\tau)\right\}. \quad (3.38)$$

Observem que as equações mestras acima são constituídas de dois termos: um termo de criação de probabilidades que representa transições de outros estados para o estado n; um termo de destruição de probabilidades que representa transições do estado n para outros estados, similarmente à equação de Boltzmann.

Exemplo

Para o decaimento redioativo, dados N átomos ativos no instante $t = 0$, qual é a probabilidade de encontrarmos $M \leq N$ átomos ativos no instante t dada a probabilidade de decaimento de γ por átomo por unidade de tempo? A equação mestra que este sistema obedece é

$$\dot{p}_M(t) = \gamma(M+1)p_{M+1}(t) - \gamma M p_M(t).$$

Esta equação pode ser resolvida pelo método da função geratriz. Definimos a mesma como

$$F(\lambda, t) = \sum_{n=0}^{N}\lambda^n\,p_n(t).$$

A função geratriz "gera" as probabilidades da seguinte forma:

$$p_n(t) = \frac{1}{n!}\,\frac{\partial^n}{\partial\lambda^n}F(\lambda,t)\bigg|_{\lambda=0}$$

Observe que $F(\lambda = 1, t) = 1$, $F(\lambda, 0) = \lambda^N$ e que $F(\lambda, \infty) = 1$.

Assim, multiplicando a EM acima por λ^M e somando de 0 a N (tomaremos $p_{N+1} = 0$ identicamente) obtemos:

$$\partial_t F(\lambda,t) = \gamma\,(1-\lambda)\,\partial_\lambda\,F(\lambda,t). \quad (3.39)$$

Para resolver o problema acima podemos fazer a mudança de variáveis $v = \ln(1-\lambda)$ que nos dá:

$$\partial_t F(v,t) = \gamma\, \partial_v\, F(v,t),$$

que é uma equação do tipo onda, cuja solução é da forma $\phi(x = \gamma\, t - v)$. Dadas as condições que esta solução deve obedecer, podemos ver que deve ser da forma

$$F(\lambda,t) = \left(1 - e^{-x}\right)^N = \left(1 - e^{-\gamma t}(1-\lambda)\right)^N.$$

É fácil verificar que esta solução satisfaz a equação 3.39 e as condições $F(\lambda = 1, t) = 1$, $F(\lambda, 0) = \lambda^N$ e $F(\lambda, \infty) = 1$.

Podemos expandir a solução acima:

$$\begin{aligned}
F(\lambda,t) &= \sum_{n=0}^{N} \frac{N!}{n!\,(N-n)!}\,(-1)^n\, e^{-n\gamma t}(1-\lambda)^n \\
&= \sum_{n=0}^{N}\sum_{k=0}^{n} \frac{N!}{n!\,(N-n)!}\,\frac{n!}{k!\,(n-k)!}\,(-1)^{n+k}\, e^{-n\gamma t}\,\lambda^k \\
&= \sum_{n=0}^{N}\sum_{k=0}^{n} \frac{N!}{(N-n)!\,k!\,(n-k)!}\,(-1)^{n+k}\, e^{-n\gamma t}\,\lambda^k \\
&= \sum_{n=0}^{N}\sum_{k=0}^{n} \theta(n,k)\,\lambda^k.
\end{aligned}$$

Portanto, podemos identificar a propriedade $p_k(t)$ como

$$\begin{aligned}
p_k(t) &= \sum_{n=k}^{N} \theta(n,k) \\
&= \sum_{n=k}^{N} \frac{N!}{(N-n)!\,k!\,(n-k)!}\,(-1)^{n+k}\, e^{-n\gamma t}.
\end{aligned}$$

Observe que $p_N(t) = e^{-N\gamma t}$, que corresponde à probabilidade de que cada átomo sobreviva pelo tempo t, ou seja $\left(e^{-\gamma t}\right)^N$. A probabilidade $p_0(t)$ é

$$p_0(t) = \sum_{n=0}^{N} \frac{N!}{(N-n)!\,n!}\,(-1)^n\, e^{-n\gamma t} = (1 - e^{-\gamma t})^N,$$

é o produto das probabilidades de que cada átomo tenha decaído no tempo t (dada por $1 - e^{-\gamma t}$).

A média de átomos sobreviventes no instante t, $\langle N \rangle (t)$ é dada por

$$\begin{aligned}
\langle N \rangle (t) &= \sum_{n=0}^{N} n\, p_n(t) \\
&= \left. \frac{\partial}{\partial \lambda} F \right|_{\lambda=1} \\
&= \left. \frac{\partial}{\partial \lambda} \left(1 - e^{-\gamma t}(1-\lambda)\right)^N \right|_{\lambda=1} \\
&= \left. N \left(1 - e^{-\gamma t}(1-\lambda)\right)^{N-1} \left(e^{-\gamma t}\right) \right|_{\lambda=1} \\
&= N\, e^{-\gamma t},
\end{aligned}$$

o que recupera um resultado esperado.

3.10.3 Classes de Matrizes-$\mathcal{W}$

É importante entender as propriedades da matriz $\mathcal{W}_{nm}$, definida por:

$$\mathcal{W}_{nm} = W_{nm} - \delta_{nm} \left(\sum_k W_{kn} \right). \tag{3.40}$$

Isto nos permite escrever a equação mestra na forma matricial:

$$\dot{p} = \mathcal{W}\, p. \tag{3.41}$$

Algumas propriedades seguem listadas abaixo.

- $\mathcal{W}_{nm} \geq 0$ para $n \neq m$.
- $\sum_n \mathcal{W}_{nm} = 0$, para cada m.

A segunda propriedade acima garante que existe um auto-vetor a esquerda, $\Psi_E = (1, 1, \ldots, 1)$, com autovetor nulo: $\Psi_E \mathcal{W} = 0$. É só lembrar que o determinante de $\mathcal{W} - \lambda \mathcal{U}$ deve ser nulo, seja para o caso left ou right! portanto, o espectro de autovalores é o mesmo para os dois casos. Então, podemos mostrar que existe um auto vetor a direita Ψ_D com autovalor nulo:

$$\mathcal{W} \Psi_D = 0,$$

o que mostra que existe um estado estacionário $p_{ss} = \Psi_D$ de modo que $\dot{p}_{ss} = 0$.

Equilíbrio.

No equilíbrio temos

$$\frac{\partial}{\partial\tau}p_n(\tau) = 0 \Rightarrow \sum_{n'} W_{nn'}\, p_{n'}(\tau) = \sum_{n'} W_{n'n}\, p_n(\tau).$$

O fluxo de probabilidades de um estado n para todos os seus vizinhos é compensado pelo pelo fluxo inverso. Existe uma condição mais forte chamada balanço detalhado, que veremos a seguir.

Balanço detalhado.

Para sistemas cujo Hamiltoniano é par nos momentos, podemos mostrar que os fluxos de probabilidade são simétricos no equilíbrio:

$$W_{nn'}\, p^{ss}_{n'} = W_{n'n}\, p^{ss}_{n}.$$

A razão é que para cada trajetória $z_F \equiv (q(t), p(t))$, para $t = 0 \to \tau$, existe uma inversão temporal $(t \to -t; q(t) \to q(\tau - t); p(t) \to -p(\tau - t))$ que refaz a tal como a figura 3.1.

Portanto o fluxo microscópico de probabilidades de estados "verdes" para "azuis" é simétrico e o balanço detalhado é obedecido.

3.10.4 Equação de Fokker-Planck

A forma diferencial da equação mestra para sistemas de estados contínuos se chama Equação de Fokker-Planck. Dada a equação 3.37

$$\frac{\partial}{\partial\tau}p_1(y,\tau) = \int dy'\, \left\{W(y|y')\, p_1(y',\tau) - W(y'|y)\, p_1(y,\tau)\right\},$$

podemos considerar passos diferenciais $r = y' - y$ e que a matriz de transição $W(y'|y) \equiv W(r, y)$ varia pouco com y.

Assim temos

$$\begin{aligned}
\frac{\partial}{\partial\tau}p_1(y,\tau) &= \int dr\, \left\{W(-r, y+r)\, p_1(y+r,\tau) - W(r,y)\, p_1(y,\tau)\right\}, \\
&= \int dr\, \left\{W(-r,y)\, p_1(y,\tau) - r\,\frac{\partial}{\partial y}W(-r,y)\, p_1(y,\tau) + \frac{r^2}{2}\,\frac{\partial^2}{\partial y^2}W(-r,y)\, p_1(y,\tau) - W(r,y)\, p_1(y,\tau)\right\} \\
&= \int dr\, \left\{-r\,\frac{\partial}{\partial y}W(-r,y)\, p_1(y,\tau) + \frac{r^2}{2}\,\frac{\partial^2}{\partial y^2}W(-r,y)\, p_1(y,\tau)\right\}, \\
&= \int dr\, \left\{r\,\frac{\partial}{\partial y}W(r,y)\, p_1(y,\tau) + \frac{r^2}{2}\,\frac{\partial^2}{\partial y^2}W(r,y)\, p_1(y,\tau)\right\},
\end{aligned}$$

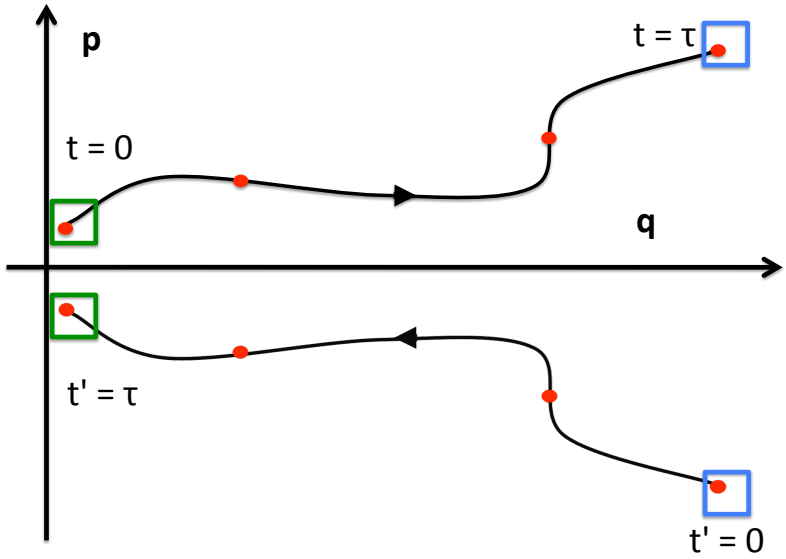


Figure 3.1: Trajetória inversa: como as variáveis conservadas (energia) são pares em p, o fluxo de probabilidades entre os estados "verde" e "azul" é conservado.

de onde obtemos a Equação de Fokker-Planck (FP):

$$\frac{\partial}{\partial\tau}p_1(y,\tau) \;=\; \frac{\partial}{\partial y}\left(a_1(y)\,p_1(y,\tau)\right) + \frac{1}{2}\frac{\partial^2}{\partial y^2}\left(a_2(y)\,p_1(y,\tau)\right),$$

onde os momentos de salto são definidos por

$$a_n(y) = \int dr\, r^n\, W(r,y). \tag{3.42}$$

A equação de FP é uma equação para o fluxo de probabilidades $J(y,\tau)$ definido pela equação de continuidade:

$$\frac{\partial}{\partial\tau}p_1(y,\tau) = -\frac{\partial}{\partial y}\,J(y,\tau),$$

onde o fluxo é dado por

$$J(y,\tau) = -a_1(y)\, p_1(y,\tau) - \frac{1}{2}\frac{\partial}{\partial y}\left(a_2(y)\, p_1(y,\tau)\right),$$

onde o primeiro termo a direita é um termo de "drift" (a_1 é uma velocidade) e o segundo termo é do tipo difusivo.