



Entropia em Redes Complexas

Palavras-Chave: Redes complexas, Entropia espectral, Comparação estrutural

Aluno: *Guilherme Massera Palmos* Orientador: *Prof. Dr. José Antônio Brum*

Departamento de Física da Matéria Condensada

Instituto de Física Gleb Wataghin - Universidade Estadual de Campinas

Campinas, 2025

1 Introdução

No cotidiano, o mundo está imerso em diversos sistemas de conexões de alta complexidade, cuja análise de cada elemento participante torna-se inviável. É possível citar alguns exemplos de redes muito presentes no dia a dia, como as redes de transportes, as redes sociais e as redes elétricas. Cada uma com sua peculiaridade, seus elementos participantes, que são chamados de nós, e suas respectivas conexões. Neste projeto, o interesse central está na rede neural da *C. Elegans*, muito estudada na área da neurociência.

Devido às limitações experimentais e teóricas, bem como à vasta extensão de redes neurais, é necessário analisá-las de forma macroscópica, a fim de facilitar sua análise. Para isso, deve-se aplicar uma renormalização da rede, de forma que não haja a perda de informações relevantes durante o processo. Entre os modelos propostos de renormalização de redes complexas, está o reescalonamento pelo Laplaciano, fundamentado na difusão de informação na rede complexa. Essas abordagens servem de motivação para este projeto, que foca no estudo do conceito de entropia e na dinâmica das redes complexas.

O presente documento apresenta um resumo do que foi estudado e realizado no projeto de Iniciação Científica, apresentado acima. O estudo apresentado aqui foi fundamentalmente apoiado nas referências [2] e [3].

2 Metodologia e discussão

Após o aprofundamento no conhecimento de redes complexas e suas caracterizações, foi realizado um estudo dos diversos tipos de entropia. A Entropia de Von Neumann teve um papel fundamental neste projeto. Portanto, nas próximas seções, tal

entropia foi definida e algumas de suas propriedades mencionadas.

2.1 Entropia de Von Neumann

A entropia de Von Neumann para um estado quântico ρ é definida por:

$$S(\rho) = -\text{Tr}(\rho \log_2(\rho)), \quad (1)$$

onde ρ é um operador Hermitiano e positivo semi-definido. Ou seja, podemos decompô-lo em termos de seus autovalores e autovetores, da forma:

$$\rho = \sum_x \lambda_x |x\rangle\langle x|, \quad (2)$$

onde λ_x é o autovalor associado ao autovetor $|x\rangle$ de ρ . Dessa forma, a entropia de Von Neumann pode ser reescrita da seguinte forma:

$$S(\rho) = - \sum_x \lambda_x \log_2 \lambda_x. \quad (3)$$

2.1.1 Entropia relativa quântica

Suponha que ρ e σ sejam operadores de densidade. A entropia relativa para os dois operadores é tal que

$$S(\rho||\sigma) = \text{tr}(\rho \log_2 \rho) - \text{tr}(\rho \log_2 \sigma). \quad (4)$$

Pela desigualdade de Klein, a entropia relativa quântica é não negativa, isto é $S(\rho||\sigma) \geq 0$.

2.2 Entropia espectral como ferramenta de comparação de redes complexas

2.2.1 Entropia de Von Neumann para redes

Seja $p(t)$ um vetor que indica a quantidade de informação em um ou mais nós em um tempo t .

Tem-se que a taxa de variação temporal de $p(t)$ é dada por:

$$p(t) = e^{-t\hat{L}}p(0), \quad (5)$$

onde $\hat{L} = \hat{D} - \hat{A}$ é o Laplaciano do grafo, definido pela subtração da matriz \hat{D} , que é uma matriz diagonal, cujos elementos são os graus de cada nó, por \hat{A} , que é a matriz adjacente da rede, cujo elemento A_{ij} é o peso do link que conecta o nó i ao nó j . Aqui, $e^{-t\hat{L}}$ denota o propagador de difusão da rede. Por definição, a matriz de densidade $\hat{\rho}$ é tal que:

$$\hat{\rho} = \frac{e^{-t\hat{L}}}{Z}, \quad (6)$$

onde $Z = \text{Tr}(e^{-t\hat{L}})$. Pode-se decompor L da forma

$$\hat{L} = \hat{Q}\hat{\Lambda}\hat{Q}^{-1}, \quad (7)$$

onde \hat{Q} é a matriz a qual suas colunas são os autovetores de \hat{L} e $\hat{\Lambda}$ é a matriz diagonal formada pelos autovalores do Laplaciano. Assim,

$$Z = \sum_{i=1}^N e^{-\beta\lambda_i}, \quad (8)$$

onde λ_i é o i -ésimo autovalor de L e β um parâmetro, atuando com um papel semelhante ao tempo. Analogamente, a entropia espectral pode ser escrita da seguinte forma:

$$S(\hat{\rho}) = - \sum_{i=1}^N \omega_i(\hat{\rho}) \log_2 \omega_i, \quad (9)$$

onde ω_i é o i -ésimo autovalor de $\hat{\rho}$, e pode-se calculá-lo em função de λ_i da forma

$$\omega_i = Z^{-1} e^{-\beta\lambda_i}. \quad (10)$$

Assim, a entropia espectral do grafo G em função dos autovalores de \hat{L} fica

$$S(G) = \log_2 Z - \beta \frac{1}{Z \ln(2)} \sum_{i=1}^N -\lambda_i e^{-\beta\lambda_i}, \quad (11)$$

que pode ser reescrita como:

$$S(G) = \log_2 Z + \frac{\beta}{\ln 2} \text{Tr}[\hat{L}\hat{\rho}]. \quad (12)$$

Abaixo, um gráfico ilustra as curvas da entropia normalizada em função do inverso do parâmetro β de uma rede neural da *C.Elegans* e de redes Erdős-Rényi, com a probabilidade de conexão p variando de 0 (grafo totalmente desconectado) até 0.12.

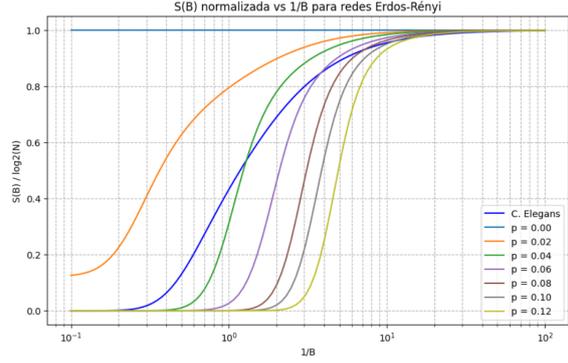


Figura 1: Gráfico com a comparação da entropia normalizada pelo inverso do parâmetro β entre uma rede *C.Elegans* e redes Erdős-Rényi com diversos parâmetros p .

Maximização da verossimilhança de um modelo: Voltando o foco para a equação (4), a entropia relativa entre uma matriz de densidade $\hat{\rho}$ e um modelo para a nossa rede, que depende de um parâmetro Θ , definido pela matriz $\hat{\sigma}(\Theta)$ é tal que

$$D(\hat{\rho}||\hat{\sigma}(\Theta)) = \text{Tr}[\hat{\rho} \log_2 \hat{\rho}] - \text{Tr}[\hat{\rho} \log_2 (\hat{\sigma}(\Theta))]. \quad (13)$$

A equação acima é a divergência de Kullback-Leibler, e nos indica o quão "distante" um modelo está de descrever a rede a ser analisada. A fim de minimizar essa divergência, mudando os parâmetros do modelo, basta maximizar o segundo termo do lado direito da equação, ou seja,

$$\min_{\Theta} [D(\hat{\rho}||\hat{\sigma}(\Theta))] = \max_{\Theta} [\text{Tr}[\hat{\rho} \log_2 (\hat{\sigma}(\Theta))]]. \quad (14)$$

Assim, vem a definição da função *log-likelihood* de uma rede:

$$\log_2 \mathcal{L}(\Theta) = \text{Tr}[\hat{\rho} \log_2 (\hat{\sigma}(\Theta))]. \quad (15)$$

Ou seja,

$$\mathcal{L}(\Theta) = e^{\text{Tr}[\hat{\rho} \ln (\hat{\sigma}(\Theta))]} = \det[e^{\hat{\rho} \ln (\hat{\sigma}(\Theta))}]. \quad (16)$$

Tem-se alguns critérios de seleção do modelo a ser utilizado. Por exemplo, o Critério de Informação de Akaike (*AIC - Akaike Information Criterion*) tende a ver o maior benefício entre a minimização do número de parâmetros a serem utilizados e a maximização de $\mathcal{L}(\Theta)$, ou seja, é de interesse utilizar o modelo que minimiza a equação a seguir:

$$AIC = 2k - 2\log_2 \mathcal{L}(\Theta^*), \quad (17)$$

onde k é o número de parâmetros utilizados pelo modelo e Θ^* é o melhor arranjo de parâmetros de um modelo.

A Figura abaixo mostra a divergência de Kullback-Leibler entre uma rede de referência Erdős-Rényi ($p^* = 0.5, N = 200$) e redes modelo Erdős-Rényi com mesmo N mas diferentes probabilidades de conexão p .

Cada curva corresponde a um valor fixo de entropia espectral normalizada (c), representando a média de 20 amostras por valor de p .

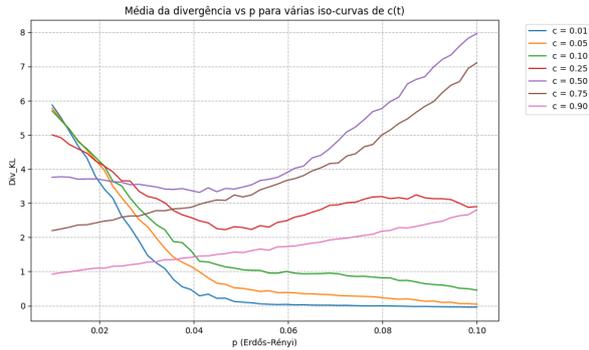


Figura 2: Divergência KL em função da probabilidade de conexão (p) para diferentes entropias normalizadas (c). Rede de referência: Erdős-Rényi com $p^* = 0.5$, $N = 200$.

Agrupamento de sistemas: É comum utilizar a divergência para investigar sistemas complexos com partes semelhantes, podendo agrupá-las, de forma inteligente, em grupos que atuarão de forma análoga, reduzindo a complexidade da rede e mantendo seu funcionamento.

2.2.2 Gap espectral

O gap espectral de um grafo é definido por:

$$\Delta\omega = \omega_1 - \omega_0, \quad (18)$$

onde ω_0 e ω_1 são, respectivamente, o menor e o segundo menor autovalores do operador de densidade ρ .

O gap espectral possui uma propriedade interessante: seja β^* o valor de β cujo gap espectral é máximo, é possível notar que o valor $\frac{1}{\beta^*}$ no gráfico da entropia normalizada corresponde a uma inflexão, isto é, uma rápida mudança da derivada da entropia em função de $\frac{1}{\beta}$. Os gráficos a seguir ilustram tal propriedade.

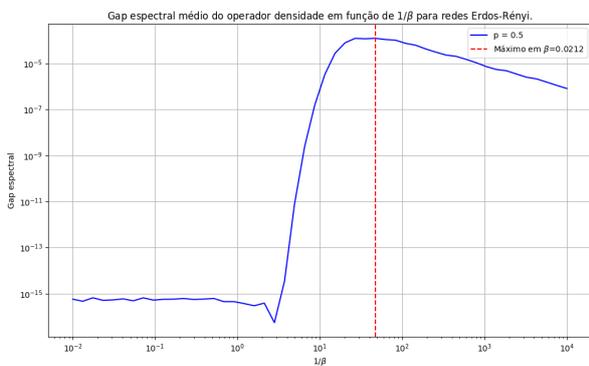


Figura 3: Gap espectral médio. Média realizada entre 50 redes Erdős-Rényi.

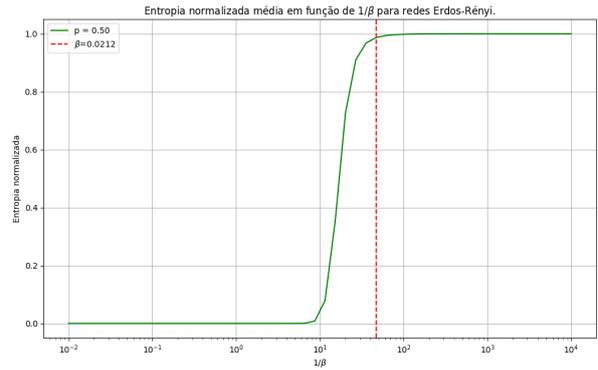


Figura 4: Entropia normalizada média. Média realizada entre 50 redes Erdős-Rényi.

A mesma análise foi realizada para a rede neural da *C. Elegans*, obtendo os seguintes resultados:

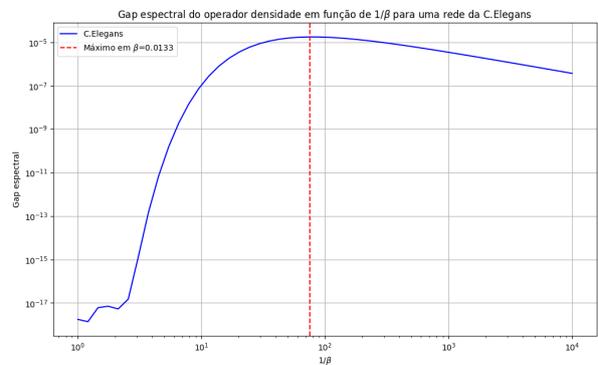


Figura 5: Gap espectral da rede neural da *C. Elegans*. Valor de $1/\beta$ cujo gap é máximo marcado em vermelho.

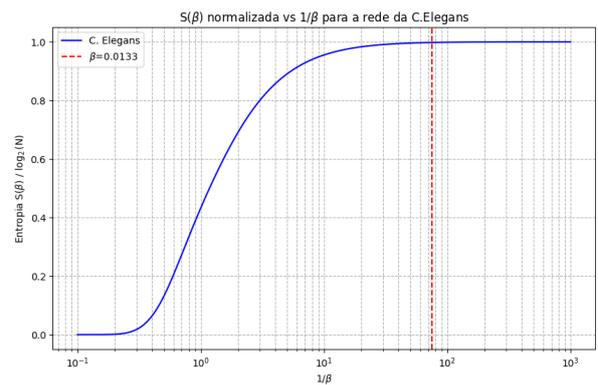


Figura 6: Entropia normalizada da rede neural da *C. Elegans*. Valor de $1/\beta$ cujo gap é máximo marcado em vermelho.

Esses resultados mostram que a rede neural da *C. Elegans* possui uma transição mais rápida na dinâmica de informação do que as redes Erdős-Rényi presentes na Figura 1 (basta observar os pontos de inflexão das curvas), indicando melhor eficiência na transmissão de informação do que redes aleatórias.

2.3 Dinâmica de Informação Complexa

As seções anteriores trataram do caso específico da difusão contínua, onde a entropia de Von Neumann se reduz à entropia espectral, e quantifica a não uniformidade dos fluxos informacionais, capturando efeitos dinâmicos invisíveis às métricas estruturais. Nessa seção, o formalismo de entropia será generalizado para outras dinâmicas de informação.

Pode-se visualizar os nós de uma rede como vetores canônicos $|i\rangle$, com $i = 1, 2, \dots, N$, onde N é o número de nós da rede. Além disso, seja $\hat{W}(t)$ a matriz adjacente da rede, que varia com o tempo. O peso do link, em um instante de tempo t , que sai do nó i e chega ao nó j , é dado pelo elemento $W_{ij}(t)$ da matriz adjacente, onde

$$W_{ij}(t) = \langle i | \hat{W}(t) | j \rangle \quad (19)$$

2.3.1 Campo de Informação

Introduz-se o campo de informação $\langle \phi(t) |$, de forma que a quantidade de campo no nó i em um instante t é dada por $\langle \phi(t) | i \rangle$.

Evolução do campo de informação: A evolução do campo de informação, na sua forma mais geral, é dada por

$$\partial_t \langle \phi(t) | = F(t, \hat{W}(t), \langle \phi(t) |). \quad (20)$$

é possível realizar uma aproximação por linearização da equação acima, de forma que $\partial_t \langle \phi(t) | = \langle \phi(t) | F(t, \hat{W}(t))$, que tem solução na forma

$$\langle \phi(t) | = \langle \phi(0) | \hat{S}(t, \hat{W}(t)), \quad (21)$$

onde $\langle \phi(0) |$ é o estado inicial do campo e \hat{S} é um propagador, elemento que diz como o sistema evolui.

Normalmente, tem-se uma distribuição probabilística, que diz qual a probabilidade de cada nó ser a fonte do campo de informação. Assim, define-se:

$$\langle \phi(0) | = \sum_i p_i \phi_o \langle i |, \quad (22)$$

onde p_i é a probabilidade do nó i ser a fonte do campo. Dessa forma, a informação esperada para sair do nó i é $\langle \phi(t) | = p_i \phi_o \langle i | \hat{S}(t)$, e a informação esperada que o nó j recebe é $p_i \phi_o \langle i | \hat{S}(t) | j \rangle$.

Exemplo: Assumindo que tem-se a mínima informação sobre as condições iniciais, tendo assim uma distribuição uniforme de probabilidade na forma $p_i = \frac{1}{N}$. Assim, o fluxo proveniente de i é $\frac{1}{N} \phi_o \langle i | \hat{S}(t)$ e a informação recebida por j é $\frac{1}{N} \phi_o \langle i | \hat{S}(t) | j \rangle$.

2.3.2 Transmissões de Informação

Um propagador diagonalizável pode ser escrito na forma decomposta

$$\hat{S}(t) = \sum_{l=1}^N s_l(t) \hat{\sigma}^{(l)}(t), \quad (23)$$

onde $s_l(t)$ é o l -ésimo autovalor do propagador e $\hat{\sigma}^{(l)}(t)$ é o produto externo entre os autovetores $|\alpha_l\rangle$ e $\langle \beta_l |$, onde $\hat{S}(t) |\alpha_l\rangle = s_l(t) |\alpha_l\rangle$ e $\langle \beta_l | \hat{S}(t) = s_l(t) \langle \beta_l |$. $|\alpha_l\rangle$ e $\langle \beta_l |$ são chamados de autovetores da direita e da esquerda, respectivamente.

Assim, considerando a distribuição equiprovável, a informação esperada que é trocada entre os pares i e j é dada por

$$\frac{1}{N} \phi_o \langle i | \hat{S}(t) | j \rangle = \sum_{l=1}^N \frac{1}{N} \phi_o s_l(t) \langle i | \hat{\sigma}^{(l)}(t) | j \rangle. \quad (24)$$

É possível observar que a informação trocada entre os nós i e j foi escrita como um somatório de N diferentes fluxos, que são direcionados por um conjunto de operadores $[\hat{\sigma}^{(l)}(t)]$, atuando como vias de transmissão de informação. Cada via é multiplicada por um coeficiente correspondente $\frac{\phi_o}{N} s_l(t)$, que pode ser interpretado como o tamanho da via de transmissão, que é considerada inativa se $\frac{\phi_o}{N} s_l(t) = 0$.

Ativação da via de transmissão de informação

É possível demonstrar que a quantidade de campo que cada transmissão de informação prende é igual ao tamanho da via de transmissão; conseqüentemente, o campo preso total esperado pode ser obtido por

$$\sum_{l=1}^N \frac{\phi_o}{N} s_l(t) = \frac{\phi_o}{N} Z(t), \quad (25)$$

onde $Z(t) = \sum_{l=1}^N s_l(t) = \text{tr}(\hat{S}(t))$. Pode-se interpretar esse resultado da seguinte maneira: o campo preso esperado regula o tamanho das vias de transmissão de informação, e pode ser considerado o responsável pela ativação das transmissões. Portanto, isso sugere que a dinâmica do campo de informações pode ser reduzida na dinâmica do campo preso. Pode-se definir o operador de densidade como

$$\hat{\rho}(t) = \sum_{l=1}^N \rho_l(t) \hat{\sigma}^{(l)}(t) = \frac{\hat{S}(t)}{\text{tr}(\hat{S}(t))}, \quad (26)$$

onde a transmissão l tem peso

$$\rho_l = \frac{\frac{\phi_o}{N} s_l(t)}{\frac{\phi_o}{N} Z(t)} = \frac{s_l(t)}{Z(t)}. \quad (27)$$

Com isso, tem-se outra descrição do fluxo esperado de informação proveniente do nó i :

$$\frac{\phi_o}{N} \langle i | \hat{S}(t) = \frac{\phi_o}{N} Z(t) \langle i | \hat{\rho}(t). \quad (28)$$

Ensemble de transmissões de informação Suponha que o campo é discretizado em um grande número de quantas infinitesimais, carregando um valor h . Estas quantas podem ser bits de informação, pequenos pacotes de energia, etc.. O número de quantas que participam na ativação das vias de transmissão é dado por

$$n(t) = \frac{\phi_0}{hN} Z(t). \quad (29)$$

Assim, dividindo a informação provinda do nó i igualmente para todas as quantas, isto é, dividindo a equação (28) por $n(t)$ obtém-se o fluxo unitário que cada quanta gera:

$$\frac{1}{n(t)} \frac{\phi_0}{N} Z(t) \langle i | \hat{\rho}(t) = h \langle i | \hat{\rho}(t) \quad (30)$$

Repare que esse processo atua de forma probabilística, cada modo de transmissão com um peso ρ_l . Assim, é possível visualizar um conjunto estatístico, formado pelos fluxos na rede, que codifica todos os fluxos possíveis entre os nós e suas respectivas probabilidades.

2.3.3 Diversidade funcional

Mesmo que as unidades de um sistema sejam similares, elas atuam de forma distinta no sistema. Tem-se que o fluxo unitário provindo de i é dado pela equação (30) e o fluxo que chega em i é dado por $h \hat{\rho}(t) | i \rangle$. O produto interno entre esses vetores é calculado por

$$h \langle i | \hat{\rho}(t) h \hat{\rho}(t) | i \rangle = h^2 \langle i | \hat{\rho}^2(t) | i \rangle, \quad (31)$$

ou seja, depende dos elementos da diagonal da matriz $\hat{\rho}^2(t)$. Quando é analisado um sistema quase misto, pode-se aproximar a entropia de Von Neumann a uma entropia linear, na forma

$$S(t) \approx 1 - \text{tr}(\hat{\rho}^2(t)), \quad (32)$$

onde $\text{tr}(\hat{\rho}^2(t)) = \sum_{i=1}^N \langle i | \hat{\rho}^2(t) | i \rangle$.

Dissimilaridade dos cossenos Pode-se calcular a similaridade dos fluxos de entrada e de saída de um nó como

$$M(t) = \frac{\langle i | \hat{\rho}(t) \hat{\rho}(t) | i \rangle}{\| \langle i | \hat{\rho}(t) \| \cdot \| \hat{\rho}(t) | i \rangle \|}. \quad (33)$$

Repare que $M(t)$ é análogo ao cosseno entre dois vetores, onde $\cos(\theta) = 0$ se os vetores são colineares e $\cos(\theta) \neq 0$ caso não sejam colineares.

Portanto, a entropia depende da similaridade e magnitude dos fluxos de entrada e de saída. Além disso, a entropia favorece a assimetria da função do componente como emissor ou receptor.

3 Conclusão

O presente trabalho explorou a aplicação de conceitos da física estatística e da teoria da informação no estudo de redes complexas, com ênfase especial na entropia espectral. Através da formulação de operadores de densidade associados ao Laplaciano da rede (fundamentada em [1], [4] e [5]), foi possível extrair medidas informacionais que refletem propriedades estruturais das redes analisadas.

A entropia de Von Neumann forneceu uma ferramenta poderosa para comparação entre diferentes topologias de rede, evidenciando transições e características específicas, como o ponto de máximo do gap espectral e suas implicações na dinâmica de informação.

Os resultados obtidos estão sendo testados em redes modelo do tipo Erdős-Rényi e comparados com a rede neural da *C.Elegans*, revelando semelhanças e distinções relevantes. As análises mostraram-se promissoras para aplicações futuras em classificação e modelagem de redes reais. No caso desse projeto, as informações provenientes da região de máximo gap espectral serão utilizadas no reescalonamento da rede neural da *C.Elegans*, preservando parte do seu funcionamento, dentro de uma margem de erro controlada.

Referências

- [1] Albert-László Barabási. *Network Science*. Cambridge: Cambridge University Press, 2016.
- [2] Manlio De Domenico e Jacob Biamonte. “Spectral Entropies as Information-Theoretic Tools for Complex Network Comparison”. Em: *Physical Review X* 6.041062 (2016), pp. 01–13.
- [3] Arsham Ghavasieh, Carlo Nicolini e Manlio De Domenico. “Statistical physics of complex information dynamics”. Em: *Physical Review E* 102 (2020), p. 052304.
- [4] Mark E. J. Newman. *Networks*. Oxford; New York: Oxford University Press, 2010.
- [5] Greiner N. Stöcker. *Thermodynamics and Statistical Mechanics*. New York: Springer Science e Business Media, LLC, 2004.