

Introdução aos Processos Estocásticos e Aplicações

Daniel A Stariolo

Departamento de Física
Universidade Federal do Rio Grande do Sul

2006

Sumário

1	Variáveis aleatórias e distribuições de probabilidade	1
1.1	O conceito de probabilidade e propriedades fundamentais . . .	1
1.1.1	Axiomas básicos	2
1.1.2	Probabilidades Condicionadas e Teorema de Bayes . . .	3
1.1.3	Independência estatística	4
1.1.4	Variáveis aleatórias	4
1.2	Distribuição conjunta	6
1.3	Algumas distribuições importantes	6
1.4	Valor esperado e momentos de uma distribuição	8
1.5	Função característica	9
1.6	Soma de variáveis independentes	10
1.7	Transformação de variáveis aleatórias	12
1.8	Lei dos Grandes Números e o Teorema do Limite Central . . .	13
1.9	Função geratriz	14
1.10	Lista de exercícios	16
2	Movimento Browniano I: Caminhadas aleatórias e a equação de difusão	18
2.1	Caminhada aleatória numa rede hipercúbica	19
2.2	A equação de difusão	22
3	Processos Markovianos e a Equação Mestra	26
3.1	Processos Markovianos	26
3.2	Matriz Estocástica	27
3.2.1	Algumas propriedades de matrizes não negativas	28
3.2.2	Método algébrico	30
3.2.3	<i>Sistema de dois estados</i>	32
3.2.4	<i>O modelo de urnas dos Ehrenfest</i>	34

3.2.5	<i>A caminhada aleatória, tempo de primeira passagem, probabilidade de recorrência</i>	39
3.3	A Equação Master	45
3.3.1	<i>Caminhada aleatória em tempo contínuo</i>	47
3.3.2	A matriz de evolução W	48
3.3.3	<i>Sistema de dois estados</i>	49
3.3.4	Propriedades das matrizes W	49
3.3.5	Balanco Detalhado	51
3.3.6	Processos “One step” ou “Nascimento-Morte”	51
3.3.7	<i>Decaimento Radiativo</i>	52
3.3.8	<i>Reações químicas</i>	53
3.4	Lista de exercícios	55
4	O Modelo de Ising	58
4.1	Dinâmica de Glauber [13]	61
4.2	Aproximação de campo médio	62
4.3	O modelo de Ising unidimensional	64
4.3.1	Relaxação da Magnetização	65
4.3.2	Correlações a tempos iguais	67
5	Movimento Browniano II: a equação de Langevin	71
5.1	Caminhada aleatória nas velocidades: a equação de Langevin	71
5.1.1	De volta as coordenadas: o deslocamento quadrático médio	75
5.1.2	Balanco energético	77
5.2	Funções de correlação	78
5.3	Análise espectral de um processo estocástico: o teorema de Wiener-Kintchine	78
5.4	Conjunto de equações de Langevin acopladas	83
5.5	Equações de Langevin Markovianas e Não-Markovianas	84
5.6	Partícula browniana num banho térmico de osciladores harmônicos	86
5.7	Lista de Exercícios	90
6	A equação de Fokker-Planck	91
6.1	De Langevin a Fokker-Planck	91
6.2	Solução estacionária	93
6.3	Operador de evolução	94

6.4 O problema de Kramers	97
Referências Bibliográficas	101

Capítulo 1

Variáveis aleatórias e distribuições de probabilidade

1.1 O conceito de probabilidade e propriedades fundamentais

Faremos, neste primeiro capítulo, uma introdução rápida dos conceitos fundamentais da teoria das probabilidades, que é a metodologia matemática básica para poder desenvolver os conceitos físicos que iremos estudar neste curso. A seguinte discussão está baseada no clássico livro de Gardiner [1], nas notas do curso de teoria de probabilidades de Marinari e Parisi [2] e no livro de Cáceres [4]. Todo ao longo do curso iremos fazer referência também ao Feller [3], que é um dos melhores livros sobre probabilidades e aplicações.

A noção de probabilidade de um evento é intuitiva, e nós fazemos uso dela no nosso dia a dia. No entanto, para fins de sistematizar a aplicação de um conceito, é necessário definir ele de forma mais ou menos precisa. A definição mais tradicional e intuitiva de probabilidade é a *freqüencial* que diz mais ou menos o seguinte: *quando jogamos uma moeda a cara ou cruz, se diz que o resultado “cara” tem probabilidade $1/2$ se f_N , o número de vezes nos quais “cara” apareceu dentre N lances, dividido pelo número total de lances efetuados, tende a $1/2$ quando N tende a infinito.*

Outro ponto de vista é o chamado *subjetivista*, segundo o qual a probabilidade de um evento indica uma “previsão” sobre o mesmo, condicionada sobre a nossa ignorância. Consideremos por exemplo a seguinte afirmação: *A probabilidade de chuvas nos próximos dias é de 30%*. A interpretação freqüencial

não é aqui tão intuitiva, em tanto que a interpretação subjetivista seria: “Temos dados suficientes para afirmar que, se fizermos uma previsão assim 10 vezes e acertarmos em 3 delas, o resultado é considerado aceitável”. É claro que ambas interpretações devem ser equivalentes, e por tanto devem nos levar as mesmas conclusões.

1.1.1 Axiomas básicos

Podemos fazer uma formulação axiomática da teoria das probabilidades, de forma que podemos deduzir toda a teoria a partir de um conjunto pequeno de axiomas, que definiremos a seguir. A linguagem da teoria de conjuntos também é uma ferramenta muito útil para formular princípios básicos da teoria. Consideremos um caso simples no qual os eventos podem ser caracterizados por um número inteiro n . Suponhamos que estes eventos sejam:

- mutuamente excludentes, e que
- em cada realização um e somente um deles se verifique.

Indicando com p_n a probabilidade do evento n -ésimo, se deve verificar que:

1. $p_n \geq 0 \quad \forall n$
2. $\sum_n p_n = 1$
3. A cada subconjunto A dos números inteiros, a probabilidade de que algum evento pertencente a A se verifique é dada por:

$$P(A) = \sum_{n \in A} p_n$$

As seguintes conseqüências destes três axiomas são importantes:

- Se \bar{A} é o conjunto complementar de A , então:

$$P(\bar{A}) = 1 - P(A) \tag{1.1}$$

-

$$P(\emptyset) = 0$$

•

$$P(I) = 1$$

•

$$P(A \cup B) = P(A) + P(B) - P(A \cap B),$$

onde \emptyset é o conjunto vazio e I é o conjunto de todos os inteiros. $P(A \cup B)$ é a probabilidade de que aconteça um evento do conjunto A , **ou** um evento do conjunto B , **ou** ambos. $P(A \cap B)$ é a probabilidade de que aconteça um evento do conjunto A e um evento do conjunto B (probabilidade conjunta).

1.1.2 Probabilidades Condicionadas e Teorema de Bayes

As vezes estaremos interessados em conhecer a probabilidade de um conjunto L , *condicionado* a formar parte de um outro conjunto M . Consideremos uma população de N indivíduos, dos quais M são mulheres e L são loiras. As probabilidades de que um indivíduo particular seja mulher e a probabilidade de que um indivíduo seja uma mulher loira são, respectivamente: $P(M) = M/N$ e $P(L) = L/N$. Agora consideremos a probabilidade de que um indivíduo, que sabemos é uma mulher, seja loira. Esta probabilidade é L/M . Assim podemos definir a probabilidade condicionada de um evento L , sabendo que outro evento M é certo, na forma:

$$P(L|M) = \frac{P(L \cap M)}{P(M)} \quad (1.2)$$

onde deve acontecer necessariamente que $P(M) \neq 0$ e a probabilidade conjunta $P(L \cap M) = L/N$. Para dois conjuntos qualquer L e M , também é válida a relação

$$P(M|L) = \frac{P(M \cap L)}{P(L)} \quad (1.3)$$

Das duas últimas relações surge o Teorema de Bayes que diz que

$$P(M|L) = \frac{P(L|M) P(M)}{P(L)} \quad (1.4)$$

e define uma espécie de “reversibilidade” nas probabilidades condicionadas.

1.1.3 Independência estatística

No contexto da teoria das probabilidades, dizemos que dois conjuntos de eventos A e H são independentes se a probabilidade de qualquer evento em um dos conjuntos não está condicionado pelos eventos do segundo conjunto. A partir de (1.2) obtemos, neste caso:

$$P(A \cap H) = P(A) P(H). \quad (1.5)$$

No caso mais geral em que temos uma série de conjuntos A, B, C, \dots estatisticamente independentes, devemos estender a condição anterior para todos os possíveis pares de conjuntos. Facilmente podemos mostrar que a condição de independência estatística é em geral dada por:

$$P(A \cap B \cap C \cap \dots) = P(A)P(B)P(C) \dots \quad (1.6)$$

1.1.4 Variáveis aleatórias

De forma mais geral, se os eventos A_1, A_2, \dots de um conjunto A são identificados com um número real x , podemos interpretar esses eventos como os possíveis valores de uma **variável aleatória X** . A introdução de variáveis aleatórias simplifica consideravelmente o cálculo de valores médios e outras propriedades estatísticas que serão definidas através de *funções de variáveis aleatórias*.

Exemplo: Vidros de spin

Um vidro de spin é um material magnético, no qual momentos de dipolo estão localizados em posições ao acaso numa matriz não magnética. Um modelo básico de vidro de spin (conhecido como modelo de Edwards-Anderson) está definido pela seguinte função energia:

$$H = \sum_{\langle i,j \rangle} J_{ij} s_i s_j \quad (1.7)$$

onde s_i é a variável que representa a componente z do momento de dipolo no sítio i -ésimo de uma rede cristalina e J_{ij} representa a intensidade da interação entre o par de dipolos ou spins s_i e s_j . Como os dipolos estão localizados em posições aleatórias, os valores das energias de interação J_{ij} podem ser consideradas como variáveis aleatórias. Elas são determinadas no modelo de acordo com uma certa probabilidade, por exemplo

$$P(J_{ij}) \propto \exp -\frac{J_{ij}^2}{2\sigma^2} \quad (1.8)$$

A possibilidade de que a interação J_{ij} possa tomar valores positivos ou negativos introduz um efeito interessante no sistema, chamada *frustração*, porque dada a matriz de interações de pares de spin, não é possível minimizar simultaneamente a energia de todos os pares. Sendo assim o sistema fica *frustrado*.

Se os possíveis valores da variável aleatória X são numerables $\{x_1, x_2, \dots\}$ se diz que a variável aleatória é **discreta**. Quando o conjunto de valores possíveis é um intervalo contínuo, por exemplo $x \in [a, b]$ se diz que a variável aleatória é **contínua**. No caso discreto, de acordo aos axiomas vistos antes, cada evento x_1, x_2, \dots terá associada uma probabilidade $P(x_1), P(x_2), \dots$, etc, de forma que

$$\sum_n P(x_n) = 1. \quad (1.9)$$

No caso em que somente um evento x_p seja certo, enquanto nenhum outro pode ocorrer, temos que

$$P(x_n) = \delta_{n,p},$$

onde $\delta_{n,p}$ é a Delta de Kronecker.

Quando a variável aleatória é contínua não podemos mais aplicar a condição (1.9) e é necessário estender o formalismo e passar a um limite contínuo definindo uma **probabilidade diferencial** $P(x_j) dx_j$ que se interpreta como a probabilidade da variável x tomar um valor no intervalo diferencial dx_j . A probabilidade assim definida tende a zero quando o intervalo $dx_j \rightarrow 0$, ou em outras palavras, a probabilidade de um número isolado x_0 é zero quando a variável aleatória é contínua. Com esta definição de probabilidade contínua podemos interpretar a quantidade $P(x)$ como a **distribuição de probabilidade** da variável x . A quantidade $P(x)$ não é uma função usual, mas uma função generalizada ou distribuição, já que a variável é aleatória. De acordo com a definição, se a variável x está definida num domínio \mathcal{D} a condição de normalização é

$$\int_{x \in \mathcal{D}} P(x) dx = 1. \quad (1.10)$$

Neste caso contínuo, se apenas um evento x_p é certo e nenhum outro pode acontecer, a densidade de probabilidade de este evento se escreve

$$P(x) = \delta(x - x_p),$$

onde $\delta(x - x_p)$ é a Delta de Dirac. A Delta de Dirac é uma função generalizada, e dentre as propriedades que a distinguem das funções normais uma

das mais importantes é que as funções generalizadas somente fazem sentido quando consideradas debaixo do símbolo integral, por exemplo

$$f(x_p) = \int \delta(x - x_p) f(x) dx. \quad (1.11)$$

A matemática envolvida ao considerar distribuições de probabilidade são tratados no ramo da matemática conhecido como *teoria da medida*.

1.2 Distribuição conjunta

Sejam x e y duas variáveis aleatórias. A probabilidade de que x se encontre no intervalo $[a, b]$ e y no intervalo $[c, d]$ é dada por:

$$\int_a^b \int_c^d P(x, y) dx dy \quad (1.12)$$

onde $P(x, y)$ é a *densidade de probabilidade conjunta* de x e y (probabilidade da interseção dos dois conjuntos de eventos, na linguagem de teoria de conjuntos). Integrando respeito de uma das variáveis podemos obter a *distribuição marginal* da outra:

$$P(x) = \int P(x, y) dy \quad (1.13)$$

e

$$P(y) = \int P(x, y) dx \quad (1.14)$$

Se as variáveis aleatórias x e y são independentes então:

$$P(x, y) = P(x) P(y) \quad (1.15)$$

Se $z = f(x, y)$ e $P(x, y)$ é a distribuição conjunta de x e y , então a distribuição de z é dada por:

$$P(z) = \int \int \delta(z - f(x, y)) P(x, y) dx dy \quad (1.16)$$

1.3 Algumas distribuições importantes

- *Distribuição binomial*

Consideremos uma moeda assimétrica, na qual a probabilidade de tirar

cara seja p e probabilidade de obter cruz seja $q = 1 - p$. A probabilidade de obter m caras depois de N jogadas da moeda será:

$$p_m = \binom{N}{m} p^m q^{N-m} \quad (1.17)$$

onde

$$\binom{N}{m} = \frac{N!}{m! (N-m)!} \quad (1.18)$$

(1.17) é a distribuição binomial.

- *Distribuição de Poisson*

Muitas vezes é necessário considerar a probabilidade de ocorrência de eventos tais que p é muito pequena, mas no entanto quando $N \rightarrow \infty$ o produto $Np \equiv \lambda$ é uma constante finita. Considerando esse limite se obtém uma aproximação da distribuição binomial:

$$p_m = e^{-\lambda} \frac{\lambda^m}{m!} \quad (1.19)$$

onde $\lambda > 0$ é um parâmetro. (1.19) é a distribuição de Poisson.

Nas duas distribuições anteriores a variável aleatória m pode tomar valores inteiros não negativos $m = 0, 1, 2 \dots N$, é uma *variável aleatória discreta*. Quando a variável aleatória pode assumir valores contínuos reais já vimos que podemos definir uma **densidade de probabilidade** $P(x)$, sendo $P(x) dx$ a probabilidade da variável x assumir um valor entre x e $x + dx$. Dessa forma a probabilidade da variável tomar qualquer valor num intervalo $[a, b]$ é dada por:

$$\int_a^b P(x) dx \quad (1.20)$$

Exemplos de distribuições contínuas são:

- *Distribuição gaussiana*

$$P(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right) \quad (1.21)$$

onde σ é um parâmetro.

- *Distribuição de Laplace*

É definida por:

$$P(x) = \frac{1}{2\alpha} \exp\left(-\frac{|x|}{\alpha}\right) \quad (1.22)$$

- *Distribuição de Lorentz*

A distribuição de Lorentz, ou *lorentziana* é definida por:

$$P(x) = \frac{a}{\pi(a^2 + x^2)} \quad (1.23)$$

Uma característica importante desta distribuição é o decaimento algébrico com x , a diferença dos decaimentos tipo exponencial da gaussiana e da distribuição de Laplace. Como consequência disto o segundo momento ou largura da distribuição é infinito, diverge (ver seção seguinte para a definição de momentos de uma distribuição).

Com a ajuda da função Delta de Dirac podemos expressar uma distribuição discreta por meio de uma densidade continua:

$$P(x) = \sum_m p_m \delta(x - m) \quad (1.24)$$

1.4 Valor esperado e momentos de uma distribuição

Seja x uma variável aleatória discreta que pode tomar valores $\{x_i, i = 1 \dots\}$ com probabilidade $P_i \equiv P(x_i)$. Quando a série é absolutamente convergente, se define o *valor esperado* de x como:

$$\langle x \rangle = \sum_i x_i P_i \quad (1.25)$$

Em geral, podemos definir os valores esperados de funções da variável aleatória x :

$$\langle f(x) \rangle = \sum_i f(x_i) P_i \quad (1.26)$$

O caso particular $f(x) = x$ nos dá o valor esperado da própria variável. Estas definições se generalizam de forma trivial para o caso no qual a variável aleatória pode tomar valores num intervalo contínuo. Vamos continuar nesta

seção considerando apenas o caso contínuo, mas tendo presente que o caso discreto é imediato.

O momento de ordem n de uma distribuição de probabilidades é dado por:

$$\langle x^n \rangle = \int x^n P(x) dx \quad (1.27)$$

O momento de ordem 1, $\langle x \rangle$ corresponde ao valor esperado ou **valor médio** da variável x . O **desvio quadrático médio** ou **variância** é dado por $\sigma^2 = \langle (x - \langle x \rangle)^2 \rangle$, e representa uma medida da largura da distribuição. A raiz quadrada desta quantidade, ou seja, σ é conhecida como **desvio padrão**.

Se $f(x)$ é qualquer função da variável aleatória contínua x , então a média de $f(x)$ é definida por analogia com (1.26):

$$\langle f(x) \rangle = \int f(x) P(x) dx \quad (1.28)$$

1.5 Função característica

Alternativamente, podemos caracterizar uma variável aleatória por meio do desenvolvimento de Fourier da distribuição de probabilidades. A função característica $g(k)$ de uma variável aleatória x é definida como:

$$\begin{aligned} g(k) &= \int_{\mathcal{D}} P(x) e^{ikx} dx \\ &= \langle e^{ikx} \rangle \end{aligned} \quad (1.29)$$

Alternativamente, ela pode ser interpretada como o valor médio da função e^{ikx} . Se os momentos da distribuição $\langle x^n \rangle$ existem, a função característica admite desenvolvimento em série de Taylor no entorno de $k = 0$:

$$g(k) = \sum_{n=0}^{\infty} \frac{(ik)^n}{n!} \langle x^n \rangle \quad (1.30)$$

A função característica permite determinar todos os momentos da distribuição, que são dados por:

$$\langle x^n \rangle = \frac{1}{i^n} \frac{d^n}{dk^n} g(k) \Big|_{k=0}. \quad (1.31)$$

Desta relação fica estabelecido que nos casos em que $g(k)$ não seja diferenciável em $k = 0$, um ou alguns momentos podem não estar definidos. Exemplos

de distribuições onde momentos não estão definidos são a distribuição de Lorentz e distribuições do tipo Lévy, cuja característica fundamental é um decaimento algébrico da densidade de probabilidade para grandes valores da variável, do tipo $x^{-\alpha}$.

Exemplo: Distribuição de Weierstrass

Dada uma variável aleatória discreta s , se define a probabilidade de Weierstrass na forma [4]:

$$P(s) = \frac{a-1}{2a} \sum_{n=0}^{\infty} \frac{1}{a^n} (\delta_{s,b^n} + \delta_{s,-b^n}) \quad (1.32)$$

onde $a > 1$ e $b > 1$. Esta distribuição é do tipo Lévy, e tem propriedades interessantes no contexto de caminhadas aleatórias, nas quais, se interpretarmos a variável s como o comprimento de um salto, vemos que a probabilidade de dar um salto de comprimento $s = b^n$ decai como $1/a^n$, e por tanto decairá muito rapidamente, algebricamente, se o fator $1/a$ for muito pequeno ($a \gg 1$). Uma caminhada aleatória deste tipo estará dominada por eventos ou saltos muito largos de probabilidade muito baixa.

Podemos verificar que o segundo momento da distribuição diverge se $b^2 > a$:

$$\langle s^2 \rangle = \sum_s s^2 P(s) = \frac{a-1}{a} \sum_{n=0}^{\infty} (b^2/a)^n. \quad (1.33)$$

A função característica da distribuição de Weierstrass está bem definida, é contínua em todo o eixo real e se conhece como *função de Weierstrass* [5].

1.6 Soma de variáveis independentes

Muitas vezes temos que fazer uma análise de variáveis que são a soma de variáveis aleatórias independentes, por exemplo, na repetição de um experimento idêntico. Consideremos o processo $y = x_1 + x_2$, onde x_1 e x_2 são variáveis independentes. É fácil mostrar que a função característica do processo y é dada por:

$$G(k) = g_1(k)g_2(k) \quad (1.34)$$

onde $g_1(k)$ e $g_2(k)$ são as funções características dos processos x_1 e x_2 respectivamente. Partindo da definição de função característica:

$$G(k) = \int e^{iky} P(y) dy = \langle e^{ikx_1} e^{ikx_2} \rangle. \quad (1.35)$$

Mas como x_1 e x_2 são independentes:

$$\langle e^{ikx_1} e^{ikx_2} \rangle = \langle e^{ikx_1} \rangle \langle e^{ikx_2} \rangle = g_1(k) g_2(k) \quad (1.36)$$

Este resultado é trivialmente generalizável no caso de uma soma de N variáveis aleatórias independentes $y = \sum_{i=1}^N x_i$, em cujo caso:

$$G(k) = g_1(k) g_2(k) \dots g_N(k). \quad (1.37)$$

A partir das propriedades anteriores e das propriedades da função característica é fácil mostrar que a média e a variância de uma variável que é soma de variáveis independentes são dadas por:

$$\langle y \rangle = \sum_{i=1}^N \langle x_i \rangle \quad (1.38)$$

e

$$\langle y^2 \rangle - \langle y \rangle^2 = \sum_{i=1}^N \{ \langle x_i^2 \rangle - \langle x_i \rangle^2 \} \quad (1.39)$$

Exemplo: Ensaios de Bernoulli

O exemplo utilizado no início, da moeda assimétrica, é um caso particular de um experimento muito geral e comum, chamado ensaios de Bernoulli. Consideremos N experimentos idênticos, nos quais podem ocorrer apenas dois eventos por ensaio, A ou B. Seja p a probabilidade de ocorrência do evento A e $q = 1 - p$ a probabilidade do evento B. Uma forma de determinar a probabilidade $P_N(m)$ de que A ocorra m vezes é considerar N destes eventos e definir N variáveis aleatórias independentes x_i , $i = 1 \dots N$ que podem valer 1 caso ocorra A ou 0 caso ocorra B. Então o problema é determinar a distribuição de probabilidades da variável:

$$m = x_1 + x_2 + \dots + x_N \quad (1.40)$$

A função característica da variável m é:

$$G(k) = \sum_m P_N(m) e^{ikm} \quad (1.41)$$

Como as variáveis x_i são independentes:

$$G(k) = [g(k)]^N \quad (1.42)$$

onde $g(k)$ é a função característica de cada uma das variáveis x_i , dada por:

$$g(k) = \langle e^{ikx_i} \rangle = p e^{ik} + q \quad (1.43)$$

Então,

$$G(k) = (p e^{ik} + q)^N \quad (1.44)$$

Usando a expansão binomial obtemos:

$$G(k) = \sum_{m=0}^N \binom{N}{m} p^m e^{ikm} q^{N-m} \quad (1.45)$$

que comparando com (1.41) da

$$P_N(m) = \binom{N}{m} p^m q^{N-m} \quad (1.46)$$

1.7 Transformação de variáveis aleatórias

Estamos interessados em saber como a distribuição de probabilidade $P(x)$ da variável aleatória x se transforma frente a uma transformação da variável, do tipo $y = f(x)$. Vamos supor $f(x)$ conhecida e cuja inversa pode não ser única. Se a transformação for monótona, ou seja, $x = h(y)$, $h = f^{-1}$, a lei de transformação para a distribuição é dada simplesmente pelo Jacobiano da transformação:

$$P(y) = P(h(y)) |dx/dy| \quad (1.47)$$

No caso geral, se $x = h_j(y)$, onde h_j é uma das transformações inversas, se pode mostrar que

$$P(y) = \sum_{j=1}^r P(h_j(y)) |dx/dy|_{x=h_j(y)} \quad (1.48)$$

no caso em que a transformação f tem r funções inversas.

Exemplo: Distribuição de Boltzmann

Seja $P(v)$ a distribuição das velocidades de uma partícula livre num ambiente a temperatura T :

$$P(v) = \sqrt{m/2\pi k_B T} \exp\left(-\frac{mv^2}{2k_B T}\right) \quad (1.49)$$

Queremos calcular a distribuição de probabilidades da energia: $P(E)$. Como $E = \frac{1}{2}mv^2 \equiv f(v)$, observamos que temos duas raízes:

$$v_{1,2}(E) = f_{1,2}^{-1}(E) = \pm\sqrt{2E/m} \quad (1.50)$$

Então $|f'(v)|_{v=v_j} = |mv_j| = +\sqrt{2mE}$. Usando este resultado e (1.49) e (1.48), podemos escrever

$$P(E) = \sum_{j=1,2} P(f_j^{-1}(E)|f'(v)|_{v_j}^{-1}) \quad (1.51)$$

e substituindo obtemos finalmente:

$$P(E) = 2\sqrt{m/2\pi k_B T} \exp\left(-\frac{E}{k_B T}\right) \frac{1}{\sqrt{2mE}} = \frac{1}{\sqrt{\pi E k_B T}} \exp\left(-\frac{E}{k_B T}\right) \quad (1.52)$$

que é a conhecida distribuição de Boltzmann das energias no equilíbrio termodinâmico.

1.8 Lei dos Grandes Números e o Teorema do Limite Central

Consideremos uma sequência de N variáveis aleatórias x_i independentes e *identicamente distribuídas*. A lei dos grandes números diz que:

$$\frac{1}{N} \sum_{i=1}^N x_i \rightarrow a \quad \text{para } N \rightarrow \infty \quad (1.53)$$

onde $a = \langle x_i \rangle$ é a média da distribuição comum das variáveis x_i . A lei dos grandes números permite interpretar uma probabilidade como a frequência

de um evento. No caso dos ensaios de Bernoulli a frequência do evento A é $f = m/N$. Por outro lado $a = \langle x_i \rangle = p$ é a probabilidade de ocorrência de A . A lei dos grandes números garante então que $f = p$ quando $N \rightarrow \infty$.

O Teorema do Limite Central (TLC) afirma que a variável aleatória z definida por:

$$z = \frac{1}{\sqrt{N\sigma^2}} \left[\sum_{i=1}^N x_i - Na \right] \quad (1.54)$$

possui uma densidade de probabilidade *gaussiana*:

$$P(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \quad (1.55)$$

no limite $N \rightarrow \infty$. As únicas condições para a validade do teorema são que existam a média a e a variância σ . O teorema do limite central não se aplica, por tanto, a uma distribuição como a de Lorentz. No entanto existe uma generalização do TLC que define famílias de variáveis aleatórias cuja soma tende a distribuições do tipo Lévy. Isto leva a considerar bacias de atração no espaço de distribuições, com distribuições atratoras, como a Gaussiana e as distribuições de Lévy.

1.9 Função geratriz

Para distribuições de probabilidade discretas, a função geratriz $G(z)$ se define como:

$$G(z) = \sum_{m=0}^{\infty} p_m z^m \quad (1.56)$$

A série converge pelo menos para $-1 \leq z \leq 1$. Derivando respeito de z e calculando as derivadas em $z = 1$ podemos obter os momentos da distribuição:

$$G'(1) = \sum_{m=1}^{\infty} m p_m = \langle m \rangle \quad (1.57)$$

e

$$G''(1) = \sum_{m=2}^{\infty} m(m-1) p_m = \langle m^2 \rangle - \langle m \rangle \quad (1.58)$$

Por exemplo, a função geratriz da distribuição binomial é:

$$G(z) = \sum_{m=0}^{\infty} \binom{N}{m} p^m q^{N-m} z^m = (pz + q)^N \quad (1.59)$$

Fazendo as derivadas e avaliando em 1 obtemos a média $\langle m \rangle = Np$ e a variância $\langle m^2 \rangle - \langle m \rangle^2 = Npq$.

1.10 Lista de exercícios

1. Considere uma caixa C_1 que tem 3 bolas azuis e 1 branca, e outra caixa C_2 com 2 bolas azuis e 1 branca. Se escolhe ao acaso uma das caixas, e em seguida se extrai uma bola dela. Qual a probabilidade de extrair uma bola azul ?
2. Considere as famílias que têm exatamente dois filhos. Dado que uma família tem um filho homem, qual a probabilidade do segundo filho ser também homem ?
3. Considere agora as famílias com três filhos. Qual a probabilidade da família ter filhos de ambos os sexos ? Qual a probabilidade de ter, no máximo, uma menina ? Qual a probabilidade de ambos eventos acontecerem simultaneamente? Que pode dizer sobre a independência estatística destes eventos ?
4. Responda as mesmas perguntas do exercício anterior no caso de famílias com dois filhos. Quais conclusões você pode tirar da comparação dos dois casos ?
5. Calcule a normalização de uma distribuição gaussiana com suporte não negativo e valor mais provável x_p .
6. Suponha que a taxa normal de infecção de uma doença seja de 25%. Para testar uma nova vacina são infectados n animais saudáveis. Qual a probabilidade de exatamente k animais ficarem livres da infecção ? Numa população de 10 animais, qual a probabilidade de todos ficarem livres ? e de 1 apenas pegar a doença ? Quais seriam as respostas se a população fosse de 100 animais ?
7. A probabilidade de que uma máquina tipográfica cometa um erro de sintaxe é baixa, mas fixa. Um livro de n páginas contém, em média, λ erros por página. Estime a probabilidade de que pelo menos uma página contenha mais de k erros.
8. Mostre que a função característica da distribuição gaussiana:

$$P(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) \quad (1.60)$$

onde $x \in [-\infty, +\infty]$ é dada por:

$$G(k) = \exp\left(ik\mu - \frac{\sigma^2}{2}k^2\right) \quad (1.61)$$

9. Mostre que a função característica da distribuição gamma:

$$P(x) = \frac{c^b}{\Gamma(b)} x^{b-1} e^{-cx}, \quad x \in [0, \infty], \{b, c\} > 0 \quad (1.62)$$

é

$$G(k) = \frac{c^b}{(c - ik)^b} \quad (1.63)$$

Note que $P(x)$ tem suporte não negativo. O caso $b = 1$ corresponde à distribuição exponencial. A partir de $G(k)$ mostre que o momento de ordem n é dado por:

$$\langle x^n \rangle = \frac{b(b+1) \cdots (b+n-1)}{c^n} \quad (1.64)$$

e a variância é dada por:

$$\sigma^2 = \langle (x - \langle x \rangle)^2 \rangle = \frac{b}{c^2} \quad (1.65)$$

10. Calcule a função característica da distribuição de Weierstrass:

$$P(s) = \frac{a-1}{2a} \sum_{n=0}^{\infty} \frac{1}{a^n} (\delta_{s, b^n} + \delta_{s, -b^n}) \quad (1.66)$$

onde $a > 1$ e $b > 1$. Qué pode dizer sobre a continuidade da função? Calcule as derivadas primeira e segunda e determine os dois primeiros momentos.

Capítulo 2

Movimento Browniano I: Caminhadas aleatórias e a equação de difusão

Uma vez assumido o caráter estatístico da descrição do sistema é muitas vezes útil pensar que o mesmo (uma partícula suspensa na superfície de um fluido ou um ponto num espaço de fases complexo) está sujeito a forças de origem determinista (as forças usuais) e outras de origem estocástica. Estas últimas surgem de fazer uma descrição fenomenológica de uma série de efeitos microscópicos complexos e cujo conhecimento em detalhe não é necessário para a análise que se quer fazer do sistema. O caso mais simples é o de uma partícula Browniana, que sofre sucessivas colisões com outras partículas num líquido, por exemplo, e descreve como consequência uma trajetória errática, chamada **caminhada aleatória**. É imediato perceber que esta partícula, como consequência das colisões irá **difundir** no espaço de uma forma muito mais lenta do que se não estivesse sujeita a tais colisões. Vamos ver que através de uma descrição estatística é possível obter uma descrição precisa e rica da difusão sem necessidade de considerar os detalhes das colisões particulares a nível microscópico.

O movimento Browniano deve seu nome ao botânico Robert Brown, que em torno de 1827 fez observações ao microscópio de partículas de polen suspensas em água. Ele notou um movimento errático e muito rápido das partículas, e suspeitou que elas fossem algum organismo vivo. Após fazer a mesma experiência com outras substâncias, inclusive inorgânicas, ele se convenceu que aquele movimento não tinha uma origem orgânica. A origem do movi-

mento Browniano ficou inexplicada até Einstein, que em um dos seus famosos trabalhos de 1905, *Concerning the motion, as required by the molecular-kinetic theory of heat, of particles suspended in liquids at rest*, introduziu os conceitos fundamentais que deram lugar a moderna teoria dos processos estocásticos e seu enorme impacto na física e outras ciências (ver os artigos em [6] de comemoração dos 100 anos do trabalho de Einstein).

2.1 Caminhada aleatória numa rede hipercúbica

Nas palavras de Einstein: *Deve ser claramente assumido que cada partícula executa um movimento que é independente do movimento de todas as outras partículas; também será considerado que o movimento de uma e a mesma partícula em diferentes intervalos de tempo são processos independentes, sempre que estes intervalos de tempo não sejam muito curtos.*

Introduzimos um intervalo de tempo τ , que é muito pequeno se comparado com os intervalos de tempo de observação, mas ao mesmo tempo suficientemente grande para que quando considerados dois intervalos de tempo τ sucessivos, o movimento executado pela partícula pode ser considerado como eventos independentes um do outro.

Vamos supor que a partícula pode ocupar os vértices de uma rede hipercúbica em d dimensões. Passa um tempo num dado vértice e de repente sofre uma colisão e *salta* a um vértice vizinho. Vamos supor ainda que as colisões são tais que só pode saltar para um vizinho imediato do vértice no qual se encontra no instante t : os movimentos são *locais*. Em d dimensões cada vértice possui $2d$ vizinhos próximos. Se h é a distância entre qualquer par de vizinhos próximos, então podemos descrever a posição da partícula no instante t pelo vetor posição:

$$\mathbf{R}(t) = h(n_1, n_2, \dots, n_d) \quad (2.1)$$

onde os $\{n_i, i = 1 \dots d\}$ são inteiros. Para descrever a evolução temporal deste vetor vamos considerar uma seqüência de saltos. Um salto do vértice i é descrito pelo vetor $\vec{\xi}_i = \pm h\mathbf{e}_\mu$, onde o \mathbf{e}_μ é o vetor unitário na direção μ . Vamos supor que as sucessivas colisões, que dão lugar aos saltos, são **eventos independentes**, ou seja, a partícula *não guarda memória* das diferentes colisões. Nestas condições temos que:

$$\begin{aligned} \langle \vec{\xi}_i \rangle &= 0 \\ \langle \vec{\xi}_i \cdot \vec{\xi}_j \rangle &= h^2 \delta_{ij} \end{aligned} \quad (2.2)$$

Nestes dois resultados está contida toda a informação matemática que necessitamos para determinar as propriedades estatísticas da trajetória da partícula. Depois de N saltos o vetor posição vem dado por:

$$\mathbf{R}_N = \sum_i^N \vec{\xi}_i \quad (2.3)$$

e, então

$$\langle \mathbf{R}_N \rangle = \sum_i^N \langle \vec{\xi}_i \rangle = 0 \quad (2.4)$$

com o qual comprovamos que, em média, a partícula não avança. Isto é devido à simetria nas direções das colisões individuais. No entanto ela vai “difundindo” no espaço como sugerido pelo valor do *desvio quadrático médio*:

$$\langle \mathbf{R}_N \cdot \mathbf{R}_N \rangle = \sum_{i,j=1}^N \langle \vec{\xi}_i \cdot \vec{\xi}_j \rangle = h^2 \sum_{i,j=1}^N \delta_{ij} = Nh^2 \quad (2.5)$$

Notamos que se τ é o intervalo de tempo entre dois saltos sucessivos, após N saltos a partícula conseguiu difundir ou explorar a vizinhança até distâncias da ordem de \sqrt{N} , ou seja, muito mais lentamente do que se as colisões acontecessem em uma direção preferencial, em cujo caso o deslocamento aumentaria proporcionalmente a N . O expoente $1/2$ num processo difussivo é característico da chamada *difusão simples*. Outros expoentes podem caracterizar processos difussivos mais complexos em cujo caso temos a chamada *difusão anômala*.

Vamos agora calcular a probabilidade de observar a partícula na posição r ao tempo t . Formalmente:

$$\rho(\vec{r}, t) = \langle \delta_{\mathbf{R}(t), \vec{r}} \rangle_{R(0)} \quad (2.6)$$

onde $\mathbf{R}(t)$ é uma variável aleatória representativa da posição da partícula no instante t e \vec{r} um ponto dado da rede d -dimensional, e os colchetes indicam uma média sobre diferentes trajetórias com a mesma condição inicial $R(0)$.

Para simplificar a análise vamos nos limitar ao caso unidimensional. Como a dinâmica progride via saltos discretos de comprimento $\pm h$, vamos ter que após um tempo $t = m\tau$ a partícula estará em $r = nh$ e vamos utilizar os índices n e m para descrever a probabilidade $\rho_{n,m}$. Ainda vamos supor que a partícula pode ir para direita com probabilidade p e para esquerda com probabilidade $q = 1 - p$. Claramente após m passos $n = \xi_1 + \xi_2 + \dots + \xi_m$,

então n é uma soma de variáveis aleatórias independentes. Redefinindo $\xi_j/h \rightarrow \xi_j = \pm 1$, a média e a variância de ξ_j são:

$$a = \langle \xi_j \rangle = p - q \quad (2.7)$$

e

$$b = \langle \xi_j^2 \rangle - \langle \xi_j \rangle^2 = 1 - (p - q)^2 = 4pq \quad (2.8)$$

A função característica da variável ξ_j é:

$$g(k) = \langle e^{ik\xi_j} \rangle = pe^{ik} + qe^{-ik} \quad (2.9)$$

A função característica do processo n será:

$$G_m(k) = [g(k)]^m = (pe^{ik} + qe^{-ik})^m \quad (2.10)$$

Fazendo a expansão binomial:

$$G_m(k) = \sum_{l=0}^m \binom{m}{l} p^l q^{m-l} e^{ik(2l-m)} \quad (2.11)$$

e comparando com

$$G_m(k) = \sum_{n=-m}^m \rho_{n,m} e^{ikn} \quad (2.12)$$

onde n pode tomar os valores $-m, -m+2, \dots, m-2, m$ vemos que:

$$\rho_{n,m} = \frac{m!}{\left(\frac{m+n}{2}\right)! \left(\frac{m-n}{2}\right)!} p^{(m+n)/2} q^{(m-n)/2} \quad (2.13)$$

que é uma distribuição binomial com média e variância dadas por:

$$\langle n \rangle = ma = m(p - q) \quad (2.14)$$

e

$$\langle n^2 \rangle - \langle n \rangle^2 = mb = 4mpq \quad (2.15)$$

No limite de um número muito grande de passos $m \gg 1$, podemos usar o teorema central do limite para obtermos:

$$\rho_{n,m} = \frac{1}{\sqrt{2\pi mb}} \exp\left\{-\frac{(n - ma)^2}{2mb}\right\} \quad (2.16)$$

No limite contínuo, a densidade de probabilidade de observar a partícula no ponto x ao tempo t é dado por $\rho(x, t) = \rho_{n,m}/h$:

$$\rho(x, t) = \frac{1}{\sqrt{2\pi Dt}} \exp\left\{-\frac{(x - ct)^2}{2Dt}\right\} \quad (2.17)$$

onde

$$c = \frac{ha}{\tau} = \frac{h(p - q)}{\tau} \quad (2.18)$$

e

$$D = \frac{h^2b}{\tau} = \frac{h^2 4pq}{\tau} \quad (2.19)$$

Podemos notar que:

$$\langle x \rangle = ct \quad (2.20)$$

e

$$\langle x^2 \rangle - \langle x \rangle^2 = Dt \quad (2.21)$$

que permitem identificar c com a velocidade média e D com o coeficiente de difusão da partícula.

2.2 A equação de difusão

Vamos ver agora que o problema da caminhada aleatória é equivalente ao da difusão de partículas ou ainda ao da difusão da probabilidade de encontrar uma partícula num dado ponto no espaço e no tempo. Como o sistema não possui memória, a probabilidade de observar uma partícula no ponto \vec{n} depois de $m + 1$ passos de tempo será basicamente a soma das probabilidades de encontrar a partícula em um dos $2d$ sítios vizinhos no instante m :

$$\rho_{\vec{n}, m+1} = \frac{1}{2d} \sum_{\xi} \rho_{\vec{n}-\xi, m} \quad (2.22)$$

Esta equação é um caso particular da famosa *Equação de Chapman-Kolmogorov* [1]. O postulado básico é a ausência de memória entre movimentos sucessivos, propriedade que também é conhecida como *Processo de Markov ou Processos Markovianos*. No próximo capítulo serão estudadas em detalhe diversas propriedades de processos markovianos, que jogam um papel fundamental na teoria dos processos estocásticos. A partir da expressão anterior é fácil ver como as probabilidades de ocupação se propagam. No primeiro passo a

probabilidade de ocupação de qualquer vizinho será $1/2d$, no segundo passo os vizinhos mais externos serão ocupados com probabilidade $(1/2d) \cdot (1/2d)$, e assim sucessivamente (ver [7]), de forma que podemos notar que após um número grande de iterações a probabilidade terá uma variação suave no tempo e no espaço de forma que poderemos transformar o processo discreto (2.22) numa equação diferencial parcial. Ao fazer esta passagem assumimos implicitamente que não estamos interessados na dinâmica discreta, microscópica, mas na evolução da probabilidade numa região “mesoscópica”. Esta região deverá ser grande respeito a distância \vec{h} entre sítios da rede, mas ainda pequena o suficiente para que a probabilidade não varie apreciavelmente no seu interior. Se assumirmos que esta região ocupa um volume Ω então:

$$\sum_{n \in \Omega} \rho_{n,m} \rightarrow \int_{\vec{r} \in \Omega} d\vec{r} \rho(\vec{r}, t) \quad (2.23)$$

Como cada ponto \vec{r} corresponde a um volume unitário h^d , da relação anterior temos que $\rho(\vec{r}, t) = h^{-d} \rho_{n,m}$. Agora se multiplicamos (2.22) por h^{-d} obtemos:

$$\rho(\vec{r}, t + \tau) - \rho(\vec{r}, t) = \sum_{\xi} \frac{\rho(\vec{r} - \xi, t) - \rho(\vec{r}, t)}{2d} \quad (2.24)$$

Até aqui o cálculo é exato. Se agora consideramos τ e ξ pequenos e expandimos em série de Taylor nas ordens mais baixas obtemos:

$$\tau \frac{\partial}{\partial t} \rho(\vec{r}, t) = \sum_{\xi} \frac{-\xi \cdot \nabla + (\xi \cdot \nabla)^2}{2d} \rho(\vec{r}, t) \quad (2.25)$$

O primeiro termo da parte da direita se anula ao fazer a soma. Lembrando que $\vec{\xi} = h\mathbf{e}_{\mu}$ obtemos:

$$\tau \frac{\partial}{\partial t} \rho(\vec{r}, t) = \frac{h^2}{2d} \nabla^2 \rho(\vec{r}, t) \quad (2.26)$$

Vemos que a probabilidade nas condições descritas, variação suave no espaço e no tempo, obedece uma equação diferencial parcial conhecida como **equação de difusão**. Em particular, o coeficiente

$$D = \frac{h^2}{d\tau} \quad (2.27)$$

é a *difusividade* ou *coeficiente de difusão* (comparar com a equação (2.19) para o caso $d = 1$).

Existem diversas formas de resolver a equação (2.26) para obter a distribuição de probabilidades. Uma possível é introduzindo transformadas de Fourier:

$$G(\vec{q}, t) = \int d\vec{r} \rho(\vec{r}, t) e^{-i\vec{q}\cdot\vec{r}} \quad (2.28)$$

$$\rho(\vec{r}, t) = \int \frac{d\vec{q}}{(2\pi)^d} G(\vec{q}, t) e^{i\vec{q}\cdot\vec{r}} \quad (2.29)$$

Temos que escolher uma condição inicial: supomos que inicialmente a partícula está localizada na origem: $\vec{r} = 0$,

$$\rho(\vec{r}, 0) = \delta(\vec{r}) \quad (2.30)$$

de forma que

$$G(\vec{q}, 0) = 1 \quad (2.31)$$

Transformando Fourier ambos lados da equação (2.26) obtemos:

$$\frac{\partial}{\partial t} G(\vec{q}, t) = -\frac{1}{2} D q^2 G(\vec{q}, t) \quad (2.32)$$

cuja solução é:

$$G(\vec{q}, t) = e^{-\frac{1}{2} D q^2 t} \quad (2.33)$$

Finalmente fazendo a anti-transformada (2.29) obtemos:

$$\rho(\vec{r}, t) = (2\pi D t)^{-d/2} \exp\left[-\frac{r^2}{2D t}\right]. \quad (2.34)$$

Podemos notar que o desvio quadrático médio é dado por:

$$\langle r^2 \rangle = D t \quad (2.35)$$

que permite interpretar o coeficiente de difusão como o desvio quadrático médio por unidade de tempo.

A lei de Fick

Uma forma clássica e mais fenomenológica de chegar à equação de difusão é via um argumento puramente contínuo, hidrodinâmico [8]. Consideremos, por simplicidade, que a probabilidade corresponda a densidade de partículas

$\rho(\vec{r}, t) \rightarrow n(\vec{r}, t)$. Se o número de partículas no sistema é conservado, ele deve obedecer uma equação de continuidade do tipo:

$$\frac{\partial n(\vec{r}, t)}{\partial t} + \nabla \cdot \vec{j}(\vec{r}, t) = 0 \quad (2.36)$$

onde

$$\vec{j}(\vec{r}, t) = \sum_i \vec{v}_i(t) \delta(\vec{r} - \vec{r}_i(t)) \quad (2.37)$$

é a corrente de partículas, com $\vec{v}_i(t)$ a velocidade da partícula i . No equilíbrio termodinâmico as partículas se distribuem uniformemente no espaço, de forma que a densidade média $\langle n(\vec{r}, t) \rangle$ não depende do espaço nem do tempo. Mas se, em algum instante, por causa de alguma flutuação ou por uma força externa, a densidade se tornar inhomogênea, quando a causa da inhomogeneidade passar, o sistema deverá tentar voltar ao estado de equilíbrio homogêneo. Esse processo será governado pela presença de correntes de partículas que tenderão a restabelecer a homogeneidade espacial. Se a variação espacial da densidade não for muito grande, podemos assumir que as correntes serão proporcionais aos gradientes de partículas:

$$\vec{j}(\vec{r}, t) = -D \vec{\nabla} n. \quad (2.38)$$

Esta equação é conhecida como *lei de Fick*. É uma relação fenomenológica que nos diz que são os gradientes os que dão origem as correntes. O coeficiente de difusão também pode ser definido através da lei de Fick. Quando substituirmos a lei de Fick na equação de continuidade, obtemos a equação de difusão para a distribuição espacial e temporal da densidade de partículas.

Capítulo 3

Processos Markovianos e a Equação Mestra

Neste capítulo vamos descrever os conceitos fundamentais dos Processos Markovianos, como calcular a evolução temporal da distribuição de probabilidades e determinar se a mesma converge a um valor estacionário para tempos grandes (equilíbrio estatístico). Começando a descrição com processos em tempo discreto, numa segunda parte vamos introduzir um limite de tempo contínuo, conhecido como Equação Mestra. Além da bibliografia já introduzida nos capítulos anteriores, irão ser úteis como referências bibliográficas dos assuntos deste capítulo o clássico livro de van Kampen [9] e o livro de Tomé e de Oliveira [10].

3.1 Processos Markovianos

Consideremos um processo estocástico u_t no qual a variável estocástica possa tomar um conjunto discreto de valores e o tempo também seja discretizado. Um processo com estas características fica definido pela distribuição de probabilidades conjunta

$$P_l(n_0, n_1, n_2, \dots, n_l) \quad (3.1)$$

de que u_t tome o valor n_0 no instante $t = 0$, n_1 no instante $t = 1$ e assim por diante até n_l no instante $t = l$. Consideremos agora a probabilidade condicional:

$$P_{l+1}(n_{l+1}|n_0, n_1, \dots, n_l) \quad (3.2)$$

de que a variável estocástica u_t tome o valor n_{l+1} no instante $t = l + 1$ **dado que** ela tenha tomado os valores n_0 em $t = 0$, n_1 em $t = 1$ e assim por diante até n_l em $t = l$. Se esta probabilidade for igual à probabilidade condicional

$$P_{l+1}(n_{l+1}|n_l) \quad (3.3)$$

ou seja, se ela só depende do estado do sistema no instante imediatamente anterior, então o processo estocástico se chama markoviano ou *Processo de Markov*. Então para um processo markoviano a probabilidade conjunta (3.1) se pode escrever na forma:

$$\begin{aligned} P_l(n_0, n_1, n_2, \dots, n_l) &\equiv P_0(n_0)P_1(n_1|n_0)P_2(n_2|n_0, n_1) \dots P_l(n_l|n_0, n_1, \dots, n_{l-1}) \\ &= P_0(n_0)P_1(n_1|n_0)P_2(n_2|n_1) \dots P_l(n_l|n_{l-1}) \end{aligned} \quad (3.4)$$

A probabilidade (marginal) de que a variável estocástica tome o valor n_l em $t = l$ independentemente dos valores tomados em tempos anteriores é dada por:

$$P_l(n_l) = \sum_{n_0, n_1, \dots, n_{l-1}} P_l(n_0, n_1, n_2, \dots, n_l) \quad (3.5)$$

Utilizando a relação (3.4) para um processo markoviano obtemos a relação de recorrência:

$$P_l(n_l) = \sum_{n_{l-1}} P_l(n_l|n_{l-1})P_{l-1}(n_{l-1}) \quad (3.6)$$

As probabilidades condicionais $P_l(n_l|n_{l-1})$ podem ser interpretadas como *probabilidades de transição* do estado n_{l-1} para o estado n_l no instante $l - 1$. Nós vamos nos limitar a considerar casos nos quais a probabilidade condicional ou de transição entre dois estados quaisquer independe do instante particular t . Nesse caso escrevemos:

$$P_l(n_l|n_{l-1}) = T(n_l, n_{l-1}) \quad (3.7)$$

de forma que a recorrência (3.6) se escreve:

$$P_l(n_l) = \sum_{n_{l-1}} T(n_l, n_{l-1})P_{l-1}(n_{l-1}) \quad (3.8)$$

3.2 Matriz Estocástica

As probabilidades de transição e a condição inicial determinam completamente um processo estocástico markoviano. Por tanto é do maior interesse

conhecer as propriedades da probabilidade de transição que pode ser escrita em forma de matriz $T(n, m)$ tal que:

$$P_l(n) = \sum_m T(n, m) P_{l-1}(m) \quad (3.9)$$

onde $T(n, m)$ é chamada **matriz estocástica**. Uma matriz estocástica possui as seguintes propriedades:

- $T(n, m)$ é uma matriz não negativa, ou seja, $T(n, m) \geq 0 \forall n, m$, já que T é uma probabilidade condicional.
- A soma dos elementos de cada coluna é igual a um: $\sum_n T(n, m) = 1$

A segunda propriedade corresponde a normalização das probabilidades. Se consideramos P_l como sendo uma matriz coluna com elementos iguais a $P_l(n)$ podemos escrever (3.9) na forma:

$$P_l = T P_{l-1} \quad (3.10)$$

Dadas as propriedades de um processo markoviano obtemos

$$P_l = T^l P_0 \quad (3.11)$$

Aqui vemos claramente como a matriz estocástica determina completamente o processo. Dada uma condição inicial P_0 o problema de obter as probabilidades P_l ao tempo t se reduz a conhecer a potência l -ésima da matriz estocástica. A relação anterior pode ser escrita em termos dos elementos das matrizes:

$$P_l(n) = \sum_m T^l(n, m) P_0(m) \quad (3.12)$$

Aqui interpretamos o elemento de matriz $T^l(n, m)$ como a probabilidade de transição do estado m para o estado n depois de l passos, ou em outras palavras, como a probabilidade da variável u_t tomar o valor n no instante t sendo que tomou o valor m no instante $t - l$. Notar que os dois estados podem ter sido conectados dinamicamente por diferentes caminhos.

3.2.1 Algumas propriedades de matrizes não negativas

Um problema fundamental é descobrir quais propriedades deve satisfazer a matriz estocástica de forma que as probabilidades atinjam a solução assintótica desejada, isto é, para que

$$\lim_{t \rightarrow \infty} P_t = P \quad (3.13)$$

onde P é a distribuição assintótica ou estacionária que satisfaz

$$TP = P, \quad (3.14)$$

e saber quando esta solução é única. Vamos agora enunciar uma série de definições e resultados importantes de matrizes estocásticas :

- Uma matriz quadrada $T(n, n)$ se diz *reduzível* se o conjunto de índices $1, 2, \dots, n$ pode ser separada em dois conjuntos complementares (sem índices comuns), $i_1, i_2, \dots, i_\mu; k_1, k_2, \dots, k_\nu$ com $(\mu + \nu = n)$, tais que:

$$T(i_\alpha, k_\beta) = 0, \quad (\alpha = 1, 2, \dots, \mu; \beta = 1, 2, \dots, \nu) \quad (3.15)$$

Caso contrário a matriz se diz *irreduzível*.

- Se uma matriz T é *irreduzível* então qualquer par de estados n, m podem ser alcançados com probabilidade finita, ou seja, $\forall n, m$ existe um inteiro positivo l tal que $T^l(n, m) > 0$. Em física dizemos que um sistema que satisfaz esta propriedade é **ergódico**.
- *Teorema de Perron-Frobenius:*

Teorema 1 *Uma matriz não negativa, irreduzível T sempre tem um autovalor real positivo r que é raiz simples da equação característica, ou seja, é não degenerado. Os módulos de todos os outros autovalores são menores ou iguais a r . Ao autovalor r corresponde um autovetor com componentes positivas. Se T possui k autovalores $\lambda_0 = r, \lambda_1, \dots, \lambda_{k-1}$ de módulo igual a r , então estes autovalores são todos diferentes e são as raízes da equação característica $\lambda^k - r^k = 0$.*

Para a demonstração do teorema ver, por exemplo [11]. Se a matriz T é uma matriz estocástica se pode mostrar que o autovalor máximo corresponde a $r = 1$.

- *Definição:*
Se uma matriz não negativa, irreduzível, estocástica, T possui k autovalores $\lambda_1, \lambda_2, \dots, \lambda_k$ com módulo máximo $r = 1$ ($\lambda_1 = |\lambda_2| = \dots = |\lambda_k| = 1$), então T se chama *regular* se $k = 1$ e não regular se $k > 1$.

Teorema 2 *Uma matriz estocástica T é regular se e somente se alguma potência l de T é estritamente positiva quaisquer sejam n e m [11]:*

$$T^l(n, m) > 0 \quad \forall n, m \quad (3.16)$$

Se T é regular se pode mostrar que quando $l \rightarrow \infty$, T^l converge a uma matriz com todas as colunas iguais a P . Então como P_0 está normalizado $P_l = T^l P_0 \rightarrow P$ independente de P_0 .

- Para que $\lim_{l \rightarrow \infty} P_l = P$ exista é necessário que não haja nenhum autovalor complexo sobre o círculo unitário, exceto $\lambda = 1$.
- Para que o limite seja independente da probabilidade inicial o autovalor $\lambda = 1$ deve ser não degenerado, ou seja, a probabilidade estacionária deve ser única.
- Se a matriz T é regular estas propriedades se verificam. No entanto podem existir matrizes não regulares com estas propriedades, por exemplo um sistema com um estado absorvente [10].

3.2.2 Método algébrico

Vamos agora ver um método para determinar a matriz T^l e que por tanto permite calcular as probabilidades de transição entre qualquer par de estados para qualquer instante num processo markoviano. O método algébrico consiste em determinar os autovalores e autovetores da matriz T . A matriz T geralmente será não simétrica já que as probabilidades de transição entre dois estados n e m em geral vão depender da direção. Por tanto a matriz poderá ter autovalores complexos. Vamos considerar o caso em que os autovalores sejam não degenerados e o número de estados finito. Como consequência da falta de simetria temos que considerar autovetores à esquerda e à direita diferentes. Sejam $\{\Psi_k\}$ e $\{\phi_k\}$ os autovetores à direita e à esquerda respectivamente e $\{\lambda_k\}$ os correspondentes autovalores. Temos

$$T\Psi_k = \lambda_k\Psi_k \quad (3.17)$$

e

$$\phi_k T = \lambda_k\phi_k \quad (3.18)$$

É fácil verificar a partir destas equações que os autovetores formam um conjunto completo ortogonal e é possível normalizá-lo, de forma que:

$$\sum_n \phi_j(n) \Psi_k(n) = \delta_{jk} \quad (3.19)$$

e

$$\sum_k \Psi_k(n) \phi_k(m) = I(n, m) \quad (3.20)$$

sendo I a matriz identidade. Notar que os Ψ 's são matrizes coluna enquanto que os ϕ 's são matrizes linha. Se definimos:

$$H = \left(\Psi_1, \Psi_2, \dots, \Psi_n \right) \quad (3.21)$$

e

$$H^{-1} = \begin{pmatrix} \phi_1 \\ \phi_2 \\ \cdot \\ \cdot \\ \cdot \\ \phi_n \end{pmatrix} \quad (3.22)$$

as propriedades de ortonormalidade implicam que H e H^{-1} são matrizes inversas: $HH^{-1} = I$. Então podemos escrever:

$$T = H\Lambda H^{-1} \quad (3.23)$$

onde Λ é uma matriz diagonal cujos elementos da diagonal são os autovalores de T :

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \cdot & \cdot & \dots & \cdot \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} \quad (3.24)$$

Agora a partir de (3.23) é fácil ver que:

$$T^l = H\Lambda^l H^{-1} \quad (3.25)$$

onde Λ^l é uma matriz diagonal com os elementos da diagonal iguais a $\lambda_1^l, \lambda_2^l, \dots, \lambda_n^l$. A partir destas expressões é fácil determinar as probabilidades $T^l(n, m)$. Os elementos de matriz são dados por:

$$T^l(n, m) = \sum_k \Psi_k(n) \lambda_k^l \phi_k(m) \quad (3.26)$$

Vamos supor agora que $\lim_{l \rightarrow \infty} P_l = P$ exista e seja único (por exemplo se T é regular). A solução estacionária P é, por definição, um autovetor com autovalor um. Seja $\Psi_1 = P$. Devido à condição de ortonormalidade o correspondente autovetor à esquerda ϕ_1 será a matriz linha com todos os elementos iguais a um, ou seja, $\phi_1(n) = 1$ para qualquer n . Consideremos agora a potência l -ésima de T . Podemos escrevê-la em termos dos autovetores e autovalores na forma:

$$T^l = T^l I = T^l \sum_k \Psi_k \phi_k = \sum_k T^l \Psi_k \phi_k = \sum_k \lambda_k^l \Psi_k \phi_k \quad (3.27)$$

onde usamos mais uma vez a condição de ortonormalização para escrever a matriz identidade. Agora podemos escrever:

$$\begin{aligned} P_l &= T^l P_0 = \sum_k \lambda_k^l \Psi_k (\phi_k P_0) \\ &= P + \sum_{k \neq 1} \lambda_k^l \Psi_k (\phi_k P_0) \end{aligned} \quad (3.28)$$

onde usamos os resultados $\lambda_1^l = 1$, $\Psi_1 = P$ e $\phi_1 P_0 = 1$ porque P_0 está normalizado. Como $\lim_{l \rightarrow \infty} P_l = P$ existe e é único, os autovalores satisfazem $|\lambda_k| < 1$ para $k \neq 1$, então $|\lambda_k|^l \rightarrow 0$ quando $l \rightarrow \infty$. Aqui vemos explicitamente que $P_l \rightarrow P$ quando $l \rightarrow \infty$ para qualquer condição inicial P_0 . Isto mostra que, nestas condições, o estado estacionário P é um atrator da dinâmica definida pela matriz estocástica.

3.2.3 Sistema de dois estados

Consideremos a seguinte matriz estocástica:

$$T = \begin{pmatrix} 1-p & q \\ p & 1-q \end{pmatrix} \quad (3.29)$$

Ela pode representar as probabilidades de transição num sistema que pode estar em um dentre dois estados possíveis, A e B . Supomos que conhecemos as probabilidades de transição: de A para B é $T(2,1) = p$ e de B para A é $T(1,2) = q$. Notar que as colunas estão normalizadas com as probabilidades do sistema ficar no próprio estado $T(1,1) = 1-p$ e $T(2,2) = 1-q$. Os autovalores são dados pela equação característica:

$$|T - \lambda I| = \begin{vmatrix} 1-p-\lambda & q \\ p & 1-q-\lambda \end{vmatrix} = (1-p-\lambda)(1-q-\lambda) - pq = 0 \quad (3.30)$$

ou

$$\lambda^2 - (2 - p - q)\lambda + (1 - p)(1 - q) - pq = 0 \quad (3.31)$$

Resolvendo obtemos:

$$\lambda_{1,2} = \frac{(2 - p - q) \pm \sqrt{(2 - p - q)^2 - 4[(1 - p)(1 - q) - pq]}}{2}. \quad (3.32)$$

Então:

$$\lambda_1 = 1 \quad (3.33)$$

e

$$\lambda_2 = 1 - p - q \quad (3.34)$$

Das equações para os autovetores (3.17) e (3.18) obtemos $p\Psi_1(1) = q\Psi_1(2)$, $\phi_1(1) = \phi_1(2)$, $\Psi_2(1) = -\Psi_2(2)$ e $p\phi_2(2) = -q\phi_2(1)$, que junto com as condições de normalização dão:

$$\Psi_1 = \frac{1}{p+q} \begin{pmatrix} q \\ p \end{pmatrix} \quad \phi_1 = \begin{pmatrix} 1 & 1 \end{pmatrix} \quad (3.35)$$

$$\Psi_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad \phi_2 = \frac{1}{p+q} \begin{pmatrix} p & -q \end{pmatrix} \quad (3.36)$$

Agora podemos calcular as probabilidades de transição T^l a partir da (3.26) e obtemos:

$$T^l = \frac{1}{p+q} \begin{pmatrix} q & q \\ p & p \end{pmatrix} + \frac{(1-p-q)^l}{p+q} \begin{pmatrix} p & -q \\ -p & q \end{pmatrix} \quad (3.37)$$

Dada uma probabilidade inicial:

$$P_0 = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} \quad (3.38)$$

podemos calcular a probabilidade em qualquer instante posterior:

$$P_l = T^l P_0 = \frac{1}{p+q} \begin{pmatrix} q \\ p \end{pmatrix} + \frac{(1-p-q)^l}{p+q} \begin{pmatrix} pp_1 - qp_2 \\ -pp_1 + qp_2 \end{pmatrix} \quad (3.39)$$

Se $p+q \neq 0$ vemos que no limite $l \rightarrow \infty$ a probabilidade P_l se aproxima da probabilidade estacionária Ψ_1 independentemente da condição inicial.

3.2.4 O modelo de urnas dos Ehrenfest

P. e T. Ehrenfest descreveram uma experiência imaginária na qual N moléculas são distribuídas entre duas caixas A e B. As moléculas são distinguíveis e em cada instante uma molécula é sorteada e transferida da caixa na qual se encontra para a outra. Como é a evolução temporal das probabilidades de ocupação das caixas? O sistema atinge uma distribuição estacionária? Em caso afirmativo, qual é a forma desta distribuição? Quais são as probabilidades de, no instante l ter n moléculas na caixa A e $N - n$ na caixa B?

Seja $P_l(n)$ a probabilidade de haver n moléculas na caixa A no instante l . Suponha ainda que neste instante a caixa A possua exatamente n moléculas. A probabilidade de ter $n - 1$ no instante $l + 1$ é igual a probabilidade de escolher uma molécula da caixa A (que será transferida à caixa B). Esta probabilidade é n/N . De forma equivalente a probabilidade de ter $n + 1$ moléculas em A no instante $l + 1$ será igual a probabilidade de escolher uma molécula da caixa B, isto é, $(N - n)/N$. As probabilidades de transição serão então:

$$\begin{aligned} T(n - 1, n) &= \frac{n}{N}, & n = 1, 2, \dots, N \\ T(n + 1, n) &= \frac{N - n}{N}, & n = 0, 1, \dots, N - 1 \end{aligned} \quad (3.40)$$

Os outros elementos de matriz são zero. Note que T é uma matriz $(N + 1) \times (N + 1)$ tridiagonal. Se permitirmos que o número de moléculas em um determinado instante não se altere, então:

$$\begin{aligned} T(n - 1, n) &= q \frac{n}{N}, & n = 1, 2, \dots, N \\ T(n, n) &= p, & n = 0, 1, 2, \dots, N \\ T(n + 1, n) &= q \frac{N - n}{N}, & n = 0, 1, \dots, N - 1 \end{aligned} \quad (3.41)$$

sendo p a probabilidade da molécula permanecer na caixa na qual se encontra e $q = 1 - p$ a probabilidade de mudar de caixa.

A equação para a evolução temporal das probabilidades é:

$$P_{l+1}(n) = \sum_{m=0}^N T(n, m) P_l(m) \quad (3.42)$$

que neste caso se reduz a:

$$P_{l+1}(n) = q \left(1 - \frac{n-1}{N}\right) P_l(n-1) + pP_l(n) + q \left(\frac{n+1}{N}\right) P_l(n+1), \quad (3.43)$$

válida para $n = 0, 1, 2, \dots, N$ com as condições de contorno $P_l(N+1) = P_l(-1) = 0$. A probabilidade estacionária $P(n)$ satisfaz:

$$P(n) = \left(1 - \frac{n-1}{N}\right) P(n-1) + \left(\frac{n+1}{N}\right) P(n+1). \quad (3.44)$$

cuja solução é

$$P(n) = 2^{-N} \binom{N}{n} \quad (3.45)$$

Vemos na forma desta solução que a distribuição estacionária equivale a distribuir as N moléculas aleatoriamente entre as duas caixas. Vamos ver ainda que este resultado independe da condição inicial. Podemos calcular a evolução no tempo dos momentos, por exemplo, o número médio de moléculas na caixa A em função de t (1) será:

$$\langle n \rangle_l = \sum_{n=0}^N n P_l(n) \quad (3.46)$$

Substituindo o resultado (3.43) na equação anterior obtemos uma recorrência para $\langle n \rangle_l$:

$$\langle n \rangle_{l+1} = \langle n \rangle_l + q \left(1 - \frac{2}{N} \langle n \rangle_l\right) \quad (3.47)$$

Propondo uma solução da forma $\langle n \rangle_l = ar^l + b$ com a condição inicial $\langle n \rangle_0 = N$ de que todas as N moléculas estejam na caixa A em $l = 0$ obtemos:

$$\langle n \rangle_l = \frac{N}{2} + \frac{N}{2} \left(1 - \frac{2}{N}q\right)^l \quad (3.48)$$

onde vemos que o número médio de moléculas tende a $N/2$ como é de se esperar fisicamente. O segundo momento é dado por:

$$\langle n^2 \rangle_l = \sum_{n=0}^N n^2 P_l(n) \quad (3.49)$$

Substituindo (3.43) obtemos:

$$\langle n^2 \rangle_{l+1} = \left(1 - \frac{4}{N}q\right) \langle n^2 \rangle_l + 2q \langle n \rangle_l + q \quad (3.50)$$

que pode ser resolvida da mesma forma que o caso anterior com a condição inicial $\langle n^2 \rangle_0 = N^2$ para obter:

$$\langle n^2 \rangle_l = \frac{N^2}{4} + \frac{N}{4} + \frac{N^2}{2} \left(1 - \frac{2}{N}q\right)^l + \left(\frac{N^2}{4} - \frac{N}{4}\right) \left(1 - \frac{4q}{N}\right)^l. \quad (3.51)$$

E com estes resultados obtemos a variância:

$$\langle n^2 \rangle_l - \langle n \rangle_l^2 = \frac{N}{4} - \frac{N}{4} \left(1 - \frac{2}{N}q\right)^l + \left(\frac{N^2}{4} - \frac{N}{4}\right) \left(1 - \frac{4q}{N}\right)^l. \quad (3.52)$$

que se aproxima assintoticamente ao valor $N/4$.

Vamos agora resolver o modelo de Ehrenfest com o método algébrico, ou seja, analisando a estrutura de autovalores e autovetores da matriz estocástica. Neste caso a matriz T tem a forma (ver (3.41)):

$$T = \begin{pmatrix} p & q/N & 0 & 0 & 0 & \cdots & 0 \\ q & p & 2q/N & 0 & 0 & \cdots & 0 \\ 0 & q(1-1/N) & p & 3q/N & 0 & \cdots & 0 \\ 0 & 0 & q(1-2/N) & p & 4q/N & \cdots & 0 \\ \vdots & \cdots & & & \cdots & & \vdots \\ \vdots & \cdots & & & & & \vdots \\ 0 & \cdots & & & & p & q \\ 0 & \cdots & & & & q/N & p \end{pmatrix} \quad (3.53)$$

Agora poderíamos proceder com no exemplo do sistema de dois estados, diagonalizar a matriz, obter os autovalores e autovetores e reconstruir as probabilidades. No entanto, como a matriz é maior e mais complexa vamos seguir outro possível caminho. Escrevamos a equação para os autovetores a direita:

$$\sum_{m=0}^N T(n, m) \Psi(m) = \lambda \Psi(n) \quad (3.54)$$

Desenvolvendo a soma esta equação equivale a (ver (3.43)):

$$q \left(1 - \frac{n-1}{N}\right) \Psi(n-1) + p \Psi(n) + q \left(\frac{n+1}{N}\right) \Psi(n+1) = \lambda \Psi(n) \quad (3.55)$$

que vale para $n = 0, 1, 2, \dots, N$ com as condições $\Psi(-1) = \Psi(N+1) = 0$. Agora definimos uma função geratriz:

$$f(x) = \sum_{n=0}^N \Psi(n) x^n \quad (3.56)$$

Seguidamente multiplicamos (3.55) por x^n e somamos em n . Vamos ter t erminos da forma:

$$\begin{aligned} \sum_0^N (n-1)\Psi(n-1)x^n &= x^2 \frac{d}{dx} \sum_0^N \Psi(n-1)x^{n-1} \\ &= x^2 \left(\frac{df}{dx} - N\Psi(N)x^{N-1} \right), \end{aligned} \quad (3.57)$$

$$\sum_0^N (n+1)\Psi(n+1)x^n = \frac{d}{dx} \left(\sum_0^N \Psi(n+1)x^{n+1} \right) = \frac{df}{dx} \quad (3.58)$$

e

$$\sum_0^N \Psi(n-1)x^n = x \sum_0^N \Psi(n-1)x^{n-1} = x(f - \Psi(N)x^N), \quad (3.59)$$

onde fizemos uso das condi oes de contorno. Juntando estes resultados parciais obtemos uma equa o diferencial para $f(x)$:

$$q(1-x^2) \frac{df}{dx} = N(\lambda - p - qx)f \quad (3.60)$$

que   integr vel e da como resultado:

$$f(x) = C(1+x)^{N-k}(1-x)^k \quad (3.61)$$

com C uma constante de integra o e $k = -N(\lambda - 1)/2q$. Da defini o de f pode-se concluir que k dever  ser um n mero inteiro e $0 \leq k \leq N$. Estes valores determinar o os autovalores:

$$\lambda_k = 1 - \frac{2q}{N}k \quad k = 0, 1, 2, \dots, N. \quad (3.62)$$

Ent o a fun o geratriz fica parametrizada pelo n mero k , e para cada k teremos uma expans o da forma:

$$f_k(x) = \sum_{n=0}^N \Psi_k(n)x^n \quad (3.63)$$

sendo $\Psi_k(n)$ o autovetor correspondente ao autovalor λ_k . Vemos ent o que os autovetores   direita s o os coeficientes da expans o da fun o geratriz.

Se identificamos o autovetor Ψ_0 com a probabilidade estacionária, correspondendo ao autovalor $\lambda_0 = 1$, da definição de f e da normalização da probabilidade obtemos o valor da constante $C = 2^{-N}$. Desta forma:

$$f_k(x) = 2^{-N}(1+x)^{N-k}(1-x)^k \quad (3.64)$$

Em particular temos que $f_0(x) = 2^{-N}(1+x)^N$. Expandindo o binômio obtemos o resultado para a probabilidade estacionária (3.45). Utilizando as condições de ortonormalização entre os autovetores a esquerda e direita é possível mostrar que:

$$\phi_k(n) = 2^N \Psi_n(k) \quad (3.65)$$

Se escolhemos como condição inicial que todas as moléculas estejam na caixa A, então $P_0(n) = \delta_{n,N}$ e:

$$P_l(n) = \sum_{m=0}^N T^l(n, m) P_0(m) = T^l(n, N) = \sum_{k=0}^N \lambda_k^l \Psi_k(n) \phi_k(N) \quad (3.66)$$

Para calcular os momentos da distribuição determinamos primeiro os momentos da variável x da função geratriz:

$$\langle x^n \rangle_l = \sum_{n=0}^N x^n P_l(n) = \sum_{k=0}^N \lambda_k^l f_k(x) \phi_k(N) \quad (3.67)$$

De (3.63), (3.64) e (3.65) obtemos:

$$\phi_k(N) = (-1)^k \binom{N}{k} \quad (3.68)$$

e:

$$\langle x^n \rangle_l = 2^{-N}(1+x)^N + \sum_{k=1}^N \lambda_k^l (-1)^k \binom{N}{k} 2^{-N}(1+x)^{N-k}(1-x)^k \quad (3.69)$$

Finalmente derivando respeito de x e fazendo $x = 1$ obtemos:

$$\langle n \rangle_l = \frac{N}{2} + \frac{N}{2} \lambda_1^l = \frac{N}{2} + \frac{N}{2} \left(1 - \frac{2}{N} q\right)^l \quad (3.70)$$

que é o mesmo resultado que obtivemos antes com outro método. Da mesma forma podemos obter os outros momentos.

Para concluir com a análise do modelo de Ehrenfest notamos que se associamos o estado n com uma posição em uma dimensão $x = n$, é possível interpretar o problema com uma caminhada aleatória unidimensional na qual as probabilidades de saltar para sítios vizinhos dependem da posição da partícula. Temos em total N sítios e a probabilidade de ir para a direita ou esquerda depende de que a coordenada seja $x < N/2$ ou $x > N/2$ respectivamente. Por tanto a caminhada aleatória é tal que tende a levar sempre a partícula para o ponto médio $x = N/2$ e então o problema pode ser interpretado como uma caminhada aleatória com uma força elástica que cresce proporcionalmente com a distância ao centro, que é o ponto de equilíbrio estacionário.

3.2.5 A caminhada aleatória, tempo de primeira passagem, probabilidade de recorrência

Vamos agora voltar a analisar o processo da caminhada aleatória com as técnicas da matriz estocástica aplicada a processos markovianos. Consideremos inicialmente a caminhada unidimensional generalizada para o caso em que o caminhante possa permanecer num determinado sítio com probabilidade p . Neste caso os elementos não nulos da matriz estocástica são:

$$\begin{aligned} T(n, n-1) = T(n, n+1) &= \frac{1}{2}q \\ T(n, n) &= p \end{aligned} \quad (3.71)$$

onde $q = 1 - p$ e $n = 0, 1, 2, \dots, N-1$. Consideremos uma condição inicial com a partícula na origem, ou seja, $P_0(n) = \delta_{n0}$. A equação para a evolução temporal das probabilidades é:

$$\begin{aligned} P_{l+1}(n) &= \sum_{m=0}^{N-1} T(n, m)P_l(m) \\ &= \frac{1}{2}qP_l(n+1) + \frac{1}{2}qP_l(n-1) + pP_l(n) \end{aligned} \quad (3.72)$$

Consideramos condições periódicas de contorno $P_l(n+N) = P_l(n)$. Notamos que a matriz T é tridiagonal com os elementos das diagonais sendo todos iguais. Matrizes com estas características se chamam de *Toeplitz*. Formalmente, uma matriz de Toeplitz satisfaz:

$$T(n, m) = f(n - m) \quad (3.73)$$

com $f(n)$ uma função periódica $f(n+N) = f(n)$. A $f(n) \geq 0$ e $\sum_n f(n) = 1$. No caso da caminhada aleatória unidimensional a função f é $f(1) = f(N-1) = q/2$, $f(0) = p$ e zero para outros valores do argumento. Em termos de f a probabilidade pode ser escrita como:

$$P_{l+1}(n) = \sum_m f(n-m)P_l(m) \quad (3.74)$$

Para calcular as probabilidades é útil definir neste caso a correspondente função característica:

$$G_l(k) = \sum_n P_l(n)e^{ikn} \quad (3.75)$$

A equação (3.74) representa uma convolução e então:

$$G_{l+1}(k) = F(k)G_l(k) \quad (3.76)$$

onde $F(k)$ é a função característica da $f(n)$:

$$F(k) = \sum_n f(n)e^{ikn} \quad (3.77)$$

A partir destas definições obtemos:

$$G_l(k) = F^l(k)G_0(k) \quad (3.78)$$

Para a condição inicial escolhida $G_0(k) = 1$ e então $G_l(k) = F^l(k)$. Para o caso da caminhada aleatória e usando a periodicidade da função f obtemos $F(k) = p + q \cos k$ e então:

$$G_l(k) = (p + q \cos k)^l. \quad (3.79)$$

A partir da definição da função característica podemos identificar as probabilidades $P_l(n)$ como os coeficientes da expansão do binômio (3.79).

Outra forma de obter as probabilidades é pelo método algébrico, calculando os autovalores e autovetores da matriz estocástica. Se pode mostrar que toda matriz de Toplitz possui o seguinte conjunto de autovetores:

$$\psi_k(n) = \frac{1}{N}e^{ikn} \quad k = \frac{2\pi}{N}j, \quad j = 0, 1, 2, \dots, N-1 \quad (3.80)$$

É fácil verificar que estas funções satisfazem a equação de autovalores

$$\sum_m T(n, m)\psi_k(m) = \lambda_k\psi_k(n) \quad (3.81)$$

com os autovalores dados por:

$$\lambda_k = \sum_n f(n)e^{ikn} = F(k) \quad (3.82)$$

e neste caso $\lambda_k = p + q \cos k$. Os autovetores a esquerda são $\phi_k(n) = e^{-ikn}$. De acordo a estes resultados obtemos:

$$\begin{aligned} P_l(n) &= \sum_m T^l(n, m)P_0(m) = T^l(n, 0) \\ &= \sum_k \lambda_k^l \psi_k(n) \phi_k(0) = \frac{1}{N} \sum_k \lambda_k^l e^{ikn} \end{aligned} \quad (3.83)$$

Substituindo o resultado para os autovalores obtemos:

$$P_l(n) = \frac{1}{N} \sum_k (p + q \cos k)^l e^{ikn} \quad (3.84)$$

Destes resultados (ver (3.80)) podemos concluir que **a caminhada aleatória terá uma distribuição estacionária somente no caso no qual o número de estados N seja finito**. Este fato se torna evidente se analisarmos o problema da **recorrência** ou a probabilidade da partícula voltar a uma posição pela qual já passou num instante anterior. Seja $R_l(n, m)$ a probabilidade da partícula atingir *pela primeira vez* o estado n após l passos tendo começado no estado m . O tempo correspondente l é chamado de *tempo da primeira passagem* (first passage time). A probabilidade da primeira recorrência está relacionada com a matriz de transição pela expressão:

$$T^l(n, m) = \sum_{j=1}^l R_j(n, m) T^{l-j}(n, n) \quad (3.85)$$

válida para $l \geq 1$ sendo $T^l(n, m)$ a probabilidade de passar do estado m para o n após l passos com $T^0(n, m) = \delta(n, m)$. A equação da recorrência (3.85) tem a seguinte interpretação: podemos supor que existem l formas de atingir o estado n após l passos começando pelo m dependendo do tempo que levou para atingir n pela primeira vez. Este último tempo é rotulado pelo índice j (ver equação). Então dado j a probabilidade $T^l(n, m)$ é dada pelo produto de $R_j(n, m)$ (probabilidade de atingir n pela primeira vez após j passos) com a probabilidade de voltar a n após os $l-j$ passos restantes dada por $T^{l-j}(n, n)$. Como $1 \leq j \leq l$ temos que somar em todas estas possibilidades para obter a probabilidade final (3.85).

Vamos calcular a probabilidade de recorrência e o tempo médio correspondente para a caminhada aleatória. Para isso definimos as seguintes funções geratrizes:

$$\begin{aligned} g(n, m, z) &= \sum_{l=1}^{\infty} T^l(n, m) z^l + \delta(n, m) \\ h(n, m, z) &= \sum_{l=1}^{\infty} R_l(n, m) z^l \end{aligned} \quad (3.86)$$

As funções anteriores têm o seguinte significado: a $g(n, m, 1)$ é o tempo médio que o sistema passa no estado n partindo do estado m e a $h(n, m, 1)$ é a probabilidade de o sistema passar ao menos uma vez pelo estado n partindo do m . Fazendo uso da (3.85) e das definições anteriores, e após inverter as somas ($\sum_{l=1}^{\infty} \sum_{j=1}^l \rightarrow \sum_{j=1}^{\infty} \sum_{l=j}^{\infty}$), obtemos:

$$g(n, m, z) = \delta(n, m) + h(n, m, z)g(n, n, z) \quad (3.87)$$

Para $n = m$:

$$h(n, n, z) = 1 - \frac{1}{g(n, n, z)} \quad (3.88)$$

Definimos a *probabilidade de recorrência do estado n* , $R(n)$, como a probabilidade de retorno da partícula ao estado n após começar desde o mesmo estado n :

$$R(n) = \sum_{l=1}^{\infty} R_l(n, n) = h(n, n, 1) = 1 - \frac{1}{g(n, n, 1)} \quad (3.89)$$

Se $R(n) = 1$ o estado correspondente n é chamado *recorrente*. Para que um estado seja recorrente é necessário que $g(n, n, 1)$ diverja. Se o estado é recorrente se pode definir o *tempo médio de recorrência* como:

$$\langle l \rangle = \sum_{l=1}^{\infty} l R_l(n, n) = \lim_{z \rightarrow 1} \frac{d}{dz} h(n, n, z) \quad (3.90)$$

Utilizando a relação espectral (3.26) obtemos:

$$g(n, m, z) = \sum_k \frac{1}{1 - z\lambda_k} \psi_k(n) \phi_k(m) \quad (3.91)$$

Então, para saber se um estado é recorrente devemos calcular $g(n, n, z)$.

No caso em que a probabilidade estacionária $P(n) = \psi_0(n)$ não for identicamente nula, a relação anterior fica

$$g(n, n, z) = \frac{P(n)}{1-z} + \sum_{k \neq 0} \frac{1}{1-z\lambda_k} \psi_k(n) \phi_k(n) \quad (3.92)$$

já que $\lambda_0 = 1$, $\psi_0(n) = P(n)$ e $\phi_0(n) = 1$. Desta última expressão vemos que sempre que $P(n) \neq 0$ obtemos que $\lim_{z \rightarrow 1} g(n, n, z) = \infty$. Neste caso concluímos que $R(n) = 1$ e portanto o estado n é recorrente. Agora é simples calcular o tempo médio de recorrência $\langle l \rangle$. Para $z \simeq 1$ a (3.92) fica dominada pelo primeiro termo, e então podemos escrever $h(n, n, z) \approx 1 - (1-z)/P(n)$. Logo da (3.90) resulta $\langle l \rangle = 1/P(n)$. No caso da caminhada aleatória em $d = 1$, $P(n) = 1/N$ para qualquer estado, e então todos têm o mesmo tempo médio de recorrência $\langle l \rangle = N$.

No caso em que a probabilidade estacionária se anule, devemos considerar todos os termos da soma. No caso da caminhada aleatória em $d = 1$ obtemos:

$$g(n, m, z) = \frac{1}{N} \sum_k \frac{e^{ik(n-m)}}{1-z\lambda_k} \quad (3.93)$$

e por tanto,

$$g(n, n, z) = \frac{1}{N} \sum_k \frac{1}{1-z\lambda_k} \quad (3.94)$$

A probabilidade estacionária va a zero quando $N \rightarrow \infty$. Neste caso passando ao contínuo obtemos:

$$g(n, n, z) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{1-z\lambda_k} dk \quad (3.95)$$

Para determinar se esta função diverge para $z \rightarrow 1$ devemos analisar a integral

$$g(n, n, 1) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{1-\cos k} dk \quad (3.96)$$

Para ver se esta integral diverge podemos analisar o comportamento perto da origem. Para isto escrevemos:

$$\frac{1}{2} \int_{-\pi}^{\pi} \frac{1}{1-\cos k} dk = \int_0^{\pi} \frac{1}{1-\cos k} dk = \int_0^{\epsilon} \frac{1}{1-\cos k} dk + \int_{\epsilon}^{\pi} \frac{1}{1-\cos k} dk \quad (3.97)$$

com $\epsilon \ll 1$. Agora podemos aproximar o integrando da primeira integral por $2/k^2$ cuja integral diverge em $k = 0$. Logo $R(n) = 1$ e o estado é recorrente. No entanto neste caso em $N \rightarrow \infty$ o tempo médio de recorrência é infinito!

Em d dimensões a equação (3.93) fica (com $q = 1/2$ e fazendo uso das condições de contorno periódicas):

$$g(\vec{n}, \vec{m}, z) = \int_{-\pi}^{\pi} \frac{d^d \vec{k}}{(2\pi)^d} \frac{e^{i\vec{k}(\vec{n}-\vec{m})}}{1 - \frac{z}{d} \sum_{\mu=1}^d \cos k_{\mu}} \quad (3.98)$$

Neste caso temos que analisar o comportamento perto da origem da integral:

$$g(\vec{n}, \vec{n}, 1) = \frac{1}{(2\pi)^d} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} \frac{d}{d - \cos k_1 - \cos k_2 - \cdots - \cos k_d} dk_1 dk_2 \cdots dk_d \quad (3.99)$$

Fazendo uma expansão num entorno da origem somos levados a analisar a integral:

$$\int \frac{1}{k_1^2 + k_2^2 + \cdots + k_d^2} dk_1 dk_2 \cdots dk_d = c \int_0^{\pi} \frac{1}{k^2} k^{d-1} dk \quad (3.100)$$

Em $d = 2$ o resultado da integral é $c \ln k$ que apresenta uma divergência logarítmica na origem. $d = 2$ é um exemplo de *dimensão crítica*. Para $d \geq 3$ o resultado da integral é $\frac{c}{d-2} k^{d-2}$, e por tanto $g(\vec{n}, \vec{n}, 1)$ é finita. Em particular para $d = 3$ (e $p = q = 0.5$) a probabilidade de escape $P_{escape} = 1/g(n, n, 1) \approx 0.66$ ou a probabilidade de recorrência $R(n) = 1 - P_{escape} \approx 0.44$ [12]. Concluimos que para $d \geq 3$ a probabilidade de recorrência é menor do que um, $R(n) < 1$, e então a partícula pode não voltar nunca ao ponto de partida. A probabilidade de recorrência decresce com a dimensão espacial. Para ver como decresce podemos fazer uma expansão da $g(\vec{n}, \vec{n}, 1)$, equação (3.99), para d grande. O resultado é:

$$g(\vec{n}, \vec{n}, 1) = 1 + \frac{1}{2d} + 3 \left(\frac{1}{2d}\right)^2 + 12 \left(\frac{1}{2d}\right)^3 + 60 \left(\frac{1}{2d}\right)^4 + 355 \left(\frac{1}{2d}\right)^5 + \mathcal{O}\left(\left(\frac{1}{2d}\right)^6\right) \quad (3.101)$$

Como o tempo está discretizado, esta equação nos diz que para d grande o tempo que o sistema passa na origem tende ao intervalo de tempo entre dois saltos consecutivos, ou seja, em dimensões grandes a partícula nunca volta

ao sítio inicial. A probabilidade de recorrência tende a zero e é dada por:

$$h(\vec{n}, \vec{n}, 1) = \frac{1}{2d} + 2 \left(\frac{1}{2d}\right)^2 + 7 \left(\frac{1}{2d}\right)^3 + 35 \left(\frac{1}{2d}\right)^4 + 215 \left(\frac{1}{2d}\right)^5 + \mathcal{O}\left(\left(\frac{1}{2d}\right)^6\right) \quad (3.102)$$

O fato de que $R(n) = 1$ para $d \leq 2$ nos diz que as trajetórias do caminhante aleatório serão densas em uma e duas dimensões, e não o serão mais em dimensões maiores ou iguais a três. O espaço visitado por uma caminhada aleatória em $d \geq 3$ é fractal.

A função $g(n, m, z)$ pode ser considerada uma matriz e pode ser escrita formalmente como (ver definição (3.86) e (3.91)):

$$g(z) = \frac{1}{I - zT} \quad (3.103)$$

Também a partir da equação (3.98) se pode verificar que a $g(\vec{n}, \vec{m}, 1)$ satisfaz:

$$g(\vec{n}, \vec{m}, 1) = \delta(\vec{n}, \vec{m}) + \frac{1}{2d} \sum_{\mu} g(\vec{n} - \vec{m} + \vec{e}_{\mu}, 1) + g(\vec{n} - \vec{m} - \vec{e}_{\mu}, 1) \quad (3.104)$$

que é a forma discreta do operador Laplaciano, ou seja,

$$\Delta_r g(\vec{n} - \vec{m}, 1) = -\delta(\vec{n}, \vec{m}) \quad (3.105)$$

Por tanto esta função que interpretamos com o tempo médio de permanência num estado dado é uma **função de Green**. É fácil mostrar que a função geratriz $g(\vec{n}, \vec{m}, z)$ também é uma função de Green.

Finalmente, se identificamos o estado \vec{n} com uma coordenada \vec{x} , passamos ao limite contínuo $\vec{x} \rightarrow \vec{x}/a$, $\vec{k} \rightarrow \vec{k}a$, fazendo $a \rightarrow 0$ e identificando $1/z \equiv 1 + m^2 a^2 / 2d$, obtemos

$$g(\vec{x} - \vec{x}_0, m^2) = \int_{-\infty}^{\infty} \frac{d^d \vec{k}}{(2\pi)^d} \frac{e^{i\vec{k} \cdot (\vec{x} - \vec{x}_0)}}{k^2 + m^2} \quad (3.106)$$

que é o propagador massivo numa teoria de campos escalares livres [7].

3.3 A Equação Master

A aproximação de considerar o intervalo de tempo entre dois eventos como sendo fixo e finito não é adequada para descrever muitos processos, como

por exemplo o decaimento radioativo, onde os instantes em que acontecem as transições entre estados podem tomar quaisquer valores reais e os intervalos não são necessariamente iguais. Vamos generalizar nossa análise para estes casos considerando que o intervalo de tempo para que aconteça uma transição, τ , possa tomar valores arbitrariamente pequenos. Quando $\tau \rightarrow 0$ podemos escrever:

$$T(n, m) = \tau W(n, m) \quad m \neq n \quad (3.107)$$

Interpretamos a matriz $W(n, m)$ como a probabilidade de transição por unidade de tempo, ou *taxa de transição* do estado m para o n para $m \neq n$. O termo $m = n$ pode ser escrito da forma:

$$T(n, n) = 1 - \tau \Omega(n) \quad (3.108)$$

e corresponde a probabilidade de não haver transição no intervalo τ , que será muito próxima de um para $\tau \rightarrow 0$. Utilizando a normalização das colunas da matriz estocástica obtemos:

$$\begin{aligned} \sum_m T(m, n) &= T(n, n) + \sum_{m \neq n} T(m, n) \\ &= 1 - \tau \Omega(n) + \tau \sum_{m \neq n} W(m, n) = 1 \end{aligned} \quad (3.109)$$

e por tanto

$$\Omega(n) = \sum_{m \neq n} W(m, n) \quad (3.110)$$

A probabilidade do sistema estar no estado n ao tempo $l + 1$ é dada por

$$\begin{aligned} P_{l+1}(n) &= \sum_m T(n, m) P_l(m) = \sum_{m \neq n} T(n, m) P_l(m) + T(n, n) P_l(n) \\ &= \tau \sum_{m \neq n} W(n, m) P_l(m) + P_l(n) - \tau \Omega(n) P_l(n) \end{aligned} \quad (3.111)$$

Se definimos $t = l\tau$ e a probabilidade $P(n, t) = P_l(n)$ temos:

$$\frac{P(n, t + \tau) - P(n, t)}{\tau} = \sum_{m \neq n} W(n, m) P(m, t) - \Omega(n) P(n, t) \quad (3.112)$$

Tomando o limite $\tau \rightarrow 0$ e utilizando a relação (3.110) obtemos

$$\frac{dP(n, t)}{dt} = \sum_{m \neq n} [W(n, m) P(m, t) - W(m, n) P(n, t)] \quad (3.113)$$

que é a **equação master**. Notamos que a equação master é uma equação diferencial para a variação temporal das probabilidades e aparece como elemento novo a matriz $W(n, m)$ interpretada como a *taxa de transição entre os estados m e n* . Reconhecemos no primeiro termo da equação master o aumento da probabilidade do estado n devido a transições para o estado, enquanto que o segundo termo representa a diminuição da probabilidade por transições do estado n para outros estados. É claro que o termo com $m = n$ está ausente da soma por não contribuir para a variação da probabilidade. Toda a informação sobre o sistema está na matriz $W(n, m)$. Esta matriz pode ser calculada com qualquer método para determinar taxas de transição num sistema, como por exemplo teoria de perturbações dependentes do tempo. Num sistema quântico isto leva a famosa *regra de ouro de Fermi*:

$$W(n, m) = \frac{2\pi}{\hbar} |H'(n, m)|^2 \rho(E_n) \quad (3.114)$$

onde $H'(n, m)$ é o elemento de matriz do termo de perturbação no Hamiltoniano e $\rho(E_n)$ a densidade de estados do sistema não perturbado.

3.3.1 Caminhada aleatória em tempo contínuo

Uma interessante generalização da caminhada aleatória consiste em considerar as probabilidades de transição dependentes do tempo, da forma $\gamma\Delta t/2$. O caminhante ainda pode dar um salto a direita ou esquerda mas agora a probabilidade de transição aumenta com o aumento do tempo transcorrido desde o último salto. Desta forma as taxas de transição são dadas por

$$W(n, n+1) = W(n, n-1) = \frac{\gamma}{2} \quad (3.115)$$

sendo nulas todas as outras. Para este problema a equação master toma a forma:

$$\frac{dP(n, t)}{dt} = \frac{\gamma}{2} P(n+1, t) + \frac{\gamma}{2} P(n-1, t) - \gamma P(n, t) \quad (3.116)$$

Notamos que o lado direito corresponde à versão discreta do operador Laplaciano o que põe em evidência mais uma vez a relação entre a caminhada aleatória e a difusão.

3.3.2 A matriz de evolução W

Na seção anterior definimos as taxas de transição $W(n, m)$ para $m \neq n$. No entanto não foi necessário definir o elemento diagonal $W(n, n)$. Dada esta liberdade vamos escolher estes elementos de forma que

$$\sum_m W(m, n) = 0 \quad (3.117)$$

ou seja

$$W(n, n) = - \sum_{m \neq n} W(m, n) = -\Omega(n) \quad (3.118)$$

Desta forma fica completamente definida a matriz W chamada *matriz de evolução* e que tem as seguintes propriedades:

- Todos os elementos fora da diagonal são maiores ou iguais a zero:

$$W(n, m) \geq 0, \quad m \neq n \quad (3.119)$$

- A soma dos elementos de qualquer coluna é nula.

Se definimos uma matriz coluna de elementos $\psi(n)$ podemos escrever:

$$\sum_m W(n, m)\psi(m) = \sum_{m \neq n} [W(n, m)\psi(m) - W(m, n)\psi(n)] \quad (3.120)$$

Isto nos permite reescrever a equação master na forma:

$$\frac{d}{dt}P(n, t) = \sum_m W(n, m)P(m, t) \quad (3.121)$$

cuja solução formal com condição inicial $P(n, 0)$ é:

$$P(n, t) = e^{tW}P(n, 0) \quad (3.122)$$

onde $\exp(tW)$ é a matriz definida por

$$e^{tW} = I + tW + \frac{t^2}{2!}W^2 + \frac{t^3}{3!}W^3 + \dots \quad (3.123)$$

sendo I a matriz identidade.

3.3.3 Sistema de dois estados

Consideremos o sistema de dois estados dado pela matriz de evolução:

$$W = \begin{pmatrix} -\gamma/2 & \gamma/2 \\ \gamma/2 & -\gamma/2 \end{pmatrix} \quad (3.124)$$

É fácil mostrar que:

$$W^l = (-\gamma)^{l-1}W \quad l = 1, 2, 3, \dots \quad (3.125)$$

Com este resultado obtemos:

$$\begin{aligned} \exp(tW) &= I + \sum_{l=1}^{\infty} \frac{t^l}{l!} W^l \\ &= I + W \sum_{l=1}^{\infty} \frac{t^l}{l!} (-\gamma)^{l-1} \\ &= I + \frac{1}{\gamma} (1 - e^{-t\gamma})W \end{aligned} \quad (3.126)$$

Logo, escrevendo em forma matricial:

$$P(t) = P(0) + \frac{1}{\gamma} (1 - e^{-t\gamma})WP(0) \quad (3.127)$$

onde $P(t)$ é a matriz coluna de elementos $P(n, t)$. Explicitamente:

$$\begin{aligned} P(1, t) &= \frac{1}{2}(1 - e^{-t\gamma}) + e^{-t\gamma}P(1, 0) \\ P(2, t) &= \frac{1}{2}(1 - e^{-t\gamma}) + e^{-t\gamma}P(2, 0) \end{aligned} \quad (3.128)$$

Notamos nesta solução que para $t \rightarrow \infty$, $P(1, t) = P(2, t) = 1/2$ independentemente da condição inicial, ou seja, os dois estados ficam igualmente populados.

3.3.4 Propriedades das matrizes W

A normalização das colunas de uma matriz W nos diz que uma tal matriz possui um autovetor a esquerda com todas as componentes iguais a um e cujo autovalor correspondente é zero:

$$\phi = (1, 1, 1, \dots, 1) \quad \phi W = 0 \quad (3.129)$$

Existe o correspondente autovetor a direita com o mesmo autovalor:

$$W\psi = 0 \quad (3.130)$$

O autovetor ψ é uma solução independente do tempo da equação master e pode ser uma probabilidade estacionária se suas componentes são não negativas e convenientemente normalizadas. As matrizes W podem ser classificadas de forma semelhante a classificação das matrizes T .

Podemos obter as soluções da equação master se determinamos todos os seus autovetores e autovalores. Chamando Λ_k aos autovalores, os autovetores a esquerda e direita satisfazem as equações:

$$W\psi_k = \Lambda_k\psi_k \quad \phi_k W = \Lambda_k\phi_k \quad (3.131)$$

Os autovetores formam um conjunto completo ortonormalizado de forma análoga ao que acontece com os autovetores das matrizes T :

$$\sum_n \phi_j(n)\psi_k(n) = \delta_{jk} \quad \sum_k \psi_k(n)\phi_k(m) = I(n, m) \quad (3.132)$$

Já vimos que o vetor de probabilidade estacionária P_e é um autovetor com autovalor nulo. Seja $\Lambda_0 = 0$ o correspondente autovalor. Então $P_e = \psi_0$. O correspondente autovetor a esquerda tem todas as componentes iguais a um.

Consideremos agora a seguinte expansão:

$$e^{tW} = e^{tW}I = e^{tW} \sum_k \psi_k\phi_k = \sum_k e^{\Lambda_k t} \psi_k\phi_k \quad (3.133)$$

Utilizando este resultado obtemos:

$$P(t) = \sum_k a_k e^{\Lambda_k t} \psi_k, \quad (3.134)$$

onde $a_k = \phi_k P(0)$ é um escalar. Utilizando agora as propriedades associadas ao autovalor zero obtemos:

$$P(t) = P_e + \sum_{k \neq 0} a_k e^{\Lambda_k t} \psi_k \quad (3.135)$$

3.3.5 Balanço Detalhado

No estado estacionário a equação master se reduz a:

$$\sum_{m \neq n} W(n, m)P_e(m) = \sum_{m \neq n} W(m, n)P_e(n). \quad (3.136)$$

Esta equação representa o fato de que no estado estacionário a soma de todas as transições por unidade de tempo para qualquer estado n deve ser equilibrada pela soma das transições do estado n para todos os outros.

Para sistemas físicos em equilíbrio termodinâmico uma identidade mais forte se satisfaz:

$$W(n, m)P_e(m) = W(m, n)P_e(n). \quad (3.137)$$

Esta identidade implica que o balanço das probabilidades acontece para *cada par de estados m e n* e é conhecida como *balanço detalhado*.

O balanço detalhado implica *reversibilidade microscópica* da dinâmica do sistema. Ainda é possível mostrar que a condição de balanço detalhado depende só das taxas de transição do sistema e não da forma particular da probabilidade de equilíbrio: consideremos três estados n, n' e n'' distintos. No estado estacionário a probabilidade de observar a trajetória fechada $n \rightarrow n' \rightarrow n'' \rightarrow n$ em três intervalos de tempo consecutivos Δt é:

$$\Delta t W(n, n'') \Delta t W(n'', n') \Delta t W(n', n) P_e(n) \quad (3.138)$$

enquanto que a probabilidade de observar a trajetória inversa é:

$$\Delta t W(n, n') \Delta t W(n', n'') \Delta t W(n'', n) P_e(n) \quad (3.139)$$

Se o sistema apresenta reversibilidade microscópica essas duas probabilidades são iguais e então:

$$W(n, n'') W(n'', n') W(n', n) = W(n, n') W(n', n'') W(n'', n) \quad (3.140)$$

que é independente da probabilidade estacionária. Ainda é possível mostrar que as matrizes W que satisfazem balanço detalhado possuem todos os autovalores reais.

3.3.6 Processos “One step” ou “Nascimento-Morte”

Os processos de um passo ou *one step* são definidos como processos markovianos em tempo contínuo com estados discretos nos quais só são permitidas transições entre estados vizinhos. Assim consideremos uma caminhada aleatória generalizada na qual a taxa de transição para a direita é

$W(n+1, n) = a_n$ e para esquerda $W(n-1, n) = b_n$ enquanto que todas as outras são nulas. A equação master correspondente é:

$$\frac{d}{dt}P(n, t) = b_{n+1}P(n+1, t) + a_{n-1}P(n-1, t) - (a_n + b_n)P(n, t) \quad (3.141)$$

Em muitos casos esta equação não pode ser resolvida exatamente ¹, no entanto as vezes é possível obter valores médios de funções dos estados, $f(n)$. Os valores médios se obtêm multiplicando a eq.(3.141) por $f(n)$ e somando sobre n :

$$\frac{d}{dt} \langle f(n) \rangle = \langle [f(n+1) - f(n)]a_n \rangle + \langle [f(n-1) - f(n)]b_n \rangle \quad (3.142)$$

O caso mais elementar corresponde aos momentos de n :

$$\frac{d}{dt} \langle n \rangle = \langle a_n \rangle - \langle b_n \rangle \quad (3.143)$$

$$\frac{d}{dt} \langle n^2 \rangle = \langle (2n+1)a_n \rangle + \langle (-2n+1)b_n \rangle \quad (3.144)$$

Notamos que se as taxas de transição não forem lineares em n os momentos irão depender de momentos de ordem superior.

3.3.7 Decaimento Radiativo

Suponhamos que temos um material radiativo que, ao tempo t , possui n átomos. Em um intervalo Δt um átomo do tipo A pode decair em outro de tipo B . O decaimento por átomo é governado pela taxa de decaimento γ de forma que a probabilidade de transição $n \rightarrow n-1$ num intervalo pequeno de tempo Δt é $\Delta t \gamma n$. Neste problema $a_n = 0$, $b_n = \gamma n$ e a equação master para as probabilidades de observar n átomos do material A no instante t é:

$$\frac{d}{dt}P(n, t) = \gamma(n+1)P(n+1, t) - \gamma n P(n, t) \quad (3.145)$$

A equação para a evolução temporal do primeiro momento é:

$$\frac{d}{dt} \langle n \rangle = -\gamma \langle n \rangle, \quad (3.146)$$

¹Para uma discussão da teoria geral de processos one-step e numerosas aplicações ver [9] capítulo VI

e do segundo momento:

$$\frac{d}{dt} \langle n^2 \rangle = -2\gamma \langle n^2 \rangle + \gamma \langle n \rangle \quad (3.147)$$

cujas soluções são:

$$\langle n(t) \rangle = n_0 e^{-\gamma t} \quad (3.148)$$

e

$$\langle n^2(t) \rangle = n_0 e^{-\gamma t} + (n_0^2 - n_0) e^{-2\gamma t} \quad (3.149)$$

com as condições iniciais $\langle n(0) \rangle = n_0$ e $\langle n^2(0) \rangle = n_0^2$. Desta forma obtivemos o conhecido resultado do decaimento exponencial do número de átomos num material radiativo. Notar que o estado $n = 0$ é um estado absorvente e a solução assintótica da dinâmica a tempos grandes (infinito).

3.3.8 Reações químicas

É possível fazer modelos simplificados da cinética das reações químicas considerando-as como processos estocásticos. Os casos mais simples correspondem a processos de um passo. Consideremos por exemplo a reação:



com taxas de reação k_1, k_2 e k_3 respectivamente. A primeira é uma reação catalítica onde se assume que o número de átomos B é muito grande e praticamente constante. A segunda é a reação reversa e a terceira uma aniquilação espontânea. A taxa de transição do primeiro processo que corresponde a $n \rightarrow n + 1$ é $k_1 n$; do segundo que corresponde a $n \rightarrow n - 1$ é $k_2 n^2$ e do terceiro que corresponde de novo a $n \rightarrow n - 1$ é $k_3 n$. Por tanto $a_n = k_1 n$, $b_n = k_2 n^2 + k_3 n$ e a equação master fica:

$$\frac{d}{dt} P(n, t) = [k_2(n+1)^2 + k_3(n+1)]P(n+1, t) + k_1(n-1)P(n-1, t) - (k_1 n + k_2 n^2 + k_3 n)P(n, t) \quad (3.151)$$

A evolução do primeiro momento é:

$$\frac{d}{dt} \langle n \rangle = (k_1 - k_3) \langle n \rangle - k_2 \langle n^2 \rangle \quad (3.152)$$

e do segundo:

$$\frac{d}{dt} \langle n^2 \rangle = (k_1 + k_3) \langle n \rangle + (2k_1 - 2k_3 + k_2) \langle n^2 \rangle - 2k_2 \langle n^3 \rangle \quad (3.153)$$

onde notamos que os momentos de cada ordem dependem dos momentos da ordem imediata superior, o que produz uma hierarquia de equações acopladas. Para poder resolver as equações devemos fazer alguma aproximação. A mais simples consiste em eliminar flutuações e considerar $\langle n^2 \rangle = \langle n \rangle^2$ que nos permite resolver a equação para o primeiro momento. Se $k_1 \neq k_3$ obtemos nesta aproximação:

$$\langle n \rangle (t) = \frac{k_1 - k_3}{k_2 + c \exp\{-(k_1 - k_3)t\}} \quad (3.154)$$

sendo c uma constante a determinar pela condição inicial. Este modelo apresenta duas possíveis soluções estacionárias dependendo dos valores relativos de k_1 e k_3 :

$$\begin{aligned} \langle n \rangle &= 0 & k_1 < k_3 \\ \langle n \rangle &= \frac{k_1 - k_3}{k_2} & k_1 > k_3 \end{aligned} \quad (3.155)$$

Notamos que o sistema apresenta uma mudança no comportamento assintótico dependendo de que a taxa da reação catalítica k_1 seja menor que a da reação de aniquilação espontânea k_3 em cujo caso o número final de átomos do tipo A é nulo ou então que $k_1 > k_3$ em cujo caso o número final de átomos A é finito. Este é um exemplo simples de *transição de fase*. A relaxação ao valor assintótico é exponencial em ambos os casos com taxa de relaxação $\tau = |k_1 - k_3|^{-1}$. Notamos que a relaxação se torna mais e mais lenta a medida que k_1 e k_3 se aproximam e o tempo diverge quando ambas taxas são iguais. Neste caso a equação de evolução do primeiro momento é:

$$\frac{d}{dt} \langle n \rangle = -k_2 \langle n \rangle^2 \quad (3.156)$$

cuja solução é:

$$\langle n \rangle (t) = \frac{1}{c + k_2 t} \quad (3.157)$$

Notamos que neste caso o número de átomos do tipo A decai a zero de forma algébrica.

3.4 Lista de exercícios

1. Escreva um programa que calcule o deslocamento e o deslocamento quadrático em função do tempo (número de iterações) de uma caminhada aleatória em uma dimensão espacial.
 - 1) Grafique o deslocamento $x(t)$ para uma caminhada.
 - 2) Calcule o deslocamento quadrático médio em função do tempo fazendo uma média em 1000 caminhadas (diferentes números aleatórios). Grafique $\langle x^2(t) \rangle$. Comparar com os resultados analíticos.
2. Mostre que a equação (2.22) é um caso particular da (3.6).
3. Mostre que as probabilidades condicionais de um processo de Markov satisfazem a equação de Chapman-Kolmogorov:

$$P(n_3|n_1) = \sum_{n_2} P(n_3|n_2) P(n_2|n_1) \quad (3.158)$$

4. Mostre que a caminhada aleatória de Cauchy, definida pela distribuição de Lorentz:

$$P(x, t) = \frac{t/\pi}{(x^2 + t^2)} \quad (3.159)$$

e probabilidades condicionais

$$P(x_2, t_2|x_1, t_1) = \frac{(t_2 - t_1)/\pi}{(x_2 - x_1)^2 + (t_2 - t_1)^2} \quad (3.160)$$

é um processo markoviano.

5. Considere a cadeia de Markov de dois estados E_1 e E_2 com probabilidades de transição $T(1, 1) = T(2, 2) = p$, $T(1, 2) = T(2, 1) = q$, $p + q = 1$ e condição inicial $P_0(1) = \alpha$, $P_0(2) = \beta$ ($\alpha + \beta = 1$). Determinar as probabilidades de transição no l -ésimo passo $T^l(n, m)$, as probabilidades $P_l(n)$ e as probabilidades assintóticas ou estacionárias $P(n)$. Interprete fisicamente o sistema e compare-o com o exemplo de sistema de dois estados da seção 3.2.3.
6. A equação para a evolução temporal das probabilidades no modelo dos Ehrenfest é:

$$P_{l+1}(n) = q \left(1 - \frac{n-1}{N}\right) P_l(n-1) + p P_l(n) + q \left(\frac{n+1}{N}\right) P_l(n+1). \quad (3.161)$$

O número médio de moléculas na caixa A no instante l é dado por:

$$\langle n \rangle_l = \sum_{n=0}^N n P_l(n). \quad (3.162)$$

Substituindo esta definição na equação da evolução temporal podemos obter uma recorrência para $\langle n \rangle_l$. Propondo uma solução da forma $\langle n \rangle_l = ar^l + b$, e uma condição inicial $\langle n \rangle_0 = N$, calcule os dois primeiros momentos da distribuição no instante l e a variância.

7. Seja

$$f(x) = \sum_{n=0}^N \Psi(n) x^n \quad (3.163)$$

uma função geratriz dos autovetores a direita do modelo dos Ehrenfest. Mostrar que

$$f_k(x) = 2^{-N} (1+x)^{N-k} (1-x)^k \quad (3.164)$$

onde $k = -N(\lambda - 1)/2q$ parametriza os autovalores λ_k e autovetores $\Psi_k(n)$.

8. Mostre que os autovalores de uma matriz de Toeplitz satisfazem a relação:

$$\lambda_k = \sum_n f(n) e^{ikn} = F(k) \quad (3.165)$$

onde $f(n - m) = T(n, m)$ define uma matriz de Toeplitz.

9. Considere um processo estocástico de nascimento-morte linear, equação (3.141), com $b_n = \gamma n$ e $a_n = \beta n$, sendo β e γ constantes. O processo resultante representa uma primeira aproximação para a evolução do número de bactérias numa população. Defina a função geratriz:

$$F(z, t) = \sum_{n=-\infty}^{\infty} z^n P(n, t). \quad (3.166)$$

Mostre que a função F satisfaz a seguinte equação a derivadas parciais:

$$\frac{\partial F}{\partial t} = (z - 1)(\beta z - \gamma) \frac{\partial F}{\partial z}. \quad (3.167)$$

Esta equação diferencial pode ser resolvida pelo método das características. Supondo que no instante inicial existem exatamente m bactérias, verifique que:

$$F(z, t) = \left(\frac{\gamma(z-1)e^{(\beta-\gamma)t} - \beta z + \gamma}{\beta(z-1)e^{(\beta-\gamma)t} - \beta z + \gamma} \right)^m \quad (3.168)$$

é solução da equação diferencial. Usando expressões para os momentos da distribuição em termos de derivadas da função geratriz, calcule o primeiro momento e a variância em função do tempo.

10. Repita o problema 1 para uma caminhada com distribuição de Weierstrass. O que pode concluir da comparação de ambas caminhadas?

Capítulo 4

O Modelo de Ising

Neste capítulo vamos começar com o estudo de sistemas de partículas em interação sob a perspectiva dos processos estocásticos associados. Um dos modelos mais universais e importantes, tanto pela diversidade de aplicações como pela relativa simplicidade que permite obter soluções exatas em muitos casos é o **modelo de Ising**.

O modelo de Ising pode ser definido sobre uma rede hipercúbica em d dimensões espaciais. Em cada vértice da rede é definida uma variável **binária**, ou seja a variável dinâmica relevante pode assumir somente dois valores $\sigma_i = \{\pm 1\}$ ou $\{0, 1\}$. O índice $i = 1, \dots, N$ corresponde aos N sítios da rede. Um estado do sistema corresponde a uma configuração completa das N variáveis definidas nos sítios da rede $\sigma = (\sigma_1, \sigma_2, \sigma_3, \dots, \sigma_{N-1}, \sigma_N)$. A evolução temporal das probabilidades é dada pela equação master:

$$\frac{d}{dt}P(\sigma, t) = \sum_{\sigma' \neq \sigma} [W(\sigma, \sigma')P(\sigma', t) - W(\sigma', \sigma)P(\sigma, t)] \quad (4.1)$$

Notamos que dado o caráter binário das variáveis σ_i o sistema tem 2^N estados possíveis, ou configurações. O próximo passo é determinar a forma das taxas de transição entre os estados $W(\sigma, \sigma')$. Primeiramente vamos nos limitar a considerar uma dinâmica na qual a cada passo somente um sítio muda de estado. Esta dinâmica é conhecida como **single spin flip dynamics** ou dinâmica de reversão de um único spin. Para este tipo de dinâmica as taxas são dadas por:

$$W(\sigma', \sigma) = \sum_i \delta(\sigma'_1, \sigma_1) \delta(\sigma'_2, \sigma_2) \dots \delta(\sigma'_i, -\sigma_i) \dots \delta(\sigma'_N, \sigma_N) w(\sigma_i) \quad (4.2)$$

onde $\delta(\sigma'_j, \sigma_j)$ são deltas de Kronecker e o fator $w(\sigma_i)$ é a taxa de inversão do estado do i -ésimo sítio de σ_i para $-\sigma_i$. Com esta definição a equação master se pode escrever da seguinte forma:

$$\frac{d}{dt}P(\sigma, t) = \sum_i [w(-\sigma_i)P(\mathbf{R}_i\sigma, t) - w(\sigma_i)P(\sigma, t)] \quad (4.3)$$

onde \mathbf{R}_i é um operador que inverte o estado do i -ésimo sítio. A evolução temporal da média de uma função de estado $f(\sigma)$:

$$\langle f(\sigma) \rangle = \sum_{\sigma} f(\sigma)P(\sigma, t) \quad (4.4)$$

fica da forma:

$$\frac{d}{dt} \langle f(\sigma) \rangle = \sum_{i=1}^N \langle [f(\mathbf{R}_i\sigma) - f(\sigma)]w(\sigma_i) \rangle \quad (4.5)$$

Como casos particulares importantes analizaremos a evolução temporal da média de $f(\sigma) = \sigma_j$:

$$\frac{d}{dt} \langle \sigma_j \rangle = -2 \langle \sigma_j w(\sigma_j) \rangle \quad (4.6)$$

e da correlação de dois pontos $f(\sigma) = \sigma_j \sigma_k$ para $j \neq k$:

$$\frac{d}{dt} \langle \sigma_j \sigma_k \rangle = -2 \langle \sigma_j \sigma_k w(\sigma_j) \rangle - 2 \langle \sigma_j \sigma_k w(\sigma_k) \rangle . \quad (4.7)$$

O modelo de Ising foi inicialmente formulado para descrever a transição de fase num material ferromagnético. Nas seções seguintes vamos ver como as taxas $w(\sigma_i)$ devem ser especificadas para corresponder ao Hamiltoniano de um sistema ferromagnético. Mas desde essa aplicação inicial o modelo tem demonstrado ser de utilidade para estudar sistemas tão diversos como ligas binárias, sistemas magnéticos desordenados, fluídos na rede ou gases de rede, sistemas de redes de neurônios, o sistema imunológico, crescimento de estruturas fractais como o DLA (diffusion limited aggregation) e redes de genes entre muitas outras aplicações em diversas áreas das ciências. A característica binária das variáveis do modelo em analogia com a forma de armazenamento de dados num computador fazem dele um modelo perfeito para simulações computacionais de uma grande variedade de sistemas.

De aqui em diante vamos utilizar a linguagem de sistemas de spins para discutir o comportamento dinâmico do modelo de Ising. Podemos definir uma função energia, ou Hamiltoniano do modelo, da forma:

$$\mathcal{H} = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j \quad (4.8)$$

onde J é uma constante que representa a magnitude da interação entre as variáveis de spin $\sigma_i = \pm 1$. Notamos que a forma da energia corresponde a interações de pares de spins. Se $J > 0$ a interação corresponde a um comportamento *ferromagnético* e favorece o alinhamento dos spins vizinhos (a energia diminui neste caso) enquanto que se $J < 0$ a interação é do tipo *antiferromagnético* e os spins tendem a ficar antialinhados ou antiparalelos. Este modelo apresenta, para dimensões $d \geq 2$, uma transição de fase a uma certa **temperatura crítica** T_c , de uma fase desordenada a temperaturas altas para uma fase ordenada ou ferromagnética a temperaturas baixas ($T < T_c$) quando $J > 0$ ou para uma fase com ordem antiferromagnético quando $J < 0$. Em $d = 1$ o modelo não apresenta transição a temperatura finita tendo um ponto crítico a $T = 0$.

A presença de um banho térmico com o qual o sistema está em contato induz uma dinâmica estocástica no mesmo que pode ser analisada com as ferramentas estudadas até agora. Em particular podemos tentar responder a perguntas como qual é a evolução das probabilidade dos estados fora do equilíbrio?, com qué velocidade o sistema se aproxima do equilíbrio nas fases de alta e baixa temperatura?, acontece alguma coisa de especial na dinâmica exatamente no ponto crítico T_c ?, como se comportam as quantidades termodinâmica em função do tempo? a energia interna, a magnetização, a susceptibilidade, as correlações?

Notamos que o modelo de Ising não possui uma dinâmica “natural”, ou seja não temos equações de Hamilton para ele, já que o Hamiltoniano está definido num espaço discreto. No entanto a forma do Hamiltoniano e o conhecimento da forma das probabilidades estacionárias, dadas pela mecânica estatística, permitem introduzir regras dinâmicas razoáveis para levar o sistema no caminho do equilíbrio termodinâmico conhecido. Uma destas regras, ou “dinâmicas” mais conhecidas foi introduzida por Glauber [13].

4.1 Dinâmica de Glauber [13]

Como já vimos, uma condição necessária para a dinâmica de qualquer sistema convergir à distribuição de equilíbrio é o balanço detalhado, que no modelo de Ising corresponde a exigir:

$$w(-\sigma_i)P(\mathbf{R}_i\sigma, t) = w(\sigma_i)P(\sigma, t) \quad (4.9)$$

Definimos o *campo local* sobre o sítio i :

$$h_i = \sum_j \sigma_j, \quad (4.10)$$

com a soma extendendo-se a todos os vizinhos com os quais o spin i -ésimo interage. Assim, podemos escrever a energia do modelo na forma:

$$\mathcal{H} = -J \sum_i \sigma_i h_i. \quad (4.11)$$

Como as probabilidades de equilíbrio devem ser dadas pela distribuição de Boltzmann, a condição de balanço detalhado pode ser escolhida como:

$$\frac{P(\mathbf{R}_i\sigma, t)}{P(\sigma, t)} = \frac{w(\sigma_i)}{w(-\sigma_i)} = \frac{\exp(-\beta J \sigma_i h_i)}{\exp(\beta J \sigma_i h_i)} \quad (4.12)$$

onde $\beta = 1/kT$. Utilizando a identidade $\exp \pm x = \cosh x \pm \sinh x$ obtemos

$$\frac{\exp(-\beta J \sigma_i h_i)}{\exp(\beta J \sigma_i h_i)} = \frac{1 - \sigma_i \tanh(\beta J h_i)}{1 + \sigma_i \tanh(\beta J h_i)} \quad (4.13)$$

Finalmente assumindo que a taxa de inversão de um spin livre seja $\Gamma/2$, a taxa de inversão do modelo de Glauber corresponde a:

$$w(\sigma_i) = \frac{\Gamma}{2} [1 - \sigma_i \tanh(\beta J h_i)] \quad (4.14)$$

Esta forma particular das taxas de transição são conseqüência do caráter binário das variáveis de spin. No entanto o resultado pode ser generalizado para qualquer sistema que apresente um conjunto de estados discretos caracterizados por energias E_i entre as quais o sistema pode realizar transições que tendam a diminuir a energia. A expressão geral é da forma:

$$w(E_i \rightarrow E_j) = \Gamma \frac{1}{1 + \exp[\beta(E_j - E_i)]} \quad (4.15)$$

É fácil mostrar que para uma energia da forma (4.8) as taxas (4.15) se reduzem às correspondentes (4.14).

4.2 Aproximação de campo médio

Vamos discutir nesta seção uma aproximação para a solução da dinâmica do modelo de Ising-Glauber conhecida como *aproximação de campo médio*. A mesma corresponde fisicamente a desprezar ou suprimir flutuações nas variáveis dinâmicas ou spins. A partir das equações (4.6) e (4.14) obtemos uma equação para a evolução do valor médio do spin ou magnetização:

$$\frac{1}{\Gamma} \frac{d}{dt} \langle \sigma_i \rangle = - \langle \sigma_i \rangle + \langle \tanh(\beta J h_i) \rangle \quad (4.16)$$

Esta equação não pode ser resolvida exatamente em dimensão arbitrária. A aproximação de campo médio equivale a considerar:

$$\langle \tanh(\beta J h_i) \rangle \rightarrow \tanh(\beta J \langle h_i \rangle) \quad (4.17)$$

onde claramente estamos eliminando as flutuações no valor dos campos locais h_i . Como o sistema apresenta invariância translacional podemos escrever $\langle \sigma_i \rangle = m$ independentemente do sítio i e a equação diferencial para a dinâmica de m se reduz a:

$$\frac{1}{\Gamma} \frac{dm}{dt} = -m + \tanh(\beta J z m) \quad (4.18)$$

onde z é o número de vizinhos de um dado sítio ou *número de coordenação*. Em equilíbrio se deve satisfazer:

$$m = \tanh(\beta J z m) \quad (4.19)$$

que é a conhecida equação de Curie-Weiss para a magnetização. Esta equação transcendental apresenta dois tipos de soluções diferentes dependendo do valor do parâmetro $\beta J z$ que corresponde a inclinação da tangente hiperbólica na origem. Se $\beta J z \leq 1$ a solução é única $m = 0$. Esta região corresponde a temperaturas altas e é chamada *fase desordenada ou paramagnética*. Já se $\beta J z > 1$ existem duas soluções com $m \neq 0$. Esta região corresponde a temperaturas baixas e é chamada *fase ordenada ou ferromagnética*. Vemos então que na *temperatura crítica* $T_c = Jz/k_B$ acontece uma transição de fase. Na realidade análises melhores que a aproximação de campo médio permitem verificar que o modelo de Ising apresenta uma transição de fase somente para dimensões $d \geq 2$, o modelo unidimensional com interações de vizinhos

próximos só apresenta um ponto crítico em $T_c = 0$. A aproximação de campo médio equivale a considerar o modelo em dimensões altas.

Podemos estudar a dinâmica do parâmetro de ordem na vizinhança do ponto crítico fazendo uma expansão em série da tangente hiperbólica e truncando a ordem m^3 (já que $m \ll 1$ perto da transição):

$$\frac{1}{\Gamma} \frac{dm}{dt} = -\epsilon m - \frac{1}{3} m^3 \quad (4.20)$$

onde $\epsilon = (T - T_c)/T_c$ é uma *temperatura reduzida*, uma temperatura adimensional relativa à temperatura crítica. Multiplicando ambos os membros por m obtemos uma equação diferencial para m^2 :

$$\frac{1}{2\Gamma} \frac{dm^2}{dt} = -\epsilon m^2 - \frac{1}{3} m^4 \quad (4.21)$$

A solução para $\epsilon \neq 0$ é:

$$m^2(t) = \frac{3\epsilon}{ce^{2\epsilon\Gamma t} - 1} \quad (4.22)$$

onde c é uma constante. Para $t \rightarrow \infty$ obtemos soluções diferentes dependendo do sinal de ϵ . Para $\epsilon > 0$ ($T > T_c$) $m \rightarrow 0$ que corresponde a fase paramagnética ou desordenada. Notamos que o decaimento é exponencial com um tempo característico $\tau \propto \epsilon^{-1}$. Se $\epsilon < 0$ ($T < T_c$) $m \neq 0$:

$$m_{eq} = \sqrt{3|\epsilon|} \quad (4.23)$$

Notamos que existem duas soluções simétricas para m com sinais opostos. A magnetização na vizinhança da transição se comporta como $m_{eq} \propto |\epsilon|^{1/2}$. O expoente $1/2$ é um dos chamados *expoentes críticos*, neste caso o β que caracteriza o comportamento do parâmetro de ordem perto da transição $m_{eq} \propto |\epsilon|^\beta$. Também notamos que o decaimento aos valores de equilíbrio, tanto acima como abaixo da transição são exponenciais. Exatamente no ponto crítico $\epsilon = 0$ e a equação dinâmica se reduz a:

$$\frac{1}{\Gamma} \frac{dm}{dt} = -\frac{1}{3} m^3 \quad (4.24)$$

cuja solução é:

$$m(t) = \frac{\sqrt{3}}{\sqrt{2\Gamma t + c}} \quad (4.25)$$

Neste caso o decaimento deixa de ser exponencial para ser algébrico.

4.3 O modelo de Ising unidimensional

Nesta seção vamos analisar a dinâmica do modelo de Ising em uma dimensão espacial [14]. Este caso é exatamente solúvel já que as equações dinâmicas são lineares. No entanto a física do sistema é altamente não trivial sendo um excelente exemplo para estudar a dinâmica fora do equilíbrio de um modelo ferromagnético. Já comentamos antes que em uma dimensão o modelo de Ising só apresenta um ponto crítico em $T = 0$. Em $d = 1$ a taxa de transição (4.14) se pode escrever na forma:

$$w(\sigma_i) = \frac{\Gamma}{2} \{1 - \sigma_i \tanh [\beta J (\sigma_{i-1} + \sigma_{i+1})]\} \quad (4.26)$$

Fazendo uso da identidade

$$\tanh [\beta J (\sigma_{i-1} + \sigma_{i+1})] = \frac{1}{2} (\sigma_{i-1} + \sigma_{i+1}) \tanh 2\beta J \quad (4.27)$$

reescrevemos a taxa na forma:

$$w(\sigma_i) = \frac{1}{2} [1 - \frac{1}{2} \gamma \sigma_i (\sigma_{i-1} + \sigma_{i+1})] \quad (4.28)$$

onde fixamos $\Gamma = 1$ e definimos

$$\gamma = \tanh 2\beta J = \frac{1}{\cosh \mu}, \quad \mu = -\ln \tanh \beta J. \quad (4.29)$$

Estes parâmetros estão relacionados com o tempo de relaxação e o comprimento de correlação de equilíbrio [15] na forma:

$$\tau_{eq} = \frac{1}{1 - \gamma} \quad \xi_{eq} = \frac{1}{\mu} \quad (4.30)$$

que divergem exponencialmente ao se aproximar da temperatura crítica:

$$\tau_{eq} \approx \frac{\exp 4\beta J}{2} \quad \xi_{eq} \approx \frac{\exp 2\beta J}{2} \quad (4.31)$$

Conseqüentemente observamos o seguinte comportamento de escala:

$$\tau_{eq} \approx 2\xi_{eq}^2 \quad (4.32)$$

A forma desta lei de escala revela o comportamento difusivo do sistema a temperaturas baixas. O expoente $z = 2$ é outro exemplo de expoente crítico, neste caso o expoente que relaciona o tempo de relaxação com o comprimento de correlação de equilíbrio $\tau_{eq} \propto \xi_{eq}^z$. O expoente z é conhecido como *expoente dinâmico* e da informação da velocidade de relaxação do sistema ao equilíbrio.

4.3.1 Relaxação da Magnetização

A equação diferencial para a evolução da magnetização local (4.16) fica agora da forma:

$$\frac{d}{dt} \langle \sigma_i \rangle = - \langle \sigma_i \rangle + \frac{\gamma}{2} (\langle \sigma_{i-1} \rangle + \langle \sigma_{i+1} \rangle) \quad (4.33)$$

Agora está claro o caráter linear da equação dinâmica o que permite uma solução exata da mesma. O mesmo acontece com momentos de ordem superior. Para resolver a equação diferencial vamos utilizar as técnicas de transformadas de Fourier, Laplace e Fourier-Laplace definidas a seguir:

- *Transformada temporal de Laplace*

$$f_n^L(z) = \int_0^\infty f_n(t) e^{-zt} dt \quad f_n(t) = \int \frac{dz}{2\pi i} f_n^L(z) e^{zt} \quad (4.34)$$

- *Transformada espacial discreta de Fourier*

$$f^F(q, t) = \sum_n f_n(t) e^{-inq} \quad f_n(t) = \int_0^{2\pi} \frac{dq}{2\pi} f^F(q, t) e^{inq} \quad (4.35)$$

- *Transformada de Fourier-Laplace*

$$f^{FL}(q, z) = \sum_n \int_0^\infty f_n(t) e^{-(zt+inq)} dt \quad f_n(t) = \int \frac{dz}{2\pi i} \int_0^{2\pi} \frac{dq}{2\pi} f^{FL}(q, z) e^{zt+inq} \quad (4.36)$$

Chamando $\langle \sigma_n \rangle = m_n(t)$ e utilizando estas transformadas, a equação para a magnetização se pode escrever como:

$$\frac{dm^F(q, t)}{dt} = (\gamma \cos q - 1) m^F(q, t) \quad (4.37)$$

ou

$$z m^{FL}(q, z) = (\gamma \cos q - 1) m^{FL}(q, z) + m^F(q, t = 0) \quad (4.38)$$

então

$$m^{FL}(q, z) = \frac{m^F(q, t = 0)}{z + 1 - \gamma \cos q} \quad (4.39)$$

Vamos obter agora a função de Green do problema correspondente a uma condição inicial com o spin na origem apontando para cima $\sigma_0 = 1$ e os

outros em orientações completamente aleatórias. A solução então satisfaz $m_n(0) = G_n(t=0) = \delta(n,0)$ de forma que $G^F(q, t=0) = 1$. Assim a equação (4.39) fica da forma:

$$G^{FL}(q, z) = \frac{1}{z + 1 - \gamma \cos q} \quad (4.40)$$

Antitransformando Laplace:

$$G^F(q, t) = e^{(\gamma \cos q - 1)t} \quad (4.41)$$

e antitransformando Fourier:

$$G_n(t) = e^{-t} \int_0^{2\pi} \frac{dq}{2\pi} e^{\gamma t \cos q + inq} = e^{-t} I_n(\gamma t) \quad (4.42)$$

onde I_n é uma função de Bessel modificada. Agora, invertendo a equação (4.39) obtemos a solução geral para a magnetização em função do tempo como uma convolução no espaço:

$$m_n(t) = m_n(t=0) * G_n(t) = \sum_m m_m(t=0) G_{n-m}(t) \quad (4.43)$$

Em particular se o sistema está inicialmente magnetizado de forma homogênea $m_n(t=0) = m_0 \forall n$ a relaxação é exponencial:

$$m_n(t) = m_0 \sum_n G_n(t) = m_0 G^F(q=0, t) = m_0 e^{-t/\tau_{eq}} \quad (4.44)$$

com τ_{eq} dado por (4.30). Notamos que o tempo de relaxação diverge quando $T \rightarrow 0$. A seguir vamos considerar o comportamento para tempos e distâncias grandes tais que se satisfaz:

$$t \sim \tau_{eq} \gg 1 \quad n \sim \xi_{eq} \gg 1 \quad (4.45)$$

com a razão entre estas grandezas um número finito arbitrário. Neste regime a lei de escala (4.32) implica:

$$t \sim \tau_{eq} \sim n^2 \sim \xi_{eq}^2 \gg 1 \quad (4.46)$$

Como conseqüência $z \approx q^2 \approx \mu^2$ são pequenas e a equação (4.40) se simplifica na forma:

$$G^{FL}(q, z) \approx \frac{2}{2z + q^2 + \mu^2} \quad (4.47)$$

Antitransformando obtemos:

$$G_n(t) \approx \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{t}{\tau_{eq}} - \frac{n^2}{2t}\right) \quad (4.48)$$

que é a aproximação ao comportamento de escala da função de Green (4.42). Notamos que as variáveis relevantes neste regime são t/τ_{eq} e n/\sqrt{t} , o que está de acordo com (4.46). Notamos que a $T = 0$ o primeiro termo da exponencial se anula e o processo de relaxação não atinge jamais o equilíbrio se comportando como um processo difusivo auto-similar como se observa da variável de escala n/\sqrt{t} .

4.3.2 Correlações a tempos iguais

Vamos analisar a seguir as correlações de dois pontos a tempos iguais definidas na forma:

$$C(i, j, t) = \langle \sigma_i(t) \sigma_j(t) \rangle \quad (4.49)$$

Como assumimos invariância por translações espaciais podemos escrever:

$$C(i, j, t) = C_{j-i}(t). \quad (4.50)$$

Vamos então resolver as equações dinâmicas dadas por (4.7) assumindo uma condição inicial completamente aleatória, o que equivale a estudar um resfriamento rápido desde um estado de temperatura infinita. A equação diferencial fica da forma:

$$\frac{dC_n(t)}{dt} = -2C_n(t) + \gamma(C_{n-1}(t) + C_{n+1}(t)), \quad n \neq 0 \quad (4.51)$$

com a condição inicial $C_n(0) = \delta_{n,0}$ e $C_0(t) = 1$. Para poder resolver a equação pela técnica das transformadas vamos completar o conjunto com o caso $n = 0$ somando um termo da forma $v(t)\delta_{n,0}$ onde $v(t)$ é uma função a determinar com a condição $C_0(t) = 1$, ou seja,

$$\frac{dC_n(t)}{dt} = -2C_n(t) + \gamma(C_{n-1}(t) + C_{n+1}(t)) + v(t)\delta_{n,0} \quad (4.52)$$

Transformando Fourier-Laplace:

$$zC^{FL}(q, z) = 2(\gamma \cos q - 1)C^{FL}(q, z) + v^L(z) + 1 \quad (4.53)$$

de onde

$$C^{FL}(q, z) = \frac{v^L(z) + 1}{z + 2 - 2\gamma \cos q} \quad (4.54)$$

A condição $C_0(t) = 1$ transformada Laplace equivale a:

$$C_0^L(z) = \frac{1}{z} = \int_0^{2\pi} \frac{dq}{2\pi} C^{FL}(q, z) = \frac{v^L(z) + 1}{\sqrt{(z + 2)^2 - 4\gamma^2}} \quad (4.55)$$

e então

$$v^L(z) = \frac{\sqrt{(z + 2)^2 - 4\gamma^2}}{z} - 1. \quad (4.56)$$

Com este resultado obtemos

$$C^{FL}(q, z) = \frac{\sqrt{(z + 2)^2 - 4\gamma^2}}{z(z + 2 - 2\gamma \cos q)} \quad (4.57)$$

e anti-transformando Fourier,

$$C_n^L(z) = \frac{1}{z} \left(\frac{z + 2 - \sqrt{(z + 2)^2 - 4\gamma^2}}{2\gamma} \right)^{|n|}. \quad (4.58)$$

Correlações de equilíbrio

A solução de equilíbrio para qualquer temperatura finita se pode obter considerando um tempo suficientemente grande $t \gg \tau_{eq}$ que equivale o tomar o limite $\lim_{z \rightarrow 0} z C_n^L(z)$ obtendo:

$$C_{n,eq} = \left(\frac{1 - \sqrt{1 - \gamma^2}}{\gamma} \right)^{|n|} = e^{-\mu|n|} = e^{-|n|/\xi_{eq}} \quad (4.59)$$

que é um resultado conhecido da solução de equilíbrio do modelo [15]. A qualquer temperatura finita as correlações espaciais decaem exponencialmente com uma distância característica, ou comprimento de correlação: ξ_{eq} .

Aproximação ao equilíbrio, temperaturas baixas, tempos longos

Vamos analisar agora a solução fora do equilíbrio no regime de tempos longos e temperaturas baixas tal que $z \sim q^2 \sim \mu^2 \ll 1$. Neste regime a função de correlação escala como:

$$C^{FL}(q, z) \simeq \frac{2\sqrt{z + \mu^2}}{z(z + q^2 + \mu^2)} \quad (4.60)$$

Antitransformando outra vez:

$$C_n^L(z) \simeq \frac{e^{-|n|\sqrt{z+\mu^2}}}{z} \quad (4.61)$$

Uma última transformada inversa [14] nos permite obter a expressão das correlações neste regime dinâmico:

$$C_n(t) \simeq \frac{1}{2} \left\{ e^{\mu|n|} \operatorname{erfc} \left(\frac{|n|}{2\sqrt{t}} + \mu\sqrt{t} \right) + e^{-\mu|n|} \operatorname{erfc} \left(\frac{|n|}{2\sqrt{t}} - \mu\sqrt{t} \right) \right\} \quad (4.62)$$

onde a função erro é definida por

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-x^2} dx \quad (4.63)$$

e a função erro complementar

$$\operatorname{erfc}(x) = 1 - \operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-x^2} dx \quad (4.64)$$

Do último resultado podemos obter comportamentos interessantes em diversas aproximações. Por exemplo a $T = 0$ no regime de crescimento de domínios ($1 \ll t \ll \tau_{eq}$), $\mu = 0$ e a correlação se simplifica para:

$$C_n(t) \simeq \operatorname{erfc} \left(\frac{|n|}{2\sqrt{t}} \right) \quad (4.65)$$

onde observamos uma estrutura auto semelhante dos domínios na variável de escala n/\sqrt{t} .

Finalmente vamos analisar a *susceptibilidade magnética* definida por:

$$\chi(t) = \sum_n C_n(t) = C^F(q = 0, t) \quad (4.66)$$

A partir da (4.57) obtemos:

$$\chi^L(z) = \frac{1}{z} \sqrt{\frac{z+2+2\gamma}{z+2-2\gamma}} \quad (4.67)$$

Fazendo o limite $\lim_{z \rightarrow 0} z\chi^L(z)$ obtemos a susceptibilidade de equilíbrio:

$$\chi_{eq} = \sqrt{\frac{1+\gamma}{1-\gamma}} = \coth \frac{\mu}{2} = \frac{\gamma+1-\sqrt{1-\gamma^2}}{\gamma-1+\sqrt{1-\gamma^2}} = e^{2\beta J} \quad (4.68)$$

Vemos que a susceptibilidade de equilíbrio diverge em $T_c = 0$ indicando a presença da transição de fase.

Capítulo 5

Movimento Browniano II: a equação de Langevin

5.1 Caminhada aleatória nas velocidades: a equação de Langevin

Na descrição do movimento browniano feito até agora as variáveis relevantes utilizadas foram as coordenadas da partícula. Para ter uma descrição mais próxima do que seria uma formulação newtoniana teríamos que incluir as mudanças nos momentos ou velocidades devidas as colisões entre a partícula browniana e as partículas do fluido no qual está imersa. Em 1908 P. Langevin propôs uma descrição alternativa do movimento browniano, mais próxima das equações de Newton. A partir das observações do movimento browniano de partículas suspensas na superfície de fluidos era evidente que a partícula realizava uma espécie de movimento errático muito rápido, sem uma direção definida, provavelmente consequência das colisões das partículas do fluido com a partícula browniana. No entanto observando numa escala de tempo maior a partícula sofria um “arraste” produzido pela viscosidade do fluido. Langevin propôs uma equação de Newton *fenomenológica* para a partícula, da forma:

$$m \frac{dv}{dt} = -\frac{v}{B} + F(t) \quad (5.1)$$

onde o primeiro termo corresponde a uma força devida a *viscosidade* do meio, sendo B a *mobilidade* e $F(t)$ uma força aleatória devida as contínuas colisões das partículas do meio com a partícula browniana. Duas suposições

fundamentais são feitas:

1. A força média devida as colisões é nula:

$$\langle F(t) \rangle = 0 \quad (5.2)$$

2. As colisões sucessivas são independentes, estão descorrelacionadas, o que é expressado matematicamente da forma:

$$\langle F(t)F(t') \rangle = C\delta(t - t'). \quad (5.3)$$

A média é um promedio temporal, C é uma constante que mede a intensidade do ruído estocástico. Da primeira condição concluímos que a força média exercida sobre a partícula corresponde a força viscosa. A segunda condição implica que a partícula não guarda memória das sucessivas colisões.

Dividindo (5.1) por m a equação de Langevin se escreve:

$$\frac{dv}{dt} = -\gamma v + \eta(t) \quad (5.4)$$

onde $\gamma = 1/(mB)$ é o coeficiente de atrito e $\eta(t) = F(t)/m$ é um *ruído estocástico branco*. Esta variável estocástica, o ruído branco, por ser descorrelacionado tem uma distribuição gaussiana. Mais adiante veremos que se chama “branco” porque seu espectro de frequências contém todas as frequências possíveis por igual. Ele possui as propriedades:

$$\begin{aligned} \langle \eta(t) \rangle &= 0 \\ \langle \eta(t)\eta(t') \rangle &= \Gamma\delta(t - t') \end{aligned} \quad (5.5)$$

onde $\Gamma = C/m^2$. Podemos notar então que, por causa do termo de ruído estocástico, a equação de Langevin é um exemplo das chamadas *equações diferenciais estocásticas*. O termo de ruído por ser descorrelacionado varia de forma descontínua com o tempo e então a variação temporal da posição (ou velocidade) da partícula browniana não é suave, em geral as funções não são diferenciáveis. Isto exige um cálculo especial para tratar com estas equações. A evolução temporal obedecida pelo processo estocástico no movimento browniano, ou mais geralmente na equação de Langevin, se chama *processo de Wiener*. O processo de Wiener está relacionado ao ruído $\eta(t)$ na forma [16]:

$$W(t) = \int_0^t \eta(t') dt'. \quad (5.6)$$

Algumas propriedades importantes do processo de Wiener que são consequência da sua definição são:

1. $\langle W(t) \rangle = 0$.
2. $\langle W(t) W(t') \rangle = \Gamma \min(t, t')$.
3. $W(t)$ é um processo estocástico markoviano e gaussiano. O caráter markoviano segue por ser uma soma (integral) de processos markovianos, e a distribuição gaussiana segue também da definição e do Teorema Central do Limite. Como consequência, além de ter média nula sua variância resulta $\sigma^2 = \Gamma t$.
4. Chamando $W(t_0) = w_0$, das propriedades anteriores segue que a distribuição do incremento de Wiener

$$W(t) - w_0 = \int_{t_0}^t \eta(t') dt' \quad (5.7)$$

é gaussiana:

$$P(w - w_0, t - t_0) = \frac{1}{\sqrt{2\pi(t - t_0)}} \exp \left[-\frac{(w - w_0)^2}{2\Gamma(t - t_0)} \right], \quad (5.8)$$

e também pode ser considerada uma probabilidade condicional ou probabilidade de transição $P(w, t | w_0, t_0)$. Por tanto a largura da distribuição do incremento de Wiener $\Delta W = W(t + \Delta t) - W(t)$ será $\sigma_{\Delta W}^2 = \Gamma \Delta t$.

Estas propriedades fazem com que o processo de Wiener tenha características matemáticas singulares. Por exemplo, sua derivada temporal diverge:

$$\frac{dW}{dt} = \lim_{\Delta t \rightarrow 0} \frac{\Delta W(\Delta t)}{\Delta t} \simeq \lim_{\Delta t \rightarrow 0} \frac{\sqrt{\Delta t}}{\Delta t} \rightarrow \infty \quad (5.9)$$

Isto reflete o fato de que um processo estocástico não pode ser considerado como uma função usual. Wiener formalizou as técnicas de cálculo que permitem definir corretamente as propriedades matemáticas destes processos estocásticos. Para uma introdução aos processos de Wiener se pode ver [9, 17, 16]. Também fizeram importantes contribuições ao formalismo do movimento browniano Stratonovitch, Ito, Ornstein e Uhlenbeck, Langevin, Einstein e Perrin, dentre outros [18].

Para obter a solução da equação (5.4) escrevemos $v(t) = u(t)e^{-\gamma t}$ com $u(t)$ uma função a ser determinada. Substituindo em (5.4) obtemos que u deve satisfazer a equação:

$$\frac{du}{dt} = e^{\gamma t} \eta(t) \quad (5.10)$$

cujas solução é:

$$u = u_0 + \int_0^t e^{\gamma t'} \eta(t') dt' \quad (5.11)$$

e por tanto:

$$v(t) = v_0 e^{-\gamma t} + \int_0^t e^{-\gamma(t-t')} \eta(t') dt' \quad (5.12)$$

com condição inicial $v(t=0) = u(t=0) = u_0 = v_0$. Notamos que esta é uma equação determinista e vale para qualquer realização particular do ruído. No entanto, pela discussão acima, a forma mais correta de interpretar estes resultados é em forma probabilística. Vamos utilizar as propriedades do ruído (5.5) para obter a velocidade média e a velocidade quadrática média (v.q.m.) da partícula, que são grandezas bem definidas a partir das propriedades estatísticas do ruído:

$$\langle v \rangle = v_0 e^{-\gamma t} \quad (5.13)$$

Vemos que a velocidade média decai a zero exponencialmente rápido num tempo característico $\tau = 1/\gamma$. τ é o *tempo de relaxação*. Este resultado é consequência da viscosidade ou em geral da presença de *forças dissipativas*.

Para calcular a v.q.m. usamos os resultados (5.12) e (5.13) para obter:

$$(v - \langle v \rangle)^2 = e^{-2\gamma t} \int_0^t \int_0^t \eta(t') \eta(t'') e^{\gamma(t'+t'')} dt' dt'' \quad (5.14)$$

Usando agora (5.5)

$$\langle (v - \langle v \rangle)^2 \rangle = e^{-2\gamma t} \int_0^t \Gamma e^{2\gamma t'} dt' \quad (5.15)$$

Resolvendo a integral obtemos finalmente a variância da velocidade:

$$\langle v^2 \rangle - \langle v \rangle^2 = \frac{\Gamma}{2\gamma} (1 - e^{-2\gamma t}). \quad (5.16)$$

Vemos que ao contrário da velocidade média, a variância não se anula a tempos longos e se aproxima exponencialmente a:

$$\langle v^2 \rangle = \frac{\Gamma}{2\gamma} \quad (5.17)$$

Da teoria cinética dos gases sabemos que, para um sistema clássico em equilíbrio, vale o princípio de equipartição da energia, que para cada grau de liberdade equivale a:

$$\frac{1}{2}m \langle v^2 \rangle = \frac{1}{2}k_B T \quad (5.18)$$

onde k_B é a constante de Boltzmann e T a temperatura absoluta. Comparando os dois últimos resultados obtemos

$$\Gamma = \frac{2\gamma k_B T}{m} \quad (5.19)$$

que é conhecida como *relação de Einstein* ou *relação de flutuação-dissipação* já que prediz uma relação entre a intensidade das flutuações contidas em Γ com a intensidade da dissipação, contida em γ , quando o sistema se encontra em equilíbrio com o um banho térmico à temperatura T .

5.1.1 De volta as coordenadas: o deslocamento quadrático médio

Para obter o d.q.m. da partícula integramos a velocidade

$$x = x_0 + \int_0^t v(t') dt' \quad (5.20)$$

e fazendo uso de (5.12):

$$x = x_0 + v_0 \int_0^t e^{-\gamma t'} dt' + \int_0^t e^{-\gamma t'} \int_0^{t'} \eta(t'') e^{\gamma t''} dt'' dt' \quad (5.21)$$

Integrando o segundo termo e invertendo a ordem das integrais do terceiro obtemos

$$x = x_0 + v_0 \frac{1}{\gamma} (1 - e^{-\gamma t}) + \int_0^t \eta(t'') e^{\gamma t''} \int_{t''}^t e^{-\gamma t'} dt' dt'' \quad (5.22)$$

Integrando em t' obtemos

$$x = x_0 + v_0 \frac{1}{\gamma} (1 - e^{-\gamma t}) + \frac{1}{\gamma} \int_0^t \eta(t'') (1 - e^{\gamma(t''-t)}) dt'' \quad (5.23)$$

Até aqui o resultado é válido para qualquer seqüência do ruído $\eta(t)$. Usando as propriedades (5.5) obtemos o deslocamento médio da partícula:

$$\langle x(t) \rangle = x_0 + v_0 \frac{1}{\gamma} (1 - e^{-\gamma t}) \quad (5.24)$$

Este resultado reflete o fato de que $\langle x \rangle$ depende somente das condições iniciais, ou mais exatamente da evolução em tempos $t \ll 1/\gamma$. De fato nesta escala de tempos a velocidade média é $\langle v \rangle \approx v_0$ (ver 5.13) e a partícula se comporta como uma partícula livre.

Para calcular o deslocamento quadrático médio (d.q.m) procedemos como no caso das velocidades:

$$x - \langle x \rangle = \frac{1}{\gamma} \int_0^t \eta(t'') (1 - e^{\gamma(t''-t)}) dt'' \quad (5.25)$$

Elevando ao quadrado esta expressão e usando (5.5) obtemos:

$$\langle (x - \langle x \rangle)^2 \rangle = \frac{\Gamma}{\gamma^2} \int_0^t (1 - e^{\gamma(t'-t)})^2 dt' \quad (5.26)$$

Fazendo a integral obtemos finalmente:

$$\langle x^2 \rangle - \langle x \rangle^2 = \frac{\Gamma}{\gamma^2} \left\{ t - \frac{2}{\gamma} (1 - e^{-\gamma t}) + \frac{1}{2\gamma} (1 - e^{-2\gamma t}) \right\} \quad (5.27)$$

Para tempos muito grandes esta expressão é dominada pelo termo linear em t e se obtém o conhecido resultado:

$$\langle x^2 \rangle - \langle x \rangle^2 \approx Dt \quad (5.28)$$

onde $D = \frac{\Gamma}{\gamma^2}$ é a *constante de difusão*. Este último resultado põe em evidência a relação entre o movimento browniano de uma partícula e uma caminhada aleatória. A física por trás dos dois processos é essencialmente a mesma. No entanto, vemos que no formalismo de Langevin este resultado sai essencialmente das equações de Newton, assumindo a presença de uma força efetiva de caráter estocástico. Utilizando a relação (5.19) obtemos que, quando a partícula entra em equilíbrio com o meio:

$$D = \frac{2k_B T}{\gamma m} \quad (5.29)$$

que é outra forma da relação de flutuação-dissipação vista antes. No seu trabalho sobre o movimento browniano, Einstein propôs uma expressão para medir o *número de Avogadro*. Se a partícula browniana for esférica, a força de viscosidade obedece a lei de Stokes e a mobilidade é igual a $B = 1/6\pi\eta a$ onde η é o *coeficiente de viscosidade* do fluido e a o raio da partícula. Neste

caso, usando as relações anteriores temos que a constante de difusão de uma partícula esférica é dada por $D = k_B T / 3\pi\eta a$. A constante de Boltzmann é dada por $k_B = R/N$, sendo R a constante universal dos gases e N o número de Avogadro. Supondo R conhecida, e como D , T , η e a são acessíveis experimentalmente, a partir das relações anteriores é possível estimar o número de Avogadro. Em 1908, três anos após o paper de Einstein, Perrin realizou uma série de experimentos detalhados e obteve $N \approx 6.85 \cdot 10^{23} \text{ mol}^{-1}$, dando uma evidência definitiva da existência dos átomos e moléculas, coisa que até então continuava a ser debatida. O valor atualmente aceito para o número de Avogadro é $N \approx 6.02 \cdot 10^{23} \text{ mol}^{-1}$. Por este trabalho Perrin ganhou o premio Nobel em 1926 [18].

5.1.2 Balanço energético

Se multiplicamos (5.1) por v obtemos:

$$\frac{m}{2} \frac{d}{dt} v^2 = -\frac{v^2}{B} + vF \quad (5.30)$$

ou

$$\frac{d}{dt} \left(\frac{m}{2} v^2 \right) + vF_{at} = vF \quad (5.31)$$

onde $F_{at} = v/B$ é a força de atrito ou viscosa. Tomando a média:

$$\frac{dK}{dt} + P_{dis} = P \quad (5.32)$$

sendo K a energia cinética, P_{dis} a potência dissipada e P a potência transferida para a partícula. Vemos que existe um balanço entre a variação da energia cinética e as potências transferida para e dissipada pela partícula. Usando o valor já calculado do desvio quadrático médio da velocidade podemos calcular explicitamente o lado esquerdo da (5.32) e obtemos:

$$P = \frac{m\Gamma}{2} = \frac{k_B T}{mB} \quad (5.33)$$

o que indica que a potência transferida para a partícula é independente do tempo. Ainda podemos notar que em geral a variação de energia cinética pode ser positiva, negativa ou nula dependendo da relação entre a K inicial e a de equilíbrio $k_B T/2$. No entanto no regime estacionário a energia cinética é constante e então a potência transferida deve ser igual a potência dissipada.

5.2 Funções de correlação

Para um intervalo de tempo dado a seqüência de valores das variáveis aleatórias $\{\eta(t)\}$, $\{x(t)\}$ e $\{v(t)\}$ representa um processo estocástico. Uma característica importante de um processo estocástico é o grau de correlação estatística entre valores das variáveis em diferentes instantes ao longo do processo. Para uma variável genérica $u(t)$ definimos a **função de autocorrelação** de dois tempos:

$$C(t, t') = \langle u(t)u(t') \rangle \quad (5.34)$$

Algumas propriedades importantes são:

- Se o processo estocástico é *estacionário*, o que quer dizer que as suas propriedades estatísticas não variam com o tempo, a autocorrelação depende só da diferença dos dois tempos $C(t'+s, t') = C(s)$, independente de t' .
- A autocorrelação a tempos iguais $C(t, t) = C(s=0)$ que é igual ao valor quadrático médio da variável u , é uma quantidade *positiva definida*. Em particular para um processo estacionário é uma constante independente de t : $C(0) = \text{const} > 0$.
- Se pode provar [19] que para qualquer valor de s , $C(s) \leq C(0)$.
- Se pode provar [19] que a autocorrelação é simétrica respeito de s : $C(s) = C(-s) = C(|s|)$

As funções de (auto)-correlação ocupam um papel fundamental tanto na mecânica estatística quanto em geral na teoria de processos estocásticos. Elas permitem obter informação importante sobre as mudanças na evolução temporal de um sistema físico qualquer. As funções de correlação são uma das quantidades fundamentais na formulação do Teorema de Flutuação-Dissipação, que veremos mais adiante.

5.3 Análise espectral de um processo estocástico: o teorema de Wiener-Kintchine

Uma forma muito comum de analisar a estrutura temporal de um processo dependente do tempo consiste em descrever as possíveis *frequências* que po-

dem dar lugar uma dada evolução temporal. Esta análise pode permitir distinguir, por exemplo, dois processos que embora possam apresentar a mesma média e desvio quadrático, sejam no entanto qualitativamente diferentes na sua estrutura temporal. Em sistemas lineares a técnica de análise do espectro de freqüências é conhecido como *análise de Fourier*. Vamos considerar processos estacionários nos quais a variável relevante $u(t)$ apresenta um desvio quadrático médio independente do tempo. É o caso, por exemplo, da velocidade de uma partícula browniana após equilibrar com o meio que a rodeia. Neste caso a variável é chamada *estatisticamente estacionária*. Não é o caso da posição da partícula browniana livre, já que o d.q.m. desta aumenta continuamente proporcionalmente ao tempo como já vimos.

Então consideremos o processo estacionário $\{u(t)\}$ durante um intervalo de tempo $0 \leq t \leq T$. Podemos desenvolver a função $u(t)$ em uma série de Fourier da forma:

$$u(t) = \sum_{n=-\infty}^{\infty} a_n e^{i w_n t} \quad (5.35)$$

onde as freqüências são:

$$w_n = \frac{2\pi n}{T} \quad (5.36)$$

correspondentes a um período T . A função $u(t)$ é considerada real, e então,

$$a_n = a'_n + i a''_n \quad a_{-n} = a_n^* = a'_n - i a''_n. \quad (5.37)$$

Os coeficientes de Fourier são dados por:

$$a_n = \frac{1}{T} \int_0^T u(t) e^{-i w_n t} dt \quad (5.38)$$

Se a função $u(t)$ fosse perfeitamente periódica os coeficientes a_n seriam números bem definidos e definiriam sem incerteza nenhuma o espectro de freqüências de $u(t)$. No entanto como $u(t)$ é uma variável aleatória os coeficientes a_n também o serão. Neste caso se pode pensar que o intervalo de periodicidade da função é infinito. Definimos então:

$$a_0 = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T u(t) dt, \quad (5.39)$$

onde a_0 é igual a média temporal da variável u . Vamos assumir no seguinte que $a_0 = 0$ ou, de forma mais geral, que na variável u já foi subtraída a

média. Também se verifica que, para todo T , $\langle a_n \rangle = 0$. O quadrado médio dos coeficientes de Fourier é definido por:

$$\langle |a_n|^2 \rangle = \langle |a'_n|^2 \rangle + \langle |a''_n|^2 \rangle \quad (5.40)$$

Se forem filtradas as frequências de forma a deixar passar uma banda de largura $\Delta\omega$ se define a *intensidade média* da banda como:

$$I(\omega)\Delta\omega = \sum_{\omega_n \in \Delta\omega} \langle |a_n|^2 \rangle \quad (5.41)$$

O número de modos na banda de largura $\Delta\omega$ é $\Delta\omega/(\omega_{n+1}-\omega_n) = \Delta\omega/(2\pi/T) = \frac{T}{2\pi}\Delta\omega$. Assumindo que temos um contínuo de frequências, passamos ao contínuo e definimos o *espectro de potência* do processo $u(t)$ como:

$$I(\omega) = \lim_{T \rightarrow \infty} \frac{T}{2\pi} \langle |a_n|^2 \rangle \quad (5.42)$$

Existe uma relação importante entre o espectro de potências de um processo estocástico e a função de autocorrelação do processo. Esta relação se conhece como *Teorema de Wiener-Kintchine*:

Teorema 1 *Dado um processo estacionário $u(t)$ se verifica que:*

$$I_u(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} C_u(t) e^{-i\omega t} dt \quad (5.43)$$

Demonstração

De (5.38) temos que:

$$\langle |a_n|^2 \rangle = \frac{1}{T^2} \int_0^T dt_1 \int_0^T dt_2 \langle u(t_1)u(t_2) \rangle \exp[-i\omega_n(t_1 - t_2)] \quad (5.44)$$

Como o integrando depende só da diferença $t_1 - t_2$ para resolver a dupla integral é conveniente fazer um cambio de variáveis e integrar em duas etapas: para $t_1 > t_2$ integramos na nova variável $t = t_1 - t_2$ e em t_2 . A integral em t_2 tem por limites 0 e $T - t$, do que resulta:

$$\int_0^T (T - t) C(t) e^{-i\omega_n t} dt \quad (5.45)$$

Na região $t_1 < t_2$ obtemos

$$\int_0^T (T - t) C(-t) e^{i\omega_n t} dt \quad (5.46)$$

Introduzimos estes resultados em (5.42) e tomamos o limite $T \rightarrow \infty$. Assumindo que as integrais que restam são convergentes obtemos:

$$I(\omega) = \frac{1}{2\pi} \left[\int_0^\infty C(t)e^{-i\omega t} dt + \int_0^\infty C(-t)e^{i\omega t} dt \right] \quad (5.47)$$

Finalmente convertendo a última integral a uma no intervalo $(-\infty, 0)$ obtemos (5.43).

O teorema de Wiener-Kintchine nos diz que se conhecemos o espectro de potência de um processo estocástico conhecemos a função de correlação, e viceversa. Antitransformando Fourier obtemos:

$$C(t) = \int_{-\infty}^\infty I(\omega)e^{i\omega t} d\omega \quad (5.48)$$

Este é um resultado muito poderoso para o cálculo de propriedades dinâmicas de processos estocásticos estacionários. Vamos ver alguns exemplos:

- Suponhamos que temos um processo cuja variável $u(t)$ é muito irregular, de forma que a evolução dela é praticamente imprevisível. Neste caso as correlações decairão imediatamente e podemos supor que no caso extremo:

$$C(t) = \Gamma\delta(t) \quad (5.49)$$

É o caso do ruído estocástico considerado na equação de Langevin (5.4). Para o espectro de potências obtemos

$$I(\omega) = \frac{\Gamma}{2\pi} = cte \quad (5.50)$$

A intensidade é independente da frequência. Um espectro de potência com esta característica se chama “ruído branco”. No entanto na prática um espectro estritamente branco é inaceitável já que o valor inicial da função de correlação diverge $C(0) = \infty$. Em lugar deste comportamento extremo na prática $C(t)$ será uma função concentrada em torno da origem mas que definirá uma *banda de frequências* permitidas. Por exemplo se

$$C(t) = C(0) \frac{\sin(at)}{at}, \quad (a > 0) \quad (5.51)$$

obtemos

$$I(\omega) = \frac{2\pi}{a} C(0) \quad se \quad w < a/2\pi \quad (5.52)$$

$$= 0 \quad se \quad w > a/2\pi \quad (5.53)$$

No limite $a \rightarrow \infty$ recuperamos o caso ideal da delta na origem.

- O caso contrario ao anterior no qual temos um processo muito regular, ou seja, preeditível, a função de correlação irá se estender para valores grandes de t e o espectro de frequências estará caracterizado por uma série de picos nos valores de algumas poucas frequências características. O caso mais simples é o de um espectro com uma frequência só, ou “monocromático”. Neste caso a correlação será do tipo:

$$C(t) = C(0) \cos(\omega_0 t) \quad (5.54)$$

e o espectro será:

$$I(\omega) = C(0)\delta(\omega - \omega_0) \quad (5.55)$$

- Consideremos a equação de Langevin básica (5.4). Podemos resolver a equação fazendo uma análise de Fourier. Desenvolvamos o ruído em série de Fourier:

$$\eta(t) = \sum_{n=-\infty}^{\infty} \eta_n e^{i\omega_n t} \quad (5.56)$$

e também a velocidade:

$$v(t) = \sum_{n=-\infty}^{\infty} v_n e^{i\omega_n t} \quad (5.57)$$

Substituindo na equação diferencial transformamos ela numa equação algébrica:

$$v_n = \frac{1}{i\omega_n + \gamma} \eta_n \quad (5.58)$$

de forma que o espectro de potências da velocidade fica (ver 5.42):

$$I_v(\omega) = \frac{1}{|i\omega_n + \gamma|^2} I_\eta(\omega) \quad (5.59)$$

Como $\eta(t)$ é um ruído branco $I_\eta(\omega) = \Gamma/(2\pi)$ e então:

$$C_v(t) = \frac{\Gamma}{2\pi} \int_{-\infty}^{\infty} \frac{e^{i\omega t} d\omega}{\omega^2 + \gamma^2} \quad (5.60)$$

Resolvendo a integral pelo teorema dos resíduos obtemos:

$$C(\tau) = \langle v(t)v(t + \tau) \rangle = \frac{\Gamma}{2\gamma} e^{-\gamma\tau}. \quad (5.61)$$

Vemos que a função de correlação de dois tempos da velocidade de uma partícula browniana livre decai exponencialmente com o tempo. Deste resultado podemos obter como caso particular quando $\tau = 0$ o desvio quadrático médio da velocidade:

$$C(0) = \langle v(t)^2 \rangle = \frac{\Gamma}{2\gamma} \quad (5.62)$$

que coincide com o resultado (5.17) obtido da solução direta da equação de Langevin. É importante notar que embora este caso simples da caminhada aleatória livre possa ser resolvido de diversas formas, isto não será assim para um sistema mais geral e então será nestes casos onde o resultado do teorema de Wiener- Kintchin é de grande utilidade.

Finalmente, é importante notar que a análise de Fourier é uma ferramenta poderosa para o estudo de **sistemas lineares no regime estacionário**. No entanto deveremos apelar para outros métodos no caso de sistemas não-lineares ou processos não estacionários.

5.4 Conjunto de equações de Langevin acopladas

Em um sistema formado por um número grande de graus de liberdade, como por exemplo o conjunto de variáveis microscópicas de um sistema de muitos corpos, a dinâmica do mesmo pode ser descrita pelo conjunto de equações de Langevin:

$$\frac{d\phi_i}{dt} = f_i(\phi_1, \phi_2, \dots, \phi_N) + \xi_i(t) \quad (5.63)$$

com $i = 1, 2, \dots, N$ e as variáveis do ruído branco obedecendo as seguintes propriedades:

$$\begin{aligned} \langle \xi_i(t) \rangle &= 0 \\ \langle \xi_i(t) \xi_j(t') \rangle &= \Gamma_{ij} \delta(t - t'), \end{aligned} \quad (5.64)$$

onde Γ_{ij} são constantes.

Um exemplo simples é dado por uma partícula que se move ao longo do eixo x realizando um movimento browniano e sujeita a uma força externa

elástica. As equações de Langevin que descrevem o movimento serão:

$$\begin{aligned}\frac{dx}{dt} &= v \\ \frac{dv}{dt} &= -\gamma v - \omega^2 x + \eta(t)\end{aligned}\quad (5.65)$$

onde $\omega^2 = k/m$ e k é a constante elástica. Podemos identificar $x_1 = x$, $f_1(x, v) = v$, $x_2 = v$ e $f_2(x, v) = -\gamma v - \omega^2 x$. Fixando $\Gamma_{11} = \Gamma_{12} = \Gamma_{21} = 0$ e $\Gamma_{22} = \Gamma$, podemos notar que as equações anteriores se ajustam à forma (5.63) com $N = 2$. Casos muito comuns são dados quando queremos descrever a dinâmica de relaxação dos spins em sistemas magnéticos ou de variáveis dinâmicas microscópicas em sistemas em contato com um banho térmico.

5.5 Equações de Langevin Markovianas e Não-Markovianas

O caráter Markoviano no contexto da equação de Langevin faz referência ao fato de que o atrito ao tempo t é proporcional à velocidade no mesmo instante e que o ruído é branco ou suas correlações são dadas por uma função Delta. Muitos problemas reais não têm este caráter Markoviano: a força de atrito que atua no tempo t pode depender da história da velocidade $v(t)$ em tempos anteriores a t . Ou seja, o atrito pode ter “memória”. Neste caso o coeficiente de atrito γ é substituído por uma função de memória $K(t)$, de forma que a força de atrito no tempo t é dada por:

$$-\gamma v(t) \rightarrow -\int_{-\infty}^t ds K(t-s) v(s) \quad (5.66)$$

ou cambiando variáveis de s para $t-s$:

$$-\gamma v(t) \rightarrow -\int_0^{\infty} ds K(s) v(t-s). \quad (5.67)$$

Este efeito de memória faz com que o ruído não seja mais branco e a relação de flutuação-dissipação (5.5) deve ser modificada. Este tipo de problemas se chamam Não-Markovianos.

Um exemplo de como pode surgir um comportamento não markoviano é eliminando a velocidade nas equações do movimento browniano de um

oscilador harmônico (5.65). Vamos supor uma condição inicial tal que a velocidade é nula no pasado infinito $v(-\infty) = 0$. Integrando a segunda das equações (5.65):

$$\begin{aligned} v(t) &= \int_{-\infty}^t ds e^{-\gamma(t-s)} (-\omega^2 x(s) + \eta(s)) \\ &= \int_0^{\infty} ds e^{-\gamma s} (-\omega^2 x(t-s) + \eta(t-s)) \end{aligned} \quad (5.68)$$

Agora substituindo este resultado na primeira equação (5.65) obtemos:

$$\frac{dx(t)}{dt} = - \int_0^{\infty} ds K(s) x(t-s) + F(t), \quad (5.69)$$

onde o núcleo de memória $K(s)$ e a força estocástica $F(t)$ são agora dados por:

$$\begin{aligned} K(t) &= \omega^2 e^{-\gamma|t|} \\ F(t) &= \int_0^{\infty} ds e^{-\gamma s} \eta(t-s). \end{aligned} \quad (5.70)$$

Usando as propriedades do ruído branco $\eta(t)$ e notando que, em equilíbrio, o segundo momento da variável x satisfaz:

$$\langle x^2 \rangle_{eq} = \frac{k_B T}{m\omega^2} \quad (5.71)$$

podemos expressar as correlações do novo ruído na forma:

$$\langle F(t)F(t') \rangle = \langle x^2 \rangle_{eq} K(|t-t'|). \quad (5.72)$$

Esta última é uma versão não markoviana do Teorema de Flutuação-Dissipação. Notamos que o novo ruído não é branco, a função de correlação é proporcional ao núcleo de memória da força de atrito.

No limite de atrito muito grande e para tempos muito maiores que $1/\gamma$ a função $K(t)$ pode ser aproximada por uma função Delta de Dirac preservando a área:

$$K(s) \approx 2 \frac{\omega^2}{\gamma} \delta(s) \quad (5.73)$$

que corresponde a um atrito markoviano. Neste limite a equação (5.69) se torna uma equação de Langevin aproximadamente markoviana para a posição $x(t)$ da partícula.

Neste exercício notamos que quando eliminamos alguma variável de um sistema de equações markovianas, o resultado é não markoviano. A inversa é válida e é importante lembrar: um processo não markoviano pode ser transformado num processo markoviano acrescentando uma variável ao problema.

Equações de Langevin não markovianas surgem naturalmente no estudo de sistemas desordenados, devido à dinâmica lenta que se reflete em fortes efeitos de memória no sistema. É possível mostrar, por exemplo, que partindo de um sistema de equações de Langevin com ruído branco para os spins num modelo de vidro de spin arrivamos a um modelo efetivo onde as correlações do novo ruído não são mais dadas por funções Delta mas por núcleos de memória que se originam na desordem das interações do Hamiltoniano do sistema. Estes sistemas apresentam em geral características de difusão anômala, também consequência da memória presente nos termos dissipativos.

5.6 Partícula browniana num banho térmico de osciladores harmônicos

Uma forma ilustrativa de como pode surgir uma equação de Langevin a partir de uma dinâmica Hamiltoniana consiste em considerar um sistema arbitrário caracterizado por coordenadas e momentos $\{x, p\}$ interagindo com um sistema de osciladores harmônicos independentes através de uma interação bilinear. Vamos ver como é possível derivar de forma exata uma equação de Langevin e ainda qual é neste sistema a origem da força estocástica e o caráter dos termos dissipativos que podem dar lugar a uma dinâmica markoviana ou não-markoviana.

O banho de osciladores é descrito por um conjunto de coordenadas $\{q_i\}$ e seus momentos conjugados $\{p_i\}$. Para simplificar a notação fixamos as massas dos osciladores iguais a um. O Hamiltoniano do sistema é:

$$\mathcal{H}_S = \frac{p^2}{2m} + U(x). \quad (5.74)$$

O Hamiltoniano do banho térmico inclui os osciladores e um acoplamento com o sistema na forma:

$$\mathcal{H}_B = \sum_j \left[\frac{p_j^2}{2} + \frac{1}{2} \omega_j^2 \left(q_j - \frac{\gamma_j}{\omega_j^2} x \right)^2 \right], \quad (5.75)$$

no qual ω_j são as frequências dos osciladores e γ_j mede a intensidade do acoplamento do sistema com o j -ésimo oscilador. O Hamiltoniano do banho térmico possui três termos: o primeiro é o Hamiltoniano usual do oscilador harmônico, o segundo é um acoplamento bilinear da forma $(\sum_j \gamma_j q_j)x$ entre os osciladores e o sistema, e o último depende só do sistema e pode ser considerado como parte do potencial arbitrário $U(x)$. Considerando o Hamiltoniano completo $\mathcal{H} = \mathcal{H}_S + \mathcal{H}_B$ do sistema+banho térmico, as equações de Hamilton do sistema serão:

$$\begin{aligned}\frac{dx}{dt} &= \frac{p}{m} \\ \frac{dp}{dt} &= -U'(x) + \sum_j \gamma_j \left(q_j - \frac{\gamma_j}{\omega_j^2} x \right),\end{aligned}\quad (5.76)$$

e as do banho térmico:

$$\begin{aligned}\frac{dq_j}{dt} &= p_j \\ \frac{dp_j}{dt} &= -\omega_j^2 q_j + \gamma_j x.\end{aligned}\quad (5.77)$$

Para $x(t)$ dada é fácil resolver as equações para os osciladores:

$$q_j(t) = q_j(0) \cos \omega_j t + p_j(0) \frac{\sin \omega_j t}{\omega_j} + \gamma_j \int_0^t ds x(s) \frac{\sin \omega_j(t-s)}{\omega_j} \quad (5.78)$$

Integrando por partes podemos expressar o resultado anterior da forma:

$$q_j(t) - \frac{\gamma_j}{\omega_j^2} x(t) = \left(q_j(0) - \frac{\gamma_j}{\omega_j^2} x(0) \right) \cos \omega_j t + p_j(0) \frac{\sin \omega_j t}{\omega_j} - \gamma_j \int_0^t ds \frac{p(s)}{m} \frac{\cos \omega_j(t-s)}{\omega_j^2}. \quad (5.79)$$

Substituindo este resultado na equação (5.76) para o sistema obtemos uma equação de Langevin:

$$\frac{dp(t)}{dt} = -U'(x(t)) - \int_0^t ds K(s) \frac{p(t-s)}{m} + F_p(t), \quad (5.80)$$

na qual podemos identificar o núcleo de memória como:

$$K(t) = \sum_j \frac{\gamma_j^2}{\omega_j^2} \cos \omega_j t \quad (5.81)$$

e o "ruído" $F_p(t)$ é dado explicitamente por:

$$F_p(t) = \sum_j \gamma_j p_j(0) \frac{\sin \omega_j t}{\omega_j} + \sum_j \gamma_j \left(q_j(0) - \frac{\gamma_j}{\omega_j^2} x(0) \right) \cos \omega_j t. \quad (5.82)$$

Notamos que fazendo uma escolha adequada do espectro $\{\omega_j\}$ dos osciladores e das constantes de acoplamento com o sistema $\{\gamma_j\}$ podemos obter um núcleo de memória de forma arbitrária. Por exemplo, se o espectro é contínuo e a soma em j é substituída por uma integral $\sum_j \rightarrow \int dw g(w)$, com $g(w)$ a densidade de estados do banho térmico, e supondo em geral que $\gamma = \gamma(w)$ é uma função de w , a função de memória é dada por uma integral de Fourier:

$$K(t) = \int_0^\infty dw g(w) \frac{\gamma(w)^2}{\omega^2} \cos \omega t. \quad (5.83)$$

O termo de memória em geral decai numa escala de tempo característica τ . Experimentalmente, se o tempo de observação for muito menor que o tempo característico $t \ll \tau$, teremos um forte efeito de memória e o ruído neste caso se chama *colorido*. No extremo oposto, se o tempo de observação for muito maior que o tempo característico $t \gg \tau$, o efeito de memória terá se perdido completamente, e o ruído então será *branco* ou descorrelacionado. Um exemplo simples é considerar o modelo de Debye de sólido cristalino elástico. A distribuição de freqüências de Debye é:

$$g(w) = \frac{3w^2}{\omega_D^3} \theta(\omega_D - w). \quad (5.84)$$

Considerando $\gamma(w) = \gamma$ constante, obtemos para o núcleo de memória:

$$K(t) = \frac{3\gamma^3}{m\omega_D^2} \frac{\sin \omega_D t}{\omega_D t}. \quad (5.85)$$

Se a freqüência de Debye ω_D for suficientemente grande podemos aproximar $K(t)$ por uma função Delta de Dirac $K(t) \approx 2\beta\delta(t)$ com $\beta = 3\gamma^3\pi/(2m\omega_D^2)$ obtendo explicitamente o caso correspondente a um ruído branco.

Finalmente, notemos que a expressão do ruído (5.82) depende exclusivamente das coordenadas e momentos iniciais dos osciladores, e então é uma função do tempo conhecida. No entanto como os osciladores são muitos e independentes entre si, pelo teorema central do limite esperamos que apresentem propriedades estatísticas simples. Dependendo das propriedades estatísticas das coordenadas e momentos iniciais serão as propriedades do ruído

$F_p(t)$. O caso mais simples é quando o ruído é gaussiano, que seria o caso quando as coordenadas e momentos iniciais dos osciladores são variáveis aleatórias independentes e identicamente distribuídas, então o teorema central do limite nos diz que a soma delas se aproxima de uma gaussiana quando o número for muito grande.

5.7 Lista de Exercícios

1. Calcule a função de autocorrelação de dois tempos das velocidades de uma partícula livre em movimento browniano. Mostre que, em geral:

$$\langle v(t_w)v(t_w+t) \rangle = e^{-\gamma|t|} \langle v^2(t_w) \rangle \quad (5.86)$$

onde t_w (waiting time) é o tempo de espera. A partir deste resultado determine a autocorrelação de equilíbrio definida por:

$$C(t) = \lim_{t_w \rightarrow \infty} \langle v(t_w)v(t_w+t) \rangle \quad (5.87)$$

e o correspondente espectro de potências: $C(\omega) = \frac{1}{2\pi} \int e^{-i\omega t} C(t) dt$.

2. Determine explicitamente como funções do tempo a variação da energia cinética média dE_c/dt e a potência média dissipada P_{dis} para o caso da partícula livre que executa movimento browniano. Mostre que a soma $dE_c/dt + P_{dis} = P$, onde a potência transferida P é independente do tempo. Faça um gráfico dessas três grandezas versus t .
3. Pospisil reportou em Ann. Phys. **83**, (1927) 735 a observação do movimento browniano de partículas de fumaça de raio $a = 0.4 \times 10^{-4}$ cm submersas em uma solução de água-glicerina, de viscosidade 0.0278 poise, a uma temperatura de 18.8°C. O valor observado de $\langle x^2 \rangle$ num intervalo de 10 seg foi 3.3×10^{-8} cm². Com estes dados determine a constante de Boltzmann e compare com valores atuais tabelados.
4. A equação de Langevin para uma partícula de massa m em movimento browniano sujeita a uma força elástica $-m\omega_0^2 x$ é dada por:

$$\frac{d^2x}{dt^2} + \gamma \frac{dx}{dt} + \omega_0^2 x = \eta(t) \quad (5.88)$$

Calcular $\langle x^2(t) \rangle$ e $\langle v^2(t) \rangle$ e mostrar que, no limite $\omega_0 \rightarrow 0$ os resultados se reduzem as expressões (5.27) e (5.16) no caso $x_0 = v_0 = 0$. Discutir os resultados para o caso de uma partícula de massa desprezível $m \rightarrow 0$ (caso súper- amortecido ou não-inercial).

5. Faça uma análise de Fourier da equação de Langevin do problema anterior e calcule o espectro de potências e a função de autocorrelação das coordenadas. A partir do resultado obtido calcule o desvio quadrático médio.

Capítulo 6

A equação de Fokker-Planck

Vamos ver neste capítulo um formalismo equivalente e alternativo ao de Langevin para descrever a dinâmica de sistemas estocásticos. Veremos que a toda equação de Langevin corresponde uma equação diferencial que determina a evolução temporal da distribuição de probabilidades $P(u, t)$ da variável estocástica u . Esta equação diferencial estocástica, conhecida como equação de Fokker-Planck é de fundamental importância na análise da evolução da distribuição de probabilidades de um sistema estatístico para o equilíbrio assintótico ou estado estacionário. Em particular, para sistemas físicos, a solução estacionária da equação de Fokker-Planck corresponde à distribuição de Boltzmann-Gibbs para o sistema em equilíbrio termodinâmico. No entanto com a equação de Fokker-Planck temos acesso a toda a história da evolução temporal das probabilidades, permitindo analisar a forma como um sistema caminha para o equilíbrio final. A equação de Fokker-Planck pode ser considerada como um caso particular de equação mestra, para processos markovianos e quando as taxas de transição são pequenas [9]. No entanto aqui vamos deduzir a equação de Fokker-Planck por sua relação com a equação de Langevin.

6.1 De Langevin a Fokker-Planck

Nas seções seguintes vamos acompanhar de perto a exposição de T. Tomé e M. de Oliveira [10]. Começemos escrevendo a equação de Langevin no caso não-inercial na forma:

$$\frac{dx}{dt} = f(x) + \eta(t) \quad (6.1)$$

com $\langle \eta(t) \rangle = 0$ e $\langle \eta(t)\eta(t') \rangle = \Gamma \delta(t - t')$. Se discretizamos o tempo de forma que $t = n\tau$, e considerando as propriedades de um processo de Wiener vistas na seção 5.1, a versão discreta da equação pode ser escrita na forma:

$$x_{n+1} = x_n + \tau f(x_n) + \sqrt{\tau \Gamma} \xi_n, \quad (6.2)$$

com ξ_n variáveis aleatórias independentes com $\langle \xi_n \rangle = 0$ e $\langle \xi_n^2 \rangle = 1$. Desta forma podemos, por exemplo, gerar no computador uma “trajetória” particular do processo definido pela variável x_n , dada uma condição inicial x_0 e uma seqüência de números aleatórios ξ_n e “simular” desta forma uma dinâmica de Langevin.

Vamos mostrar agora que a esta equação de Langevin corresponde uma equação diferencial para a distribuição de probabilidades $P(x, t)$. Seja $P_n(x_n)$ a distribuição de probabilidades da variável x_n e $g_n(k)$ a correspondente função característica, dada por:

$$g_n(k) = \langle e^{ikx_n} \rangle = \int e^{ikx_n} P_n(x_n) dx_n \quad (6.3)$$

Então,

$$g_{n+1}(k) = \langle e^{ikx_{n+1}} \rangle = \langle e^{ik[x_n + \tau f(x_n) + \tau \zeta_n]} \rangle, \quad (6.4)$$

onde $\langle \zeta_n \rangle = 0$ e $\langle \zeta_n^2 \rangle = \Gamma/\tau$. Tendo em vista que x_n e ζ_n são independentes, podemos escrever

$$g_{n+1}(k) = \langle e^{ik[x_n + \tau f(x_n)]} \rangle \langle e^{ik\tau \zeta_n} \rangle. \quad (6.5)$$

Expandindo $g_{n+1}(k)$ até primeira ordem em τ o primeiro fator resulta:

$$\langle e^{ikx_n} \{1 + ik\tau f(x_n)\} \rangle = \langle e^{ikx_n} \rangle + ik\tau \langle f(x_n) e^{ikx_n} \rangle \quad (6.6)$$

e o segundo

$$1 + ik\tau \langle \zeta_n \rangle - \frac{1}{2} k^2 \tau^2 \langle \zeta_n^2 \rangle = 1 - \frac{1}{2} k^2 \tau \Gamma. \quad (6.7)$$

Logo, até primeira ordem,

$$g_{n+1}(k) = g_n(k) + \tau \left\{ ik \langle f(x_n) e^{ikx_n} \rangle - \frac{\Gamma}{2} k^2 \langle e^{ikx_n} \rangle \right\}. \quad (6.8)$$

Usando agora as propriedades:

$$ik \langle f(x_n) e^{ikx_n} \rangle = \langle f(x) \frac{d}{dx} e^{ikx} \rangle = - \int e^{ikx} \frac{d}{dx} [f(x) P_n(x)] dx \quad (6.9)$$

e

$$-k^2 \langle e^{ikx} \rangle = \left\langle \frac{d^2}{dx^2} e^{ikx} \right\rangle = \int e^{ikx} \frac{d^2}{dx^2} P_n(x) dx \quad (6.10)$$

concluimos que

$$\frac{P_{n+1}(x) - P_n(x)}{\tau} = -\frac{d}{dx}[f(x)P_n(x)] + \frac{\Gamma}{2} \frac{d^2}{dx^2} P_n(x). \quad (6.11)$$

Tomando o limite para $\tau \rightarrow 0$ obtemos finalmente

$$\frac{\partial}{\partial t} P(x, t) = -\frac{\partial}{\partial x}[f(x)P(x, t)] + \frac{\Gamma}{2} \frac{d^2}{dx^2} P(x, t). \quad (6.12)$$

Esta é uma equação diferencial para a evolução temporal da distribuição de probabilidades da variável x e é conhecida como *Equação de Fokker-Planck*. Nesta forma particular a equação de Fokker-Planck também é conhecida como *Equação de Smoluchowski*, já que foi M. von Smoluchowski o primeiro a estudar, em 1906, o caso de uma partícula browniana sujeita à ação de um campo externo, efeito que aparece no primeiro termo da direita da equação (6.12).

6.2 Solução estacionária

Podemos escrever a equação de Fokker-Planck (6.12) na forma:

$$\frac{\partial}{\partial t} P(x, t) = -\frac{\partial}{\partial x} J(x, t) \quad (6.13)$$

onde $J(x, t)$ é uma corrente de probabilidade definida por:

$$J(x, t) = f(x)P(x, t) - \frac{\Gamma}{2} \frac{\partial}{\partial x} P(x, t). \quad (6.14)$$

Na forma (6.13) a equação de Fokker-Planck é uma equação de continuidade. Vamos supor que a variável x tome valores num intervalo $[a, b]$, onde um ou ambos limites podem ser infinitos.

Integrando ambos os lados de (6.13) em x obtemos:

$$\frac{d}{dt} \int_a^b P(x, t) dx = J(a, t) - J(b, t). \quad (6.15)$$

A densidade de probabilidade $P(x, t)$ deve estar normalizada para todo tempo t , ou seja

$$\int_a^b P(x, t) dx = 1, \quad (6.16)$$

e então o lado esquerdo de (6.15) deve se anular, e por tanto $J(a, t) = J(b, t)$. Consideremos por simplicidade apenas o caso em que a corrente de probabilidade se anule em ambos extremos $x = a$ e $x = b$ em qualquer instante t , ou seja $J(a, t) = J(b, t) = 0$.

No regime estacionário a densidade de probabilidade será independente de t , e por tanto, por sua definição a corrente de probabilidade também deve ser independente de t neste regime. Como o lado esquerdo da (6.13) se anula, então a corrente também deverá ser independente de x . Logo ela deve se anular para todo valor de x dada a condição de contorno que escolhemos.

Por tanto a distribuição estacionária deve satisfazer:

$$f(x)P(x, t) - \frac{\Gamma}{2} \frac{d}{dx} P(x, t) = 0 \quad (6.17)$$

ou

$$\frac{d}{dx} \ln P(x) = \frac{2}{\Gamma} f(x). \quad (6.18)$$

Chamando $V(x)$ o potencial correspondente à força $f(x) = -dV/dx$ então

$$\ln P(x) = -\frac{2}{\Gamma} V(x) + \text{const}, \quad (6.19)$$

ou

$$P(x) = A \exp\left(-\frac{2}{\Gamma} V(x)\right) \quad (6.20)$$

com A uma constante de normalização. Notamos que identificando corretamente as constantes A e Γ a solução estacionária para o caso de um sistema físico num potencial externo $V(x)$ corresponde à distribuição Boltzmann-Gibbs. Dada a equivalência entre a equação de Fokker-Planck e a equação de Langevin, este resultado garante também que assintoticamente o processo descrito pela equação de Langevin com uma força externa derivada de um potencial leva o sistema para o equilíbrio termodinâmico.

6.3 Operador de evolução

Vamos introduzir nesta seção um operador de evolução \mathcal{W} de forma semelhante ao feito quando estudamos a equação mestra. Com ele é possível

mostrar que a densidade de probabilidade $P(x, t)$ tende ao valor estacionário obtido na seção anterior quando $t \rightarrow \infty$.

A equação de Fokker-Planck (6.12) pode ser escrita na forma:

$$\frac{\partial}{\partial t} P(x, t) = \mathcal{W} P(x, t), \quad (6.21)$$

onde o operador de evolução \mathcal{W} age sobre funções reais $\phi(x)$ e é definido por:

$$\mathcal{W} \phi(x) = -\frac{\partial}{\partial x} [f(x)\phi(x)] + \frac{\Gamma}{2} \frac{d^2}{dx^2} \phi(x). \quad (6.22)$$

De acordo com as condições de contorno impostas nas seções anteriores, a classe de funções $\phi(x)$ sobre as quais atua o operador \mathcal{W} devem satisfazer que a densidade de corrente de probabilidade $-f(x)\phi(x) + (\Gamma/2)\phi'(x) = 0$ nos dois extremos $x = a$ e $x = b$. Por tanto por (6.22) estas funções satisfazem a seguinte propriedade:

$$\int_a^b \mathcal{W} \phi(x) dx = 0. \quad (6.23)$$

Esta propriedade garante que a densidade $P(x, t)$ estará normalizada para todo tempo t desde que seja normalizada no instante inicial.

Com o recurso do operador de evolução podemos escrever a solução formal da equação de Fokker-Planck na forma:

$$P(x, t) = e^{t\mathcal{W}} P(x, 0) \quad (6.24)$$

onde o operador $e^{t\mathcal{W}}$ é definido por:

$$e^{t\mathcal{W}} = 1 + t\mathcal{W} + \frac{t^2}{2!}\mathcal{W}^2 + \frac{t^3}{3!}\mathcal{W}^3 + \dots \quad (6.25)$$

Assumimos agora que \mathcal{W} possua um espectro de autovalores discreto, de forma que:

$$\mathcal{W}\phi_l(x) = \Lambda_l \phi_l(x) \quad (6.26)$$

para $l = 0, 1, 2, \dots$, onde $\phi_l(x)$ são as autofunções de \mathcal{W} e Λ_l seus autovalores. Assumamos ainda que $P(x, 0)$ admita uma expansão em autofunções da forma:

$$P(x, 0) = \sum_{l=0}^{\infty} a_l \phi_l(x). \quad (6.27)$$

Então usando (6.24) obtemos:

$$P(x, t) = \sum_{l=0}^{\infty} a_l e^{t\Lambda_l} \phi_l(x). \quad (6.28)$$

Usando agora a propriedade (6.23) na equação de autovalores (6.26) podemos mostrar que:

$$\Lambda_l \int_a^b \phi_l(x) dx = 0 \quad (6.29)$$

A densidade de probabilidade estacionária satisfaz a equação:

$$\mathcal{W} P(x) = 0. \quad (6.30)$$

A equação (6.30) nos diz que a densidade de probabilidade estacionária é autofunção do operador de evolução com autovalor nulo. Vamos definir então $\phi_0(x) = P(x)$ e $\Lambda_0 = 0$. Assim podemos escrever:

$$P(x, t) = P(x) + \sum_{l=1}^{\infty} a_l e^{t\Lambda_l} \phi_l(x). \quad (6.31)$$

onde $a_0 = 1$ que se obtém integrando ambos os lados de (6.27) e usando (6.29).

É possível mostrar que todos os autovalores exceto Λ_0 são estritamente negativos, e por tanto toda a soma do lado direito de (6.31) se anula no limite $t \rightarrow \infty$. Então $P(x, t) \rightarrow P(x)$ assintoticamente. Por tanto, a dependência de $P(x, t)$ com o tempo, para tempos longos, será exponencial, e caracterizada pelo autovalor Λ_1 , ou seja:

$$P(x, t) \rightarrow P(x) + a_1 \phi_1(x) e^{-t\Lambda_1}, \quad (6.32)$$

desde que $a_1 \neq 0$. Por tanto o decaimento para o estado estacionário se dá numa escala de tempo característica $\tau = 1/|\Lambda_1|$. Pode acontecer que $\tau \rightarrow \infty$ em cujo caso o decaimento não é mais exponencial mas algébrico, caracterizando um sistema com dinâmica lenta na relaxação da densidade de probabilidade.

Exemplo

Consideremos o caso de uma partícula confinada no intervalo $-L/2 \leq x \leq$

$L/2$, na ausência de forças externas. Neste caso $f(x) = 0$ e temos que resolver a equação de autovalores:

$$\frac{\Gamma}{2} \frac{d^2}{dx^2} \phi(x) = \Lambda \phi(x) \quad (6.33)$$

com as condições de contorno $\phi'(-L/2) = \phi'(L/2) = 0$. As soluções para essas condições são:

$$\phi_l(x) = \begin{cases} L^{-1} \cos(kx) & l = 0, 2, 4, \dots \\ L^{-1} \sin(kx) & l = 1, 3, 5, \dots \end{cases} \quad (6.34)$$

$\Lambda_l = -\frac{\Gamma}{2} k^2$, $k = \pi l/L$ e a normalização foi escolhida de forma que $\phi_0(x) = P(x)$. A solução da equação de Fokker-Planck para o caso com condição inicial com a partícula na origem, isto é, $P(x, 0) = \delta(x)$, é:

$$P(x, t) = \frac{1}{L} + \frac{2}{L} \sum_{l=2(\text{pares})}^{\infty} e^{-t\Gamma k^2/2} \cos(kx). \quad (6.35)$$

Enquanto L for finito o tempo característico de relaxação será $\tau = 1/|\Lambda_2| = \frac{2}{\Gamma}(L/2\pi)^2$. Esse tempo diverge para $L \rightarrow \infty$. No entanto, nesse caso a solução da equação de Fokker-Planck é:

$$P(x, t) = \frac{1}{\pi} \int_0^{\infty} e^{-t\Gamma k^2/2} \cos(kx) dk = \frac{1}{\sqrt{2\pi\Gamma t}} e^{-x^2/2\Gamma t} \quad (6.36)$$

que corresponde ao conhecido processo de difusão normal num espaço ilimitado, e por tanto o decaimento da densidade de probabilidade será algébrico, $P(x, t) \simeq t^{-1/2}$.

6.4 O problema de Kramers

Um problema clássico, estudado pela primeira vez no contexto da equação de Fokker-Planck por Kramers em 1940, é o cálculo da probabilidade por unidade de tempo para uma partícula escapar de um poço de potencial passando através de uma barreira.

Este problema tem múltiplas aplicações e realizações, como por exemplo na cinética das reações químicas, na teoria de nucleação em transições de fases de primeira ordem, e em geral na relaxação em superfícies potenciais

com múltiplos mínimos, como em vidros, sistemas desordenados e problemas de otimização [20].

Vamos analisar aqui o mais simples dos casos citados acima, no qual a relaxação de uma partícula num potencial unidimensional $V(x)$ acontece no limite não-inercial e então obedece a equação de Smoluchowski (6.12). No problema em questão a partícula se encontra confinada no entorno de um mínimo do potencial, de coordenada x_A , e para escapar da atração desse mínimo deve superar uma barreira que no ponto máximo x_C tem uma altura U . Para além de x_C a partícula escapa por exemplo para o entorno de outro mínimo em $x_B > x_A$ (esta é somente uma possibilidade). O escape através da barreira se dá por ativação térmica, mas supomos que a energia térmica $k_B T \ll U$, e então a probabilidade de escape por unidade de tempo será pequena.

Vamos mostrar que, no limite não-inercial ou súper amortecido $\gamma/m \gg 1$, relevante por exemplo para estudar processos de relaxação, a taxa de escape é dada por:

$$\frac{1}{\tau} = \frac{\omega_A \omega_C}{2\pi} \frac{m e^{-U/kT}}{\gamma} = \nu e^{-U/kT}. \quad (6.37)$$

Os parâmetros que aparecem nesta equação são: a altura da barreira $U = V(x_C) - V(x_A)$, a frequência angular de vibração no mínimo metaestável $\omega_A^2 = m^{-1} V''(x_A)$, e a frequência angular de vibração no estado de transição $\omega_C^2 = m^{-1} |V''(x_C)|$.

Agora definimos a taxa de escape do estado metaestável em x_A para x_B como sendo $\kappa = 1/\tau$ tal que satisfaz:

$$J(x) = \kappa P(x) \quad (6.38)$$

onde $J(x)$ é a densidade de corrente de probabilidade e $P(x)$ a probabilidade de encontrar a partícula num entorno de x . Expressando as constantes na equação de Fokker-Planck através da relação de Einstein, a corrente pode ser escrita na forma:

$$\begin{aligned} J(x, t) &= -D \left(\frac{P(x, t)}{D\gamma} \frac{\partial V}{\partial x} + \frac{\partial P(x, t)}{\partial x} \right) \\ &= -D e^{-\frac{V(x)}{kT}} \left[\frac{\partial}{\partial x} e^{\frac{V(x)}{kT}} P(x, t) \right]. \end{aligned} \quad (6.39)$$

Seguidamente assumimos que a energia térmica disponível é pequena com-

parada com a altura da barreira de potencial, ou seja,

$$\frac{U}{kT} = \frac{V(x_C) - V(x_A)}{kT} \quad (6.40)$$

é muito grande. Logo a densidade de corrente através da barreira em x_C é muito pequena e a taxa de variação da densidade de probabilidade $P(x, t)$ também é pequena. Sendo assim podemos considerar $J(x, t) = J$ independente de x . Integrando (6.39) entre x_A e x_B obtemos:

$$J \int_{x_A}^{x_B} e^{\frac{V(x)}{kT}} dx = D \left[e^{\frac{V(x_A)}{kT}} P(x_A, t) - e^{\frac{V(x_B)}{kT}} P(x_B, t) \right]. \quad (6.41)$$

Agora assumindo que muito poucas partículas conseguem escapar para x_B podemos aproximar a equação anterior como:

$$J = D \frac{e^{\frac{V(x_A)}{kT}} P(x_A, t)}{\int_{x_A}^{x_B} e^{\frac{V(x)}{kT}} dx}. \quad (6.42)$$

Se a altura da barreira for muito grande, então a distribuição de probabilidade no entorno de x_A será aproximadamente a distribuição de Boltzmann-Gibbs:

$$P(x, t) = P(x_A, t) e^{-\frac{V(x) - V(x_A)}{kT}}. \quad (6.43)$$

Agora, a probabilidade P de encontrar a partícula no entorno $x_1 < x_A < x_2$ será:

$$P = \int_{x_1}^{x_2} P(x, t) dx = P(x_A, t) e^{\frac{V(x_A)}{kT}} \int_{x_1}^{x_2} e^{-\frac{V(x)}{kT}} dx \quad (6.44)$$

A partir de (6.42) e (6.44) obtemos a taxa de escape:

$$D\kappa^{-1} = \int_{x_1}^{x_2} e^{-\frac{V(x)}{kT}} dx \int_{x_A}^{x_B} e^{\frac{V(x)}{kT}} dx = I_1 I_2 \quad (6.45)$$

A contribuições principais nas integrais anteriores são nas regiões no entorno de x_A e x_C respectivamente. Podemos então fazer uma expansão dos potenciais $V(x)$ em série de Taylor na forma:

$$V(x) \approx V(x_A) + \frac{1}{2}(x - x_A)^2 V''(x_A) \quad (6.46)$$

e

$$V(x) \approx V(x_C) - \frac{1}{2}(x - x_C)^2 |V''(x_C)| \quad (6.47)$$

Agora estendendo os limites de integração para $\pm\infty$ obtemos:

$$I_1 = e^{-\frac{V(x_A)}{kT}} \int_{-\infty}^{\infty} e^{-\frac{(x-x_A)^2}{2kT}} V''(x_A) dx = \sqrt{2\pi} e^{-\frac{V(x_A)}{kT}} \sqrt{\frac{kT}{V''(x_A)}} \quad (6.48)$$

e

$$I_2 = \sqrt{2\pi} e^{\frac{V(x_C)}{kT}} \sqrt{\frac{kT}{|V''(x_C)|}}. \quad (6.49)$$

Por tanto finalmente obtemos:

$$\begin{aligned} \kappa &= \frac{1}{2\pi} \frac{\sqrt{V''(x_A)} \sqrt{|V''(x_C)|}}{\gamma} e^{-\frac{|V(x_C)-V(x_A)|}{kT}} \\ &= \frac{1}{2\pi} \frac{m}{\gamma} \omega_A \omega_C e^{-\frac{U}{kT}} \end{aligned} \quad (6.50)$$

Este resultado nos diz que os tempos característicos de relaxação num sistema caracterizado por barreiras de energia de altura U é dado pela *lei de Arrhenius*:

$$\tau \propto e^{U/kT} \quad (6.51)$$

ou seja, cresce exponencialmente com a altura da barreira.

O problema de Kramers pode ser generalizado para o caso com inércia e para sistemas com um número arbitrário de graus de liberdade. O problema geral estudado por Kramers ainda é um problema em aberto, cujas soluções só se conhecem em algumas aproximações.

Referências Bibliográficas

- [1] C. W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*, (Springer-Verlag, Berlin, 1985).
- [2] E. Marinari e G. Parisi, *Trattatello di Probabilità*, preprint 2000.
- [3] W. Feller, *An Introduction to Probability Theory and its Applications*, (J. Wiley, New York, 1950) volume I.
- [4] M. O. Cáceres, *Elementos de estadística de no equilibrio y sus aplicaciones al transporte en medios desordenados*, (Editorial Reverté, Barcelona, 2003).
- [5] B. D. Hughes, W. W. Montroll e M. F. Schlesinger, *Journal of Statistical Physics* **30**, 273 (1983).
- [6] *Einstein's random walk*, em *Physics World*, January 2005, pag. 19; P. Hanggi e F. Marchesoni, *100 Years of Brownian motion*, *Chaos* **15**, 026101 (2005).
- [7] C. Itzykson e J-M Drouffe, *Statistical Field Theory, vol.I*, (Cambridge University Press, 1989).
- [8] P. M. Chaikin e T. C. Lubensky, *Principles of Condensed Matter Physics*, (Cambridge University Press, 1995).
- [9] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry*, (North-Holland, 1992).
- [10] T. Tomé, M. J. de Oliveira, *Dinâmica Estocástica e Irreversibilidade*, (EDUSP, Editora da Universidade de São Paulo, 2001).

- [11] F. R. Gantmacher, *The Theory of Matrices*, vol.II, (Chelsea Publishing Co., 1971)
- [12] Watson, M. L. Glasser , I. J. Zucker, Proc. Natl. Acad. Sci. USA, **74**, (1977) 1800.
- [13] R. J. Glauber, J. Math. Phys. **4**, (1963) 294.
- [14] C. Godrèche, J. M. Luck, J. Phys. A: Math. Gen. **33**, (2000) 1151.
- [15] R. J. Baxter, *Exactly Solved Models in Statistical Mechanics*, (Academic Press, London, 1982).
- [16] C. Scherer, *Métodos Computacionais da Física*, Editora do Instituto de Física, UFRGS, 2002.
- [17] R. Kubo, M. Toda, N. Hashitsume, *Statistical Physics II, Nonequilibrium Statistical Mechanics*, (Springer Verlag, 1995).
- [18] W. T. Coffey, Y. P. Kalmykov e J. T. Waldron, *The Langevin Equation*, (World Scientific, Singapore, 1998).
- [19] R. K. Pathria, *Statistical Mechanics*, (Pergamon Press, 1977).
- [20] P. Hänggi, P. Talkner e M. Borkovec, *Reaction Rate Theory 50 Years After Kramers*, Rev. Mod. Phys. **62**, (1990) 251.