

O Teorema de Perron-Frobenius e a Ausência de Transição de Fase em Modelos Unidimensionais da Mecânica Estatística

Marcelo Richard Hilário

Aluno do Curso de Graduação em Física - UFMG,

mhilario@gold.com.br

Gastão Braga

Departamento de Matemática - UFMG

Caixa Postal 1621

30161-970 - Belo Horizonte - MG

gbraga@mat.ufmg.br

29 de setembro de 2004

Resumo

Nesse trabalho, apresentamos o Teorema de Perron-Frobenius e algumas de suas aplicações em Mecânica Estatística do Equilíbrio. Resumidamente, o teorema afirma que o maior autovalor de matrizes quadradas com todas as entradas positivas tem multiplicidade algébrica igual a 1. A sua conexão com a Mecânica Estatística é imediata, pois matrizes com todas as entradas positivas surgem, naturalmente, quando se avalia a função partição utilizando-se o método da matriz transferência. Mostraremos que, para uma grande classe de modelos de curto alcance, a pressão é dada pelo logaritmo natural do maior autovalor da matriz transferência. Depois mostraremos que, se a energia do sistema for uma função suave dos seus parâmetros então esse autovalor também o será. Como resultado, mostraremos que a pressão é uma função tão suave quanto a energia, com respeito a esses parâmetros; o que impossibilita a presença de transições de fase para esses modelos.

1 Introdução

Esse trabalho versa sobre o Teorema de Perron-Frobenius [1], [2], [3] e suas conseqüências em Mecânica Estatística do Equilíbrio. Resumidamente, o teorema afirma que o maior autovalor de uma matriz $n \times n$, com entradas positivas, é simples. Esse fato gera vários resultados importantes sobre a termodinâmica de sistemas físicos. Vamos argumentar que, para uma grande classe de sistemas de spins unidimensionais, a pressão (função a partir da

qual se determinam as grandezas fisicamente mensuráveis) é igual ao logaritmo natural do maior autovalor de uma matriz de entradas positivas (a matriz transferência). A conexão entre o resultado matemático e a descrição física desses sistemas se dá pelo seguinte: o fenômeno de transição de fase se manifesta como algum tipo de descontinuidade da pressão com respeito aos parâmetros do sistema. Contudo, o maior autovalor da matriz transferência é simples e recebe de "herança" a suavidade da energia de interação do sistema (se a energia de interação for de classe C^∞ então a energia livre também o será) e, portanto, nenhuma transição de fase ocorrerá no sistema.

Desde a sua publicação, em 1907, várias provas do Teorema de Perron-Frobenius foram produzidas. Neste texto, fixaremos atenção nas provas apresentadas por B. Simon [8] e D. Ruelle [9] e em algumas de suas conseqüências que também são apresentadas nessas referências. As aplicações serão concentradas na descrição matemática de sistemas de spin unidimensionais, na qual aparecem matrizes com entradas positivas (matrizes transferência). Um exemplo particular que será tratado é o Modelo de Ising [4]. Foi graças ao método da matriz transferência que L. Onsager [5] determinou a energia livre do Modelo de Ising bi-dimensional, sem campo externo, e concluiu que o mesmo apresentava o fenômeno de transição de fase ao contrário do que acontece no caso unidimensional. O Teorema de Perron-Frobenius também apresenta aplicações em Sistemas Dinâmicos [11] e em Probabilidades [12].

Modelos de spin podem ser utilizados, por exemplo, para tratar sistemas ferromagnéticos. Esses sistemas ferromagnéticos, por sua vez, são corriqueiros no dia-a-dia. Várias vezes, temos à mão um objeto magnetizado que é capaz de produzir forças de atração em metais como o ferro. Suponha que esse objeto seja constituído por átomos que ocupam posições bem definidas em \mathbb{R}^d . Cada ponto no qual se encontra um átomo é chamado de *sítio* e o conjunto de todos os sítios é uma *rede*. Para entender o comportamento de tais sistemas, vamos imaginar que cada átomo tenha, associado a si, um momento de dipolo magnético de spin \vec{s} . A origem desse momento de dipolo é quântica. A orientação do vetor \vec{s} em um sítio influencia diretamente na orientação do vetor nos sítios vizinhos. Dizemos, então, que há um acoplamento entre os sítios do sistema. Como um sistema ferromagnético ordinário possui um número de átomos da ordem de 10^{23} , é necessário que uma descrição de seu comportamento seja feita em termos da Mecânica Estatística do Equilíbrio.

Observa-se, experimentalmente, que, a baixas temperaturas, um sistema ferromagnético pode se manter magnetizado indefinidamente. Contudo, se aumentarmos gradualmente a temperatura, observa-se que existe um valor crítico T_c a partir do qual a magnetização é nula. Tal comportamento indica a presença de transição de fase nesses sistemas. De uma maneira mais refinada, uma transição de fase é definida como uma descontinuidade em alguma derivada de um potencial termodinâmico. Nesse caso, a susceptibilidade magnética, que é uma derivada da energia livre, é infinita quando a temperatura atinge o valor $T = T_c$.

Intuitivamente, podemos pensar que, em baixas temperaturas, há uma cooperação entre os spins de sítios vizinhos, devido ao acoplamento entre eles, que resulta em um alinhamento preferencial para o momento de dipolo. Há assim uma fase ordenada para o sistema em temperaturas baixas. Já em temperaturas altas, a agitação térmica consegue destruir o comportamento ordenado do sistema. Os momentos de dipolo assumem direções aleatórias, sendo que, na média, não há uma magnetização resultante.

Neste trabalho, nós trataremos de modelos de spin unidimensionais e provaremos que, sob certas hipóteses, tais modelos se comportam como o Modelo de Ising, não apresentando transição de fase.

Esse trabalho está assim dividido: na Seção 2, nós provamos o Teorema de Perron-Frobenius. Na Seção 3, fazemos uma breve apresentação de algumas funções importantes na construção e tratamento matemático de modelos de Mecânica Estatística. Na Seção 4, utilizamos o Modelo de Ising para mostrar como matrizes com entradas positivas (matrizes transferência) surgem ao se estudar esses modelos. Ficará estabelecida, nessa seção, a relação entre a pressão e o maior autovalor da matriz transferência. Logo em seguida, na Seção 5, mostramos que a relação entre a pressão e o maior autovalor da matriz transferência é válida para modelos um pouco mais gerais que o Modelo de Ising. Finalmente, na Seção 6, argumentamos porque uma classe bem geral de modelos, chamados modelos de interação de curto alcance, não deve apresentar transição de fase. Observamos que, para estabelecer os resultados das Seções 4, 5 e 6, nos inspiramos na referência [8]. Nos Apêndices B, C e D, mostramos alguns resultados auxiliares que são utilizados nas provas de alguns teoremas ou proposições.

2 O Teorema de Perron-Frobenius

Nesta seção, vamos enunciar e provar o Teorema de Perron-Frobenius 2.1. Esse teorema se refere a matrizes $d \times d$ cujos elementos são positivos. Resumidamente, o teorema afirma que tal matriz possui um autovalor positivo que tem multiplicidade algébrica um e que é o maior autovalor em módulo.

Teorema 2.1 (Perron-Frobenius) *Seja A uma matriz $d \times d$ de entradas positivas, isto é, $A_{ij} > 0 \forall i, j = 1, \dots, d$. Então:*

1. *A possui um único autovetor x , de norma 1, cujas componentes são, todas elas, positivas;*
2. *o autovalor λ_+ , associado ao autovetor x , é positivo e, para qualquer outro autovalor $\lambda \in \mathbb{C}$, temos que $|\lambda| < \lambda_+$;*
3. *o autovalor λ_+ é simples.*

Antes de apresentar a prova do Teorema 2.1, vamos enunciar e provar alguns resultados preliminares que serão utilizados no decorrer da dedução.

Lema 2.1.1 *Seja $f(v) : D \rightarrow \mathbb{R}$ uma função contínua, limitada inferiormente e definida no aberto limitado $D \subset \mathbb{R}^n$. Suponha que $\lim_{v \rightarrow v_0} f(v) = \infty$ sempre que $v_0 \in \partial D$. Então $f(v)$ tem um ponto de mínimo global em D .*

Prova: Para $n = 1, 2, \dots$, considere o conjunto

$$D_n \equiv \{v \in D : \text{dist}(v, \partial D) \geq \frac{1}{n}\},$$

onde $\text{dist}(v, \partial D)$ é a distância de v à fronteira ∂D de D . D_n é claramente fechado e como $D_n \subset D$, D_n também é limitado. Portanto, D_n é um conjunto compacto e, como $f(v)$ é contínua em D_n , podemos concluir que

$$\exists v_n \in D_n \quad \text{tal que} \quad f(v_n) \leq f(x), \quad \forall x \in D_n.$$

Observe que $v_n \in D_k$ e que $f(v_k) \leq f(v_n) \leq f(v_1)$ para todo $k \geq n \geq 1$. Como $f(v)$ é limitada inferiormente, concluímos que a seqüência $\{f(v_n)\}$ é limitada. Afirmamos agora que:

$$\exists n_0 \in \mathbb{N} \quad \text{tal que} \quad v_n \in D_{n_0} \quad \forall n.$$

De fato, se não existisse tal n_0 , então seríamos capazes de determinar uma subseqüência $\{v_{n_k}\}$ tal que $\lim_{k \rightarrow \infty} \text{dist}(v_{n_k}, \partial D) = 0$, implicando que a subseqüência $\{f(v_{n_k})\}$ não é limitada superiormente, o que é um absurdo. \square

Lema 2.1.2 *Seja A uma matriz $d \times d$ com todas as entradas positivas. Suponha que A tenha dois autovetores x e y linearmente independentes, com autovalores λ_+ e λ respectivamente. Suponha, ainda, que x tenha somente componentes positivas, isto é, $x_i > 0 \forall i \in 1, \dots, d$. Então $|\lambda| < \lambda_+$.*

Prova: Sendo $x_i > 0$ e $A_{ij} > 0 \forall i, j = 1, \dots, d$; observamos que $\lambda_+ > 0$, pois $\lambda_+ = \frac{1}{x_i} \sum_j A_{ij} x_j$.

Seja $w_j = \frac{y_j}{x_j}$, para $j = 1, \dots, d$. Reindexando, se necessário for, afirmamos que pode-se escrever

$$w_1 = |w_1| \geq |w_2| \geq \dots \geq |w_d|, \quad (1)$$

sendo que ou $w_1 > |w_d|$ ou w_d é não positivo. De fato, o reindexamento é feito da seguinte forma: tome todos os números w_j , $j = 1, \dots, d$ em valor absoluto. Se o maior deles for proveniente de um w_i negativo, então troque y por $-y$; senão deixe y inalterado¹. Feito isso, w_i será positivo. Renomeie w_i por w_1 . Reordene, agora, em forma decrescente, os demais valores absolutos de w_j . Nesse momento vale (1). Observe então que são possíveis dois casos: no primeiro deles, pelo menos uma desigualdade em (1) é estrita, implicando então que $w_1 > |w_d|$; no segundo caso, não há nenhuma desigualdade estrita e, portanto, $w_1 = |w_2| = \dots = |w_d|$. Contudo, ainda no segundo caso, as igualdades não podem ocorrer sem tomarmos o valor absoluto, isto é, não pode ser que $w_1 = w_2 = \dots = w_d$, já que y e x são linearmente independentes. Dessa forma, $\exists k \in \{2, \dots, d\}$ tal que $w_k = -w_1$. Nesse caso, trocamos w_k por w_d e então w_d será não positivo.

Vamos, então, analisar os dois casos:

- $w_1 > |w_d|$, sendo que as componentes de y e x associadas a w_1 são y_i e x_i , respectivamente;

Levando em consideração que y é um autovetor de A com autovalor λ , temos que, $\lambda y_i = \sum_j A_{ij} y_j$. Assim,

$$\begin{aligned} |\lambda| &= |y_i^{-1}| \left| \sum_j A_{ij} y_j \right| \\ &\leq |x_i w_1|^{-1} \sum_j |A_{ij} x_j w_j| \\ &= x_i^{-1} \sum_j A_{ij} x_j \frac{|w_j|}{w_1} \end{aligned}$$

¹Observe que $-y$ também é autovetor de A com autovalor λ .

$$< x_i^{-1} \sum_j A_{ij} x_j = \lambda_+,$$

onde a desigualdade estrita segue da hipótese $w_1 < |w_d|$.

- $w_1 = |w_2| = \dots = |w_d|$ e $w_d \leq 0$;

Repetindo o argumento anterior, temos que:

$$\begin{aligned} |\lambda| &= |y_i^{-1}| \left| \sum_j A_{ij} y_j \right| \\ &= |x_i w_1|^{-1} \sum_j |A_{ij} x_j w_j| \\ &< |x_i w_1|^{-1} \sum_j A_{ij} x_j |w_j| \\ &= x_i^{-1} \sum_j A_{ij} x_j = \lambda_+, \end{aligned}$$

onde a desigualdade estrita segue do fato de que $w_d < 0$. Portanto, concluímos dos dois casos acima, que $|\lambda| < \lambda_+$. \square

Corolário 2.1.1 *Sob as hipóteses do Lema 2.1.2, o autovalor λ_+ tem multiplicidade geométrica igual a 1.*

Prova do Teorema 2.1: Observamos, inicialmente, que o segundo item do enunciado é uma consequência direta do primeiro item e do Lema 2.1.2. De fato, se x é um autovetor de A , satisfazendo o primeiro item, com autovalor λ_+ , e se y é um outro autovetor de A , com autovalor λ , então segue do lema que $|\lambda| < \lambda_+$.

Provaremos, então, o primeiro item do enunciado. Para tal, vamos inicialmente mostrar que a matriz A possui um autovetor x cujas entradas são todas positivas. A unicidade deste autovetor será provada posteriormente. Seja $Q_+ = \{v \in \mathbb{R}^d \mid v_i > 0 \forall i = 1, \dots, d\}$ o “quadrante” positivo de \mathbb{R}^d . Observe que $A(Q_+) \subset Q_+$, pois $A_{ij} > 0$ para qualquer par $\langle i, j \rangle$. Além disto, se v é um autovetor de A associado ao autovalor λ , então λ e v satisfazem à relação

$$\lambda = v_i^{-1} \sum_{j=1}^d A_{ij} v_j,$$

para qualquer valor do índice i , desde que $v_i \neq 0$. Motivados pela identidade acima, definimos as seguintes funções $\alpha_i(v) : Q_+ \rightarrow \mathbb{R}^+$, $i = 1, \dots, d$

$$\alpha_i(v) \equiv v_i^{-1} \sum_{j=1}^d A_{ij} v_j. \quad (2)$$

Vamos mostrar, a seguir, que existe um vetor $w \in Q_+$ tal que $\alpha_i(w) = \alpha_j(w)$ para todo $i \neq j$, sendo isto suficiente para concluir que A tem um autovetor em que todas as entradas

são positivas. Para isso, analisaremos o ponto de mínimo da função $f(v) : Q_+ \rightarrow \mathbb{R}^+$, definida como

$$f(v) \equiv \max_j \alpha_j(v). \quad (3)$$

Observando que $\alpha_i(\cdot)$ é uma função homogênea de grau zero, isto é, $\alpha_i(\sigma v) = \alpha_i(v)$ para todo $\sigma \in \mathbb{R}$ não nulo, olharemos para as funções $\alpha_i(\cdot)$ e $f(\cdot)$ restritas à parte da esfera unitária de \mathbb{R}^d que se encontra em Q_+ , isto é, restrita à superfície $S = \{v \in Q_+ : \|v\| = 1\}$. Nessa superfície, temos um vínculo entre as coordenadas de v :

$$v \in S \Leftrightarrow v_1^2 + \dots + v_d^2 = 1 \text{ e } v_i > 0 \ \forall i = 1, \dots, d.$$

É possível, então, interpretar a superfície S como sendo o gráfico da função

$$v_d = v_d(v_1, \dots, v_{d-1}) = \sqrt{1 - [v_1^2 + \dots + v_{d-1}^2]}, \quad (4)$$

cujo domínio $D \subset \mathbb{R}^{d-1}$ é

$$D = \{(v_1, \dots, v_{d-1}) \in \mathbb{R}^{d-1} : v_1^2 + \dots + v_{d-1}^2 < 1 \text{ e } v_i > 0 \ \forall i\}.$$

Se substituirmos v_d , dado por (4), em (2), nós concluiremos que $\alpha_i(\cdot)$ e $f(\cdot)$ podem ser reescritas como funções das $d - 1$ variáveis v_1, \dots, v_{d-1} . Por exemplo, após a substituição, obteremos

$$\alpha_i(v_1, \dots, v_{d-1}) = v_i^{-1} \sum_{j=1}^{d-1} A_{ij} v_j + v_i^{-1} A_{id} \sqrt{1 - [v_1^2 + \dots + v_{d-1}^2]} \quad \text{se } i \neq d, \quad (5)$$

enquanto que

$$\alpha_d(v_1, \dots, v_{d-1}) = \left(\sqrt{1 - [v_1^2 + \dots + v_{d-1}^2]} \right)^{-1} \sum_{j=1}^{d-1} A_{dj} v_j + A_{dd}. \quad (6)$$

Denotaremos por ∂D à fronteira de D :

$$\partial D \equiv \{v \in \mathbb{R}^{d-1} : v_i \geq 0 \text{ e } \|v\| = 1 \text{ ou } v_i = 0 \text{ para algum } i = 1, \dots, d-1\}.$$

De acordo com (3), (5) e (6), $f(v)$ é uma função positiva e contínua em D , com limite $\lim_{v \rightarrow v_0} f(v) = \infty$ para qualquer $v_0 \in \partial D$. Pelo Lema 2.1.1, concluímos que $f(v)$ possui um ponto de mínimo global em D . Esse ponto tem uma imagem na superfície S que, por sua vez, está imersa em Q_+ . A conclusão imediata é que:

$$\exists x^{(0)} \in Q_+ \text{ tal que } f(x^{(0)}) \leq f(x) \quad \forall x \in Q_+. \quad (7)$$

Vamos mostrar, a seguir, que qualquer vetor $x^{(0)}$ satisfazendo a condição (7) é um autovetor da matriz A . Em outras palavras, vamos mostrar que

$$\alpha_1(x^{(0)}) = \dots = \alpha_d(x^{(0)}), \quad (8)$$

seguindo daí que $x^{(0)}$ é autovetor com autovalor $\lambda_+ \equiv f(x^{(0)})$. De fato, suponha por absurdo que (8) não se verificasse. Reindexando os α 's, se necessário for, podemos assumir que:

$$\alpha_1(x^{(0)}) \geq \alpha_2(x^{(0)}) \geq \dots \geq \alpha_{d-1}(x^{(0)}) \geq \alpha_d(x^{(0)}), \quad (9)$$

sendo que pelo menos uma das desigualdades é estrita. Fixado $\varepsilon > 0$, sejam $x_j^{(\varepsilon)} = (1+\varepsilon)x_j^{(0)}$ se $j = 1, \dots, d-1$ e $x_d^{(\varepsilon)} = x_d^{(0)}$. Claramente, $x^{(\varepsilon)}$ representa uma perturbação na posição de mínimo. Verifica-se que:

- $\alpha_j(x^{(\varepsilon)}) = \alpha_j(x^{(0)}) - \frac{\varepsilon}{1+\varepsilon} A_{jd}(x_d^0/x_j^0)$ para $j = 1, \dots, d-1$;
- $\alpha_d(x^{(\varepsilon)}) = (1+\varepsilon)\alpha_d(x^{(0)}) - \varepsilon A_{dd} = \alpha_d(x^{(0)}) + \theta(\varepsilon)$,

onde $\theta(\varepsilon) \equiv \varepsilon(\alpha_d(x^{(0)}) - A_{dd})$. Como ε e $A_{jd} > 0$, segue que $\alpha_j(x^{(\varepsilon)}) < \alpha_j(x^{(0)}) \leq f(x^{(0)})$ para $j = 1, \dots, d-1$. Já, para $j = d$, temos que $\alpha_d(x^{(\varepsilon)}) < \alpha_d(x^{(0)})$ se $\theta(\varepsilon) < 0$. Caso contrário, como pelo menos uma das desigualdades em (9) é estrita, escolhamos ε arbitrariamente pequeno tal que $\alpha_d(x^{(\varepsilon)}) = \alpha_d(x^{(0)}) + \theta(\varepsilon) < \alpha_1(x^{(0)})$. Assim, para esta escolha de ε , vale $\alpha_j(x^{(\varepsilon)}) < f(x^{(0)}) \forall j$. Tomando o máximo em j obtemos $f(x^{(\varepsilon)}) < f(x^{(0)})$, o que é um absurdo em vista de (7).

Nesse ponto, fica claro que $x^{(0)}$ é um autovetor de A com todas as componentes positivas e com autovalor $\lambda_+ = f(x^{(0)})$. A existência de tal autovalor, juntamente com o lema 2.1.2, garante que λ_+ é o maior autovalor em módulo. Fica então provado o item 2 do teorema. Ainda do lema, segue que a multiplicidade geométrica de $x^{(0)}$ é igual a 1.

Daqui em diante, vamos nos referir a $x^{(0)}$ apenas como x . Vamos, agora, completar a prova do item 1, mostrando que nenhum outro autovetor y , associado a um outro autovalor λ , tem somente componentes não negativas. Suponha, por contradição, que tal vetor y exista e escreva

$$Ay = \lambda y$$

Vamos analisar dois casos separadamente:

- O autovetor y tem pelo menos uma componente nula y_l . Nesse caso,

$$0 = y_l \lambda = \sum_{j=1}^d A_{lj} y_j.$$

Mas como $A_{lj} > 0 \forall j = 1, \dots, d$, deveríamos ter que $y_j = 0 \forall j = 1, \dots, d$, o que implicaria em y ser o vetor nulo o que contradiz a hipótese de y ser autovetor de A .

- O autovetor y tem todas as componentes positivas: Nesse caso podemos escrever

$$\lambda = y_i^{-1} \sum_{j=1}^d A_{ij} y_j = \alpha_i(y), \quad \forall i = 1, \dots, d.$$

Observa-se que $\lambda > 0$, então como λ_+ é o maior autovalor em módulo, temos:

$$\lambda < \lambda_+ \Rightarrow \alpha_i(y) < f(x), \quad \forall i = 1, \dots, d \Rightarrow f(y) < f(x),$$

uma contradição, visto que $y \in Q_+$ e x é um ponto de mínimo global de f em Q_+ .

Considerando os dois casos acima, fica claro que nenhum outro autovetor tem todas as componentes positivas. Isso conclui a prova da unicidade do autovetor x e também a prova do item 1.

Provaremos, agora, o item 3 do teorema, isto é, mostraremos que a multiplicidade algébrica m_a , de λ_+ , é 1. Como já sabemos que a sua multiplicidade geométrica é 1, basta mostrar que a equação

$$(A - \lambda_+)w = x \tag{10}$$

não é satisfeita para nenhum $w \in \mathbb{C}^d$ (ver Teorema C.1 do Apêndice C).

Seja A^* a matriz transposta de A . Segue do lema B.0.1 do Apêndice B que as matrizes A e A^* compartilham o maior autovalor λ_+ . Como A^* é uma matriz com entradas positivas, podemos usar o item 1 deste Teorema para concluir que, se z é o autovetor de A^* associado a λ_+ , então $z \in Q_+$. Supondo então, por contradição, que exista w que satisfaça (10), teremos que

$$\lambda_+ \langle z, w \rangle = \langle A^* z, w \rangle = \langle z, Aw \rangle = \langle z, \lambda_+ w \rangle + \langle z, x \rangle,$$

de onde se conclui que $\langle z, x \rangle = 0$. Contudo, $\langle z, x \rangle = 0$ é, claramente, uma contradição com o fato de que $z_i, x_i > 0 \forall i$.

□

3 Modelos de spin: Hamiltoniano, Função Partição, Energia Livre, Pressão e Transição de Fase.

Como já dissemos, anteriormente, um sistema ferromagnético (por exemplo, um pedaço de imã) pode ser representado por uma rede de sítios. Os modelos ferromagnéticos são, geralmente, definidos sobre redes d -dimensionais que são subconjuntos de \mathbb{Z}^d . Nessa seção, serão definidas as grandezas e funções importantes no tratamento estatístico de tais modelos, como o hamiltoniano, função partição e energia livre.

Considere uma *rede* (ou *volume*) em \mathbb{Z}^d dada por

$$\Lambda \equiv [-N, N] \times \dots \times [-N, N] \cap \mathbb{Z}^d, \quad (11)$$

onde $N \in \mathbb{N}$.

Um *sítio* (ou *vértice*) é, então, um ponto $i \in \Lambda$. Vamos nos referir ao número de sítios em uma determinada rede por $|\Lambda|$.

A cada sítio j é associada uma variável σ_j , denominada *variável de spin*, que pode assumir alguns valores específicos dependendo do modelo. No Modelo de Ising, geralmente, a variável de spin pode assumir apenas os valores $\sigma_j = +1$ ou $\sigma_j = -1$, representando spin para cima e spin para baixo, respectivamente. Já em modelos mais gerais, ela pode assumir também valores inteiros em um intervalo simétrico $[-S, S]$.

A *distância* entre dois sítios x e y em Λ , $\|x - y\|$ será dada pela distância euclidiana:

$$\|x - y\| \equiv \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_d - y_d)^2}.$$

Dada uma rede Λ , diremos que dois sítios i e j pertencentes a Λ são *primeiros vizinhos*, se a distância entre eles é 1. Generalizando, os *n-ésimos vizinhos* são os sítios cuja distância relativa é igual a n .

Dada uma rede Λ , definiremos sua *fronteira* $\partial\Lambda$ como o conjunto de pontos não pertencentes à rede Λ tais que a menor distância a algum ponto da rede é 1. Mais precisamente, $\partial\Lambda \equiv \{i \in \mathbb{Z}^d \setminus \Lambda : \min_{x \in \Lambda} \|i - x\| = 1\}$.

Para um modelo (definido sobre uma rede Λ) no qual as variáveis de spin possam assumir valores em um subconjunto de $S \subset \mathbb{Z}$, definimos o *espaço de configurações* Ω :

$$\Omega \equiv \underbrace{S \times S \times \dots \times S}_{|\Lambda| \text{ vezes}} = S^{|\Lambda|}.$$

Em particular, no Modelo de Ising temos o seguinte espaço de configurações.

$$\Omega = \underbrace{\{-1, 1\} \times \{-1, 1\} \times \dots \times \{-1, 1\}}_{|\Lambda| \text{ vezes}} = \{-1, 1\}^{|\Lambda|}.$$

Uma configuração, $\{\sigma\} \in \Omega$ representa o estado microscópico do sistema em um dado instante.

Podemos, também, associar uma variável de spin μ_i a um sítio i da fronteira $\partial\Lambda$ de uma rede Λ . As *condições de contorno cc* são configurações $\{\mu\}$ prescritas nos sítios $i \in \partial\Lambda$. Assim, por exemplo, no Modelo de Ising, se fixarmos todas as variáveis de spin $\mu = +1$ nos sítios da fronteira, dizemos que estamos tomando condições de contorno *positivas*. Condições de contorno *negativas* são definidas, analogamente, prescrevendo o valor $\mu = -1$ para tais sítios. Já condições de contorno *livres* são aquelas em que não são associadas variáveis de spin aos sítios da fronteira, isto é, eles não estão acoplados aos sítios da rede. As condições de contorno podem ser também *periódicas*. Para isso, o sistema deve ser definido em uma rede que é uma espécie de toro em \mathbb{Z}^d .

O conceito de energia é fundamental no tratamento termodinâmico de sistemas físicos. Dessa forma, é necessário definir, para cada modelo, uma função H que associa, a cada configuração $\{\sigma\}$, a energia correspondente $H(\{\sigma\})$. Essa função é chamada hamiltoniano.

Definido o hamiltoniano para um modelo particular, podemos iniciar o seu tratamento estatístico. Para isso define-se a Função Partição $Z_{\Lambda,cc}(\{\sigma\})$:

Definição 3.1 (Função Partição) *Dada uma rede Λ e um hamiltoniano $H_{\Lambda,cc}$, a função partição é definida como:*

$$Z_{\Lambda,cc}(\{\sigma\}) \equiv \sum_{\{\sigma\}} \exp[-\beta H_{\Lambda,cc}(\{\sigma\})]. \quad (12)$$

Observe que a soma é feita sobre todas as configurações $\{\sigma\} \in \Omega$. O termo $\beta = 1/KT$ é o inverso da *temperatura* T multiplicada pela *constante de Boltzmann* K . Já o termo $\exp(-\beta H_{\Lambda,cc}(\{\sigma\}))$ é o *fator de Boltzmann* que é tradicional em Mecânica Estatística do Equilíbrio.

Segundo o postulado de Gibbs-Boltzmann, um sistema em equilíbrio apresenta uma distribuição de probabilidade dada pela seguinte definição:

Definição 3.2 (Distribuição canônica) *A probabilidade associada à configuração $\{\sigma\}$ é definida por*

$$P(\{\sigma\}) = \frac{\exp[-\beta H_{\Lambda,cc}(\{\sigma\})]}{Z_{\Lambda,cc}}.$$

As grandezas macroscópicas são limites (com $|\Lambda| \rightarrow \infty$) de valores médios segundo essa distribuição. Isso nos leva à próxima definição:

Definição 3.3 (Valor esperado) *O valor esperado de uma função $\varphi(\{\sigma\})$ das configurações $\{\sigma\}$ a volume Λ e condições de contorno cc é definido como*

$$\langle \varphi(\{\sigma\}) \rangle_{\Lambda, cc} \equiv \frac{1}{Z_{\Lambda, cc}} \sum_{\{\sigma\}} \varphi(\{\sigma\}) \exp(-\beta H_{\Lambda, cc}).$$

A energia livre é um potencial termodinâmico cujas derivadas fornecem grandezas físicas de interesse. Dada uma rede Λ , um hamiltoniano $H_{\Lambda, cc}(\{\sigma\})$, definido sobre ela; e definida a função partição por (12), podemos definir a energia livre como segue:

Definição 3.4 (Energia Livre) *A energia livre a volume finito $|\Lambda|$ e com condições de contorno cc é*

$$f_{\Lambda, cc} \equiv -\frac{1}{\beta} \frac{\ln(Z_{\Lambda, cc})}{|\Lambda|}. \quad (13)$$

Observação: A grandeza

$$p_{\Lambda, cc} \equiv \frac{\ln(Z_{\Lambda, cc})}{|\Lambda|}$$

é denominada *pressão* a volume finito e com condições de contorno cc.

A energia livre é um potencial termodinâmico de muito interesse no contexto da Mecânica Estatística porque suas derivadas geram grandezas macroscópicas relevantes. Como exemplo, temos que a magnetização é $M_{\Lambda, cc} = -\frac{\partial f_{\Lambda, cc}}{\partial h}$.

A equação (13) nos indica que, assim como $Z_{\Lambda, cc}$, a energia livre a volume finito, é uma função analítica dos parâmetros J , h e β . Dessa forma, a volume finito, não existem as descontinuidades nas derivadas que são observadas experimentalmente. Para que esse comportamento seja notado, devemos tomar o limite $|\Lambda| \rightarrow \infty$. Quando realizamos esse processo, dizemos que estamos tomando o *limite termodinâmico*. Como exemplo, tomar o limite termodinâmico para a energia livre significa calcular o seguinte limite: $f = \lim_{\Lambda \rightarrow \infty} f_{\Lambda, cc}$. Chamamos f simplesmente de energia livre. Se o limite for tomado para a pressão a volume finito, obtemos $p = \lim_{\Lambda \rightarrow \infty} p_{\Lambda, cc}$ que é chamada pressão.

Podemos agora, definir precisamente o que se entende por transição de fase em sistemas termodinâmicos:

Definição 3.5 (Transição de fase) *Dizemos que um sistema sofre transição de fase em um ponto de seu espaço de parâmetros, se, a pressão, ou alguma de suas derivadas parciais (de ordem arbitrária) com relação a um dos seus parâmetros, for descontínua nesse ponto.*

Nas próximas seções, utilizaremos as definições dadas nessa seção para estudar o comportamento de vários tipos de modelos de spin.

4 O Modelo de Ising Unidimensional e a Matriz Transferência

Nessa seção, trataremos do problema de encontrar a função partição e a pressão para o modelo de Ising unidimensional considerando pequenas variações como, por exemplo, a ausência ou presença de campo magnético externo.

A definição abaixo diz respeito à energia associada a uma configuração no Modelo de Ising:

Definição 4.1 (Hamiltoniano Ising) *Seja Λ a rede dada por (11) e $\partial\Lambda$ sua fronteira. Suponha que sejam prescritas condições de contorno cc. Então, dada uma configuração $\{\sigma\}$, a energia associada a ela, no Modelo de Ising, é dada pelo hamiltoniano:*

$$H_{\Lambda,cc}(\{\sigma\}) \equiv - \sum_{(i,j)} J_{ij} \sigma_i \sigma_j - \sum_{i \in \Lambda} h_i \sigma_i - \sum_{(i,j): i \in \Lambda, j \in \partial\Lambda} J_{ij} \sigma_i \mu_j, \quad (14)$$

onde as variáveis de spin σ e μ , na rede Λ e na fronteira $\partial\Lambda$, respectivamente, podem assumir os valores ± 1 .

O primeiro termo nesse hamiltoniano representa o termo de interação spin-spin, sendo que a soma é realizada entre pares de primeiros vizinhos (i, j) . Para cada par de vizinhos, a constante J_{ij} é positiva e é chamada constante de acoplamento. Observe que, o fato dessa constante ser positiva, reflete o princípio físico de que, em sistemas ferromagnéticos, a configuração em que os spins estão alinhados (isto é, as variáveis de spin têm o mesmo sinal) minimiza a energia do sistema.

Já as constantes h_i , que aparecem no segundo termo, representam o campo magnético externo aplicado em cada sítio i . Novamente, o fato do campo magnético externo e a variável de spin terem o mesmo sinal minimiza a energia.

A última soma é realizada entre pares de spin vizinhos (i, j) sendo que um deles pertence à fronteira e o outro pertence à rede. Essa soma é semelhante à primeira e representa a contribuição para a energia do sistema devido às condições de contorno prescritas na fronteira. Observe que, para se obter condições de contorno livres, basta fazer todas as constantes de acoplamento J_{ij} nulas nessa soma.

Nessa seção, vamos supor que a constante de acoplamento J e o campo magnético externo são uniformes, isto é, $J_{ij} = J \forall (i, j)$ e $h_i = h \forall i$ e vamos obter, explicitamente, a função partição e a pressão relativa a esse modelo em uma dimensão.

Primeiramente, será apresentado um cálculo direto para o caso em que as condições de contorno são livres e não existe campo magnético externo. Depois será introduzido o método da matriz transferência; que é método criativo e muito conveniente para o propósito de calcular a função partição para esses modelos. Como a matriz transferência é uma matriz com todas as entradas positivas, temos uma conexão entre o Teorema 2.1 e a Mecânica Estatística. O resultado principal dessa seção é mostrar que o maior autovalor dessa matriz é muito importante na obtenção da função partição e, conseqüentemente, da pressão.

4.1 Condições de Contorno Livres sem Campo Externo: Integração Sítio-Sítio

Nessa seção, trataremos do Modelo de Ising sem campo magnético externo e com condições de contorno livres. Mostraremos que é possível obter explicitamente a função partição e a energia livre para esse caso particular. No entanto, o objetivo dessa seção é motivar a necessidade de um método que funcione em situações mais gerais como, por exemplo, no caso em que existe um campo magnético externo.

Considere o Modelo de Ising unidimensional definido sobre uma rede de $2N + 1$, $N \in \mathbb{N}$ sítios

$$\Lambda \equiv \{i \in \mathbb{Z} : -N \leq i \leq N\}. \quad (15)$$

Caso as condições de contorno sejam livres e não haja campo magnético externo, a energia associada a uma configuração é dada pelo hamiltoniano ² :

$$H_N(\{\sigma\}) = - \sum_{i=-N}^{N-1} J\sigma_i\sigma_{i+1}. \quad (16)$$

Esse caso, que apresenta a expressão mais simples possível para a energia no Modelo de Ising, pode ter a função partição avaliada através de um procedimento simples. Este consiste em somar, termo a termo, a função partição, considerando os possíveis valores para a variável de spin em cada sítio.

A função partição definida em (12) é dada por:

$$Z_N(\beta, J, h = 0) = \sum_{\{\sigma\}} \exp\left(\sum_{i=-N}^{N-1} \beta J\sigma_i\sigma_{i+1}\right) = \sum_{\{\sigma\}} \prod_{i=-N}^{N-1} \exp(\beta J\sigma_i\sigma_{i+1}) \quad (17)$$

Como o sítio $i = N$ só interage com seu primeiro vizinho, $j = N - 1$, podemos reescrever a expressão acima fatorando o termo que contém a interação entre esses dois spins:

$$Z_N(\beta, J, h = 0) = \left[\sum_{\{\sigma_i\}: i \neq N} \prod_{i=-N}^{N-2} \exp(\beta J\sigma_i\sigma_{i+1}) \right] \left[\sum_{\sigma_N = \pm 1} \exp(\beta J\sigma_{N-1}\sigma_N) \right].$$

Agora analisamos o último termo na equação acima. Somando sobre os valores $\sigma_N = \pm 1$ podemos eliminar a dependência com relação ao sítio $i = N$:

$$\left[\sum_{\sigma_N = \pm 1} \exp(\beta J\sigma_{N-1}\sigma_N) \right] = \exp(\beta J\sigma_{N-1}) + \exp(-\beta J\sigma_{N-1}) = 2 \cosh(\beta J).$$

Observe que a dependência em σ_{N-1} também é eliminada nesse termo, porque o cosseno hiperbólico é uma função par.

²Compare com o hamiltoniano (14). Observe que foi mantido apenas o termo de interação spin-spin e as constantes de acoplamento foram tomadas como uniformes.

Voltando à expressão para a função partição temos:

$$Z_N(\beta, J, h = 0) = 2 \cosh(\beta J) \left[\sum_{\{\sigma_i\}: i \neq N} \prod_{i=-N}^{N-2} \exp(\beta J \sigma_i \sigma_{i+1}) \right].$$

O termo entre colchetes na expressão acima é semelhante ao lado direito da equação (17) que foi, por sua vez, o nosso ponto de partida. Assim, pode-se repetir o processo somando sobre os possíveis valores da variável σ_{N-1} . Por indução, obtém-se, após repetir a integração nas variáveis $\sigma_N, \sigma_{N-1}, \dots, \sigma_{-N+1}$ e σ_{-N+1} :

$$Z_N(\beta, J, h = 0) = [2 \cosh(\beta J)]^{2N} \left[\sum_{\sigma_{-N}=\pm 1} 1 \right] = 2^{2N+1} [\cosh(\beta J)]^{2N} \quad (18)$$

Utilizando a expressão acima para a função partição, podemos obter a pressão $p(\beta, J, h = 0) = \lim_{N \rightarrow \infty} \frac{\ln(Z_N)}{2N+1}$. O resultado é:

$$p(\beta, J, h = 0) = \ln[2 \cosh(\beta J)]$$

Uma questão surge, naturalmente, após a apresentação do método acima. Será que esse método é prático e pode nos levar à resolução de um grande número de modelos unidimensionais? Na verdade, caso seja adicionado um campo magnético externo ao hamiltoniano, esse método deixa de ser prático. Dessa forma, será apresentado na próxima seção um método mais poderoso que se denomina *método da matriz transferência*.

4.2 Modelo Ising com Condições de Contorno Periódicas e a Matriz Transferência

Nessa seção, trataremos um sistema particular que consiste em uma linha de spins dada por (15) na qual atua um campo magnético externo uniforme de intensidade h e com condições de contorno periódicas. Será introduzida a matriz transferência para auxiliar no cálculo da função partição.

O hamiltoniano para esse sistema é dado por:

$$H_N(\{\sigma\}) = - \sum_{i=-N}^N J \sigma_i \sigma_{i+1} - h \sum_{i=-N}^{i=N} \sigma_i = - \sum_{i=-N}^N J \sigma_i \sigma_{i+1} - \frac{h}{2} \sum_{i=-N}^{i=N} (\sigma_i + \sigma_{i+1}),$$

onde tomamos a convenção de que $\sigma_{N+1} = \sigma_{-N}$.

A função partição é, então, dada por

$$Z_N(\beta, J, h) = \sum_{\{\sigma\}} \prod_{i=-N}^{i=N} \exp \left[\nu \sigma_i \sigma_{i+1} + \frac{B}{2} (\sigma_i + \sigma_{i+1}) \right]. \quad (19)$$

onde $\nu = \beta J$ e $B = \beta h$.

Notando que cada termo do produto acima depende de duas variáveis de spin σ_i e σ_{i+1} a equação (19) fica, convenientemente, reescrita como:

$$Z_N(\beta, J, h) = \sum_{\{\sigma\}} L(\sigma_{-N}, \sigma_{-N+1})L(\sigma_{-N+1}, \sigma_{-N+2}) \cdots L(\sigma_{N-1}, \sigma_N)L(\sigma_N, \sigma_{N+1}), \quad (20)$$

onde os fatores $L(\sigma_i, \sigma_j)$ são definidos por

$$L(\sigma_i, \sigma_j) \equiv \exp \left[\nu \sigma_i \sigma_j + \frac{B}{2}(\sigma_i + \sigma_j) \right]. \quad (21)$$

Fixando uma configuração particular $\{\sigma\}$, cada termo $L(\sigma_i, \sigma_{i+1})$ irá assumir um dos valores da seguinte matriz ³

$$L \equiv \begin{pmatrix} L(+, +) & L(+, -) \\ L(-, +) & L(-, -) \end{pmatrix} = \begin{pmatrix} e^{\nu+B} & e^{-\nu} \\ e^{-\nu} & e^{\nu-B} \end{pmatrix}. \quad (22)$$

Observe que a matriz L possui apenas entradas positivas. Ela é denominada *matriz transferência*. Suas linhas e colunas são indexadas por $+$ e $-$ ao invés de 1 e 2. Assim, por exemplo, o elemento L_{11} é o mesmo que $L(+, +)$.

Verifica-se que, mantendo-se fixas as configurações nos sítios $i-1$ e $i+1$ e somando-se sobre as configurações possíveis, $+$ ou $-$, no sítio i , pode-se eliminar a dependência da variável σ_i , pois,

$$\begin{aligned} \sum_{\sigma_i=\pm} L(\sigma_{i-1}, \sigma_i)L(\sigma_i, \sigma_{i+1}) &= L(\sigma_{i-1}, +)L(+, \sigma_{i+1}) + L(\sigma_{i-1}, -)L(-, \sigma_{i+1}) \\ &= L^2(\sigma_{i-1}, \sigma_{i+1}), \end{aligned}$$

onde $L^2(\sigma_{i-1}, \sigma_{i+1})$ representa o elemento devidamente indexado da matriz L^2 .

Usando o resultado acima, podemos efetuar a soma em (20), somando inicialmente sobre a variável σ_N . Repetimos o processo, sucessivamente, de forma que após a $2N$ -ésima eliminação, obtemos a expressão:

$$Z_N(\beta, J, h) = \sum_{\sigma_{-N}=\pm} L^{2N+1}(\sigma_{-N}, \sigma_{-N}) = \text{Tr}(L^{2N+1}) = \lambda_+^{2N+1} + \lambda_-^{2N+1}, \quad (23)$$

onde λ_+ e λ_- são os maior e o menor autovalor da matriz L , respectivamente. A existência do maior autovalor está assegurada pelo Teorema 2.1.

Resolvendo a equação característica para a matriz transferência (22) obtemos:

$$\lambda_{\pm} = e^{\beta J} \cosh(\beta h) \pm [e^{2\beta J} \sinh^2(\beta h) + e^{-2\beta J}]^{\frac{1}{2}} \quad (24)$$

Ao utilizar a matriz transferência para esse caso, o problema de encontrar a função partição foi reduzido ao problema de encontrar os autovalores de L . Mostraremos, agora, que nos é

³Os símbolos $+$ e $-$ devem ser entendidos como equivalentes a $+1$ e -1 , respectivamente.

permitido escrever a pressão em termos do maior autovalor. De fato, a pressão, a volume finito, $p_N(\beta, J, h)$ é dada por

$$p_N = (2N + 1)^{-1} \ln \left[\lambda_+^{2N+1} \left(1 + \frac{\lambda_-^{2N+1}}{\lambda_+^{2N+1}} \right) \right]$$

E tomando o limite termodinâmico obtemos imediatamente

$$p(\beta, J, h) = \ln(\lambda_+) = \ln \left\{ e^{\beta J} \cosh(\beta h) + [e^{2\beta J} \sinh^2(\beta h) + e^{-2\beta J}]^{\frac{1}{2}} \right\}. \quad (25)$$

Esse resultado implica que, para encontrar a pressão, e, conseqüentemente, a energia livre, basta encontrar o *maior autovalor* da matriz transferência L dada por (22). Observa-se que, nesse caso, a pressão é uma função analítica de β, J, h indicando que não há transição de fase (de nenhuma ordem) para esse modelo.

Como caso particular do sistema apresentado nessa seção, se o campo magnético externo for nulo, então, fazendo $h = 0$, obtém-se o seguinte resultado para a pressão:

$$p(\beta, J, h = 0) = \ln [2 \cosh(\beta J)]. \quad (26)$$

Note que esse resultado é exatamente igual ao encontrado na seção anterior. Justificaremos essa coincidência na próxima seção. Nela, voltaremos a trabalhar com condições de contorno livres.

4.3 Modelo de Ising Unidimensional e com Condições de Contorno Livres

No item anterior, foram consideradas condições de contorno periódicas, o que nos permitiu escrever a função partição como o traço da matriz transferência (relativa àquele modelo) elevada à potência $2N + 1$ como mostrado na equação (23). Nessa seção, o fato de que as condições de contorno são livres, impede que a função partição seja expressa dessa maneira. Vamos mostrar, entretanto, que o resultado de que a pressão é dada pelo maior autovalor da matriz transferência continua sendo válido.

Considere, novamente, o Modelo de Ising unidimensional definido sobre a rede (15). Para esse modelo, a energia é dada pelo hamiltoniano

$$H(\{\sigma\}) = - \sum_{i=-N}^{N-1} J \sigma_i \sigma_{i+1} - h/2 \sum_{i=-N}^{N-1} \sigma_i - h/2(\sigma_{-N} + \sigma_N).$$

A função partição com respeito a esse hamiltoniano, é dada por

$$Z_N(\beta, J, h) = \sum_{\{\sigma\}} \prod_{i=-N}^{N-1} \exp \left[\nu \sigma_i \sigma_{i+1} + \frac{B}{2} (\sigma_i + \sigma_{i+1}) \right] \exp \left[\frac{B}{2} (\sigma_{-N} + \sigma_N) \right].$$

Definindo $L(\sigma_i, \sigma_j)$ novamente por (21) temos a expressão:

$$Z_N(\beta, J, h) = \sum_{\{\sigma\}} \prod_{i=-N}^{N-1} L(\sigma_i, \sigma_{i+1}) \exp \left[\frac{B}{2} (\sigma_{-N} + \sigma_N) \right], \quad (27)$$

onde $L(\sigma_i, \sigma_{i+1})$ é um elemento de (22).

Na equação (27), não é possível eliminar a dependência em relação à variável de spin no sítio N da maneira realizada em (4.3). É possível, no entanto, eliminar a dependência dos sítios $i \in \mathbb{Z} : -N + 1 \leq i \leq N - 1$, fazendo com que essa equação forneça,

$$Z_N(\beta, J, h) = \sum_{\sigma_{-N}, \sigma_N = \pm} L^{2N}(\sigma_{-N}, \sigma_N) \exp \left[\frac{B}{2} (\sigma_{-N} + \sigma_N) \right],$$

o que é equivalente a escrever, a função partição como um produto escalar

$$Z_N(\beta, J, h) = \langle \Psi, L^{2N} \Psi \rangle, \quad (28)$$

onde Ψ é o vetor (e^B, e^{-B}) .

Agora, inspirados pelo resultado da seção (4.2), podemos expressar a função partição em termos dos autovalores de L , λ_+ e λ_- dados pela equação (24). Primeiramente, reescrevemos $2N - \text{ésima}$ a potência da matriz transferência em função destes:

$$L^{2N} = \lambda_+^{2N} P_+ + \lambda_-^{2N} P_-,$$

onde P_{\pm} são as projeções na direção dos autovetores associados aos autovalores λ_{\pm} , respectivamente.

A função partição é dada, nesse caso, por:

$$Z_N(\beta, J, h) = \langle \Psi, L^{2N} \Psi \rangle = \lambda_+^{2N} \langle \Psi, P_+ \Psi \rangle + \lambda_-^{2N} \langle \Psi, P_- \Psi \rangle. \quad (29)$$

Compare com a expressão análoga, com condições de contorno periódicas (23).

A equação (29) pode ser reescrita

$$Z_N(\beta, J, h) = \lambda_+^{2N} \left[\langle \Psi, P_+ \Psi \rangle + \frac{\lambda_-^{2N}}{\lambda_+^{2N}} \langle \Psi, P_- \Psi \rangle \right]. \quad (30)$$

Observando que $\lambda_+ > \lambda_-$, segue do Teorema 2.1 que o autovetor associado a λ_+ tem todas as componentes positivas. Como o vetor Ψ também tem todas as componentes positivas, então $\langle \Psi, P_+ \Psi \rangle > 0$. Dessa forma, tomando o logaritmo em (30) e considerando o limite termodinâmico obtém-se:

$$p(\beta, J, h) = \lim_{N \rightarrow \infty} \left\{ \frac{2N}{2N+1} \ln \lambda_+ + \frac{1}{2N+1} \ln \left[\langle \Psi, P_+ \Psi \rangle + \frac{\lambda_-^{2N}}{\lambda_+^{2N}} \langle \Psi, P_- \Psi \rangle \right] \right\} = \ln(\lambda_+).$$

O resultado importante é que a pressão é dada pelo logaritmo do maior autovalor da matriz transferência, tanto no caso em que as condições de contorno são periódicas, como no caso

em que elas são livres. Como as matrizes dos dois casos são iguais, temos que a pressão é independente do fato de serem as condições de contorno livres ou periódicas. Essa aparente coincidência de resultados é, na verdade, uma consequência de um resultado mais geral: *a independência das condições de contorno para o limite termodinâmico da energia livre no Modelo de Ising d-dimensional*. Para uma prova dessa independência, ver [6].

Ainda considerando esse modelo, nos concentraremos no caso particular em que o campo magnético externo é nulo. Nesse caso, fazemos $B = 0$ na matriz transferência, obtendo dessa forma,

$$L_0 \equiv \begin{pmatrix} e^{\beta J} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta J} \end{pmatrix}.$$

Os autovetores da matriz L_0 são $v_+ = 1/\sqrt{2}(1, 1)$ e $v_- = 1/\sqrt{2}(1, -1)$ com os respectivos autovalores $\lambda_+ = e^{\beta J} + e^{-\beta J}$ e $\lambda_- = e^{\beta J} - e^{-\beta J}$.

Como sabemos que a pressão é o logaritmo do maior autovalor de L_0 , temos que

$$p(\beta, J, h = 0) = \ln(\lambda_+) = \ln[2 \cosh(\beta J)].$$

Esse resultado já era esperado em virtude de ter sido obtido anteriormente em (26). Também é interessante ressaltar que o conhecimento dos autovetores e autovalores da matriz transferência nos permite calcular a função partição a volume finito. Pois sendo $(\Psi, P_+ \Psi) = 2$ e $(\Psi, P_- \Psi) = 0$ a equação (29) fornece:

$$Z_N(\beta, J, h = 0) = 2[2 \cosh(\beta J)]^{2N} = 2^{2N+1}[\cosh(\beta J)]^{2N},$$

resultado esperado em vista do que foi obtido através da integração sítio-a-sítio, equação (18).

5 Generalizações

Nessa seção, introduziremos modelos mais gerais nos quais é possível calcular a função partição através do método da matriz transferência. Serão considerados novos tipos de hamiltoniano, novos valores para as variáveis de spin e interações entre sítios cuja distância é maior que 1.

Mostraremos que é possível definir uma matriz transferência para esses modelos e expressar a função partição em termos de um produto escalar envolvendo uma potência dessa matriz semelhante à equação (28).

Em seguida, mostraremos que o resultado de que a pressão é o logaritmo do maior autovalor da matriz transferência continua sendo válido. Para isso, utilizaremos a seguinte proposição que é uma consequência do Teorema 2.1.

Proposição 5.1 *Seja A uma matriz $d \times d$ com todas as entradas positivas e seja λ_+ o maior autovalor de A com autovetor unitário associado x . Se φ e Ψ são vetores de \mathbb{R}^d tais*

que $\langle \varphi, x \rangle$ e $\langle \Psi, x \rangle$ são positivos, então a sequência $\{Z_n\}_{n \in \mathbb{N}}$, onde $Z_n = \langle \varphi, A^n \Psi \rangle$, é tal que

$$\lim_{n \rightarrow \infty} \frac{1}{n} \ln Z_n = \ln \lambda_+.$$

Prova: Seja W_1 o subespaço gerado por x . Como $A_{ij} > 0 \forall i, j$, concluímos, pelo Teorema 2.1, que W_1 tem dimensão igual a 1. O espaço vetorial \mathbb{R}^d pode ser decomposto como uma soma direta $\mathbb{R}^d = W_1 \oplus W_1^\perp$, onde W_1^\perp representa o subespaço de \mathbb{R}^d ortogonal a x . W_1 e W_1^\perp são subespaços invariantes sob a ação de A , isto é, $AW_1 \subset W_1$ e $AW_1^\perp \subset W_1^\perp$.

Os vetores φ e Ψ possuem decomposição única como: $\varphi = \alpha_1 x + w_1$ e $\Psi = \alpha_2 x + w_2$, onde w_1 e w_2 pertencem a W_1^\perp , $\alpha_1 = \langle \varphi, x \rangle$ e $\alpha_2 = \langle \Psi, x \rangle$. Segue daí que

$$\langle \varphi, A^n \Psi \rangle = \alpha_1 \alpha_2 \lambda_+^n + \langle w_1, A^n w_2 \rangle = \lambda_+^n \left[\alpha_1 \alpha_2 + \langle w_1, \left(\frac{A}{\lambda_+} \right)^n w_2 \rangle \right] \quad (31)$$

Seja, agora, $A_r = \frac{A}{\lambda_+}|_{W_1^\perp}$ a restrição da matriz $\frac{A}{\lambda_+}$ ao subespaço W_1^\perp . Pelo Teorema 2.1 o raio espectral ⁴ de A_r é menor que 1. Segue, como consequência da proposição D.1 do Apêndice D, que $\lim_{n \rightarrow \infty} \langle w_1, \left(\frac{A}{\lambda_+} \right)^n w_2 \rangle = 0$.

Sendo $\alpha_1 \alpha_2 > 0$, então, para n suficientemente grande, o termo em colchetes em (31) é positivo e seu logaritmo está bem definido. Segue que, quando $n \rightarrow \infty$,

$$\frac{1}{n} \ln Z_n = \ln \lambda_+ + \frac{1}{n} \ln \left[\alpha_1 \alpha_2 + \langle w_1, \left(\frac{A}{\lambda_+} \right)^n w_2 \rangle \right] \rightarrow \ln \lambda_+.$$

□

Essa proposição tem consequências muito interessantes, se considerarmos os modelos mais gerais que estudamos na Seção 5. Podemos afirmar que, sempre que conseguirmos escrever a função partição como um produto escalar da forma $\langle \varphi, L^n \Psi \rangle$, com L sendo a matriz transferência, a pressão será dada como o logaritmo do maior autovalor de L . Isso se aplica imediatamente aos modelos que trataremos nessa seção.

Inicialmente, vamos tratar do modelo BEG, no qual o hamiltoniano possui termos que dependem do quadrado das variáveis de spin e no qual essas variáveis podem assumir, além dos valores ± 1 , o valor 0.

Depois trataremos do Modelo de Ising com a variável de spin generalizada e do Modelo de Ising com interação entre pares de spin cuja distância é maior que 1.

5.1 Modelo de Blume-Emery-Griffiths

O modelo de Blume-Emery-Griffiths ou modelo BEG é interessante no contexto pois é possível avaliar a função partição utilizando a matriz transferência. A energia de uma configuração é dada pelo hamiltoniano

$$H\{(\sigma)\} = -J \sum_{(i,j)} \sigma_i \sigma_j - K \sum_{(i,j)} \sigma_i^2 \sigma_j^2 + D \sum_{(i)} \sigma_i^2,$$

⁴No Apêndice D está definido raio espectral

onde $i \in \mathbb{Z}^d$ é um sítio da rede, (i, j) representa um par de primeiros vizinhos. Assim como no Modelo de Ising σ_i é a variável de spin no sítio i , mas no modelo BEG ela pode assumir os valores $0, +1, -1$. A constante de acoplamento $J > 0$ é análoga à do Modelo de Ising. Existem, porém, outros dois parâmetros K e D que satisfazem $-\infty < K, D < +\infty$. Esses parâmetros regulam dois tipos de interação diferentes das estudadas até então, pois dependem do quadrado da variável de spin, não diferenciando assim se ela vale $+1$ ou -1 . Observe que, se tomarmos K e D iguais a zero, esse modelo se reduz a uma generalização do Modelo de Ising sem campo magnético externo, no sentido de que resta apenas o termo de interação spin-spin com as variáveis de spin podendo assumir, agora, os valores $0, +1, -1$. Definindo $y = K/J$ e $x = -D/(2dJ)$ o hamiltoniano se reescreve como segue

$$H\{(\sigma)\} = -J \sum_{(i,j)} [\sigma_i \sigma_j + y \sigma_i^2 \sigma_j^2 + x(\sigma_i^2 + \sigma_j^2)] \equiv \sum_{(i,j)} H_{i,j}(\sigma).$$

O objetivo é calcular a função partição para o caso unidimensional com condições de contorno livres. Se o modelo for definido sobre a rede (15), o hamiltoniano acima se reescreve de forma mais conveniente

$$H_N\{(\sigma)\} = -J \sum_{i=-N}^{N-1} [\sigma_i \sigma_{i+1} + y \sigma_i^2 \sigma_{i+1}^2 + x(\sigma_i^2 + \sigma_{i+1}^2)] - Jx(\sigma_{-N}^2 + \sigma_N^2). \quad (32)$$

Defina

$$L(\sigma_i, \sigma_j) \equiv \exp \beta J [\sigma_i \sigma_j + y \sigma_i^2 \sigma_j^2 + x(\sigma_i^2 + \sigma_j^2)]. \quad (33)$$

Fixada uma configuração particular, e utilizando $-, 0$ e $+$ como índices ao invés de $1, 2$ e 3 , respectivamente, $L(\sigma_i, \sigma_j)$ será um elemento da matriz 3×3 :

$$L \equiv \begin{pmatrix} L(-, -) & L(-, 0) & L(-, +) \\ L(0, -) & L(0, 0) & L(0, +) \\ L(+, -) & L(+, 0) & L(+, +) \end{pmatrix} = \begin{pmatrix} e^{\beta J(1+y+2x)} & e^{\beta Jx} & e^{\beta J(-1+y+2x)} \\ e^{\beta Jx} & 1 & e^{\beta Jx} \\ e^{\beta J(-1+y+2x)} & e^{\beta Jx} & e^{\beta J(1+y+2x)} \end{pmatrix} \quad (34)$$

Essa matriz será a matriz transferência ⁵para o modelo BEG. De fato,

$$\begin{aligned} Z_N(\beta, J, x, y) &= \sum_{\{\sigma\}} L(\sigma_{-N}, \sigma_N) L(\sigma_{-N+1}, \sigma_{N+2}) \cdots L(\sigma_{N-1}, \sigma_N) e^{\beta Jx(\sigma_{-N}^2 + \sigma_N^2)} \\ &= \sum_{\{\sigma_i\}: i=\pm N} L^{2N}(\sigma_{-N}, \sigma_N) e^{\beta Jx(\sigma_{-N}^2 + \sigma_N^2)} \end{aligned}$$

Agora se definirmos o vetor

$$\Psi \equiv (e^{\beta Jx}, 1, e^{\beta Jx}) \quad (35)$$

podemos escrever a função partição como o seguinte produto escalar

$$Z_N(\beta, J, x, y) = \langle \Psi, L^{2N} \Psi \rangle. \quad (36)$$

⁵Note que, o fato de que a variável de spin pode assumir agora três valores diferentes foi responsável pelo aumento das dimensões da matriz transferência.

Se λ_+ é o maior autovalor da matriz L , segue da proposição 5.1 que a pressão é igual ao logaritmo natural de λ_+ .

5.2 Modelo de Ising com Spin Generalizado e sem Campo Externo

Considere, nessa seção, o Modelo de Ising unidimensional definido sobre a rede (15) e com hamiltoniano dado por (16)

$$H(\{\sigma\}) = - \sum_{i=-N}^{N-1} J\sigma_i\sigma_{i+1}$$

em que

$$\sigma_i \in \mathbb{N} : -S \leq \sigma_i \leq S. \quad (37)$$

Nesse caso, as condições de contorno foram deixadas livres. Vamos avaliar a função partição para esse modelo utilizando novamente o método da matriz transferência.

Iniciamos fazendo $S = 1$ em (37). Na seção anterior 5.1 obtivemos a matriz transferência e através dela fomos capazes de calcular a função partição para o modelo BEG. Porém, como discutimos naquela seção, se fizermos os parâmetros x e y da equação (32) nulos, então o modelo se resume ao Modelo de Ising que trataremos agora. Nesse caso, temos pronta a matriz transferência L_1 , pois basta substituir $x = 0$ e $y = 0$ em (34) para obter

$$L_1 = \begin{pmatrix} e^{\beta J} & 1 & e^{-\beta J} \\ 1 & 1 & 1 \\ e^{-\beta J} & 1 & e^{\beta J} \end{pmatrix}.$$

Fazendo também $x = 0$ em (35), obtemos o vetor $\Psi_1 = (1, 1, 1)$. Trocamos também L por L_1 em (36). Então, segue que

$$Z_N(\beta, J, h = 0) = \langle \Psi_1, L_1^{2N} \Psi_1 \rangle.$$

Trataremos agora o caso geral em que $S \geq 1$. Seguindo o processo de generalização baseado na seção anterior, fazemos x e y nulos em (33). Isto nos fornece:

$$L(\sigma_i, \sigma_j) \equiv \exp(\beta J \sigma_i \sigma_j)$$

Se utilizarmos os números $-S, -S + 1, \dots, S - 1, S$ como índices, então, fixada uma configuração, e dois sítios da rede i e j , $L(\sigma_i, \sigma_j)$ será um dos elementos da matriz:

$$L_S = \begin{pmatrix} L(-S, -S) & L(-S, -S + 1) & \dots & L(-S, S) \\ L(-S + 1, S) & L(-S + 1, -S + 1) & \dots & L(-S + 1, S) \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ L(S, -S) & L(S, -S + 1) & \dots & L(-S, S) \end{pmatrix} \quad (38)$$

Claramente a função partição será dada por

$$Z_N(\beta, J, h = 0) = \langle \Psi_S, L_S^{2N} \Psi_S \rangle. \quad (39)$$

onde Ψ_S é o vetor de $2S + 1$ componentes, todas iguais a 1, isto é: $\Psi_S = (1, 1, \dots, 1)$.

Como exemplo, para o caso $S = 2$, substituindo os valores apropriados dos elementos de matriz em (38) e (39) obtemos:

$$Z_N(\beta, J, h = 0) = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} e^{4\beta J} & e^{2\beta J} & 1 & e^{-2\beta J} & e^{-4\beta J} \\ e^{2\beta J} & e^{\beta J} & 1 & e^{-\beta J} & e^{-2\beta J} \\ 1 & 1 & 1 & 1 & 1 \\ e^{-2\beta J} & e^{-\beta J} & 1 & e^{\beta J} & e^{2\beta J} \\ e^{-4\beta J} & e^{-2\beta J} & 1 & e^{2\beta J} & e^{4\beta J} \end{pmatrix}^{2N} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

O fato de a função partição ser dada por (39) e a proposição 5.1 nos fornecem mais uma vez que a pressão é dada pelo logaritmo do maior autovalor da matriz transferência.

5.3 Modelo de Ising, com Interação de Primeiros e Segundos Vizinhos

Considere a seguinte rede:

$$\Lambda \equiv \{i \in \mathbb{Z} : 1 \leq i \leq 2N\}, \quad (40)$$

com $N \in \mathbb{N}$. Essa rede difere ligeiramente da definida por (15) pois, agora, é considerado um número par de spins. Trataremos de uma versão do Modelo de Ising, sem campo magnético externo, que considera a interação entre todos os sítios $i, j \in \Lambda$ tais que $\|i - j\| = 1$ ou $\|i - j\| = 2$. A energia para esse modelo será dada pelo hamiltoniano:

$$H_\Lambda(\{\sigma\}) = -J \sum_{i=1}^{2N-1} \sigma_i \sigma_{i+1} - J \sum_{i=1}^{2N-2} \sigma_i \sigma_{i+2} \quad (41)$$

Nesse hamiltoniano, as parcelas $J\sigma_i\sigma_{i+1}$ são as contribuições de primeiros vizinhos e as parcelas $J\sigma_i\sigma_{i+2}$ são as contribuições de segundos vizinhos para a energia.

A rede Λ dada por (40) pode ser vista como um conjunto de $N + 1$ pares de sítios vizinhos. Então define-se o conjunto t_k :

$$t_k \equiv \{2k + 1, 2k + 2\}, \quad k = 0, \dots, N - 1;$$

como sendo o k -ésimo par de spin que forma a rede.

Definimos a auto-energia de cada par como sendo a energia de interação dentre os dois spins dos sítios que formam o par t_k , mais precisamente,

$$H(t_k) \equiv -J\sigma_{2k+1}\sigma_{2k+2}$$

Definimos também a energia de um conjunto de pares vizinhos t_k e t_{K+1} como sendo

$$H(t_k \cup t_{k+1}) \equiv H(t_k) + H(t_{k+1}) + H(t_k; t_{k+1}),$$

onde

$$H(t_k; t_{k+1}) = -J[\sigma_{2k+2}\sigma_{2k+3} + \sigma_{2k+1}\sigma_{2k+3} + \sigma_{2k+2}\sigma_{2k+4}],$$

Observe que, além das auto energias dos pares t_k e t_{k+1} , deve ser incluído o termo $H(t_k; t_{k+1})$ que representa a energia de interação entre esses dois pares. Observe que, em $H(t_k; t_{k+1})$, algumas parcelas vêm de interações entre primeiros vizinhos e outras de interações entre segundos vizinhos.

Reordenando a soma em (41), pode-se escrever:

$$\begin{aligned} H_\Lambda(\{\sigma\}) &= H(t_0) + H(t_1) + H(t_0; t_1) + H(t_2) + H(t_1; t_2) + \cdots + \\ &+ H(t_{N-1}) + H(t_{N-2}; t_{N-1}). \end{aligned} \quad (42)$$

Essa nova forma de reescrever o hamiltoniano é conveniente porque separa as energias de interação entre pares vizinhos das auto-energias de cada par.

Observe ainda que a equação (42) pode ser reescrita de forma mais simétrica:

$$H_\Lambda(\{\sigma\}) = \left[\sum_{k=0}^{N-2} \frac{1}{2} H(t_k) + \frac{1}{2} H(t_{k+1}) + H(t_k; t_{k+1}) \right] + \frac{1}{2} H(t_0) + \frac{1}{2} H(t_{N-1}) \quad (43)$$

Isso nos motiva a definição

$$L(t_k, t_s) \equiv \exp \left\{ -\beta \left[\frac{1}{2} H(t_k) + \frac{1}{2} H(t_s) + H(t_k; t_s) \right] \right\}. \quad (44)$$

Podemos escrever a função partição em termos desses fatores da maneira padrão desenvolvida nos exemplos anteriores:

$$Z_\Lambda(\beta, J) = \sum_{\{\sigma\}} L(t_0, t_1) L(t_1, t_2) \cdots L(t_{N-2}, t_{N-1}) \exp \left\{ -\frac{1}{2} [H(t_0) + H(t_{N-1})] \right\} \quad (45)$$

Fixado um par de sítios t_k , existem as seguintes possibilidades para a configuração de spins nesse par, a saber: ++, +- -+ e --. Vamos enumerar as diferentes configurações como mostrado abaixo:

- ++ \rightarrow 1
- +- \rightarrow 2
- -+ \rightarrow 3
- -- \rightarrow 4

Sendo assim existem 16 possíveis valores para $L(t_k, t_s)$. Eles poderão ser representados como elementos de uma matriz 4×4 conforme os números atribuídos acima para cada configuração.

$$L = \begin{pmatrix} L(1,1) & L(1,2) & L(1,3) & L(1,4) \\ L(2,1) & L(2,2) & L(2,3) & L(2,4) \\ L(3,1) & L(3,2) & L(3,3) & L(3,4) \\ L(4,1) & L(4,2) & L(4,3) & L(4,4) \end{pmatrix} = \begin{pmatrix} e^{4\beta J} & e^{\beta J} & e^{-\beta J} & e^{-\beta 2J} \\ e^{-\beta J} & 1 & e^{-\beta 2J} & e^{\beta J} \\ e^{\beta J} & e^{-\beta 2J} & 1 & e^{-\beta J} \\ e^{-\beta 2J} & e^{-\beta J} & e^{\beta J} & e^{\beta 4J} \end{pmatrix},$$

Obtida a matriz transferência para esse sistema e utilizando a equação (45), a função partição pode ser expressa como o seguinte produto escalar:

$$\mathbb{Z}_\Lambda(\beta, J) = \langle \Psi, L^{N-1} \Psi \rangle,$$

onde Ψ é o vetor $(e^{\frac{1}{2}\beta J}, e^{-\frac{1}{2}\beta J}, e^{-\frac{1}{2}\beta J}, e^{\frac{1}{2}\beta J})$. Como consequência, a pressão é dada por $p = \frac{1}{2} \ln(\lambda_+)$ onde λ_+ é o maior autovalor da matriz L .

5.4 Modelo com Interação de Alcance r

Nessa seção será generalizado o resultado obtido na seção anterior. Deixaremos que um determinado sítio interaja com todos os sítios da rede cuja distância relativa a ele é menor que um certo número natural r .

Fixe, então, um número $r \in \mathbb{N}$ qualquer. Esse será o alcance da interação spin-spin. Fixando outro número $N \in \mathbb{N}$, considere uma rede do tipo $\Lambda \equiv \{i \in \mathbb{N} : 1 \leq i \leq Nr\}$.

A energia associada será dada por

$$H_\Lambda(\{\sigma\}) = -J_1 \sum_{i=1}^{rN-1} \sigma_i \sigma_{i+1} - J_2 \sum_{i=1}^{rN-2} \sigma_i \sigma_{i+2} - \cdots - J_n \sum_{i=1}^{rN-r} \sigma_i \sigma_{i+r}. \quad (46)$$

Definimos o conjunto $t_k \equiv \{rk+1, rk+2, \dots, rk+r\}$ com $k = 0, \dots, N-1$. A auto energia desse conjunto será dada por:

$$H(t_k) \equiv - \sum_{i,j \in t_k, i \neq j} J_{\|i-j\|} \sigma_i \sigma_j.$$

Já a energia de interação entre conjuntos vizinhos t_k e t_{k+1} , é dada por

$$H(t_k; t_{k+1}) \equiv \sum_{\langle i,j \rangle} J_{\|i-j\|} \sigma_i \sigma_j,$$

com a soma realizada entre todos os pares de sítios $\langle i, j \rangle$ com $i \in t_k$ e $j \in t_{k+1}$ tais que $\|i-j\| \leq r$.

Utilizando as últimas duas definições, o hamiltoniano (46) pode ser reescrito na forma (43). Claramente, se definirmos os elementos $L(t_k; t_s)$ através da equação (44), a função partição será dada por (45).

Somando a função partição sobre as configurações dos conjuntos $t_1, \dots, t_k - 2$ segue que:

$$\mathbb{Z}_\Lambda(\beta, J) = \sum_{\{t_0, t_{N-1}\}} L^{N-1}(t_0, t_{N-1}) \exp \left\{ -\frac{1}{2} [H(t_0) + H(t_{N-1})] \right\}. \quad (47)$$

Nessa expressão, a soma é feita sobre as 2^r configurações possíveis em cada conjunto t_0 e t_{N-1} . Observe então que a soma possui 2^{2r} parcelas.

Se enumerarmos as 2^r configurações possíveis para um conjunto t_k como foi feito na Seção 5.3 poderemos organizar os elementos $L(t_k, t_s)$ em uma matriz L de dimensões $2^r \times 2^r$ e reescrever a soma (47) como

$$\mathbb{Z}_\Lambda(\beta, J) = \langle \Psi, L^{N-1} \Psi \rangle, \quad (48)$$

onde Ψ é o vetor de 2^n componentes: $(e^{\frac{1}{2}H(1)}, e^{-\frac{1}{2}H(2)}, \dots, e^{\frac{1}{2}H(2^r)})$. Finalmente, temos $p = \frac{1}{r} \ln(\lambda_+)$, onde λ_+ é o maior autovalor de L .

Observação: Na Seção 5.2 foram tratados modelos em que a variável de spin assume valores no conjunto $\{\sigma \in \mathbb{Z} : -S \leq \sigma \leq S\}$, mas com interação de primeiros vizinhos apenas. Já nos modelos apresentados nessa seção, a variável de spin foi restrita aos valores ± 1 ; entretanto, os sítios interagem sempre que a distância entre eles é menor que um certo número r . Utilizando o mesmo raciocínio apresentado na Seção 5.2, podemos generalizar o método apresentado nessa seção para um modelo com alcance r e com variável de spin generalizada.

6 Ausência de Transição de Fase em Modelos Unidimensionais com Interação de Curto Alcance

Nas seções anteriores, ficou estabelecido que, para uma classe bem geral de modelos de spin unidimensionais, a função partição pode ser expressa como um produto escalar que envolve vetores com componentes positivas e uma potência da matriz transferência. Foram considerados modelos em que a variável de spin assume um número finito de valores discretos e que *dois* sítios interagem sempre que a distância entre eles não ultrapassa um certo valor r .

Nos exemplos fornecidos na Seção 4, as equações (25) e (26) fornecem a pressão explicitamente. Observa-se que, nesses casos, a pressão é dada pelo logaritmo do maior autovalor da matriz transferência e que ela é uma função analítica dos parâmetros reais β , J e h , portanto, tem todas as derivadas contínuas. Sendo assim, nesses exemplos, o modelo correspondente não apresenta transição de fase.

Na Seção 5, consideramos modelos mais gerais que os estudados na Seção 4 e estabelecemos, através da proposição 5.1, que a pressão pode, ainda, ser expressa em termos do logaritmo do maior autovalor da matriz transferência. Entretanto, a tarefa de encontrar o maior autovalor da matriz transferência pode ser complicada, dependendo do hamiltoniano em questão.

Uma pergunta surge naturalmente: mesmo sem calcular explicitamente o maior autovalor da matriz transferência (e portanto a pressão), é possível obter informações sobre o comportamento termodinâmico do sistema no que diz respeito às transições de fase?

Nessa seção, vamos mostrar que para uma grande classe de modelos unidimensionais o maior autovalor da matriz transferência recebe, de herança, a "suavidade" do hamiltoniano com relação a seus parâmetros, impossibilitando a presença de transição de fase. Mais especificamente, os modelos que trataremos nessa seção possuem hamiltoniano com invariância translacional no qual aparecem produtos de variáveis de spin de *mais do que dois* sítios. Deixaremos com que essas variáveis de spin assumam valores no conjunto $\{\sigma \in \mathbb{Z} : -S \leq N \leq S\}$. Para definir, precisamente, a classe de modelos que será nosso objeto de estudo precisamos da seguinte definição:

Definição 6.1 *Seja Λ um subconjunto finito de \mathbb{Z}^d e $\mathbb{P}(\Lambda)$ o conjunto de suas partes. Se σ_i representa a variável de spin em um sítio $i \in \Lambda$, então, dado $A \in \mathbb{P}(\Lambda)$ e $n_{i,A} \in \mathbb{N}$ para cada $i \in A$, definimos σ^A*

$$\sigma^A \equiv \prod_{i \in A} \sigma_i^{n_{i,A}}$$

Sendo A um subconjunto qualquer de \mathbb{R}^d , representaremos por A_t uma translação de A por um vetor $t \in \mathbb{R}^d$. Mais precisamente, fixado um vetor t , definimos $A_t \equiv \{x + t : x \in A\}$.

Seremos capazes, agora, de definir um modelo no qual as variáveis de spin σ_i possam assumir valores arbitrários em um subconjunto finito de \mathbb{Z} e no qual os spins possam interagir entre si sempre que estão separados por uma distância menor que $r \in \mathbb{N}$. Basicamente, chamaremos os modelos que atendem a essas hipóteses de modelos de curto alcance.

Definição 6.2 (Modelo de curto alcance) *Se o hamiltoniano de um modelo é da forma:*

$$H(\{\sigma\}) \equiv \sum_{A \subset \Lambda} J_A \sigma^A,$$

onde as constantes $J_A \in \mathbb{R}$ obedecem:

- $J_A = J_{A_t}$, para todo $t \in \mathbb{Z}^d$;
- $J_A = 0$ se $\text{diâmetro}(A) > r$;

e os expoentes $n_{i,A}$ definidos em 6.1 satisfazem:

- $n_{i+t,A_t} = n_{i,A}$, se A e A_t estão contidos em Λ ,

então esse modelo é dito de curto alcance ou de alcance finito. O número r será chamado alcance.

Observe que todos os modelos que foram apresentados nas seções 4 e 5 são modelos de curto alcance. Porém existem modelos bem mais gerais que são, também, de curto alcance. Note, ainda, que estamos supondo invariância translacional em nossa definição de modelos de curto alcance. Essa suposição será muito importante para que possamos estabelecer um método para construir a matriz transferência e expressar a pressão como o logaritmo natural do maior autovalor dessa matriz. A seguinte proposição trata desse problema:

Proposição 6.1 *Considere um modelo de alcance finito r definido sobre uma rede unidimensional. É possível construir uma matriz transferência para esse modelo de tal forma que a pressão seja expressa como o logaritmo natural do seu maior autovalor multiplicado por $\frac{1}{r}$.*

Prova: : Considere, por simplicidade, uma rede cujo número de sítios seja um múltiplo do alcance r , da mesma maneira como foi feito nas seções 5.3 e 5.4. Mais precisamente, considere que $|\Lambda| = Nr$ onde N é um número natural. Podemos, então, dividir a rede em N blocos que são subconjuntos de Λ com diâmetro igual ao alcance r assim como foi feito nas seções 5.3 e 5.4. Indexaremos os blocos por t_1, t_2, \dots, t_N , onde $t_k \equiv \{rk+1, rk+2, \dots, rk+r\}$ com $k = 0, \dots, N-1$.

Fixando um bloco t_k , definimos a classe de subconjuntos de Λ que estão contidos em t_k $\gamma(k) \equiv \{A \in \mathbb{P}(\Lambda) : A \subset t_k\}$. A *auto-energia* do bloco t_k é definida como $H(t_k) \equiv \sum_{A \in \gamma(k)} J_A \sigma^A$.

Fixando, agora, dois blocos t_k e t_j contidos na rede Λ , definimos a classe de subconjuntos $\ell(k, s) = \{A \subset t_k \cup t_s : A \cap t_k \cap t_s \neq \emptyset\}$. Então a *energia de interação* entre t_k e t_s é definida $H(t_k; t_s) = \sum_{A \in \ell(k, s)} J_A \sigma^A$. É fácil se convencer de que, se t_k e t_s não são blocos vizinhos (isto é, $k \neq s+1$ e $k \neq s-1$), então $H(t_k; t_s) = 0$. De fato, nesse caso, se $A \in \ell(k, s)$ então o diâmetro de A é maior que r , o que implica que $J_A = 0$. Segue daí que, se $A \in \mathbb{P}(\Lambda)$ é tal que $J_A \neq 0$, então ou $A \in \gamma(k)$ para algum k , ou $A \in \ell(k, k+1)$ para algum k . Isso nos permite reordenar as parcelas do hamiltoniano para escrever:

$$H_\Lambda(\{\sigma\}) = \sum_{k=0}^{N-1} \left[\frac{1}{2}H(t_k) + \frac{1}{2}H(t_{k+1}) + H(t_k; t_{k+1}) \right] + \frac{1}{2}H(t_0) + \frac{1}{2}H(t_N)$$

Observe que essa forma para o hamiltoniano é a mesma que foi estabelecida na Seção 5.3. Isso nos inspira a definir os elementos $L(t_k, t_s)$ através da equação (44). O valor de cada elemento $L(t_k, t_s)$ depende, apenas, das configurações dos blocos t_k e t_s . Assim, como o sistema é invariante por translações e como t_k e t_s possuem r elementos cada, existem $(2S+1)^{2r}$ configurações possíveis para o par qualquer par (t_k, t_s) com $t_k, t_s \in \Lambda$. Construimos uma matriz L indexando as linhas pelas configurações possíveis no bloco t_k e colunas pelas configurações possíveis no bloco t_s . Então, fixada uma configuração particular na rede Λ , o elemento $L(t_k, t_s)$ poderá ser atribuído a um dos elementos dessa matriz $(2S+1)^r \times (2S+1)^r$. Esses argumentos são suficientes para concluirmos que a matriz L é a matriz transferência desejada para o modelo. Através dela, é possível mapear as configurações de cada bloco em seu vizinho, como foi feito na Seção 5.4 para expressar a função partição como na equação (48).

Finalmente, segue da proposição 5.1 que a pressão é dada pelo logaritmo do maior autovalor da matriz L (a menos de um fator $\frac{1}{r}$) definida pelos elementos $L(t_k, t_s)$. \square

O próximo teorema diz respeito à ausência de transição de fase nos modelos de curto alcance que definimos nessa seção. Antes de enunciá-lo, faremos uma observação que será útil em sua prova: Afirmamos que, para um modelo de curto alcance, não necessariamente unidimensional, o hamiltoniano depende apenas de um número finito de parâmetros não nulos $J_{A_1}, J_{A_2}, \dots, J_{A_k}$. Para se convencer disso, fixe um sítio i e considere a classe de todos

os subconjuntos de \mathbb{Z}^d que intersectam i e cujo diâmetro é menor ou igual a R . Essa classe possui um número finito de elementos. Mas todo subconjunto $A \subset \mathbb{Z}^d$ com $J_A \neq 0$ pode ser identificado como uma translação de algum conjunto dessa classe. Como J_A é invariante por translações, segue, daí, que a afirmação é verdadeira. Daqui para frente, vamos representar o conjunto desses parâmetros por $J = (J_{A_1}, J_{A_2}, \dots, J_{A_k})$.

Teorema 6.1 (Ausência de transição de fase em modelos de curto alcance) *Considere um modelo de curto alcance definido sobre uma rede unidimensional. Então, a pressão (e , conseqüentemente, a energia livre) é uma função de classe C^∞ de β e dos parâmetros $J = (J_{A_1}, J_{A_2}, \dots, J_{A_k})$.*

Prova: Como o modelo é de curto alcance, segue, da proposição acima, que existe uma matriz transferência L cujas entradas são exponenciais de fatores de Boltzmann. A pressão é dada por $p(\beta, J) = \frac{1}{\tau} \ln \lambda_+$, onde λ_+ é o maior autovalor de L .

Vamos provar que $\lambda_+ = \lambda_+(\beta, J)$ é uma função (positiva) de classe C^∞ dos parâmetros β e J sendo isso suficiente para concluir a afirmação a ser provada.

Para isso, considere o polinômio característico da matriz L : $q(\lambda)$. Como o hamiltoniano possui uma dependência linear com os parâmetros $J = (J_{A_1}, J_{A_2}, \dots, J_{A_k})$, segue que as entradas da matriz transferência, que são exponenciais da energia multiplicada por β , são funções C^∞ de β e J . Como o polinômio $q(\lambda)$ é uma composição de funções de classe C^∞ , ele também é uma função de classe C^∞ : $q = q(\beta, J, \lambda)$.

Fixados dois valores β_0 e J_0 das variáveis β e J , o Teorema 2.1 nos garante a existência do autovalor λ_{+0} correspondente a esses valores. Claramente, $q(\beta_0, J_0, \lambda_{+0}) = 0$. Segue ainda do teorema 2.1 que $\frac{\partial q}{\partial \lambda}(\beta_0, J_0, \lambda_{+0}) \neq 0$, pois λ_{+0} é um autovalor simples.

O Teorema da Função Implícita nos garante que, em uma vizinhança N de (β_0, J_0) , existe uma única função $\Phi : N \rightarrow \mathbb{R}$ de classe C^∞ tal que $\Phi(\beta, J)$ é solução de $q(\beta, J, \Phi(\beta, J)) = 0$ satisfazendo $\Phi(\beta_0, J_0) = \lambda_{+0}$.

Tudo que nos resta argumentar, nesse momento, é que a função $\Phi(\beta, J)$ corresponde ao autovalor $\lambda_+(\beta, J)$ para todo par (β, J) em N . De fato, dado um outro ponto (β_1, J_1) em N , então, como $q(\beta_1, J_1, \Phi(\beta_1, J_1)) = 0$, $\Phi(\beta_1, J_1)$ é um autovalor de L . Suponha por absurdo que seja $\Phi(\beta_1, J_1) \neq \lambda_+(\beta_1, J_1)$. Então, novamente, o Teorema da função implícita nos permite afirmar que existe uma outra função C^∞ , $\Phi'(\beta, J)$, definida suficientemente próxima de (β_1, J_1) , tal que $q(\beta, J, \Phi'(\beta, J)) = 0$ e $\Phi'(\beta_1, J_1) = \lambda_+(\beta_1, J_1) \neq \Phi(\beta_1, J_1)$. Mas isso é uma contradição com o fato de que $\Phi(\beta, J)$ é única em N .

Como o ponto β_0, J_0 foi escolhido arbitrariamente, podemos afirmar, agora, que para todo ponto (β, J) existe uma vizinhança em que a função $\lambda_+(\beta, J)$ é uma função de classe C^∞ . Segue daí que essa função é de classe C^∞ .

Como conseqüência do Teorema de Perron Frobenius 2.1 temos que $\lambda_+(\beta, J) > 0$ para todo β, J . Sendo assim, sua composta com o logaritmo natural, $\ln(\lambda_+)$ está bem definida e também é uma função de classe C^∞ . □

A Agradecimentos

Marcelo Hilário agradece ao CNPq pela bolsa de iniciação científica vinculada ao projeto de pesquisa desenvolvido pelo Prof. Gastão Braga.

B Apêndice 1

Nesse apêndice vamos provar a afirmação feita durante a prova do Teorema de Perron-Frobenius 2.1 de que a matriz A e sua transposta A^* compartilham o maior autovalor λ_+ .

Lema B.0.1 *Seja A uma matriz $d \times d$ tal que $A_{ij} > 0 \forall i, j \in \{1, \dots, d\}$ e A^* sua transposta, isto é, $A_{ij}^* = A_{ji}$. Se λ_+ é o maior autovalor de A então ele também é o maior autovalor de A^* .*

Prova: Como tanto A quanto A^* são matrizes com entradas positivas e como, nesse ponto do nosso trabalho, já provamos os itens 1 e 2 do Teorema de Perron-Frobenius, podemos afirmar que a matriz A possui um autovetor $x^{(0)}$, com todas as componentes positivas, e com valor λ_+ . De forma análoga, sejam $y^{(0)}$ e σ_+ o autovetor de componentes positivas e o maior autovalor de A^* , respectivamente. Lembrando das definições das funções α_i e f , através das equações (2) e (3), definimos suas análogas para a matriz A^* :

$$\begin{aligned}\beta_i(v) &\equiv v_i^{-1} \sum_{j=1}^d A_{ij}^* v_j; \\ g(v) &\equiv \max_i(\beta_i).\end{aligned}$$

Observe que $f(x^{(0)}) = \lambda_+$ e que $g(y^{(0)}) = \sigma_+$. Analisamos agora o seguinte produto escalar:

$$\langle y^{(0)}, Ax^{(0)} \rangle = \lambda_+ \langle y^{(0)}, x^{(0)} \rangle \quad (49)$$

Ainda sobre esse produto escalar, podemos escrever que:

$$\langle y^{(0)}, Ax^{(0)} \rangle = \langle A^* y^{(0)}, x^{(0)} \rangle = \sum_{i=1}^d x_i^{(0)} \sum_{j=1}^d A_{ij}^* y_j^{(0)} = \sum_{i=1}^d x_i^{(0)} y_i^{(0)} \beta_i(y^{(0)}).$$

Lembrando que g é obtida tomando o máximo em i para β_i , temos:

$$\langle y^{(0)}, Ax^{(0)} \rangle \leq g(y^{(0)}) \sum_{i=1}^d x_i^{(0)} y_i^{(0)} = g(y^{(0)}) \langle y^{(0)}, x^{(0)} \rangle. \quad (50)$$

Então, como $\langle y^{(0)}, x^{(0)} \rangle \neq 0$, obtemos das equações (49) e (50) que $\lambda_+ \leq g(y^{(0)})$. Em particular, como $g(y^{(0)}) = \sigma_+$, temos que $\lambda_+ \leq \sigma_+$. Para provar a desigualdade no sentido inverso, repetimos o argumento começando com

$$\langle x^{(0)}, A^* y^{(0)} \rangle = \sigma_+ \langle y^{(0)}, x^{(0)} \rangle.$$

Então,

$$\langle x^{(0)}, A^* y^{(0)} \rangle = \langle Ax, y^{(0)} \rangle = \sum_{i=1}^d y_i^{(0)} \sum_{j=1}^d A_{ij} x_j^{(0)} = \sum_{i=1}^d x_i^{(0)} y_i^{(0)} \alpha_i(x^{(0)}) \leq f(x^{(0)}) \langle y^{(0)}, x^{(0)} \rangle$$

e concluímos que $\sigma_+ \leq f(x^{(0)}) = \lambda_+$. A conclusão final é que $\lambda_+ = \sigma_+$ e o lema está provado. □

C Apêndice 2

Nesse apêndice provaremos o seguinte teorema

Teorema C.1 *Seja A uma matriz $d \times d$ e λ um autovalor de multiplicidade geométrica igual a 1, com autovetor associado x . Se a equação $(A - \lambda)w = x$ não tem solução em \mathbb{C}^d então a multiplicidade algébrica de λ também é igual a 1.*

Este teorema foi usado na prova do item 3 do Teorema de Perron-Frobenius. O teorema é colorário do seguinte lema:

Lema C.1.1 *Seja A uma matriz $d \times d$, $\lambda \in \mathbb{C}$ e $x \in \mathbb{C}^d$ tais que $Ax = \lambda x$. Seja m_g e m_a as multiplicidades geométrica e algébrica de λ , respectivamente. Se $m_g = 1 < m_a$, então a equação $(A - \lambda)w = x$ tem solução $w \in \mathbb{C}^d$.*

Prova: Definimos W_1 como o núcleo de $(A - \lambda)$ e W_2 como núcleo de $(A - \lambda)^{m_a}$. Observe que $x \in W_1 \subset W_2$, que a dimensão de W_1 é 1 (pois $m_g = 1$). Como a dimensão de W_2 é igual a m_a [10] e como $m_a \geq 2$, segue então das observações acima que existe $v \in W_2$, linearmente independente de x . Em particular, $v \notin W_1$.

Afirmamos que $\exists k \in \mathbb{N}$ com $1 \leq k < m_a$, tal que $(A - \lambda)^{m_a - k}v \in W_1$ e $(A - \lambda)^{m_a - k}v \neq 0$. De fato, como $v \in W_2$ e como $m_a \geq 2$, temos

$$0 = (A - \lambda)^{m_a}v = (A - \lambda) [(A - \lambda)^{m_a - 1}v] \Rightarrow (A - \lambda)^{m_a - 1}v \in W_1. \quad (51)$$

Se for $(A - \lambda)^{m_a - 1}v \neq 0$, então tome $k = 1$. Caso contrário, repita o procedimento dado por (51) até encontrar o número $k > 1$ desejado. A existência de tal k está garantida pois se ocorresse que $(A - \lambda)^{m_a - k}v = 0$ para todo k menor do que m_a , então tomaríamos $k = m_a - 1$ e concluiríamos que $v \in W_1$, uma contradição com o fato de v ser l.i. com x . Sendo assim, como W_1 é unidimensional, existe $\alpha \neq 0$ tal que $(A - \lambda)^{m_a - k}v = \alpha x$. Se $k = m_a - 1$, então tome $w = \frac{v}{\alpha}$. Caso contrário, fatore mais uma vez um termo $(A - \lambda)$ e obtenha $(A - \lambda)[(A - \lambda)^{m_a - k - 1}v] = \alpha x$. Fazendo $w = (A - \lambda)^{m_a - k - 1} \frac{v}{\alpha}$ temos $(A - \lambda)w = x$. \square

D Apêndice 3

Seja A uma matriz $d \times d$ e $\sigma(A) \equiv \{\lambda | \lambda \text{ é autovalor de } A\}$ o seu espectro. O raio espectral da matriz A é o número $r(A) = \sup_{\lambda \in \sigma(A)} |\lambda|$. A norma da matriz A é definida como $\|A\| \equiv \sup_{v \neq 0} \frac{\|Av\|}{\|v\|}$. O raio espectral de A e sua norma se relacionam através da seguinte identidade:

$$r(A) = \lim_{n \rightarrow \infty} \|A^n\|^{1/n} \quad (52)$$

A relação acima está provada, por exemplo, em [13] e será utilizada para provar que A^N é uma contração se $r(A) < 1$ e se n for suficientemente grande. De fato temos:

Proposição D.1 *Seja A uma matriz $d \times d$ com $r(A) < 1$. Então*

$$\lim_{n \rightarrow \infty} \|A^n\| = 0$$

Prova: Tome $\varepsilon = \frac{1-r(A)}{2}$. Segue da equação (52) que $\exists n_0 \in \mathbb{N}$ tal que para todo $n > n_0$ vale $\|A^n\|^{\frac{1}{n}} < \frac{1+r(A)}{2}$. Mas como $r(a) < 1$, então $\frac{1+r(A)}{2} < 1$ e então $0 \leq \|A^n\| < \left[\frac{1+r(A)}{2}\right]^n$. Daí segue, imediatamente, que $\lim_{n \rightarrow \infty} \|A^n\| = 0$. \square

Referências

- [1] Perron, O., Zur Theorie der Matrices, Math. Ann. 64 (1907), 248 – 264.
- [2] Perron, O., Grundlagen für eine Theorie des Jacobischen Kettenbruchalgorithmus, Math. Ann. 64 (1907), 1 – 76.
- [3] Frobenius, F. G., Über Matrizen aus positiven elementen. Sitz-Ber A Kad. Wiss.; Über Matrices aus nichtneativen elementen, i bid. (1912).
- [4] Ising, E, Beitrag zur Theorie des Ferromagnetismus, Z. Physik 31, 253, (1925).
- [5] Onsager, L., Crystal statistics, I. A two-dimensional model with an order-disorder transition, Phys. Rev. 65, 117.
- [6] Gastão A. Braga e Francisco F. Araujo Jr., Matemática Universitária, O Limite Termodinâmico e Independência das Condições de Contorno para o Modelo de Ising d-dimensional, 31, 101 – 126, (2001).
- [7] Thompson, Colin J.; Mathematical Statistical Mechanics, 1972; Princeton University Press, Princeton, N.J.; Cap 5, 116 – 142.
- [8] B. Simon, The Statistical Mechanics of Lattice Gases, Volume 1, Princeton University Press, Princeton, (1993).
- [9] D. Ruelle, Statistical Mechanics: Rigorous Results. W. A. Benjamin Inc., New York, (1969).
- [10] Halmos, P. R., Espaços vetoriais de dimensão finita, ed. Campos, Rio de Janeiro pág. 116, (1978).
- [11] Robinson, C., Dynamical Systems: Stability, Symbolic Dynamics and Chaos; CRC Press LLC, Boca Raton, Florida, (1999).
- [12] Chung, Kai Lai. Elementary probability theory with stochastic processes. New York: Springer (1974).
- [13] Reed, M., Simon, B., Methods of Modern Mathematical Physics. Academic Press, New York, San Francisco, London (1972).