Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Matemática

Um Princípio Variacional para a Entropia Específica em Dinâmica Simbólica com Alfabetos Não-enumeráveis

por

Dióscoros Brito Aguiar Júnior

Orientador: Leandro Martins Cioletti

Durante o desenvolvimento deste trabalho, o autor recebeu auxílio financeiro do CNPq.

Brasília, 2018

Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Matemática

Um Princípio variacional para Entropia Específica em Dinâmica Simbólica com Alfabetos Não-enumeráveis.

Por

Dióscoros Brito Aguiar Júnior

Tese apresentada ao Corpo Docente do Programa de Pós-Graduação em Matemática-UnB, como requisito parcial para obtenção do grau de

DOUTOR EM MATEMÁTICA

Brasília, 04 de maio de 2018.

Comissão Examinadora:

Prof. Dr. Leandro Martins Cioletti – Orientador (MAT-UnB)

Prof. Dr. Ricardo Ruviaro (MAT-UnB)

Prof. Dr. Paulo Cesar Rodrigues Pinto Varandas (UFBA)

Prof. Dr. Elismar da Rosa Oliveira (UFRGS)

Ficha catalográfica elaborada automaticamente,
com os dados fornecidos pelo(a) autor(a)

BUM52p Brito Aguiar Júnior, Dióscoros
Um Princípio Variacional para a Entropia Específica em Dinâmica Simbólica com Alfabetos Não-enumeráveis / Dióscoros Brito Aguiar Júnior; orientador Leandro Martins Cioletti. - Brasília, 2018.
109 p.

Tese (Doutorado - Doutorado em Matemática) -- Universidade de Brasília, 2018.

1. Entropia. 2. Estados de Equilíbrio. 3. Formalismo Termodinâmico. 4. Mecânica Estatística. I. Martins Cioletti, Leandro, orient. II. Título.

"Born to lose. Live to win."

—Ian Fraser Kilmister, *A.K.A. Lemmy*

Agradecimentos

Primeiramente, quero agradecer meus pais, Dióscoros e Rosângela, e minha irmã, Ana Victória, por todo o suporte dado como conselhos, orações, um ombro amigo nas horas difíceis entre outras coisas, eles foram fundamentais para chegar a este momento.

Agradeço a todos os professores que ministraram aulas, neste doutorado, para mim: Ary, Cátia, Leandro, Liliane, Luís e Zhou. Eles me deram as condições necessárias para terminar este doutorado. Em especial, quero agradecer ao Leandro Cioletti, que além de ser meu professor, foi um excelente orientador e teve papel fundamental nesta trajetória com conselhos, orientações, parceria na pesquisa e pelos puxões de orelha quando necessário. O considero como um segundo pai.

Aos professores que estiveram na minha banca da defesa da tese: Elismar, Paulo e Ricardo. Meu agradecimento pelas críticas construtivas a esta tese e pelas consequentes melhoras. Em particular, agradeço ao Ricardo, pois ele me ajudou desde o início do desenvolvimento da teoria apresentada aqui. Não posso esquecer do Eduardo, que foi suplente, mas também me deu suporte com sugestões e conselhos ternos na véspera da defesa desta.

Aos funcionários da UnB, em particular Bruna e Cacau, que me ajudaram bastante no começo e no término desta tese, me dando totais condições para a entrada e saída deste programa de doutorado. Também agradeço aos funcionários da limpeza e aos vigias da sala 24 horas, biblioteca e do Departamento de Matemática. Sem eles, não teria condições de desenvolver esta tese e, por isso, merecem meu destaque.

Também não poderia esquecer de várias pessoas especiais, cada uma da sua maneira, que me ajudaram nesta trajetória do doutorado do início ao fim: Alan Gois, Alan Mendes, Alex, Ana Karolyna, Beatriz, Bruna Bulnes, Bruna Ribeiro, Bruno, Cacau, Camila, Carol Lafetá, Carolina, Cid, Cléber, Daiane, Danilo, Dourado, Eduardo, Élis, Elson, Filipe, Fernanda, Gabriella, Gérsica, Gilberto, Henrique, Hiuri, Hudson, Jamer, Josimar, Juliana, Júnior, Laís, Lana, Layssa, Leandro Chiarini, Leonardo, Luan, Lucimeire, Luíza, Marcos, Marília, Marta, Matheus, Melinna, Mônica, Paula, Roberto, Rodrigo, Sabrina, Samuel, Susana, Valdivino, Walter, Yerko, meus outros tios e minhas amadas avós, Maria da Glória e Maria Delfina, que sempre me deram suporte quando necessário.

Dentre estas pessoas, destaco meus tios Dourado, Mônica e Susana, que confiaram seus nomes a minha pessoa para serem meus fiadores, meus amigos mais próximos da matemática, Camila, Cid, Elson, Laís e Yerko, pois me ajudaram bastante nesta trajetória, dividindo os espaços principais de convívio, seja as residências onde morei e meu principal lugar de estudo nestes quatro anos, a sala 431/11. Não posso esquecer de mencionar meus principais amigos que conheci pela UnB, Alan Mendes, Beatriz, Bruna Bulnes, Luíza e Rodrigo. Diversas vezes nos encontramos dentro e fora da UnB para falarmos sobre a vida, estudarmos juntos entre outras coisas.

A entidade espiritual suprema que, por convenção, chamamos de Deus cujas principais representações na igreja católica são Jesus e sua mãe, Nossa Senhora Aparecida. Me considero um rapaz abençoado e, portanto, merece meu agradecimento.

Por fim, agradeço a CAPES e CNPq por todo o suporte financeiro nestes quatro anos.

Resumo

Nesta tese de Doutorado, estabelecemos importantes relações entre dois objetos fundamentais que aparecem nas teorias do Formalismo Termodinâmico e da Mecânica Estatística. A primeira se refere à relação entre a *entropia específica* (Mecânica Estatística) e a *entropia variacional* (Teoria Ergódica) para medidas de probabilidade definidas na sigma-álgebra de Borel do espaço produto $\Omega = K^{\mathbb{N}}$, onde o alfabeto K é um espaço métrico compacto arbitrário. Este resultado é o conteúdo do Teorema A. A segunda contribuição desta tese é a prova da unicidade dos *estados de equilíbrio* associados a potenciais nos espaços das funções Hölder e Walters contínuas. A novidade está na validade deste resultado de unicidade para casos em que o alfabeto K é não-enumerável. Para tratar este problema, em tal generalidade, construímos uma conexão entre a teoria de Medidas de Gibbs DLR (Dobrushin-Lanford-Ruelle) e o Formalismo Termodinâmico clássico, no Capítulo 3. O principal resultado deste capítulo é o Teorema B.

Por fim mencionamos que a ferramenta principal para a teoria desenvolvida nesta tese é o famoso operador de transferência de Ruelle. Para conveniência do leitor os resultados mais importantes sobre este operador, utilizados nesta tese, são apresentados em detalhe.

Palavras-chave: Entropia, Estados de Equilíbrio, Formalismo Termodinâmico, Mecânica Estatística.

Abstract

In this thesis, we establish important relations between two fundamental objects that appear in Statistical Mechanics and Thermodynamic Formalism theories. The first one concerns to the relationship between *specific entropy* (Statistical Mechanics) and the *variational entropy* (Ergodic Theory) for probability measures defined in Borel sigma-algebra on the product space $\Omega = K^{\mathbb{N}}$, where the alphabet K is an arbitrary compact metric space. This result is the content of Theorem A. The second contribution of this thesis is the proof of the uniqueness of the equilibrium states associated to Hölder and Walters continuous potentials. The new feature is in the validity of this uniqueness result for cases where the K is an uncountable alphabet. To address this problem, in such generality, we constructed a connection between the theory of DLR-Gibbs Measures (Dobrushin-Lanford-Ruelle) and the classical Thermodynamic Formalism, in Chapter 3. The main result of that chapter is Theorem B.

Finally, we mention that the main tool for the theory developed in this thesis is the famous Ruelle transfer operator. For the readers convenience, the most important results about this operator, used in this thesis, are presented in details.

Keywords: Entropy, Equilibrium states, Statistical Mechanics, Thermodynamic Formalism.

Sumário

Lista de símbolos

M :	Espaço métrico.
$\mathcal{B}(X;\mathbb{R})$:	Conjunto das funções limitadas de X para $\mathbb{R}$.
$C(M)$:	Conjunto das funções contínuas de M para $\mathbb{R}$.
$C^\alpha(M)$:	Conjunto das funções reais α-Hölder contínuas em M.
$W(\Omega)$:	Espaço da funções reais Walters contínuas em Ω.
$\mathcal{B}(\mathbb{R})$:	Sigma-álgebra de Borel de $\mathbb{R}$.
$(\Sigma, \mathscr{F}, P)$:	Espaço de probabilidade.
$\lambda \ll \nu$:	A medida λ é absolutamente contínua com respeito a ν.
$\mu(f\|\mathcal{B})$:	Esperança condicional de f dada a sigma-algebra $\mathcal{B}$.
Λ_n :	Cubos em $\mathbb{N}$, que são da forma $\{1, \ldots, n\}$.
$\bar{\mathscr{F}}_{\Lambda_n}$:	Sigma-álgebra definida pelas projeções com base em Λ_n.
P :	Medida de probabilidade.
$C(M, N)$:	Conjunto das funções contínuas de M para N.
$\mathrm{Hol}_\alpha(f)$:	Constante de Hölder de f.
K :	Espaço métrico compacto.
Ω :	Espaço produto $K^{\mathbb{N}}$.
$\mathscr{L}_f$:	Operador de Ruelle associado ao potencial f.
σ :	Operador *shift* à esquerda.
$\bar{f}$:	Potencial normalizado cohomologo a f.

$\mathcal{T}_n$:	Sigma-álgebra definida pelas projeções com base em Λ_n^c.
$(\gamma_{\Lambda_n})_{n\in\mathbb{N}}$:	Uma especifição quasilocal.
$\mathcal{G}^{DLR}$:	Medidas de probabilidade do formalismo DLR.
$\mathcal{G}^*$:	Automedidas do dual do operador de Ruelle.
$\mathscr{L}_f^*$:	Dual do operador de Ruelle associado ao potencial f.
$\mathcal{H}$:	Hiperplano afim.
V :	Espaço Vetorial Topológico.
$\Gamma(V)$:	Conjunto das funções que são limites pontuais de funções afim.
f^* :	Conjugada convexa de f ou a Transformada de Legendre de f.
$\mathtt{h}^{\mathtt{v}}$:	Entropia variacional.
$\mathtt{h}^{\mathtt{s}}$:	Entropia específica.
H :	Entropia de Kolmogorov-Sinai.
λ_f :	Autovalor maximal de $\mathscr{L}_f$.
$\mathscr{H}_{\mathcal{A}}(\mu\|\nu)$:	Entropia relativa com respeito a sub-sigma-álgebra $\mathcal{A}$ entre as medidas μ e ν.
$\mathscr{H}_{\Lambda_n}(\mu)$:	Entropia de μ em Λ_n em relação a medida *a priori* p.
p :	Medida *a priori* sobre o espaço compacto K.
$\boldsymbol{p}$:	Medida produto das medidas *a priori*.
$\mathscr{M}_s(\Omega)$:	Medidas com sinal sobre Ω.
$\mathscr{M}_\sigma(\Omega)$:	Medidas de probabilidade *shift*-invariantes sobre Ω.
$\mathscr{M}_1(\Sigma)$:	Medidas de probabilidade sobre Σ.
$(\Sigma, \mathscr{F})$:	Espaço mensurável.
$(\overline{\Sigma}, \bar{\mathscr{F}})$:	Espaço produto genérico.
$\varphi_{\mathcal{A}}$:	Derivada de Radon-Nikodym de $\mu\|_{\mathcal{A}}$ relativa à $\nu\|_{\mathcal{A}}$.
$\mathcal{P}$:	Partição de um conjunto.
$\mathsf{P}(f)$:	Funcional Pressão aplicado a um potencial f.

Introdução

O objetivo desta tese é fazer uma conexão apropriada entre o Formalismo Termodinâmico e a Mecânica Estatística e usá-la para extrair informações importantes sobre unicidade de estados de equilíbrio associados, por exemplo, a potenciais Hölder contínuos. Vamos começar explicando como está organizado o texto bem como suas contribuições originais. Em seguida, mencionamos o contexto histórico em que os problemas mencionados abaixo estão inseridos.

O Capítulo 1 tem caráter meramente de revisão, e nele apresentamos alguns dos resultados básicos de Análise e Teoria da Medida que serão utilizados no decorrer do texto. Já no Capítulo 2, apresentamos alguns elementos da Teoria do Formalismo Termodinâmico e a Teoria das Medidas de Gibbs no sentido de Dobrushin-Lanford-Ruelle (Gibbs-DLR). Este capítulo também possui caráter introdutório, mas alguns resultados recentes de pesquisa também são apresentados. Em ambos os capítulos, a exposição é auto-contida e feita de maneira bastante gradativa. Os tópicos selecionados, foram aqueles que serão utilizados no Capítulo 3. Fornecemos também referências onde um tratamento mais completo da Teoria do Formalismo Termodinâmico é feita.

Os principais resultados originais obtidos nesta tese são apresentados no Capítulo 3. Eles são baseados no Preprint [ACR17]. A organização e as ideias gerais deste capítulo é apresentadas abaixo.

Inicialmente, mostramos a igualdade das entropias $\mathtt{h}^{\mathtt{v}}$ e $\mathtt{h}^{\mathtt{s}}$ (ver Seções 2.3.1 e 2.3.2) para estados de equilíbrio quando o Alfabeto (Espaço de Estados) K é finito. Além disso, neste caso, mostramos que estas entropias têm uma relação direta com a Entropia de Kolmogorov-Sinai. Em seguida, exibimos um exemplo onde existe uma obstrução na relação direta entre estas entropias.

Depois seguimos para o caso em que o alfabeto K é um espaço métrico compacto arbitrário. A conexão entre o Formalismo Termodinâmico e a Mecânica Estatística é feita nesta tese por meio de uma especificação apropriada compatível com as duas teorias e ela permite provar um resultado importante que é o Teorema 3.12, que pode ser visto como uma versão da identidade (15.32) de [Geo11] no nosso contexto. Com este *background* completo, partimos para a prova dos dois resultados principais da tese.

O primeiro resultado, é o Teorema A, que mostra igualdade das entropias variacional e específica para todas as medidas *shift*-invariantes sobre Ω. Com isso, podemos caracterizar todas as medidas σ-invariantes com entropia variacional finita, quando K é finito, dada uma medida *a priori*.

Por fim, o resultado mais importante desta tese é a unicidade dos estados de equilíbrio, Teorema B, para potenciais Walters contínuos, quando K é um espaço métrico compacto arbitrário. Este resultado é uma extensão do caso clássico onde K é finito. Como consequência obtemos um certo princípio variacional, que nos garante que existe apenas um estado de equilíbrio que é o único minimizador de um certo funcional semicontínuo (ou maximizador do funcional Pressão).

Por fim, na Seção 3.3.2, mostramos que o funcional Pressão e o negativo da Entropia Específica (Variacional) e que são conjugadas convexas (uma é a Transformada de Legendre-Fenchel da outra).

O conceito de entropia, de uma medida de probabilidade, foi introduzido por Shannon [Sha48], em 1948, na Teoria de Informação, para medir a falta de comunicação média entre trocas de mensagens. Tal conceito foi baseado nos estudos de dois pioneiros da Mecânica Estatística, Boltzmann e Gibbs, sendo que o primeiro estabeleceu uma fórmula, aprimorada por Planck, para medir o desordem de um sistema para distribuições de probabilidade discretas. Três anos depois, Kullback and Liebler [KL51] generalizaram o conceito de entropia de Shannon, definindo o que seria a entropia relativa de Shannon. Neste trabalho, foi considerada uma medida *a priori* sobre o sistema e se avalia o quão distantes, em um certo sentido, estão estas duas medidas.

A existência da Entropia Específica foi provada, por Shannon, em um caso particular quando a medida *a priori* é a medida de contagem sobre um alfabeto finito. McMillan [McM53] e Breiman [Bre57] provaram que os integrandos em $n^{-1}\mathcal{H}_\Lambda(\mu)$, (ver definição 2.29), convergem em $L^1(\mu)$ e $\mu-$q.c. (ver definição 1.44). A extensão para o caso geral foi feita por Perez [Per57] no mesmo ano. Além disso, a Entropia Específica tem uma forte relação com um objeto muito importante da Teoria Ergódica, a Entropia de Kolmogorov-Sinai. Com esses conceitos bem definidos, pode-se ampliar o estudo para potenciais não triviais e Dobrushin [Dob68b] e Ruelle [Rue69] provaram a existência da energia específica, enquanto Gallavotti e Miracle-Sole [GMS67] mostraram a existência do funcional Pressão.

Além das contribuições acima na Mecânica Estatística, Ruelle, dá nome a uma ferramenta muito importante em Matemática pura e aplicada que é o operador de transferência de Ruelle. Este tem suas raízes no método da matriz transferência, introduzido por Kramers e Wannier [KW41] e, independentemente, por Montroll [Mon41] para o estudo do Modelo de Ising. Tal modelo ganhou muita relevância graças a Onsager [Ons44], pois este calculou explicitamente a função partição do Modelo de Ising bi-dimensional através desta técnica.

O operador de Ruelle, como conhecemos hoje em dia, foi introduzido, em [Rue68]. Ele age em um espaço vetorial de dimensão infinita e é muito útil para fornecer uma descrição matemática rigorosa da relação entre propriedades locais e globais de um sistema unidimensional composto de infinitas partículas, sujeito a um potencial de alcance infinito. Em particular, sob uma condição local adequada, foi provada a unicidade das medida de Gibbs DLR, que é uma condição global. Tal operador é uma das ferramentas fundamentais na Teoria Ergódica/Formalismo Termodinâmico, e um dos resultados mais importantes sobre este operador é o Teorema de Ruelle-Perron-Frobenius, que generaliza o Teorema de Perron-Frobenius para matrizes, para uma classe de operadores positivos agindo sobre um espaço vetorial adequado de dimensão infinita. Desde então, este operador é uma ferramenta poderosa para se estudar Dinâmica Topológica, medidas invariantes para fluxos de Anosov, Mecânica Estatística em uma dimensão entre outros.

As medidas de Gibbs DLR, são usadas para modelar estados de equilíbrio de um sistema físico que consiste de um número bem grande de partículas que interagem entre si. Em termos probabilísticos, uma medida de Gibbs é uma distribuição de uma família infinita e enumerável de variáveis aleatórias às quais admitem probabilidades condicionais prescritas. Desde a interpretação de Dobrushin da transição de fase em termos do número de medidas de Gibbs DLR, esta noção tem recebido uma grande atenção por físico-matemáticos e probabilistas.

A Entropia Específica, ou entropia média ou taxa de entropia, mede a quantidade de aleatoriedade na descrição de um sistema infinito para um estado de equilíbrio de um sistema. Um dos primeiros resultados no contexto da Mecânica Estatística foi provado em [NG58], onde os autores estudaram o comportamento de moléculas em sistemas de fluidos infinitos. Além da Mecânica Estatística, esta noção de Entropia Específica aparece em outros contextos, por exemplo na Teoria Ergódica com a Entropia de Kolmogorov-Sinai, que também é um tipo de Entropia Média. Por outro lado, a noção de Entropia Variacional que apresentaremos aqui, tem suas raízes em ([BCL+11],[LMST09]) e tal conceito foi fundamental na extensão da ideia de estados de equilíbrio em dinâmica simbólica de alfabetos não-enumeráveis.

No que segue, movemos a discussão para um nível matemático um pouco mais rigoroso, apresentando algumas definições e conceitos preliminares, com objetivo de explicar mais detalhadamente quais são os assuntos abordados nesta tese bem como suas principais contribuições. Definiremos o espaço onde a dinâmica acontece e as ferramentas matemáticas que serão usadas nesta tese. Além disso, o leitor terá as informações sobre as referências destas ferramentas e em qual contexto foram utilizadas para elucidar a abordagem dos problemas desta tese.

Seja $\mathbb{N}$ o conjunto dos inteiros positivos, K alfabeto finito, $\Omega \equiv K^{\mathbb{N}}$ o espaço produto equipado com sua distância usual d_{Ω}, topologia produto e sigma-álgebra gerada pelos conjuntos abertos. A dinâmica neste texto é dada pela aplicação $\sigma : \Omega \to \Omega$, definida por $\sigma(x_1, x_2, \ldots) = (x_2, x_3, \ldots)$, chamada de *shift* a esquerda. Neste contexto um *estado de equilíbrio* para um potencial contínuo $f : \Omega \to \mathbb{R}$ é um elemento de $\mathscr{M}_{\sigma}(\Omega)$, o conjunto de todas as medidas de probabilidade *shift*-invariantes de Borel, que resolvem o problema variacional

$$\sup_{\mu \in \mathscr{M}_{\sigma}(\Omega)} \{H(\mu) + \int_{\Omega} f \, d\mu\}, \tag{1}$$

onde $H(\mu)$ é a entropia de Kolmogorov-Sinai de μ, (ver Capítulo 2). Tal princípio variacional foi introduzindo por Ruelle [Rue67] no contexto da Mecânica Estatística e Walters [Wal75] abordou este problema no contexto da Teoria Ergódica. Este problema clássico continua sendo um dos mais centrais na Teoria Ergódica/Formalismo Termodinâmico e a classificação completa de todos potenciais para o qual o problema (1) admite solução única permanece em aberto.

Um dos objetivos deste trabalho é estabelecer a unicidade dos estados de equilíbrio para potenciais nas classes α-Hölder e Walters (ver Definição 1.38 para a segunda classe) sobre um contexto mais geral, onde o alfabeto K é um espaço métrico compacto geral, inclusive quando K é não-enumerável. Devemos observar que tal unicidade é bem conhecida no contexto de alfabetos finitos, veja por exemplo [Wal01, Bal00, PP90] e referências contidas. O primeiro passo rumo a tal unicidade no contexto de espaços métricos compactos gerais foi dado em [LMMS15]. Os autores provaram que qualquer medida de probabilidade μ satisfazendo $\mathscr{L}_f^* \mu = \mu$ é uma solução para o problema variacional e, além disso, que o conjunto destas medidas de probabilidade é unitário. No caso do alfabeto finito isto implica, por um resultado que está em [PP90], a unicidade dos estados de equilíbrio. Mas os resultados de [PP90], pelo menos na generalidade apresentada lá, não podem ser aplicados nos casos onde K é um espaço métrico geral. Nesta tese apresentamos uma maneira de provar esta unicidade usando a teoria do operador de Ruelle e a teoria das medidas Gibbs DLR, adequadamente adaptada para nosso contexto.

No que segue, relembraremos como o operador de Ruelle, associado a um potencial α-Hölder, e os estados de equilíbrio estão relacionados. Primeiramente no contexto de alfabetos finitos. Depois, iremos focar em contextos mais gerais.

Para cada $0 < \alpha < 1$, denotamos por $C^\alpha(\Omega) \equiv C^\alpha(\Omega, \mathbb{R})$ o espaço de todas as funções reais α-Hölder contínuas as quais são definidas por

$$\left\{ f : \Omega \to \mathbb{R} : \mathrm{Hol}_\alpha(f) \equiv \sup_{x \neq y} \frac{|f(x) - f(y)|}{(d_\Omega(x,y))^\alpha} < \infty \right\}.$$

Quando o potencial f tem boas propriedades de regularidade como α-Hölder continuidade, estados de equilíbrio e sua propriedades finas podem ser obtidos usando o formalismo do operador de transferência de Ruelle. Este operador é definido da seguinte forma. A cada função contínua φ associamos a função

$$\mathscr{L}_f(\varphi)(x) = \sum_{a \in K} \exp(f(ax))\varphi(ax), \qquad \text{onde } \ ax \equiv (a, x_1, x_2, \ldots).$$

Se f é uma função α-Hölder contínua, então o teorema Ruelle-Perron-Frobenius nos garante, entre outras coisas, que λ_f, o raio espectral de $\mathscr{L}_f$ agindo sobre $C^\alpha(\Omega)$, é um autovalor positivo maximal e associado a este temos uma autofunção h_f estritamente positiva. Além disso, existe uma medida de probabilidade Boreliana ν_f sobre Ω, tal que $\mathscr{L}_f^* \nu_f = \lambda_f \nu_f$, onde $\mathscr{L}_f^*$ é o dual de $\mathscr{L}_f : C(\Omega) \to C(\Omega)$ e $C(\Omega) \equiv C(\Omega, \mathbb{R})$ é o espaço de todas as funções reais contínuas sobre Ω. Neste caso, existe um único estado de equilíbrio para f e é dado, a menos de uma multiplicação por escalar, pela medida de probabilidade $\mu_f \equiv h_f \nu_f$, ver [Bal00, PP90, Rue68].

Além de resolver o problema variacional, a informação espectral do operador de Ruelle também pode ser usada para obter uma formulação variacional da entropia de Kolmogorov-Sinai de μ_f como se segue

$$H(\mu_f) = \log |K| + \inf_{g \in C^\alpha(\Omega)} \{ - \int_\Omega g \ d\mu_f + \log \lambda_g \},$$

ver [LMMS15]. Se $f \equiv 0$, então $H(\mu_f) = \log |K|$ e portanto quando o número de símbolos em nosso alfabeto vai para o infinito, ou seja, $|K| \to \infty$, a entropia de Kolmogorov-Sinai de μ_f será infinita.

Para lidar com o caso de alfabetos infinitos, uma nova definição para entropia de uma medida de Gibbs associada a um potencial α-Hölder é proposta em [LMMS15]. Para o problema relacionado a entropia de Kolmogorov-Sinai, os autores consideraram uma medida de probabilidade boreliana *a priori* p sobre K tendo suporte total, e um novo conceito de entropia, dependente de p é dado. Este toma valores não-positivos e atinge seu supremo na medida produto $\prod_{i \in \mathbb{N}} dp$. A introdução de uma medida *a priori* é um procedimento padrão, na Mecânica estatística de equilíbrio, quando lidamos com sistemas de *spins* contínuos (ver seção 2.2), ver [vEFS93, Geo11]. Todavia, em [LMMS15] não foi mencionada nenhuma ligação entre estas duas abordagens. Em [GKLM18], os autores desenvolvem uma teoria abstrata do Formalismo Termodinâmico, onde a entropia é definida como uma espécie de transformada de Legendre da pressão topológica. Aqui, mostramos que a entropia específica coincide com uma introduzida em [LMMS15] e pode ser estendida a transformada de Legendre clássica da pressão topológica, logo fornecendo uma representação concreta (como um limite termodinâmico) desta função nos casos onde o alfabeto K é um espaço métrico compacto geral. Além disso, obtemos o Teorema 3.12, que é uma versão da importante identidade (15.32) de [Geo11]

$$\lim_{n\to\infty} \frac{1}{|\Lambda_n|} \mathscr{H}_{\Lambda_n}(\mu|\nu) = P(\Phi) + \langle \mu, P \rangle - \mathtt{h}^{\mathtt{s}}(\mu)$$

no contexto da dinâmica simbólica para potenciais na classe de Walters, veja a Definição 1.1. Também mostramos que a entropia definida em [LMMS15] é igual a entropia específica (ou taxa de entropia) recorrentemente usada na Mecânica estatística (veja [Geo11]) e como um subproduto uma formulação variacional para a entropia específica é fornecida. Depois, estes resultados são aplicados para provar a unicidade dos estados de equilíbrio para potenciais na classe de Walters, quando K é um alfabeto métrico compacto geral. Gostaríamos de salientar que para potenciais no espaço de Walters nossos resultados não podem ser deduzidos de [GKLM18]. Primeiramente, a dinâmica não é necessariamente uma aplicação *finite-to-one*. Isto é, dada uma aplicação $f : X \to Y$, dado $y \in Y$, o número de elementos que estão na pré-imagem de $f^{-1}(y)$ pode não ser finito. O que acontece, por exemplo, quando K é infinito. Além disso, não fazemos nenhuma exigência sobre existência de *gap* espectral para o operador de Ruelle.

Capítulo 1

Resultados Preliminares

Este capítulo tem o objetivo de introduzir ao leitor notações e resultados básicos que serão utilizados direta ou indiretamente no decorrer do texto. A abordagem será básica e, caso o leitor se interesse por mais detalhes sugerimos a referência [Lim05].

1.1 Espaços Métricos

1.1.1 Definição e Exemplos de Espaços Métricos

Uma *métrica* em M é uma função $d : M \times M \to \mathbb{R}$, que leva um par ordenado de elementos $x, y \in M$ a um número real $d(x, y)$, chamado de a distância de x a y, que satisfaz para quaisquer $x, y, z \in M$:

D1) $d(x, x) = 0$;

D2) se $x \neq y$ então $d(x, y) > 0$;

D3) $d(x, y) = d(y, x)$;

D4) $d(x, z) \leq d(x, y) + d(y, z)$.

A função distância é não-negativa, sendo nula apenas quando os pontos são iguais. Além disso, esta é simétrica e satisfaz a desigualdade do triangular. Com a métrica d em M bem definida, denotamos (M, d) por Espaço Métrico. Não temos restrições para os elementos deste espaço métrico, pois a única exigência é satisfazer as quatro propriedades acima. Com a métrica bem definida, citaremos alguns exemplos.

Exemplo 1.1 (Subespaço, métrica induzida)**.** *Se (M, d) é um espaço métrico, todo subconjunto $S \subset M$ pode ser considerado, de modo natural, como espaço métrico: basta considerar a restrição de d a $S \times S$, ou seja, usar entre os elementos de S a mesma distância que eles possuíam como elementos de M. Quando isso é feito, S chama-se um subespaço de M e a métrica de S diz-se induzida pela de M.*

Exemplo 1.2 (A reta)**.** *O conjunto $\mathbb{R}$ dos números reais cuja métrica neste espaço é a distância entre dois pontos x e $y \in \mathbb{R}$ dada por $d(x, y) = |x - y|$.*

Exemplo 1.3 (Um espaço de funções)**.** *Seja X um conjunto qualquer. Uma função $f : X \to \mathbb{R}$ é dita limitada se existe uma constante $k = k_f > 0$ tal que $|f(x)| \leq k$ para todo $x \in X$. Denotaremos o conjunto dessas funções limitadas por $\mathcal{B}(X; \mathbb{R})$. Se f e g são limitadas, então $f + g$, $f - g$ e fg são limitadas. Daí, podemos definir uma métrica em $\mathcal{B}(X; \mathbb{R})$ colocando, para $f, g \in \mathcal{B}(X, \mathbb{R})$ quaisquer*

$$d(f, g) = \sup_{x \in X} |f(x) - g(x)|.$$

Exemplo 1.4 (Espaços vetoriais normados)**.** *Seja E um espaço vetorial real. Uma norma em E é uma função real $\|\cdot\| : E \to \mathbb{R}$, que associa a cada vetor $x \in E$ o número real $\|x\|$ denotado a norma de x, e esta satisfaz as condições abaixo para quaisquer $x, y \in E$ e λ escalar:*

N1) Se $x \neq 0$ então $\|x\| \neq 0$;

N2) $\|\lambda x\| = |\lambda| \|x\|$;

N3) $\|x + y\| \leq \|x\| + \|y\|$.

Onde $|\lambda|$ é o valor absoluto de λ e $\|x\|$ denota a norma de x. É fácil ver que $\|x\| > 0$ se, e somente se, $x \neq 0$.

Alguns exemplos a seguir:

E1) $(\mathbb{R}^n, \|\cdot\|_2)$, com $\|x\|_2 = \sqrt{\sum_{i=1}^n x_i^2}$;

E2) $(\mathbb{R}^n, \|\cdot\|_1)$, com $\|x\|_1 = \sum_{i=1}^n |x_i|$;

E3) $(\mathbb{R}^n, \|\cdot\|_\infty)$, com $\|x\|_\infty = \max_{i=1,\ldots,n} |x_i|$;

E4) $(\mathcal{B}(X, \mathbb{R}), \|\cdot\|_\infty)$, com $\|f\|_\infty = \sup_{z \in X} |f(z)|$.

Um espaço vetorial normado $(E, \|\cdot\|)$ se torna espaço métrico definindo d por $d(x, y) = \|x - y\|$. Dizemos que, neste caso, a métrica é proveniente da norma $\|\cdot\|$.

EM1) $\Omega_1 = \{1, \ldots, d\}^{\mathbb{N}}$, com $d_\lambda : \Omega_1 \times \Omega_1 \to \mathbb{R}$ definida por:

$$d_{\Omega_1}(x, y) = \lambda^{n_0},\ n_0 = \min\{n \in \mathbb{N}; x_i \neq y_i\},\ 0 < \lambda < 1;$$

EM2) $\Omega_2 = K^{\mathbb{N}}$, onde K é espaço métrico arbitrário e $d_{\Omega_2} : \Omega_2 \times \Omega_2 \to \mathbb{R}$ é definida por

$$d_{\Omega_2}(x, y) = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{d_K(x_n, y_n)}{1 + d_K(x_n, y_n)},$$

onde x e y são, nos dois exemplos, $x = (x_1, x_2, \ldots)$ e $y = (y_1, y_2, \ldots)$.

1.1.2 Bolas e Esferas

Um dos conceitos essenciais em espaços métricos é o conceito de bola. Dado $\varepsilon > 0$ e $x \in M$, os três conjuntos a seguir: Bola aberta, bola fechada e esfera representam, respectivamente, o conjunto de pontos que estão distantes de um certo ponto x menos que ε, menos ou exatamente ε ou exatamente ε e são representados por:

$$B(x, \varepsilon) = \{y \in M; d(x, y) < \varepsilon\};$$

$$B[x, \varepsilon] = \{y \in M; d(x, y) \leq \varepsilon\};$$

$$S(x, \varepsilon) = \{y \in M; d(x, y) = \varepsilon\}.$$

Quando d é proveniente da norma do espaço vetorial E, basta trocar d pela norma $\|\cdot\|$.

Exemplo 1.5. *Com a métrica d_λ definida sobre o espaço Ω_1 como no exemplo EN1, uma bola aberta centrada em x, de raio $\varepsilon > 0$, é o conjunto dos pontos y tais que $\lambda^{n_0} < \varepsilon$. Note que o espaço inteiro está contido em uma bola aberta centrada em x e com raio $\lambda + c$, onde $c > 0$. Já a esfera centrada em x de raio ε é não vazia se, e somente se $\varepsilon = \lambda^{n_0}$, para algum $n_0 \in \mathbb{N}$.*

1.1.3 Funções Contínuas

Estando bem definido um espaço métrico, outro conceito importante é o de função contínua, que será definida a seguir.

Definição 1.6 (Aplicação contínua)**.** *Sejam M, N espaços métricos. Diz-se que a aplicação $f : M \to N$ é contínua no ponto $a \in M$ quando para todo $\varepsilon > 0$ dado, é possível obter $\delta > 0$ tal que $d(x, a) < \delta$ implica $d(f(x), f(a)) < \varepsilon$.*

$f : M \to N$ é dita contínua sobre M se, para todo $a \in M$, f é contínua. Equivalentemente, $f : M \to N$ é contínua em $a \in M$, se dada uma bola $B_1 = B(f(a); \varepsilon)$ de centro $f(a)$, existe uma bola $B = B(a, \delta)$, de centro a, tal que $f(B) \subset B_1$.

Exemplo 1.7 (Função Lipschitziana)**.** *Dada $f : M \to N$, suponhamos que exista uma constante $c > 0$ tal que $d(f(x), f(y)) \leq c\, d(x, y)$ quaisquer que sejam $x, y \in M$. Dizemos que f é uma aplicação Lipschitziana. Segue que f é contínua em todo M. De fato, dado $\varepsilon > 0$, tomemos $\delta = \varepsilon/c$. Então $d(x, a) < \delta \Rightarrow d(f(x), f(a)) \leq c\, d(x, a) < c\,\delta = \varepsilon$.*

Exemplo 1.8 (Função uniformemente contínua)**.** *Sejam M, N espaços métricos. Uma aplicação $f : M \to N$ diz-se uniformemente contínua quando, para todo $\varepsilon > 0$ dado, existir $\delta > 0$ tal que, sejam quais forem $x, y \in M$, $d(x, y) < \delta \Rightarrow d(f(x), f(y)) < \varepsilon$.*

Exemplo 1.9 (Funções α-Hölder contínuas)**.** *Seja M um espaço métrico e $\alpha \in (0, 1)$ fixado. Uma função $f : M \to \mathbb{R}$ é chamada de α-Hölder contínua se existe uma constante $C \in \mathbb{R}$ tal que*

$$|f(x) - f(y)| \leq C d(x, y)^{\alpha}$$

para todo x, y no domínio de f. Esta condição pode ser estendida para funções entre dois espaços métricos quaisquer. O número α é chamado o expoente de Hölder. Se $\alpha = 1$, então a função é dita Lipschitziana. Se $\alpha > 0$, a condição implica a continuidade da função. Se $\alpha = 0$, a função não é, necessariamente, contínua, mas é limitada.

1.1.4 Espaços Métricos Completos

Antes de falarmos do conceito de completude de um espaço métrico, definamos o conceito de um subconjunto A ser aberto em M. Um ponto $a \in X$ é ponto interior de X se, existe uma bola aberta, com centro em a, contida neste subconjunto, para algum $r > 0$. O conjunto de todos os pontos interiores a X é denotado por int X.

Dizer que um ponto $b \in X$ não está em int X significa que toda bola aberta, centrada em b, contém um ponto não pertencente a X. Neste caso, o ponto b pertence a fronteira de X. A fronteira de X em M, denotada por ∂X, é o conjunto dos pontos $b \in M$ tais que qualquer bola aberta, centrada em b, contém ao menos um ponto de X e um ponto de $M \setminus X$.

Exemplo 1.10. *Seja $\mathbb{Q}$ o conjunto dos número racionais. O interior de $\mathbb{Q}$ em $\mathbb{R}$ é vazio pois nenhum intervalo aberto pode ser formado apenas por* *número racionais. Por outro lado a fronteira de $\mathbb{Q}$ é toda a reta $\mathbb{R}$ porque qualquer intervalo aberto contém números racionais e números irracionais.*

Dizemos que um subconjunto X de M é aberto se todos seus pontos são interiores, ou seja, $\text{int } X = X$. Daí, $X \subset M$ é aberto se, e somente se, $X \cap \partial X = \emptyset$. Por outro lado, um subconjunto C de um espaço métrico M é dito fechado neste se o conjunto complementar, $M \setminus C$, é aberto. Assim $C \subset M$ é fechado se, e somente se, $C \cup \partial C = C$.

Um ponto a é dito aderente a um subconjunto X de um espaço métrico M quando $d(a, X) = 0$. Em outras palavras, dado $\varepsilon > 0$, pode-se encontrar $x \in X$ tal que $d(a, x) < \varepsilon$.

O fecho de um conjunto X em um espaço métrico M é o conjunto $\bar{X}$ dos pontos de M que são aderentes a X. Portanto, escrever $a \in \bar{X}$ é o mesmo que afirmar que o ponto a é aderente a X em M.

A proposição a seguir nos dá as propriedades que os conjuntos abertos satisfazem em um espaço métrico M arbitrário.

Proposição 1.11. *Seja $\mathfrak{U}$ a coleção dos subconjuntos abertos de um espaço métrico M. Então:*

(1) $M \in \mathfrak{U}$ e $\emptyset \in \mathfrak{U}$;

(2) se $A_1, \ldots, A_n \in \mathfrak{U}$ então $A_1 \cap \ldots \cap A_n \in \mathfrak{U}$;

(3) se $A_\lambda \in \mathfrak{U}$ para todo $\lambda \in L$ então $A = \cup_{\lambda \in L} A_\lambda \in \mathfrak{U}$.

Demonstração. Ver Capítulo 3, Proposição 2, em [Lim05]. □

Proposição 1.12. *Sejam M, N espaços métricos. A fim de que uma aplicação $f : M \to N$ seja contínua, é necessário e suficiente que a imagem inversa $f^{-1}(A_0)$ de todo subconjunto aberto $A_0 \subset N$ seja subconjunto aberto de M.*

Demonstração. Ver Capítulo 3, Proposição 3, em [Lim05]. □

Definiremos, a seguir, uma das principais ferramentas para o estudo de espaços métricos, o conceito de sequência de Cauchy.

Definição 1.13 (Sequência de Cauchy)**.** *Dizemos que uma sequência, $\{x_n\}_{n\in\mathbb{N}}$, num espaço métrico M é uma sequência de Cauchy quando, para todo $\varepsilon > 0$, existe $n_0 \in \mathbb{N}$ tal que $m, n > n_0 \Rightarrow d(x_n, x_m) < \varepsilon$.*

Toda subsequência $\{x_{n_j}\}_{j\in\mathbb{N}^*\subset\mathbb{N}}$ de uma sequência de Cauchy é uma sequência de Cauchy. A sequência $\{x_n\}_{n\in\mathbb{N}}$ é de Cauchy se, e somente se, dado $\varepsilon > 0$, existe $n_0 \in \mathbb{N}$ tal que, se $n > n_0$ então $d(x_n, x_{n+p}) < \varepsilon$.

Proposição 1.14. *Dada uma sequência $\{x_n\}_{n\in\mathbb{N}}$ num espaço métrico M, ela satisfaz as seguintes propriedades:*

(1) toda sequência convergente é de Cauchy;

(2) toda sequência de Cauchy é limitada;

(3) uma sequência de Cauchy que possui uma subsequência convergente é convergente (e tem o mesmo limite que a subsequência);

(4) toda aplicação uniformemente contínua transforma sequências de Cauchy em sequências de Cauchy.

Demonstração. Ver Capítulo 7, Proposições 1, 2, 3 e 4, em [Lim05]. □

Definição 1.15. *O espaço métrico M é dito espaço métrico completo quando toda sequência de Cauchy em M é convergente em M.*

Em particular, quando M é espaço vetorial normado completo com a métrica induzida pela norma, dizemos que M é Espaço de Banach. A reta inspirou este conceito, pois é um dos exemplos mais básicos de espaço métrico completo. Entretanto, o espaço $\mathbb{Q}$, dos números racionais, não é completo. Apesar disso, uma boa parte de subconjuntos de espaços métricos completos é completa, os subconjuntos fechados.

Proposição 1.16. *Um subespaço fechado de um espaço métrico completo é completo. Reciprocamente, um subespaço completo de qualquer espaço métrico é fechado.*

Demonstração. Ver Capítulo 7, Proposição 6, em [Lim05]. □

Queremos estudar quando o produto cartesiano $M \times N$ é completo, a proposição a seguir elucida este fato.

Proposição 1.17. *O produto cartesiano $M \times N$ é completo se, e somente se M e N são completos.*

Demonstração. Ver Capítulo 7, Proposição 7, em [Lim05]. □

Corolário 1.18. *$M_1 \times \ldots \times M_n$ é completo se, e somente se, $M_1, \ldots, M_n$ são completos.*

Por fim, o corolário acima pode ser estendido para um produto cartesiano infinito enumerável.

Proposição 1.19. *O produto cartesiano $M = \prod_{i=1}^{\infty} M_i$ é completo se, e somente se, cada um dos fatores $\{M_i\}_{i\in\mathbb{N}}$ é completo.*

Demonstração. Ver Capítulo 7, Proposição 7^{∞}, em [Lim05]. □

Sejam M espaço métrico, X um conjunto e $\alpha : X \to M$ uma aplicação. O conjunto das aplicações $f : X \to M$ tais que

$$d(f, \alpha) = \sup_{x \in X} d(f(x), \alpha(x)) < \infty,$$

com a métrica da convergência uniforme, é denotado por $\mathcal{B}_\alpha(X; M)$. A completude destes espaços também é de nosso interesse, que é o conteúdo da proposição a seguir.

Proposição 1.20. *Se o espaço métrico M é completo então $\mathcal{B}_\alpha(X; M)$ é completo, sejam quais forem X e $\alpha : X \to M$.*

Demonstração. Ver Capítulo 7, Proposição 8, em [Lim05]. □

Com isso, chegamos ao resultado desejado para $\alpha(x) \equiv 0$ e $N = \mathbb{R}$.

Corolário 1.21. *Sejam M, N espaços métricos. Se N é completo então, para toda aplicação $\alpha : M \to N$, o espaço métrico $C(M; N)$ é completo.*

1.1.5 Espaços Métricos Compactos

Outro conceito importante é o de Espaço Métrico Compacto, relembraremos tópicos pontuais nesta subseção para o desenvolvimento da nossa teoria.

Seja X um subconjunto de um espaço métrico M. Uma cobertura de X é uma família $\mathcal{C} = (C_\lambda)_{\lambda \in L}$ de subconjuntos de M tal que $X \subset \cup_{\lambda \in L} C_\lambda$. Segue que para qualquer $x \in X$, este ponto está contido em algum C_λ, com $\lambda \in L$.

Se existe um subconjunto $L^{'} \subset L$ tal que, para cada $x \in X$, ainda se pode obter $\lambda \in L^{'}$ com $x \in C_\lambda$, isto é, $X \subset \cup_{\lambda \in L^{'}} C_\lambda$, então a subfamília $\mathcal{C}^{'} = (C_\lambda)_{\lambda \in L^{'}}$ chama-se uma subcobertura de $\mathcal{C}$. Quando $L^{'}$ é subconjunto próprio de L, $\mathcal{C}^{'}$ diz-se um subcobertura própria de $\mathcal{C}$.

Uma cobertura $X \subset \cup_{\lambda \in L} A_\lambda$ é dita aberta quando cada conjunto A_λ, $\lambda \in L$, é aberto em M. A cobertura $X \subset \cup_{\lambda \in L} C_\lambda$ é dita finita quando L é um conjunto finito.

Um espaço métrico M é dito compacto quando toda cobertura aberta possui uma subcobertura finita. Já um subconjunto K de um espaço métrico M é dito subconjunto compacto quando o subespaço métrico K é compacto. A seguir, algumas propriedades básicas de espaços métricos compactos.

Proposição 1.22. *Se o espaço M é compacto, os seguintes itens são satisfeitos:*

(1) *se $F \subset M$ é subconjunto fechado, então F é compacto;*

(2) *M é limitado;*

(3) *a imagem de um conjunto compacto por uma aplicação contínua é um conjunto compacto;*

(4) *toda função real contínua de M é limitada e atinge seus extremos (máximo e mínimo) em pontos de M.*

Demonstração. Ver Capítulo 8, Proposições 1-4, em [Lim05]. □

Outra caracterização de espaço métrico compacto é a noção de conjunto totalmente limitado, um espaço métrico M é dito totalmente limitado quando, dado $\varepsilon > 0$, existe uma decomposição $M = \cup_{i=1}^{n} X_i$, onde X_i é uma bola centrada em x_i, $x_i \in M$, com raio ε. Com isso, a caracterização de um espaço métrico ser compacto é dada na proposição abaixo.

Proposição 1.23. *As seguintes afirmações a respeito de um espaço métrico M são equivalentes:*

(1) *M é compacto;*

(2) *todo subconjunto infinito de M possui um ponto de acumulação;*

(3) *toda sequência em M possui uma subsequência convergente;*

(4) *M é completo e totalmente limitado.*

Demonstração. Ver Capítulo 8, Proposição 7, em [Lim05]. □

Assim, podemos enunciar o resultado principal desta subseção, o Teorema de Cantor-Tychonov, cuja demonstração está em [Lim05].

Teorema 1.24 (Teorema de Cantor-Tychonov)**.** *O produto cartesiano $M = \prod_{i=1}^{\infty} M_i$ é compacto se, e somente se, cada fator M_i $(i = 1, 2, \ldots, n, \ldots)$ é compacto.*

Demonstração. Ver Capítulo 8, Proposição 8bis em [Lim05].

□

1.1.6 Equicontinuidade

Queremos investigar a compacidade de subconjuntos de $C(M, N)$, com M compacto. Um subconjunto X deste espaço pode ser limitado e fechado sem ser compacto. Assim, para que um subconjunto limitado $X \subset C(M)$ tenha fecho compacto, X deve ser equicontínuo. A seguir, definimos este conceito de equicontinuidade e exibiremos um contraexemplo quando esta condição não é satisfeita.

Definição 1.25. *Sejam M, N espaços métricos e E um conjunto de aplicações $f : M \to N$. O conjunto E diz-se equicontínuo no ponto $a \in M$ quando, para todo $\varepsilon > 0$, existe $\delta > 0$ tal que $d(x, a) < \delta$ em M implique $d(f(x), f(a)) < \varepsilon$, seja qual for $f \in E$.*

Exemplo 1.26. *Os elementos da bola unitária fechada $D \subset C([0, 1])$ são as funções contínuas $f : [0, 1] \to \mathbb{R}$ tais que $|f(x)| \leq 1$ para todo $x \in [0, 1]$. D, embora limitado e fechado, não é compacto porque o espaço vetorial normado $C([0, 1])$ tem dimensão infinita. De fato, defina uma sequência de funções $f_n : [0, 1] \to \mathbb{R}$ por $f_n(x) = x^n$, estas funções são linearmente independentes, para $n = 1, 2, \ldots$. Por outro lado, D não é compacto, pois considerando a sequência de funções contínuas $g_n : [0, 1] \to \mathbb{R}$, definidas por $g_n(x) = 0$ se $0 \leq x \leq 1/(n + 1)$, $g_n(x) = 1$ se $1/n \leq x \leq 1$ e $g_n(x)$ linear, do tipo $ax + b$ se $1/(n + 1) \leq x \leq 1/n$. Disto, segue que $g_n(x) = n(n + 1)x - n$ no último intervalo. Então $\|g_m - g_n\|_\infty = 1$ se $m \neq n$ e $\|g_n\|_\infty = 1$ para todo n. Logo D não é compacto.*

Observação 1.27. *Um conjunto E de aplicações $f : M \to N$ chama-se equicontínuo quando é equicontínuo em todos os pontos de M.*

Exemplo 1.28. *A sequência de funções $f_n : \mathbb{R} \to \mathbb{R}$, $f_n(x) = x/n$ é equicontínua.*

Exemplo 1.29. *Se uma sequência equicontínua de aplicações $f_n : M \to N$ converge pontualmente para $f : M \to N$ então o conjunto $\{f, f_1, \ldots, f_n, \ldots\}$ é equicontínuo. De fato, dados $a \in M$ e $\varepsilon > 0$ existe $\delta > 0$ tal que $d(x, a) < \delta \implies d(f_n(x), f_n(a)) < \varepsilon$ para todo $n \in \mathbb{N}$ e portanto $d(f(x), f(a)) \leq \varepsilon$. Em particular, f é contínua.*

O resultado principal desta subseção é o Teorema de Arzelà-Ascoli, que será enunciado após a definição de conjunto relativamente compacto.

Definição 1.30. *Um subconjunto X de um espaço métrico M é dito relativamente compacto quando seu fecho $\overline{X}$ é compacto. Isto significa que toda sequência de pontos $x_n \in X$ possui uma subsequência convergente em M.*

Teorema 1.31 (Teorema de Arzelà-Ascoli)**.** *Seja E um conjunto de aplicações contínuas $f : K \to N$, onde K é compacto. A fim de que $E \subset C(K; N)$ seja relativamente compacto, é necessário e suficiente que:*

(1) E seja equicontínuo;

(2) para cada $x \in K$, o conjunto $E(x) = \{f(x); f \in E\}$ seja relativamente compacto em N.

Demonstração. Ver Capítulo 8, Proposição 16 em [Lim05]. □

1.1.7 Teorema de Stone-Weierstrass

Seja K um espaço métrico compacto. O conjunto $C(K)$, das funções reais contínuas $f : K \to \mathbb{R}$, é um espaço vetorial cuja norma é dada por $\|f\|_\infty = \sup\{|f(x)|; x \in K\}$. Além disso, $C(K)$ possui um produto bem definido. Dadas duas funções $f, g \in C(K)$, seu produto é a função $f \cdot g \in C(K)$, definido por $(f \cdot g)(x) = f(x) \cdot g(x)$, para todo $x \in K$. Esta multiplicação é comutativa, associativa, distributiva com respeito à adição e a função constante 1 é seu elemento neutro.

Um subconjunto $A \subset C(K)$ é dito uma álgebra de funções reais contínuas, ou uma subálgebra de $C(K)$ quando é um subespaço vetorial e $f, g \in A \Rightarrow f \cdot g \in A$. Exemplos de subálgebras de $C(K)$ são o conjunto $\{0\}$ formada pela função identicamente nula e $C(K)$. O conjunto das funções constantes, também é álgebra de funções reais contínuas. Quando $K = [a, b]$, o conjunto dos polinômios $p : [a, b] \to \mathbb{R}$ e as funções deriváveis $f : [a, b] \to \mathbb{R}$ são subálgebras de $C([a, b])$.

Dizemos que m subconjunto $A \subset C(K)$ separa pontos se para cada par $x, y \in K$, existe uma função $f \in A$ tal que $f(x) \neq f(y)$.

Finalmente, enunciaremos a seguir o teorema que dá nome a esta subseção, o Teorema de Stone-Weierestrass.

Teorema 1.32 (Teorema de Stone-Weierstrass)**.** *Sejam K um espaço métrico compacto e $A \subset C(K)$ uma álgebra de funções contínuas que contém as constantes e separa pontos. Então $\bar{A} = C(K)$, isto é, toda função contínua $f : K \to \mathbb{R}$ pode ser uniformemente aproximada por funções pertencentes a A.*

Demonstração. Ver Capítulo 8, Seção 11, em [Lim05]. □

1.1.8 Funções α-Hölder e Walters contínuas

A aplicação que nos interessa deste teorema é mostrar que o espaço das funções α-Hölder contínuas é denso no espaço das funções contínuas quando o espaço M é compacto, sob a norma da convergência uniforme. O conjunto das funções α- Hölder contínuas foi definido no Exemplo 1.9. A seguir, definiremos a constante de Hölder.

Definição 1.33 (Constante de Hölder)**.** *Dado $\alpha \in (0,1)$, a constante de Hölder de uma função f α-Hölder contínua é dada por:*

$$\mathrm{Hol}_\alpha(f) \equiv \sup_{x,y\in M, x\neq y} \frac{|f(x)-f(y)|}{d(x,y)^\alpha} < +\infty.$$

O conjunto de todas as funções reais α-Hölder contínuas será denotada por $C^\alpha(K,\mathbb{R}) \equiv C^\alpha(K)$. Neste espaço, iremos adotar a norma $\|f\|_\alpha$, definida para cada $f \in C^\alpha(K)$ por

$$\|f\|_\alpha = \|f\|_\infty + \mathrm{Hol}_\alpha(f).$$

Para mostramos o resultado de interesse, devemos mostrar que $(C^\alpha(K), \|\cdot\|_\alpha)$ é espaço de Banach. Mas para chegarmos a este resultado, precisamos do conceito de funções uniformemente equicontínuas.

Um conjunto E de aplicações $f : K \to N$ diz-se uniformemente equicontínuo quando, para todo $\varepsilon > 0$ dado, existe $\delta > 0$ tal que $d(x,y) < \delta$ em K implicar $d(f(x), f(y)) < \varepsilon$ para toda $f \in E$.

Lema 1.34. *Se K é espaço métrico compacto, todo conjunto equicontínuo de aplicações $f : K \to N$ é uniformemente equicontínuo.*

Demonstração. Ver Capítulo 8, Proposição 15 em [Lim05]. □

Proposição 1.35. *O conjunto $N = \{f \in C^\alpha(K) : \|f\|_\alpha \leq 1\}$ é compacto com respeito a norma da convergência uniforme.*

Demonstração. Primeiramente, N é um conjunto $\|\cdot\|_\infty$-fechado, pois se f é α-Hölder contínua então $\|f\|_\infty \leq \|f\|_\alpha$ e como a norma é uma função contínua, segue o resultado.

Mostraremos que N é pré-compacto. De fato, dados $\varepsilon > 0$, $f \in N$ e $x, y \in M$ quaisquer, temos

$$\begin{aligned} |f(x)-f(y)| &\leq \mathrm{Hol}_\alpha(f)\, d(x,y)^\alpha \\ &\leq (\|f\|_\infty + \mathrm{Hol}_\alpha(f)) d(x,y)^\alpha \\ &\leq d(x,y)^\alpha. \end{aligned}$$

Tomando $\delta = \varepsilon^{\frac{1}{\alpha}}$ na desigualdade acima, temos que a família N é equicontínua. Pelo Lema 1.34, o conjunto N é uniformemente equicontínuo. Além disso, o conjunto $E(x) = \{f(x), f \in N\}$ está contido em $[-1, 1]$ e, portanto, tem fecho compacto. Assim, o Teorema de Arzelà-Ascoli pode ser aplicado e nos fornecer a pré-compacidade de N. Como N é fechado, este também é compacto. □

Corolário 1.36. *$C^\alpha(K)$ é espaço de Banach*

Demonstração. Note que a aplicação $F_\beta : C(K) \to C(K)$ definida por $f \mapsto \beta f, \beta \in \mathbb{R}$ é contínua para qualquer β. Assim, $F_\beta(N)$ é $\|\cdot\|_\infty$-compacto. Seja $(f_n)_{n\in\mathbb{N}}$ sequência de Cauchy em $C^\alpha(K)$, então esta sequência é limitada e está contida em algum $F_{\beta_0}(N)$. Da compacidade deste conjunto, segue o resultado. □

Teorema 1.37. *Se K é espaço métrico compacto, então $C^\alpha(K)$ é uniformemente denso em $C(K)$. Ou seja,*

$$\overline{C^\alpha(K)}^{\|\cdot\|_\infty} = C(K).$$

Demonstração. Se d é uma métrica em K, d^α, com $0 < \alpha < 1$, também é uma métrica em K. De fato, as condições D1, D2, D3 são imediatamente satisfeitas, só faltando mostrar a desigualdade triangular. Como $f(x) = x^\beta$ é crescente para todo $\beta > 0$, em particular para nosso intervalo de interesse, basta mostrar que $(a + b)^\alpha \leq a^\alpha + b^\alpha$ para quaisquer $a, b > 0$, ou equivalentemente que,

$$1 \leq \left(\frac{a}{a+b}\right)^\alpha + \left(\frac{b}{a+b}\right)^\alpha.$$

Como $0 < \alpha < 1$, temos que

$$\left(\frac{a}{a+b}\right)^\alpha \geq \frac{a}{a+b} \quad \text{e} \quad \left(\frac{b}{a+b}\right)^\alpha \geq \frac{b}{a+b}.$$

Portanto, somando as duas desigualdades acima obtemos o que se queria. Sendo assim, reduzimos a classe de funções para as Lipschitzianas, no sentido de mostrar a densidade em $C(K)$. Denote tal classe por $\mathrm{Lip}(K)$, esta contém a família das funções constantes, se, além disso, mostrarmos que separa pontos então o teorema segue do Teorema de Stone-Weierstrass. Para ver que $\mathrm{Lip}(K)$ separa pontos vamos usar o fato de que a métrica d_K é Lipschitz. Sejam $x_0, y_0 \in K$, distintos, e considere a função $f : K \to \mathbb{R}$ dada por $f(y) = d_K(x_0, y)$, temos $f(y_0) > 0 = f(x_0)$ e esta função é Lipschitziana. De fato,

$$|f(y) - f(x)| = |d_K(y, x_0) - d_K(x_0, x)| \leq d_K(x, y).$$

Temos o que queríamos. □

Outra classe de funções que nos interessa neste texto é a classe da funções contínuas sobre $\Omega = K^{\mathbb{N}}$, com K espaço métrico compacto, que satisfazem a condição de Walters, sua notação é dada por $W(\Omega, \sigma)$, onde σ denota o operador *shift* à esquerda (ver Definição 2.1) e a definição formal dessa condição é dada abaixo:

Definição 1.38 (Espaço Walters). *Uma função contínua $f : \Omega \to \mathbb{R}$ satisfaz a condição de Walters se*

$$\sup_{n\geq 1} \sup_{\mathbf{a}\in K^n} |S_n(f)(\mathbf{a}x) - S_n(f)(\mathbf{a}y)| \to 0, \quad \textit{quando } d_\Omega(x, y) \to 0, \tag{1.1}$$

onde $S_n f(x) = \sum_{k=0}^{n-1} f(\sigma^k(x))$ e K^n denota o produto cartesiano de K n vezes. O conjunto de todas tais funções é denotado por $W(\Omega, \sigma) \equiv W(\Omega)$ e é chamado Classe de Walters.

1.2 Teoria da Medida

Falaremos brevemente sobre alguns conceitos de teoria da medida e probabilidade aos quais serão essenciais no desenvolvimento do nosso trabalho. Para ver estes resultados com maior detalhe, veja, por exemplo, [Ash00], [Bar95] e [Fol99].

1.2.1 Espaços Mensuráveis

Definição 1.39 (sigma-álgebra). *Uma coleção $\mathscr{F}$ de subconjuntos de Σ é dita sigma-álgebra em Σ se satisfaz:*

(S1) Σ e $\emptyset \in \mathscr{F}$;

(S2) Se $A \in \mathscr{F}$, então A^c também está;

(S3) Se $(A_n)_{n\in\mathbb{N}}$ é uma sequência de elementos de $\mathscr{F}$, então $\cup_{n\in\mathbb{N}} A_n \in \mathscr{F}$.

Exemplo 1.40 (sigma-álgebra de Borel $\mathcal{B}(\mathbb{R})$). *Um conjunto de Borel é qualquer conjunto em um espaço topológico que pode ser formado a partir de conjuntos abertos (ou fechados) através de operações, como união ou interseção contáveis, além dos conjuntos complementares. Quando o espaço topológico é $\mathbb{R}$, considere os intervalos do tipo (a, b), $a < b \in \mathbb{R}$. A sigma-álgebra de Borel, denotada por $\mathcal{B}(\mathbb{R})$, é a menor sigma-álgebra de subconjuntos de $\mathbb{R}$ contendo todos os intervalos $(a, b), a, b \in \mathbb{R}$.*

Definição 1.41 (Espaço Mensurável). *Um par ordenado $(X, \mathscr{F})$ consistindo de um conjunto X e de uma sigma-álgebra $\mathscr{F}$ de subconjuntos de X é chamado de espaço mensurável. Qualquer conjunto em $\mathscr{F}$ é chamado um conjunto $\mathscr{F}$-mensurável, mas quando a sigma-álgebra $\mathscr{F}$ é fixa, o conjunto será dito apenas mensurável.*

Definição 1.42 (Medida). *$\mu : \mathscr{F} \to [0, \infty]$ é uma medida se satisfaz:*

a) $\mu(\emptyset) = 0$;

b) $\mu(A) \geq 0$, *para todo* $A \in \mathscr{F}$;

c) Seja $(A_n)_{n\in\mathbb{N}}$ uma coleção enumerável de conjuntos de $\mathscr{F}$ dois a dois disjuntos. Então μ satisfaz:

$$\mu\left(\cup_{n=1}^{\infty} A_n\right) = \cup_{n=1}^{\infty} \mu(A_n).$$

Tais conjuntos de $\mathscr{F}$ são denotados conjuntos mensuráveis.

Definição 1.43 (Espaço sigma-finito). *Dizemos que o conjunto Σ é sigma-finito se existe uma família enumerável $(A_n)_{n\in\mathbb{N}}$ em $\mathscr{F}$ tal que $\Sigma = \cup_{n=1}^{\infty} A_n$ e $\mu(A_n) < \infty$, para todo n.*

Os conjuntos $A \in \mathscr{F}$ com a propriedade $\mu(A) = 0$ são ditos conjuntos nulos. Tais conjuntos são importantes na definição a seguir:

Definição 1.44. *Uma condição (igualdade, convergência, etc.) é válida quase certamente com respeito a medida μ (denotada por q.c. $[\mu]$ ou simplesmente q.c. se μ é entendida) se, e somente se, existe um conjunto A $\mathscr{F}$-mensurável, com $\mu(A) = 0$, tal que a condição vale fora de B.*

Tendo a noção de espaço mensurável bem definida, enunciaremos teoremas fundamentais que serão utilizados no decorrer do texto. Para isto, definamos as funções mensuráveis.

Definição 1.45 (Função mensurável). *Sejam $(\Sigma_1, \mathscr{F}_1)$ e $(\Sigma_2, \mathscr{F}_2)$ dois espaços mensuráveis, onde os espaços Σ_1 e Σ_2 são equipados com as sigma-álgebras $\mathscr{F}_1$ e $\mathscr{F}_2$. Uma função $g : \Sigma_1 \to \Sigma_2$ é dita mensurável se a pré imagem de A sob g está em $\mathscr{F}_1$ para toda $A \in \mathscr{F}_2$, isto é,*

$$g^{-1}(A) \equiv \{x \in \Sigma_1 \,|\, g(x) \in A\} \in \mathscr{F}_2, \ \forall A \in \mathscr{F}_2.$$

Agora, temos condição de definir o que é uma função integrável. Com isto, funções que diferem apenas sobre um conjunto de medida nula podem ser identificadas como iguais, o que nos permite gerar o espaço métrico da classe de funções integráveis de Σ, denotada por $L^1(\Sigma)$.

Definição 1.46 (Função Integrável)**.** *Seja* $(\Sigma, \mathscr{F})$ *um espaço mensurável. A coleção* $L^1(\Sigma) = L^1(\Sigma, \mathscr{F}, \mu)$ *de funções integráveis consiste de todas as funções reais* $\mathscr{F}$*-mensuráveis* f *definidas sobre* X *cuja integral do módulo, com respeito a medida* μ*, é finita. Ou seja, uma função* $g : \Sigma \to \mathbb{R}$ *é dita* μ*-integrável em* $(\Sigma, \mathscr{F})$ *quando*

$$\|g\|_1 = \int_\Sigma |g| \, d\mu < +\infty.$$

O primeiro resultado de convergência para funções mensuráveis que iremos expor, não exige a integrabilidade de nenhuma das funções envolvidas, exige apenas que esta sequência de funções seja não-negativa e será enunciado a seguir.

Teorema 1.47 (Teorema da Convergência Monótona)**.** *Se* $(g_n)_{n\in\mathbb{N}}$ *é uma sequência monótona crescente de funções mensuráveis em* $(\Sigma, \mathscr{F})$*, a valores reais que convergem à* f*, então*

$$\int g \, d\mu = \lim_{n\to\infty} \int g_n \, d\mu.$$

Demonstração. Ver Teorema 4.6 em [Bar95]. □

Além desta classe de funções, outra classe de interesse é a das funções limitadas $\mu -$ q.c. sobre $(\Sigma, \mathscr{F})$. Tal classe de funções é denotada por $L^\infty(\Sigma, \mathscr{F}, \mu) \equiv L^\infty(\Sigma)$ e é descrita da seguinte forma:

Definição 1.48. *A classe das funções mensuráveis sobre* Σ *limitadas* $\mu -$ *q.c. é denotada por* $L^\infty(\Sigma)$ *e tem norma dada por:*

$$\|g\|_\infty = \inf\{B \in \mathbb{R}; |g(x)| \leq B \text{ } q.c. \text{ } sobre \text{ } \Sigma\}.$$

Observação 1.49. *Se restringirmos a norma* $L^\infty(\Sigma)$ *as funções contínuas, tal norma coincide com a norma da convergência uniforme em* $C(\Sigma)$*. Por isso, neste texto, adotaremos a notação desta norma para as funções contínuas.*

Estamos prontos para enunciar o teorema mais importante de convergência para funções integráveis, o Teorema da convergência dominada de Lebesgue.

Teorema 1.50 (Teorema da convergência dominada de Lebesgue)**.** *Sejam* $(g_n)_{n\in\mathbb{N}}$ *uma sequência de funções integráveis as quais convergem quase certamente a uma função mensurável a valores reais* g *e* μ *uma medida sobre* $(\Sigma, \mathscr{F})$*. Se existir uma função* μ*-integrável* h *tal que* $|g_n| \leq h$ *para todo* n*, então* g *é integrável e*

$$\int_\Sigma g \, d\mu = \lim_{n\to\infty} \int_\Sigma g_n \, d\mu. \tag{1.2}$$

Demonstração. Ver Teorema 5.6 em [Bar95]. □

Outro resultado que será usado no decorrer do texto é o Teorema de Fubini-Tonelli cujo conteúdo é descrito a seguir.

Teorema 1.51 (Teorema de Fubini-Tonelli)**.** *Suponha que* $(\Sigma_1, \mathscr{F}_1, \mu_1)$ *e* $(\Sigma_2, \mathscr{F}_2, \mu_2)$ *são espaços de medida sigma-finitos.*

a) [Tonelli] Se $f \in L(\Sigma_1 \times \Sigma_2)$ *é uma função positiva, então as funções* $g(x) = \int f_x d\mu_2$ *e* $h(y) = \int f^y d\mu_1$ *estão em* $L(\Sigma_1)$ *e* $L(\Sigma_2)$*, respectivamente, são positivas e*

$$\begin{aligned} \int f\, d(\mu_1 \times \mu_2) &= \int \left[\int f(x,y)\, d\mu_2(y) \right] d\mu_1(x) \\ &= \int \left[\int f(x,y)\, d\mu_1(x) \right] d\mu_2(y). \end{aligned}$$

b) [Fubini] Se $f \in L^1(\mu_1 \times \mu_2)$*, então* $f_x \in L^1(\mu_2)$ *para* $x \in \Sigma_1 - q.c.$*,* $f^y \in L^1(\mu_1)$ *para* $y \in \Sigma_2 - q.c.$*. Além disso, as funções definidas q.c.* $g(x) = \int f_x d\mu_2$ *e* $h(x) = \int f^y d\mu_2$ *estão em* $L^1(\mu_1)$ *e* $L^1(\mu_2)$*, respectivamente e a igualdade acima é válida.*

Onde f_x *e* f^y *denotam* f *com* x *e* y *fixos, respectivamente.*

Demonstração. Ver Teorema 2.37 em [Fol99]. □

Trabalharemos com um produto de uma família enumerável de espaços de medida. Precisamos definir um conceito muito importante para tais espaços, o conceito de cilindro. Sejam $(\Sigma_i, \mathcal{F}_i, \mu_i)_{i\in\mathbb{N}}$ espaços de probabilidade para todo $i \in \mathbb{N}$. Considere o produto cartesiano

$$\overline{\Sigma} = \prod_{i\in\mathbb{N}} \Sigma_i = \{(x_i)_{i\in\mathbb{N}} : x_i \in \Sigma_i\}.$$

Denotamos por cilindros de Σ os subconjuntos da forma

$$[m; A_m, \ldots, A_n] = \{(x_i)_{i\in\mathbb{N}} : x_i \in \Sigma_i \text{ para } m \leq j \leq n\}$$

com $m \in \mathbb{N}$, $n \geq m$ e $A_j \in \Sigma_j$, para $m \leq j \leq n$. Definimos a sigma-álgebra produto em $\overline{\Sigma}$ sendo a sigma-álgebra $\bar{\mathscr{F}}$ gerada pela família de todos os cilindros. Por fim, enunciaremos o seguinte resultado:

Teorema 1.52. *Existe uma única medida* μ *em* $(\overline{\Sigma}, \bar{\mathscr{F}})$ *tal que*

$$[m; A_m, \ldots, A_n] = \mu_m(A_m) \ldots \mu_n(A_n)$$

para qualquer cilindro $[m; A_m, \ldots, A_n]$*. Em particular,* μ *é uma probabilidade.*

Demonstração. Ver Teorema 0.2.13 em [VO14]. □

1.2.2 Variável Aleatória

Intuitivamente, uma variável aleatória X é uma quantidade que é medida em conexão com um experimento aleatório. Dada uma variável aleatória X sobre um espaço de probabilidade $(\Sigma, \mathscr{F}, P)$, queremos saber a probabilidade de eventos envolvendo X. Por exemplo, a probabilidade que em uma performance dada do experimento, o valor de X pertença a B, onde B é um conjunto de números reais. Em particular, estamos interessados em $P\{\omega : a < X(\omega) \leq b\}$ para todos a, b reais. Logo $\mathscr{F}$ deve conter conjuntos da forma $X^{-1}(a, b]$, e portanto todos os conjuntos $X^{-1}(B)$, B um conjunto de Borel em $\mathbb{R}$.

Definição 1.53 (Variável aleatória)**.** *Uma variável aleatória X sobre um espaço de probabilidade $(\Sigma, \mathscr{F}, P)$ é uma função Borel mensurável de Ω a $\mathbb{R}$*

Se X é uma variável aleatória sobre $(\Sigma, \mathscr{F}, P)$ a medida de probabilidade induzida por X é a medida de probabilidade P_X sobre $\mathcal{B}(\mathbb{R})$ dada por

$$P_X(B) = P\{\omega : X(\omega) \in B\}, \quad B \in \mathcal{B}(\mathbb{R})$$

Tais números caracterizam a variável aleatória X no sentido que eles fornecem as probabilidades de todos os eventos envolvendo X.

Tendo bem definido o conceito de variável aleatória, podemos falar sobre a esperança de X, um conceito fundamental na Teoria da probabilidade.

Definição 1.54. *Se X é uma variável aleatória sobre $(\Sigma, \mathscr{F}, P)$, a esperança de X é definida por*

$$E[X] = \int_{\Sigma} X \, dP,$$

se a integral existir.

O conjunto de todas as medidas de probabilidade sobre $(\Sigma, \mathscr{F})$ será denotado por $\mathscr{M}_1(\Sigma)$.

1.2.3 Teorema de Radon-Nikodym e Esperança Condicional

Suponha que λ é uma medida com sinal e μ é uma medida positiva sobre $(\Sigma, \mathscr{F})$. Dizemos que λ é absolutamente contínua com respeito a μ escrevemos

$$\lambda \ll \mu$$

se $\lambda(E) = 0$ para todo $E \in \mathscr{F}$ tal que $\mu(E) = 0$. Por outro lado, se μ é uma medida e f é uma função μ-integrável estendida, a medida com sinal λ definida por $\lambda(E) = \int_E f\,d\mu$ claramente é absolutamente contínua com respeito a μ e é finita se, e somente se $f \in L^1(\mu)$. A seguir, o teorema principal desta subseção, o que nos dá uma completa estruturação de uma medida com sinal relativa a uma medida positiva dada, O Teorema de Radon-Nikodym.

Teorema 1.55 (Teorema de Radon-Nikodym)**.** *Seja μ uma medida sigma-finita e λ uma medida com sinal sobre a sigma-álgebra $\mathscr{F}$ de subconjuntos de Σ. Assuma que λ é absolutamente contínua com respeito à μ. Então existe uma função Borel mensurável $g : \Sigma \to \overline{\mathbb{R}}$ tal que*

$$\lambda(A) = \int_A g\,d\mu \qquad \forall \;\; A \in \mathscr{F}.$$

Se h é outra função deste tipo, então $g = h$ $\mu - q.c.$

Demonstração. Ver Teorema 2.2.1 de [Ash00]. □

Este teorema nos permite abordar um dos tópicos mais importantes da teoria da probabilidade, a Esperança condicional.

De maneira intuitiva, podemos desenvolver o "teorema da esperança total" da seguinte forma. A probabilidade de X estar perto de x é $dP_X(x)$, dado que $X = x$, o valor médio de Y é que estamos procurando, denotado, $E(Y|X = x)$. É razoável esperar que a esperança total possa ser da forma:

$$E(Y) = \int_{\mathbb{R}} E(Y|X = x)\,dP_X(x).$$

Para desenvolvermos mais este resultado, trocamos Y por $YI_{\{X\in A\}}$, onde $A \in \mathcal{B}(\mathbb{R})$. Se $x \in A$, esperamos que $E(YI_{\{X\in A\}}|X = x) = E(Y|X = x)$, pois $X(\omega) = x \in A$ implica $I_{\{X\in A\}}(\omega) = 1$. Se $x \notin A$, esperamos que $E(YI_{\{X\in A\}}|X = x) = 0$. Trocando Y por $YI_{\{X\in A\}}$ na versão acima do teorema da esperança total, obtemos

$$E(YI_{\{X\in A\}}) = \int_{\mathbb{R}} E(YI_{\{X\in A\}}|X = x)\,dP_X(x)$$

ou

$$\int_{\{X\in A\}} Y\,dP = \int_A E(Y|X = x)\,dP_X(x).$$

Temos a esperança condicional $E(Y|X = x)$ determinada.

Teorema 1.56. *Sejam X, Y variáveis aleatórias sobre $(\Sigma, \mathscr{F}, P)$. Se $E(Y)$ é finita, existe uma função $g : (\mathbb{R}, \mathcal{B}(\mathbb{R})) \to (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ tal que para cada $A \in \mathcal{B}(\mathbb{R})$,*

$$\int_{\{X \in A\}} Y \, dP = \int_A g(x) \, dP_X(x).$$

Além disso, se h é outra função desta, então $g = h$ $P_X - q.c.$.

Demonstração. Seja

$$\lambda(A) = \int_{\{X \in A\}} Y \, dP = \int_{X^{-1}(A)} Y \, dP, \; A \in \mathcal{B}(\mathbb{R}).$$

Então λ é uma função sigma-aditiva sobre $\mathcal{B}(\mathbb{R})$ e é absolutamente contínua com respeito a P_X, pois $P_X(A) = P\{X \in A\}$. Assim o resultado segue do Teorema de Radon-Nikodym. □

Este resultado para variáveis aleatórias pode ser estendido a sigma-álgebras, o que nos leva ao seguinte teorema:

Teorema 1.57. *Seja Y uma variável aleatória estendida sobre $(\Sigma, \mathscr{F}, P)$, $\mathscr{S}$ uma sub-sigma-álgebra de $\mathscr{F}$. Assuma que $E(Y)$ exista. Então existe uma função $E(Y|\mathscr{S}) : (\Sigma, \mathscr{S}) \to (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$, chamada a esperança condicional de Y dada $\mathscr{S}$, tal que*

$$\int_C Y \, dP = \int_C E(Y|\mathscr{S}) \, dP \quad \text{para cada } C \in \mathscr{S}.$$

Quaisquer duas funções deste tipo coincidem $P - q.c.$

Demonstração. Ver Teorema 5.4.3. em [Ash00]. □

A esperança condicional tem diversas propriedades, condensaremos as que iremos usar neste texto em um teorema com diversos itens, enunciado com relação as sigma-álgebras. Em todos os casos abaixo, $Y, Z, Y_1, Y_2, \ldots$, são variáveis aleatórias para os números reais estendidos.

Teorema 1.58. *A esperança condicional dada uma sigma-álgebra satisfaz as seguintes propriedades:*

(E1) se Y é uma constante k quase certamente, então $E[Y|\mathscr{S}] = k$ q.c.;

(E2) se $Y_1 \leq Y_2$ q.c., então $E[Y_1|\mathscr{S}] \leq E[Y_2|\mathscr{S}]$ q.c.;

(E3) $|E[Y|\mathscr{S}]| \leq E[|Y||\mathscr{S}]$ q.c.;

(E4) *se* $a, b \in \mathbb{R}$, *e* $aE[Y_1] + bE[Y_2]$ *é bem definida (não é da forma* $\infty - \infty$*), então*

$$E[aY_1 + bY_2|\mathscr{S}] = aE[Y_1|\mathscr{S}] + bE[Y_2|\mathscr{S}] \quad q.c.$$

(E5) *se* $Y_n \geq 0$ *para todo* n *e* $Y_n \uparrow Y$ *q.c., então* $E[Y_n|\mathscr{S}] \uparrow E[Y|\mathscr{S}]$ *q.c.*

(E6) *se todas as* $Y_n \geq 0$, *então*

$$E\left[\sum_{n=1}^{\infty} Y_n|\mathscr{S}\right] = \sum_{n=1}^{\infty} E[Y_n|\mathscr{S}] \quad q.c.$$

(E7) $E[E[Y|\mathscr{S}]] = E[Y]$. *Assim se* Y *é integrável,* $E[Y|\mathscr{S}]$ *também o é;*

(E8) $E[Y|\{\emptyset, \Omega\}] = E[Y]$ *q.c.*

(E9) *se* $\mathscr{S}_1 \subset \mathscr{S}_2$, *então*

$$E[E[Y|\mathscr{S}_2]|\mathscr{S}_1] = E[Y|\mathscr{S}_1] = E[E[Y|\mathscr{S}_1]|\mathscr{S}_2] \quad q.c.$$

(E10) *Se* Z *é* $\mathscr{S}$*-mensurável e ambas* Y *e* YZ *são integráveis, então* $E[YZ|\mathscr{S}] = ZE[Y|\mathscr{S}]$ *q.c. Em particular,* $E[Z|\mathscr{S}] = Z$ *q.c.*

Demonstração. Ver Seção 5.5 de [Ash00]. □

Observação 1.59.

(O1) *As propriedades de esperança condicional podem ser vistas como as da esperança ordinária, basta tomar a sigma-álgebra como em (E8);*

(O2) *Em (E10), a mesma propriedade vale se considerarmos funções* g $\mathscr{S}$*-mensuráveis.*

1.2.4 Martingales

A Teoria de Martingales tem suas raízes nos jogos de azar, e é comumente utilizada para interpretar resultados em termos de uma situação de aposta. Por exemplo, se $X_1, \ldots, X_n$ é uma sequência de variáveis aleatórias, podemos interpretar a variável aleatória X_n como o total de vitórias após n rodadas em uma sucessão de jogos. Dado que o jogador sobreviveu a estas n jogadas, nossa fortuna esperada após a $n+1$-ésima jogada é $E(X_{n+1}|X_1, \ldots, X_n)$. Se este é igual a X_n, o jogo é justo, pois o ganho esperado na $n+1$-ésima jogada é $E(X_{n+1} - X_n|X_1, \ldots, X_n) = X_n - X_n = 0$.

Faremos uma breve introdução a sequência de variáveis aleatórias deste tipo cujos resultados tem significância além dos jogos de azar.

Definição 1.60. *Seja $(\Sigma, \mathscr{F}, P)$ um espaço de probabilidade, $\{X_1, X_2, \ldots\}$ uma sequência de variáveis aleatórias integráveis sobre $(\Sigma, \mathscr{F}, P)$, e $\mathscr{F}_1 \subset \mathscr{F}_2 \subset \ldots$ uma sequência crescente de sub-sigma-álgebras de $\mathscr{F}$ com X_n $\mathscr{F}_n$-mensurável. A sequência $\{X_n\}$ é dita ser uma martingale relativa à $\mathscr{F}_n$ se, e somente se, para todo $n = 1, 2, \ldots$, $E[X_{n+1}|\mathscr{F}_n] = X_n$ q.c.*

Antes de darmos enfoque as funções convexas, que estão na Subseção 1.4.2, iremos enunciar um resultado bem útil que usa tais funções, a desigualdade de Jensen.

Teorema 1.61 (Desigualdade de Jensen)**.** *Seja g uma função convexa de I a $\mathbb{R}$, onde I é um intervalo aberto em $\mathbb{R}$, limitado ou não. Seja X uma variável aleatória sobre $(\Sigma, \mathscr{F}, P)$, com $X(\omega) \in I$ para todo ω. Assuma $E(X)$ finito. Se $\mathcal{H}$ é uma sub-sigma-álgebra de $\mathscr{F}$, então $E[g(X)|\mathcal{H}] \geq g[E[X|\mathcal{H}]]$ q.c. Em particular, $E[g(X)] \geq g[E(X)]$.*

Antes de chegarmos ao tópico final da seção de Martingales, que são os teoremas de convergência, vamos falar sobre Integrabilidade uniforme;

Definição 1.62. *Sejam $g_1, g_2, \ldots$ funções reais ou complexas Borel mensuráveis sobre o espaço mensurável $(\Sigma, \mathscr{F}, \mu)$, μ medida finita. As g_n são ditas uniformemente integráveis se*

$$\int_{\{|g_n| \geq c\}} |g_n| \, d\mu \to 0 \quad , \text{ quando } c \to \infty$$

uniformemente em n.(Podemos considerar o conjunto de índices não-enumerável para a família de funções acima).

Dentre vários resultados de convergência para Martingales, o teorema que será usado no nosso texto será enunciado a seguir

Teorema 1.63. *$\{X_n\}$ é uma martingale uniformemente integrável se, e somente se, existe uma variável aleatória integrável Y tal que $X_n = E[Y|\mathscr{F}_n]$, $n = 1, 2, \ldots$; neste caso, $X_n \to E[Y|\mathscr{F}_\infty]$ q.c. e em L^1, onde $\mathscr{F}_\infty$ é a sigma-álgebra gerada por $\cup_{n=1}^{\infty} \mathscr{F}_n$.*

Demonstração. Ver Teorema 6.6.2 em [Ash00]. □

1.3 Topologia Fraca e Fraca-$*$

Seja X um conjunto arbitrário e $(Y_i)_{i \in I}$ uma coleção de espaços topológicos. Dada uma coleção de aplicações $(\varphi_i)_{i \in I}$ tais que para cada $i \in I$, φ é uma aplicação de X para Y_i. Queremos construir uma topologia sobre X a qual as aplicações $(\varphi_i)_{i\in I}$ são contínuas. Além disso, encontrar uma topologia $\mathfrak{T}$ que é a mais econômica no sentido de ter o número mínimo de conjuntos abertos.

Se $A_i \subset Y_i$ é um conjunto aberto qualquer, então $\varphi_i^{-1}(A_i)$ deve ser um conjunto aberto em $\mathfrak{T}$. Considerando todos os conjuntos abertos A_i em Y_i e fazendo i percorrer através do conjunto de índices I, obtemos uma família de subconjuntos de X, onde cada família deve ser aberta na topologia $\mathfrak{T}$. Denotaremos essa família por $(U_\lambda)_{\lambda\in\Lambda}$. Queremos a topologia mais grosseira $\mathfrak{T}$, sobre X, a qual U_λ é aberto para todo $\lambda \in \Lambda$, onde Λ que é associada ao conjunto $(\varphi_i)_{i\in I}$. A seguir, definiremos as topologias fraca e fraca-$*$ sobre um espaço de Banach E. Para uma abordagem mais detalhada, veja [Bre11].

Seja E um Espaço de Banach e seja $f \in E^*$, onde E^* denota o dual topológico do conjunto E. Denotamos por $\varphi : E \to \mathbb{R}$ o funcional linear $\varphi_f(x) = f(x)$. Considerando todas as funções f em E^*, obtemos uma coleção $(\varphi_f)_{f\in E^*}$ de aplicações de E para $\mathbb{R}$. Iremos ignorar a topologia induzida pela norma sobre E e definamos uma nova topologia sobre o conjunto E a seguir:

Definição 1.64. *A topologia fraca $\sigma(E, E^*)$ sobre E é a topologia mais grosseira associada a família $(\varphi_f)_{f\in E^*}$ dos funcionais lineares que as torna contínuas, onde*

$$\varphi_f(x) = f(x) \equiv \langle f, x\rangle. \tag{1.3}$$

e E^* denota o espaço de todos os funcionais lineares reais contínuos sobre E (onde $X = E$, $Y_i = \mathbb{R}$, para cada i e $I = E^*$).

Dizemos que x_n converge a x na topologia fraca se para todo f em E^*, $\langle f, x_n\rangle \to \langle f, x\rangle$. (Notação $x_n \rightharpoonup x$).

Mas nossas observações serão focadas na topologia fraca-$*$, pois iremos usar dois teoremas de representação neste texto que usam a topologia fraca-$*$. Para todo $x \in E$ considere o funcional linear $\varphi_x : E^* \to \mathbb{R}$ definido por $f \mapsto \varphi_x(f) = f(x)$. Considerando todos os pontos $x \in E$, obtemos uma coleção $(\varphi_x)_{x\in E}$ de aplicações de E^* para $\mathbb{R}$, que nos dá uma outra topologia mais grosseira de interesse, a topologia fraca-$*$ definida a seguir

Definição 1.65. *A topologia fraca-$*$ $\sigma(E^*, E)$ sobre E é a topologia mais grosseira associada a família $(\varphi_x)_{x\in E}$ dos funcionais lineares sobre E^* que as torna contínuas (com $X = E^*$, $Y_i = \mathbb{R}$, para cada i e $I = E$), com*

$$\varphi_x(f) = f(x).$$

Dizemos que f_n converge a f na topologia fraca-$*$ se para todo x em E, $\langle f_n, x\rangle \to \langle f, x\rangle$. (Notação $x_n \overset{*}{\rightharpoonup} x$).

Vamos enunciar os dois resultados importantes de interesse:

Proposição 1.66. *Seja* $\varphi : E^* \to \mathbb{R}$ *um funcional linear que é contínuo com respeito a topologia fraca-*$*$*. Então existe algum* $x_0 \in E$ *tal que*

$$\varphi(f) = \langle f, x_0 \rangle, \qquad \forall f \in E^*.$$

Demonstração. Ver Proposição 3.14 em [Bre11]. □

Considere o espaço métrico compacto K com sua sigma-álgebra de Borel $\mathcal{K}$ e $C(K)$. Se μ é uma medida finita sobre $(K, \mathcal{K})$, então a expressão $\varphi_\mu(f) = \int_K f\, d\mu, f \in C(K)$, define um funcional sobre $C(K)$. Além disso, φ_μ é funcional linear positivo, $\varphi_\mu(f) \geq 0$ para todo $f \in C(K)$ tal que $f(x) \geq 0$ para todo $x \in K$. Tomando a norma da convergência uniforme em $C(K)$, temos que φ_μ é um funcional linear contínuo sobre $C(K)$.

Teorema 1.67 (Teorema de representação de Riesz sobre $C(K)$)**.** *Seja* K *um espaço métrico compacto. Se* $\mathfrak{g}$ *é um funcional linear positivo sobre* $C(K)$*, então existe uma única medida finita* μ *sobre* $(K, \mathcal{K})$ *tal que para toda* $f \in C(K)$*,*

$$\mathfrak{g}(f) = \int_K f\, d\mu.$$

Além disso, a medida μ *é regular. Ou seja,*

$$\mu(A) = \inf\{\mu(G) : A \subset G, G \textit{ aberto}\} = \sup\{\mu(F) : F \subset A, F \textit{ fechado}\}, A \in \mathcal{K}.$$

Demonstração. Para a prova, veja o Apêndice 4 em [BW16]. □

1.4 Conjuntos Convexos

A abordagem nesta seção segue a referência [ET76].

Seja V um espaço vetorial sobre $\mathbb{R}$. Se x e y são dois pontos de V, x e y são chamados os pontos finais do segmento de reta denotado por $[x, y]$, onde

$$[x, y] = \{\lambda x + (1 - \lambda) y \,|\, 0 \leq \lambda \leq 1\}.$$

Um conjunto $C \subset V$ é dito convexo se, e somente se para todo par de elementos (x, y) de C, seu segmento $[x, y]$ está contido em V. Generalizando, C e convexo se, e somente se para todo subconjunto finito de elementos $x_1, \ldots, x_n$ de C, e para toda família de números reais positivos $\lambda_1, \ldots, \lambda_n$ com soma unitária, temos

$$\sum_{i=1}^{n} \lambda u_i \in C.$$

Seja C um subconjunto qualquer de V, a interseção de todos os conjuntos contendo C é um conjunto convexo, chamado de envoltória convexa de C e denotado por $\text{co}(C)$, que são todas as combinações convexas de C e é o menor conjunto convexo que o contém. Os elementos desta envoltória são dados por

$$\text{co}(C) = \left\{ \sum_{i=1}^{n} \lambda_i u_i | n \in \mathbb{N}, \sum_{i=1}^{n} \lambda_i = 1, \lambda \geq 0, u_i \in C, 1 \leq i \leq n \right\}.$$

Seja $\mathcal{H}$ um hiperplano com equação $f(x) = c$, com f um funcional linear não nulo sobre V e $c \in \mathbb{R}$. Os conjuntos

$$\begin{array}{ll} \{x \in V \mid f(x) < c\}, & \{x \in V \mid f(x) > c\}, \\ \{x \in V \mid f(x) \leq c\}, & \{x \in V \mid f(x) \geq c\} \end{array}$$

são chamados, respectivamente, semiespaços abertos e fechados limitados por $\mathcal{H}$. Estes são dois conjuntos convexos os quais dependem apenas de $\mathcal{H}$, e não de f e c escolhidos para esta equação.

1.4.1 Separação de Conjuntos Convexos

Relembre que um espaço vetorial topológico é definido como um espaço vetorial V equipado com uma topologia para o qual as operações de soma e multiplicação por escalar são contínuas. Tal topologia terá como base a base de vizinhanças abertas que é definida a seguir:

Definição 1.68. *Se $x \in V$, então uma base de vizinhanças (abertas) $\mathscr{B}_x$ para x é uma coleção de subconjuntos de X satisfazendo:*

(a) Cada $B \in \mathscr{B}_x$ é aberto e, além disso, é uma vizinhança de x. Ou seja, $x \in int(B)$;

(b) se U é aberto, e $x \in U$, então existe $B \in \mathscr{B}_x$ tal que $B \subset U$.

Um espaço vetorial deste é dito localmente convexo se a origem possui um sistema fundamental de vizinhanças convexas. Além disso, tal espaço é Hausdorff se pontos distintos possuem vizinhanças distintas. Queremos caracterizar o sistema fundamental de vizinhanças. Para isso, precisamos de dois conceitos: O conceito de conjunto balanceado e o conceito de conjunto absorvente.

Definição 1.69. *Seja V um espaço vetorial sobre $\mathbb{R}$. Um subconjunto $B \subset V$ é dito balanceado se para todos $x \in B$, $c \in \mathbb{R}$, com $|c| \leq 1$, temos $cx \in B$.*

Definição 1.70. *Seja V um espaço vetorial sobre $\mathbb{R}$. Um subconjunto $B \subset V$ é dito absorvente se, para todo $x \in V$, existe $t_0 > 0$ tal que $x \in cB$ sempre que $|c| \geq t_0$.*

Finalmente, especificaremos quando temos uma base de vizinhanças na origem. É o conteúdo do próximo teorema.

Teorema 1.71. *Seja V um espaço vetorial sobre $\mathbb{R}$ e seja $\mathscr{B}_0$ uma família não vazia de conjuntos convexos, balanceados e absorventes que satisfazem as duas condições:*

(a) se $B \in \mathscr{B}_0$, então $\frac{1}{2}B \in \mathscr{B}_0$.

(b) se $B_1, B_2 \in \mathscr{B}_0$, então $\exists B_3 \subset \mathscr{B}_0$ com $B_3 \subset B_1 \cap B_2$.

Então $\mathscr{B}_0$ é uma base na origem para uma (única) topologia em V, tornando V em um espaço localmente convexo. Este espaço é Hausdorff se, e somente se, $\cap \mathscr{B}_0 = \{0\}$.

Demonstração. Ver Teorema 3.2 em [Osb14]. $\square$

Para espaços vetoriais normados, é suficiente tomar o conjunto de vizinhanças formada por bolas centradas sobre a origem.

Agora, seja V um espaço vetorial topológico e $\mathcal{H}$ um hiperplano afim com equação $g(x) = c$, onde g é um funcional linear não nulo sobre V e $c \in \mathbb{R}$. Pode ser mostrado que o conjunto $\mathcal{H}$ é topologicamente fechado se, e somente se, a função g é contínua.

Um hiperplano afim $\mathcal{H}$ separa dois conjuntos C_1 e C_2 se cada um dos semiespaços limitados por $\mathcal{H}$ contém um dos conjuntos. Analiticamente, se $g(x) = c$ é a equação de $\mathcal{H}$, então

$$g(x) \leq c, \forall x \in C_1, \quad g(y) \geq c, \forall y \in C_2.$$

Quando a separação é estrita, temos

$$g(x) < c, \forall x \in C_1, \quad g(y) > c, \forall y \in C_2.$$

Relembraremos o teorema de Hahn-Banach na sua forma geométrica, e suas consequências para a separação de conjuntos convexos.

Teorema 1.72 (Teorema de Hahn-Banach)**.** *Seja V um espaço vetorial topológico, C um conjunto aberto não vazio e S um subespaço afim não vazio que não intersecta C. Então existe um hiperplano fechado afim $\mathcal{H}$ que contém S e que não intersecta C.*

Demonstração. Ver Seção 2.3 de [Sch71]. □

O corolário que será utilizado será enunciado a seguir

Corolário 1.73. *Seja V um espaço localmente convexo, C_1 e C_2 espaços convexos não vazios disjuntos com um compacto e outro fechado. Então existe um hiperplano afim fechado $\mathcal{H}$ ao qual separa estritamente C_1 e C_2.*

Outro resultado envolvendo espaços localmente convexos será utilizado, o Teorema do ponto fixo de Schauder-Tychonov e será enunciado a seguir.

Teorema 1.74 (Teorema do Ponto Fixo de Schauder-Tychonov)**.** *Seja V um espaço localmente convexo de Hausdorff, C um subconjunto convexo de V e $F : C \to E$ uma aplicação contínua tal que*

$$F(C) \subseteq K \subseteq C,$$

com K compacto. Então F tem pelo menos um ponto fixo.

Demonstração. Ver Teorema 8.2 em [AMO01]. □

1.4.2 Funções Convexas

Agora, nosso enfoque é sobre um tipo particular de função que será bem útil em nosso texto, as funções convexas. Considere V um espaço vetorial real e considere aplicações de $C \subset V$ em $\overline{\mathbb{R}}$, isto é, valores $+\infty$ e $-\infty$ são permitidos.

Definição 1.75. *Seja C um subespaço convexo de V e f uma função de C em $\overline{\mathbb{R}}$. g é dita convexa se, para todos x e y em C tivermos:*

$$g(\lambda x + (1-\lambda)y) \leq \lambda g(x) + (1-\lambda)g(y), \quad \forall \lambda \in [0,1]$$

sempre que o lado direito está definido.

Note que se $g : V \to \overline{\mathbb{R}}$ é convexa, as seções

$$\{x \,|\, g(x) \leq a\} \text{ e } \{x \,|\, g(x) < a\}$$

também são convexas. Agora, definamos o conceito de epigrafo.

Definição 1.76. *O epigrafo de uma função $g : V \to \overline{\mathbb{R}}$ é o conjunto*

$$\text{epi}(g) = \{(x, c) \in V \times \mathbb{R} \,|\, g(x) \leq c\},$$

que são os pontos de $V \times \mathbb{R}$ que estão acima do gráfico de g.

Isto será bem útil no estudo das funções convexas por causa do seguinte resultado.

Proposição 1.77. *A função $g : V \to \mathbb{R}$ é convexa se, e somente se o epigrafo é convexo.*

Demonstração. Ver Teorema 2.1. de [ET76]. □

1.4.3 Funções Semicontínuas Inferiormente

Vamos para as propriedades topológicas, então V será um espaço localmente convexo. Uma função $g : V \to \overline{\mathbb{R}}$ é dita semicontínua inferiormente(l.s.c.) sobre V, se esta satisfaz as duas condições equivalentes :

$$\forall c \in \mathbb{R}, \quad \{x \in V \,|\, g(x) \leq c\} \text{ é fechada,}$$

$$\forall \bar{x} \in V, \quad \liminf_{x \to \bar{x}} g(x) \geq g(\bar{x}).$$

Naturalmente g é semicontínua superiormente(u.s.c.) se $-g$ é l.s.c.. Por exemplo, a função indicadora 1_A de um conjunto $A \subset V$ será l.s.c. se, e somente se, A é fechado (ou aberto). Geralmente, podemos enunciar:

Proposição 1.78. *Uma função $g : V \to \overline{\mathbb{R}}$ é l.s.c. se, e somente se seu epigrafo é fechado.*

Demonstração. Proposição 2.3 de [ET76]. □

Corolário 1.79. *Toda função l.s.c. convexa g de V em $\overline{\mathbb{R}}$ permanece l.s.c. quando V é trocada por sua topologia fraca $\sigma(V, V^*)$.*

1.4.4 Γ-regularizações

Como de praxe, V é um espaço localmente convexo. As funções afim sobre V são funções do tipo $x \mapsto g(x) + c$, onde g é funcional linear contínuo sobre V e $c \in \mathbb{R}$.

Definição 1.80. *O conjunto das funções $h : V \to \overline{\mathbb{R}}$ que são supremos pontuais de uma família de funções contínuas afim será denotada por $\Gamma(V)$. $\Gamma_0(V)$ denota o subconjunto de $g \in \Gamma(V)$ fora das constantes $+\infty$ e $-\infty$.*

Segue da definição que as funções $g \in \Gamma(V)$ são convexas e l.s.c. Reciprocamente:

Proposição 1.81. *As seguintes propriedades são equivalentes :*

(a) $g \in \Gamma(V)$;

(b) *g é uma função convexa l.s.c. de V em $\overline{\mathbb{R}}$ e se g toma o valor $-\infty$, então g é identicamente igual a $-\infty$.*

Demonstração. Note que o supremo pontual de uma família vazia é $-\infty$ e que se esta família a considerar é não nula, g não pode tomar o valor $-\infty$. Logo (a) implica (b).

Reciprocamente, suponha que g é uma função convexa l.s.c. de V em $\overline{\mathbb{R}}$ que não toma o valor $-\infty$. Se g é a constante $+\infty$, então é o supremo pontual de todas as funções contínuas afim de V em $\mathbb{R}$. Se $g \in \Gamma_0(V)$, para toda $\bar{u} \in V$ e para toda $\bar{a} < g(\bar{u})$ mostraremos que existe uma função afim contínua de V em $\mathbb{R}$ cujos valores em u estão entre $\bar{a}$ e $g\bar{u}$, o qual estabelece o resultado

Temos epi(g) um conjunto convexo fechado o qual não contém o ponto $(\bar{u}, \bar{a})$. Podemos separar estritamente estes por um hiperplano fechado afim $\mathcal{H}$ de $V \times \mathbb{R}$ com a equação

$$\mathcal{H} = \{(u, a) \in V \times \mathbb{R} \,|\, h(u) + ca = d\}.$$

Onde h é um funcional linear contínuo não-nulo sobre V, c e d números reais. Portanto,

$$h(\bar{u}) + c\bar{a} < d \tag{1.4}$$

e

$$\forall (u, a) \in \text{epi}(f),\ h(u) + ca > d. \tag{1.5}$$

Se $h(\bar{u}) < +\infty$, podemos tomar $u = \bar{u}$ e $a = g(\bar{u})$ que nos dá $\beta(g(\bar{u} - \bar{a})) > 0$ quando $\beta > 0$. Quando (1.4) e (1.5) são divididas por β, obtemos:

$$\bar{a} < \frac{d}{c} - \frac{1}{c} h(\bar{u}) < g(\bar{u}).$$

A função afim contínua $\frac{d}{c} - \frac{1}{c} h(\bar{u})$ responde ao problema.

Se $g(\bar{u}) = +\infty$, podemos adaptar o raciocínio do caso 2 do Teorema A a este problema para chegarmos ao resultado de interesse. □

1.4.5 Funções Polares

Sejam V um espaço vetorial e V^* e seu espaço dual, denotaremos a sua dualidade por um pareamento bilinear denotado por $\langle \cdot, \cdot \rangle$, como em (1.3). Estes espaços terão como topologias a topologia fraca e fraca-$*$, respectivamente, o que as torna localmente convexas e Hausdorff.

Seja g uma função de V em $\overline{\mathbb{R}}$. Se $x^* \in V^*$ e $c \in \mathbb{R}$, a função afim contínua $u \mapsto \langle x, x^* \rangle - c$ é menor que g em todo ponto se, e somente se

$$\forall x \in V, \quad c \geq \langle x, x^* \rangle - g(x),$$

ou colocando $\alpha \geq g^*(x^*)$ se definirmos

$$g^*(x^*) = \sup_{u \in V} \{\langle x, x^* \rangle - g(u)\}, \tag{1.6}$$

tal igualdade nos permite definir uma função $g^* : V^* \to \overline{\mathbb{R}}$.

Definição 1.82. *Se $g : V \to \overline{\mathbb{R}}$, a fórmula (1.6) define uma função de V^* em $\overline{\mathbb{R}}$, denotado por g^*, e chamada a função polar ou conjugada convexa de g.*

Repetindo o processo, chegaremos a função bipolar g^{**}, a qual é uma função de V em $\overline{\mathbb{R}}$:

$$g^{**}(x) = \sup_{x^* \in V^*} \{\langle x, x^* \rangle - g^*(x^*)\}. \tag{1.7}$$

Temos $g^{**} \in \Gamma(V)$, e podemos comparar g e g^{**}, os quais são definidos sobre o mesmo espaço, que é o conteúdo da proposição a seguir.

Proposição 1.83. *Seja g uma função de V em $\overline{\mathbb{R}}$. Então sua função bipolar g^{**} não é nada mais que sua Γ-regularização. Em particular, se $g \in \Gamma(V)$, $g^{**} = g$.*

Demonstração. Ver Proposição 4.1 de [ET76]. □

Corolário 1.84. *Para toda $g : V \to \overline{\mathbb{R}}$, temos $g^* = g^{***}$.*

Demonstração. Ver Corolário 4.1 de [ET76]. □

Vimos que $g \in \Gamma(V)$ se, e somente se, $g^{**} = g$. Portanto, chegamos a seguinte definição:

Definição 1.85. *A polaridade estabelece uma bijeção entre $\Gamma(V)$ e $\Gamma(V^*)$. $g \in \Gamma(V)$ e $h \in \Gamma(V^*)$ são ditas estar em dualidade se eles se correspondem na bijeção:*

$$g = h^* \;\text{ e }\; h = g^*.$$

As constantes $+\infty$ e $-\infty$ sobre V e V^* estão em dualidade. Logo $g \in \Gamma_0(V)$ se, e somente se, $g^* \in \Gamma_0(V^*)$: a polaridade estabelece uma correspondência biunívoca entre $\Gamma_0(V)$ e $\Gamma_0(V^*)$.

Capítulo 2

Formalismo Termodinâmico e Teoria DLR

Neste capítulo, abordaremos resultados do Formalismo Termodinâmico e da Mecânica Estatística que são as ferramentas básicas de estudo do nosso problema. O primeiro tópico a ser abordado é o Operador de Ruelle, depois falaremos sobre as Medidas de Gibbs DLR e, por fim, falaremos sobre as entropias que trabalharemos nesta tese.

A exposição dos temas será básica e caso o leitor queira uma abordagem mais detalhada, este pode consultar, por exemplo, [LMMS15], [Geo11], [VO14], [Wal82] e [PP90].

2.1 Operador de Ruelle

Tendo em base um espaço métrico compacto (K, d_M) cuja métrica é dada por:

$$d(x, y) = \sum_{n=1}^{\infty} \frac{1}{2^n} d_M(x_n, y_n),$$

onde $x = (x_1, x_2, \ldots)$ e $y = (y_1, y_2, \ldots)$. Pelo Teorema de Tychonov, Ω é compacto. Como o alfabeto $\Omega = K^{\mathbb{N}}$ é infinito, como mencionado na introdução, os conceitos de entropias específica e variacional dependerão de uma medida *a priori* p, que terá suporte total sobre Ω. O Operador de Ruelle também dependerá desta medida p. Por outro lado, outra ferramenta importante é o operador *shift*, que também age sobre $\Omega = K^{\mathbb{N}}$. Estes dois operadores são peças chave para a conexão entre as Entropias Específica e Variacional.

A seguir, definimos estes operadores.

Definição 2.1 (Operador *shift* a esquerda)**.** *A aplicação shift* $\sigma : \Omega \to \Omega$ *definida por*

$$\sigma(x_1, x_2, x_3, \ldots) = (x_2, x_3, x_4, \ldots). \tag{2.1}$$

é dita a aplicação shift a esquerda.

Seja $C(\Omega)$ o espaço das funções contínuas de Ω à $\mathbb{R}$, e fixaremos uma medida de probabilidade *a priori* p sobre $\mathcal{B}(\mathbb{R})$. Assuma que o suporte de p é Ω. Para um potencial f em $C^\alpha(\Omega)$, que é uma função $f : \Omega \to \mathbb{R}$, definimos o Operador de Transferência (também chamado operador de Ruelle) $\mathscr{L}_f : C(\Omega) \to C(\Omega)$ por

$$\mathscr{L}_f(\phi)(x) = \int_K e^{f(ax)}\phi(ax)\, dp(a),$$

onde $x \in \Omega$ e $ax = (a, x_1, x_2, \ldots)$ denota a pré-imagem de x com $a \in K$.

Seja a^n um elemento de K^n tendo coordenadas $a^n = (a_n, a_{n-1}, \ldots, a_1)$ e considere por $a^n x \in \Omega$ a concatenação de $a^n \in K^n$ com $x \in \Omega$, ou seja, $a^n x = (a_n, \ldots, a_1, x_1, x_2, \ldots)$. No caso $n = 1$ escreveremos $a \equiv a^1 \in K$, e $ax = (a, x_1, x_2, \ldots)$.

A n-ésima iteração de $\mathscr{L}_f$ tem a seguinte expressão

$$\mathscr{L}_f^n(\phi)(x) = \int_{K^n} e^{S_n f(a^n x)}\phi(a^n x)\, dp^n(a^n), \tag{2.2}$$

onde $S_n f(a^n x) = \sum_{k=0}^{n-1} f(\sigma^k(a^n x))$ e $dp^n(a^n) = \prod_{k=1}^{n} dp(a_{n-k+1})$.

2.1.1 Teorema de Ruelle-Perron-Frobenius

O estudo de medidas de equilíbrio é um dos principais objetos de estudo do Formalismo Termodinâmico e uma das principais ferramentas deste estudo, o Teorema de Ruelle-Perron-Frobenius. O problema variacional tem suas raízes com Walters, em [Wal76], onde foi provada a existência de medidas invariantes junto a unicidade do estado de equilíbrio a partir do Operador de Perron-Frobenius.

Para o nosso caso, temos como referência [LMMS15]. Para uma exposição didática ao leitor, provaremos o teorema quando os potenciais são α-Hölder contínuos. Quando o potencial é Walters contínuo, veja [CS16].

Lema 2.2. *O operador* $\mathscr{L}_f$ *preserva o conjunto das funções* α*-Hölder. Ou seja, se* $\phi \in C^\alpha(\Omega)$*, então* $\mathscr{L}_f(\phi) \in C^\alpha(\Omega)$*.*

Demonstração. Temos

$$\frac{|\mathscr{L}_f(\phi)(x) - \mathscr{L}_f(\phi)(y)|}{d(x,y)^\alpha} = \frac{|\int_K e^{f(ax)}\phi(ax)\,dp(a) - \int_K e^{f(ay)}\phi(ay)\,dp(a)|}{d(x,y)^\alpha}.$$

Afirmamos que se $\phi, f \in C^\alpha(\Omega)$, então $e^f\phi \in C^\alpha(\Omega)$. Devemos mostrar que

$$\sup_{x,y\in\Omega, x\neq y} \frac{|e^{f(x)} - e^{f(y)}|}{d(x,y)^\alpha} < \infty.$$

Sejam $x, y \in \Omega$, temos f contínua, com Ω compacto. Segue que $f(\Omega)$ é conjunto compacto em $\mathbb{R}$. Daí, $|e^{f(x)} - e^{f(y)}| \leq \|e^f\|_\infty |f(x) - f(y)|$ e

$$\frac{|e^{f(x)} - e^{f(y)}|}{d(x,y)^\alpha} \leq \|e^f\|_\infty \mathrm{Hol}_\alpha(f).$$

Com esta informação, temos

$$\int_K |e^{f(ax)}\phi(ax)\,dp(a) - e^{f(ay)}\phi(ay)|\,dp(a) \leq \mathrm{Hol}_\alpha(e^f\phi)p(K) = \mathrm{Hol}_\alpha(e^f\phi)$$

e o lema está provado. □

Lema 2.3. *Defina o operador $\mathcal{T}_{s,f}$ sobre $C(\Omega)$ por*

$$\mathcal{T}_{s,f}(\phi)(x) = \log\left(\int_K e^{f(ax)+s\phi(ax)}\,dp(a)\right).$$

Então para cada $0 < s < 1$, $\mathcal{T}_{s,f}$ é uma contração uniforme.

Demonstração.

$$\begin{aligned}
|\mathcal{T}_{s,f}(\phi_1)(x) - \mathcal{T}_{s,f}(\phi_2)(x)| &= \left|\log\left(\frac{\int_K e^{f(ax)+s\phi_1(ax)}\,dp(a)}{\int_K e^{f(ax)+s\phi_2(ax)}\,dp(a)}\right)\right| \\
&= \left|\log\left(\frac{\int_K e^{f(ax)+s\phi_2(ax)+s\phi_1(ax)-s\phi_2(ax)}\,dp(a)}{\int_K e^{f(ax)+s\phi_2(ax)}\,dp(a)}\right)\right| \\
&\leq \left|\log\left(\frac{\int_K e^{f(ax)+s\phi_2(ax)+s\|\phi_1-\phi_2\|_\infty}\,dp(a)}{\int_K e^{f(ax)+s\phi_2(ax)}\,dp(a)}\right)\right| \\
&= \left|\log\left(\frac{e^{s\|\phi_1-\phi_2\|_\infty}\int_K e^{f(ax)+s\phi_2(ax)}\,dp(a)}{\int_K e^{f(ax)+s\phi_2(ax)}\,dp(a)}\right)\right| \\
&= s\|\phi_1 - \phi_2\|_\infty.
\end{aligned}$$

□

Seja u_s o único ponto fixo para $\mathcal{T}_{s,f}$. Defina

$$\log\left(\int_K e^{f(ax)+su_s(ax)}\,dp(a)\right) \equiv u_s(x).$$

Lema 2.4. *A família* $\{u_s\}_{0<s<1}$ *é uma família equicontínua de funções. Em particular, é uma família* α*-Hölder contínua.*

Demonstração. Seja $H_s(x,y) = u_s(x) - u_s(y)$. Da definição acima, temos

$$\begin{aligned} e^{u_s(x)} &= \int_K e^{f(ax)+su_s(ax)}\,dp(a) \\ &= \int_K e^{f(ay)+su_s(a)} e^{f(ax)-f(ay)+s[u_s(ax)-u_s(ay)]}\,dp(a) \\ &\leq e^{u_s(y)} \max_{a\in K}\{e^{f(ax)-f(ay)+s[u_s(ax)-u_s(ay)]}\}. \end{aligned}$$

Segue que

$$e^{u_s(x)-u_s(y)} \leq \max_{a\in K}\{f(ax)-f(ay)+s|u_s(ax)-u_s(ay)|\},$$

e isto implica

$$H_s(x,y) = u_s(x)-u_s(y) \leq \max_{a\in K}[f(ax)-f(ay)+sH_s(ax,ay)].$$

Repetindo este processo, novamente, temos

$$H_s(x,y) \leq \max_{a\in K}[f(ax)-f(ay)+s\max_{b\in K}[f(bax)-f(bax)+H_s(bax,bay)]].$$

Note que estamos tomando o máximo dentre os elementos de K, podemos reescrever o lado direito da desigualdade acima por

$$\max_{(a,b)\,\in K^2}[f(ax)-f(ay)+s[f(bax)-f(bax)+H_s(bax,bay)]].$$

Procedendo por indução e considerando $\theta = (\theta_0, \theta_1, \ldots) \in \Omega$ um elemento arbitrário, temos

$$\begin{aligned} H_s(x,y) &\leq \max_{\theta\in\Omega}\sum_{n=0}^{\infty} s^n[f(\theta_n\ldots\theta_0 x) - f(\theta_n\ldots\theta_0 y)] \\ &\leq \mathrm{Hol}_\alpha(f)\max_{\theta\in\Omega}\sum_{n=0}^{\infty} s^n d((\theta_n\ldots\theta_0 x),(\theta_n\ldots\theta_0 y))^\alpha \\ &\leq \mathrm{Hol}_\alpha(f)\sum_{n=0}^{\infty}\left(\frac{s}{2^\alpha}\right)^n d(x,y)^\alpha \\ &\leq \frac{2^\alpha}{2^\alpha-1}\mathrm{Hol}_\alpha(f)d(x,y)^\alpha. \end{aligned}$$

Temos que a família $\{u_s\}_{s\in(0,1)}$ é α-Hölder contínua e, em particular, equicontínua. Além disso, a constante de Lipschitz para tal família não depende de s. $\square$

Teorema 2.5. *Considere a medida de probabilidade a priori p. Fixado $f \in C^\alpha(\Omega)$, existe uma autofunção estritamente positiva α-Hölder ψ_f para $\mathscr{L}_f : C(\Omega) \to C(\Omega)$ associada a um autovalor maximal estritamente positivo λ_f. O autovalor é simples, o que significa que a autofunção é única (a menos de multiplicações por constantes).*

Demonstração. Para cada $0 < s < 1$, definimos o operador $\mathcal{T}_{s,f}$ sobre $C(\Omega)$, dado por

$$\mathcal{T}_{s,f}(\phi)(x) = \log\left(\int_K e^{f(ax)+s\phi(ax)}\,dp(a)\right).$$

Pelo Lema 2.3, temos $\mathcal{T}_{s,f}(\phi)(x)$ uma contração uniforme. Seja u_s o único ponto fixo para $\mathcal{T}_{s,f}(\phi)(x)$, então u_s satisfaz

$$\log\left(\int_K e^{f(ax)+s\phi(ax)}\,dp(a)\right) = u_s(x).$$

Pelo Lema 2.4, esta família $\{u_s\}_{s\in(0,1)}$ de funções é equicontínua. Segue da equação acima que

$$-\|f\|_\infty + s\min u_s \le u_s(x) \le \|f\|_\infty + s\max u_s.$$

Assim, $-\|f\|_\infty \le (1-s)\min u_s \le (1-s)\max u_s \le \|f\|_\infty$, para $0 < s < 1$.

A família $\{u_s^* = u_s - \max u_s\}_{0<s<1}$ é equicontínua e uniformemente limitada. Fixada uma subsequência $s_n \to 1$ tal que $[(1-s_n)\max u_{s_n}] \to k$, e usando o Teorema de Arzelà-Ascoli, $\{u_{s_n}^*\}_{n\ge1}$ tem um ponto de acumulação em $C(\Omega)$, ao qual chamaremos de u.

Observe que para qualquer s

$$\begin{aligned} e^{u_s^*} &= e^{u_s(x)-\max u_s} = e^{-(1-s)\max u_s + u_s(x) - s\max u_s} \\ &= e^{-(1-s)\max u_s}\int_K e^{f(ax)+(su_s(ax)-s\max u_s)}\,dp(a). \end{aligned}$$

Tomando o limite de n ao infinito para a sequência s_n, temos que u satisfaz

$$e^{u(x)} = e^{-k}\int_K e^{f(ax)+u(ax)}\,dp(a) = e^{-k}\mathscr{L}_f(e^u)(x).$$

Isto, junto aos Lemas 2.2 e 2.4, mostra que $\psi_f \equiv e^u$ é uma autofunção α-Hölder positiva para $\mathscr{L}_f$ associada ao autovalor $\lambda_f = e^k$. Suponha que o autovalor não seja simples, então existem duas autofunções ψ_1 e ψ_2. Tome $t = \min\{\psi_1/\psi_2\}$. Então $\psi_3 = \psi_1 - t\psi_2$ é uma autofunção não-negativa tal que esta se anula em algum ponto $z \in \Omega$. Portanto

$$0 = \lambda_f^n \psi_3(z) = \int_{b \in K^n} e^{S_n(f)(bz)} \psi(bz)\, dp(b),$$

o que implica $\psi_3(bz) = 0$, $\forall b \in K^n$, $\forall\, n$, implicando $\psi_3 = 0$. □

A autofunção é única a menos de um fator multiplicativo. Dividindo pela norma, esta se torna normalizada, vista como um vetor em $C(\Omega)$. Assim, podemos considerar que seu máximo é 1.

Outro conceito de função normalizada é o seguinte. Dizemos que um potencial f está normalizado se $\mathscr{L}_f(1)(x) = 1$, para todo $x \in \Omega$, ou seja

$$\int_K e^{f(ax)}\, dp(a) = 1,\ \forall x \in \Omega.$$

Seja f α-Hölder, ψ_f e λ_f dados pelo teorema anterior, temos

$$\int_K \frac{e^{f(ax)\psi_f(ax)}}{\lambda_f \psi_f(x)}\, dp(a) = 1,\ \forall x \in \Omega.$$

Daí, podemos definir o potencial normalizado $\bar{f}$ associado a f, por

$$\bar{f} \equiv f + \log \psi_f - \log \psi_f \circ \sigma - \log \lambda_f,$$

onde $\sigma : \Omega \to \Omega$ é a aplicação *shift*. Como $\psi_f \in C^\alpha(\Omega)$, temos que $\bar{f} \in C^\alpha(\Omega)$.

Definimos a sigma-algebra de Borel $\mathscr{F}$ sobre Ω considerando os cilindros como conjuntos abertos. Ou seja, os conjuntos da forma $B_1 \times \ldots \times B_n \times \Omega^{\mathbb{N}}$, onde $n \in \mathbb{N}$, e B_j, $j \in \{1, \ldots, n\}$ são conjuntos abertos em Ω.

Dizemos que uma medida de probabilidade μ sobre $\mathscr{F}$ é σ-invariante, se para qualquer conjunto Boreliano B, temos $\mu(B) = \mu(\sigma^{-1}(B))$. Denotamos por $\mathscr{M}_\sigma(\Omega)$ o conjunto das medidas σ-invariantes.

Note que Ω é um espaço métrico compacto e pelo Teorema da Representação de Riesz, uma medida de probabilidade sobre a sigma-álgebra de Borel é identificada como um funcional linear positivo $L : C(\Omega) \to \mathbb{R}$ que leva a função constante 1 ao número real 1. Note que $\mu \in \mathscr{M}_\sigma(\Omega)$ se, e somente se, para qualquer $\psi \in C(\Omega)$

$$\int_\Omega \psi\ d\mu = \int_\Omega \psi \circ \sigma\ d\mu.$$

Definimos o operador dual $\mathscr{L}_f^*$ sobre o espaço das medidas de Borel sobre Ω como o operador que leva uma medida μ a medida $\mathscr{L}_f^*(\mu)$, definida por

$$\int_\Omega \psi \, d\mathscr{L}_f^*(\mu) = \int_\Omega \mathscr{L}_f(\psi) \, d\mu,$$

para qualquer $\psi \in C(\Omega)$. Para finalizar a subseção, provaremos o Teorema de Ruelle-Perron-Frobenius.

Teorema 2.6 (Teorema de Ruelle-Perron-Frobenius)**.** *Seja f um potencial α-Hölder contínuo, ψ_f e λ_f a autofunção e autovalor dados pelo Teorema 2.5. Associamos a f o potencial normalizado $\bar{f} = f + \log \psi_f - \log \psi_f \circ \sigma - \log \lambda_f$. Então*

(a) Existe um único ponto fixo μ_f para $\mathscr{L}_{\bar{f}}^*$, que é medida de probabilidade σ-invariante;

(b) A medida

$$\rho_f = \frac{1}{\psi_f} \mu_f$$

satisfaz $\mathscr{L}_f^*(\rho_f) = \lambda_f \rho_f$. Portanto, ρ_f é uma auto-medida para $\mathscr{L}_f^*$;

(c) Para qualquer potencial α-Hölder contínuo $g : \Omega \to \mathbb{R}$, temos que, na topologia da convergência uniforme,

$$\mathscr{L}_{\bar{f}}^n g \to \int_\Omega g \, d\mu_f$$

e

$$\frac{\mathscr{L}_f^n(g)}{(\lambda_f)^n} \to \psi_f \int_\Omega g \, d\rho_f,$$

onde $\mathscr{L}_f^n$ denota a n-ésima iteração do operador $\mathscr{L}_f : C^\alpha(\Omega) \to C^\alpha(\Omega)$.

Demonstração.

(a) Observamos que a propriedade de normalização implica que os conjuntos convexos e compactos dos conjuntos das medidas de probabilidade sobre Ω é preservada pelo operador $\mathscr{L}_{\bar{f}}^*$. Portanto, usando o Teorema de Tychonoff-Schauder concluímos a existência de um ponto fixo μ_f para o operador $\mathscr{L}_{\bar{f}}^*$. Agora provemos que μ_f é σ-invariante: Se $\psi \in C(\Omega)$, temos

$$\int_\Omega \psi \circ \sigma \, d\mu_f = \int_\Omega \psi \circ d\mathscr{L}_{\bar{f}}(\mu_f) = \int_\Omega \mathscr{L}_f(\psi \circ \sigma) \, d\mu_f = \int_\Omega \psi \, d\mu_f,$$

onde na última igualdade usamos a hipótese de normalização para $\bar{f}$. A unicidade será obtida na prova do item (c).

(b) $\mathscr{L}^*_{\bar{f}} = \mu_f$ implica que, para qualquer $\psi \in C(\Omega)$,

$$\int_\Omega \psi \; d\mu_f = \int_\Omega \psi \, d\mathscr{L}^*_{\bar{f}}(\mu_f) = \int_\Omega \mathscr{L}_{\bar{f}}(\psi) \; d\mu_f$$
$$= \int_\Omega \left(\int_K \psi(ax) \frac{e^{f(ax)} \psi_f(ax)}{\lambda_f \psi_f(x)} \, dp(a) \right) \, d\mu_f(x).$$

Agora, se $\varphi \in C(\Omega)$, tomando $\psi = \varphi/\psi_f$ na última equação, temos

$$\int_\Omega \frac{\varphi}{\psi_f} \, d\mu_f = \frac{1}{\lambda_f} \int_\Omega \left(\int_K \varphi(ax) \frac{e^{f(ax)}}{\psi_f(x)} \, dp(a) \right) \, d\mu_f(x),$$

o que é equivalente a

$$\lambda_f \int_\Omega \varphi \, d\rho_f = \int_\Omega \mathscr{L}_f(\varphi) \, d\rho_f,$$

isto é, $\mathscr{L}^*_f(\rho_f) = \lambda_f \rho_f$.

(c) A prova será dividida em duas partes. Considere qualquer potencial α-Hölder f, seja $\|g\|_\infty$ a norma uniforme da função α-Hölder $g : C(\Omega) \to \mathbb{R}$, temos

$$|\mathscr{L}^n_f(g)(x) - \mathscr{L}^n_f(g)(y)| \leq \left[C_{e^f} \|g\|_\infty \left(\frac{1}{2^\alpha} + \ldots + \frac{1}{2^{n\alpha}} \right) + \frac{C_g}{2^{n\alpha}} \right] d(x,y)^\alpha,$$

onde C_{e^f} e C_g são constantes de Hölder de e^f e g, respectivamente.

Como consequência disto, o conjunto $\{\mathscr{L}^n_{\bar{f}} g\}_{n\geq 0}$ é equicontínuo. Para mostrarmos a limitação uniforme desta família, usamos a condição de normalização, a qual implica $\|\mathscr{L}^n_{\bar{f}} g\|_\infty \leq \|g\|_\infty$, para todo $n \geq 1$.

Portanto, pelo Teorema de Arzelà-Ascoli existe um ponto de acumulação, $\bar{g}$, para $\{\mathscr{L}^n_{\bar{f}} g\}_{n\geq 0}$, ou seja, existe uma subsequência $\{n_k\}_{k\geq 0}$ tal que

$$\bar{g}(x) = \lim_{k\geq 0} \mathscr{L}^{n_k}_{\bar{f}} g(x). \tag{2.3}$$

Vamos mostrar que esta função, $\bar{g}$, é constante. Primeiramente, observe que

$$\sup \bar{g} \geq \sup \mathscr{L}_f \bar{g}. \tag{2.4}$$

Por outro lado, (2.3) implica

$$\bar{g}(x) = \lim_{k\geq 0} \mathscr{L}^{n_k}_{\bar{f}} \bar{g}(x),$$

e isto mostra que devemos ter em (2.4) é uma igualdade. De fato, temos

$$\sup \bar{g} = \sup \mathscr{L}_{\bar{f}}^{n} \bar{g}, \; n \geq 0.$$

Agora, tome x_M^n um ponto de máximo para $\mathscr{L}_{\bar{f}}^{n} \bar{g}$, para qualquer $n \geq 0$. Temos

$$\bar{g}(x_M^0) = \mathscr{L}_{\bar{f}}^{n} \bar{g}(x_M^n)$$

e isto prova a segunda parte, pois a propriedade de normalização implica que $\mathscr{L}_{\bar{f}}^{n} \bar{g}(x_M^n)$ é uma combinação convexa das ω nas pré-imagens de x_M^n.

Agora que $\bar{g}$ é uma função constante, podemos provar que

$$\bar{g} = \int_\Omega \bar{g} \; d\mu_f = \lim_k \int_\Omega \mathscr{L}_{\bar{f}}^{n_k} g \; d\mu_f = \lim_k \int_\Omega g \, d(\mathscr{L}_{\bar{f}}^{*})^{n_k}(\mu_f) = \int_\Omega g \; d\mu_f.$$

O que mostra que $\bar{g}$ não depende da subsequência escolhida. Portanto, para qualquer $x \in \Omega$, temos

$$\mathscr{L}_{\bar{f}}^{n} g(x) \to \bar{g} = \int_\Omega g \; d\mu_f.$$

O último limite mostra que o ponto fixo μ_f é único.

Para finalizar a prova do item (c), como $f = \bar{f} - \log \psi_f + \log \psi_f \circ \sigma + \log \lambda_f$, temos

$$S_n f(z) = \sum_{k=0}^{n-1} f \circ \sigma^k(z) = S_n \bar{f}(z) - \log \psi_f + \log \psi_f \circ \sigma^n + n \log \lambda_f,$$

e portanto

$$\begin{aligned}
\frac{\mathscr{L}_f^n(g)(x)}{\lambda_f^n} &= \frac{1}{\lambda_f^n} \int_K e^{S_n f(a^n x)} g(a^n x) \; dp^n(a^n) \\
&= \psi_f(x) \int_{K^n} \frac{e^{S_n \bar{f}(a^n x)}}{\psi_f(a^n x)} g(a^n x) \; dp^n(a^n) \\
&= \psi_f(x) \mathscr{L}_{\bar{f}}^{n} \left(\frac{g}{\psi_f} \right) \to \psi_f(x) \int_\Omega \frac{g}{\psi_f} \; d\mu_f \\
&= \psi_f(x) \int_\Omega g \, d\rho_f.
\end{aligned}$$

□

Chamamos μ_f de medida de Gibbs para f. Estas medidas também são chamadas de estados de equilíbrio, pois elas atingem o supremo do problema variacional

$$\sup_{\mu \in \mathscr{M}_\sigma(\Omega)} \left\{ \mathtt{h}^{\mathtt{v}}(\mu) + \int_\Omega f \, d\mu \right\}.$$

Conforme o Teorema 2.25, o estado de equilíbrio é onde o funcional Pressão (ver Definição 2.24) é maximizado. O conjunto de todas as medidas de Gibbs para um potencial f α-Hölder ou Walters contínuo será denotado por $\mathcal{G}^*(\bar{f})$. Por fim, enunciaremos o Teorema de Ruelle-Perron-Frobenius para quando os potenciais são Walters-contínuos.

Teorema 2.7 (Ruelle-Perron-Frobenius: Caso Walters)**.** *Seja f um potencial satisfazendo a condição de Walters e considere o operador de Ruelle $\mathscr{L}_f : C(\Omega) \to C(\Omega)$ associado à f. Então existem um número real $\lambda_f > 0$, uma função estritamente positiva h_f e uma única medida de probabilidade Boreliana μ_f tal que*

i) $\mathscr{L}_f h_f = \lambda_f h_f, \mathscr{L}_f^* \mu_f = \lambda_f \mu_f.$

ii) Para qualquer $\varphi \in C(\Omega)$, temos

$$\|\lambda_f^{-n} \mathscr{L}_f^n \varphi - \lambda_f \int_\Omega \varphi \, d\mu_f\|_\infty \to 0, \text{ quando } n \to \infty.$$

Demonstração. Ver Teorema 4.4 em [CS16]. □

2.2 Medidas de Gibbs DLR

Agora vamos introduzir as medidas de Gibbs no contexto DLR, e mostrar algumas propriedades deste conjunto de medidas. No fim desta seção, enunciaremos uma relação entre as medidas de Gibbs DLR e as medidas de Gibbs como definidas acima.

Sejam S um conjunto infinito enumerável e $(\Sigma, \mathscr{F})$ um espaço mensurável arbitrário.

Definição 2.8. *Uma família $(X_i)_{i \in S}$ de variáveis aleatórias definidas sobre $(\overline{\Sigma}, \bar{\mathscr{F}}, \mu)$ e que tomam valores em $(\Sigma, \mathscr{F})$ é chamada um campo aleatório. Neste caso, S é chamado o conjunto de parâmetros e $(\Sigma, \mathscr{F})$ é chamado de estados de espaços do campo aleatório.*

Cada elemento i de S é chamado de sítio e a variável aleatória associada X_i é chamada de *spin* no sítio i. Faremos a escolha canônica para o campo aleatório. Ou seja, $\overline{\Sigma} = \Sigma^S$, $\bar{\mathscr{F}} = \mathscr{F}^S$ e μ uma medida de probabilidade qualquer sobre $(\Sigma, \mathscr{F})$. Vamos considerar o caso quando $S = \mathbb{N}$ a partir de agora. Sabendo o campo aleatório a se trabalhar, temos alguns conceitos importantes a destacar como, por exemplo, a projeção sobre a i-ésima coordenada, denotada por $\pi_i(x) = x_i$.

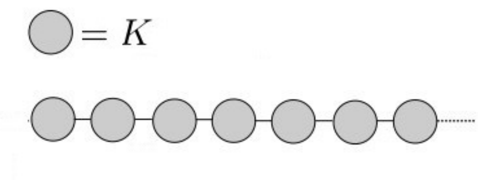

Figura 2.1: *O caso* $\overline{\Sigma} = \Omega = K^{\mathbb{N}}$*, com* K *espaço métrico compacto.*

Quando o conjunto $\Lambda \subset \mathbb{N}$ é finito, a projeção sobre as coordenadas em Λ é denotada por π_Λ e é definida por $\pi_\Lambda(x) = (x_i)_{i\in\Lambda}$. Se $\Lambda = \Lambda_n \equiv \{1, \ldots, n\}$, definimos $\pi_{\Lambda_n} = (x_i)_{i\in\{1,\ldots,n\}}$. Nesta abordagem "canônica", um campo aleatório é simplesmente uma medida de probabilidade sobre o espaço produto $(\overline{\Sigma}, \bar{\mathscr{F}})$. Em outras palavras, qualquer $\mu \in \mathscr{M}_1(\overline{\Sigma})$ é um campo aleatório.

Observação 2.9. *Em sistemas dinâmicos, estes espaços de estados são denotados por alfabetos. Como estamos fazendo uma conexão entre Formalismo Termodinâmico e Mecânica Estatística, os conceitos de espaços de estados e alfabetos neste texto, serão equivalentes, no sentido de que eles denotam os sítios de* $\overline{\Sigma}$.

Definição 2.10. *Sejam* $(\Sigma_1, \mathscr{F}_1)$ *e* $(\Sigma_2, \mathscr{F}_2)$ *dois espaços mensuráveis. Uma função* $\gamma : \mathscr{F}_1 \times \Sigma_2 \to [0, \infty]$ *é dita um núcleo de medida de* $(\Sigma_2, \mathscr{F}_2)$ *para* $(\Sigma_1, \mathscr{F}_1)$ *se*

(a) $\gamma(\cdot|y)$ *é uma medida sobre* $(\Sigma_1, \mathscr{F}_1)$ *para todo* $y \in \Sigma_2$ *e ;*

(b) $\gamma(A|\cdot)$ *é* $\mathscr{F}_2$*-mensurável para cada* $A \in \mathscr{F}_1$.

Em particular, se $\gamma(\Sigma_1|\cdot) = 1$*, então* γ *é chamado um núcleo de probabilidade de* $\mathscr{F}_2$ *a* $\mathscr{F}_1$.

Por exemplo, se g é uma aplicação mensurável de Σ_1 a Σ_2 então a função $(A, y) \to 1_A \circ g(y)$ é um núcleo de probabilidade de $\mathscr{F}_2$ a $\mathscr{F}_1$.

Um núcleo de medida γ de $\mathscr{F}_2$ a $\mathscr{F}_1$ leva uma medida μ sobre Σ_2 a uma medida $\mu\gamma$ sobre Σ_1 a qual é denotada por

$$\mu\gamma(A) = \int \gamma(A|\cdot)\ d\mu,\ (A \in \mathscr{F}_1).$$

Quando se trabalha com núcleos de probabilidade, as vezes queremos trabalhar em uma sub-sigma-álgebra especifica de $\mathscr{F}_1$. Seja $\mathscr{B}$ uma sub-sigma-álgebra de interesse. Dizemos que um núcleo de medida é próprio se satisfaz:

$$\gamma(A \cap B|\cdot) = \gamma(A|\cdot)1_B, \quad (A \in \mathscr{F}_1, B \in \mathscr{B}).$$

Assim como na esperança condicional, quando o núcleo é próprio, pode-se "retirar" a função do núcleo de medida, se a mesma for mensurável com respeito a sub-sigma-álgebra de interesse. Além disso, tal conceito pode ser estendido a funções, ou seja, se g é $\mathscr{F}_1$-mensurável e h é $\mathscr{B}$-mensurável e limitada, vale o seguinte:

$$\gamma(gh|\cdot) = \gamma(g|\cdot)h.$$

Quando o núcleo de medida é de probabilidade, a condição acima é equivalente a

$$\gamma(B|\cdot) = 1_B \ (B \in \mathscr{B}).$$

A vantagem de se trabalhar com núcleos próprios é que eles podem ser vistos como versões de probabilidades condicionais.

Proposição 2.11. *Sejam $(\Sigma, \mathscr{F})$ um espaço mensurável, $\mathscr{B}$ uma sub-sigma-álgebra de $\mathscr{F}$, γ um núcleo de probabilidade próprio de $\mathscr{B}$ a $\mathscr{F}$, e $\mu \in \mathscr{M}_1(\Sigma)$. Então $\mu(A|\mathscr{B}) = \gamma(A|\cdot)$ $\mu - q.c.$ para toda $A \in \mathscr{F}$ se, e somente se $\mu\gamma = \mu$.*

Demonstração. Ver [Geo11], Observação (1.20). □

2.2.1 Especificações

Vamos retornar a discussão dos campos aleatórios. De acordo com as nossas intenções, estamos interessados em campos aleatórios aos quais as variáveis X_i apresentem um tipo particular de dependência. Seguiremos a abordagem feita por [Dob68a] e [LR69] aos quais prescrevem as distribuições condicionais de todas as coleções finitas de spins. Em outras palavras, olharemos as distribuições condicionais com respeito às sigma-álgebras externas que são geradas a partir das projeções externas $\pi_{\Lambda_n^c}(x) = (x_{n+1}, x_{n+2}, \ldots)$.

$$\mathcal{T}_{\Lambda_n} \equiv \bar{\mathscr{F}}_{\mathbb{N}\setminus\Lambda_n} = \{(\pi_{\Lambda_n^c})^{-1}(A) : A \in \mathscr{F}^{\Lambda_n^c}\},$$

onde $\Lambda_n = \{1, \ldots, n\}$. Tal sigma-álgebra contém todas informações sobre as "condições de fronteira", isto é, a configuração de todos os spins fora da região finita Λ_n. Esta dependência é especificada por uma família $(\gamma_n)_{n\in\mathbb{N}}$ de núcleos de probabilidade de $(\overline{\Sigma}, \mathcal{T}_{\Lambda_n})$ a $(\overline{\Sigma}, \bar{\mathscr{F}})$, e um campo aleatório μ ao qual iremos considerar tal dependência se esta satisfaz a condição

$$\mu(A|\mathcal{T}_{\Lambda_n}) = \gamma_n(A|\cdot)\ \mu - \text{q.c. para todo } A \in \bar{\mathscr{F}} \text{ e } n \in \mathbb{N}. \tag{2.5}$$

Voltemos a espaços métricos compactos, considere K um espaço métrico compacto e $\Omega = K^{\mathbb{N}}$. O exemplo a seguir contém o núcleo de probabilidade base para o nosso texto.

Exemplo 2.12. *Considere $f \in C(\Omega)$ um potencial e para cada $n \in \mathbb{N}$, $x \in \Omega$ e $A \in \bar{\mathscr{F}}$, considere a aplicação $\gamma_n : \bar{\mathscr{F}} \times \Omega \to [0, 1]$ dada por*

$$\begin{aligned}\gamma_n(A|x) &\equiv \frac{\mathscr{L}_f^n(1_A)(\sigma^n(x))}{\mathscr{L}_f^n(1)(\sigma^n(x))} \\ &= \frac{\int_{K^n} e^{S_n f(a^n\sigma^n(x))} 1_A(a^n\sigma^n(x))\, dp^n(a^n)}{\int_{K^n} e^{S_n f(a^n\sigma^n(x))}\, dp^n(a^n)}.\end{aligned}$$

Onde $\mathscr{L}_f$ é o operador de transferência para o potencial f e σ é a aplicação shift. Afirmamos que para cada $n \in \mathbb{N}$ a aplicação definida acima é um núcleo de probabilidade.

Demonstração. Mostraremos que as propriedades de um núcleo de probabilidade são satisfeitas. Note que basta analisar o numerador, pois o denominador influencia apenas na segunda propriedade de núcleo de probabilidade, mas como veremos, é um caso particular quando $A = \Omega$.

Primeiramente, pela definição, $\gamma_n(\emptyset|x) = 0$. Por outro lado. Sejam $A_1, \ldots, A_i, \ldots$ conjuntos dois a dois disjuntos $\bar{\mathscr{F}}$-mensuráveis. Então:

$$\begin{aligned}\mathscr{L}_f^n(1_{\cup_{i=1}^\infty A_i})(\sigma^n(x)) &= \lim_{m\to\infty} \mathscr{L}_f^n(1_{\cup_{i=1}^m A_i})(\sigma^n(x)) \\ &= \lim_{m\to\infty} \int_{K^n} e^{S_n f(a^n\sigma^n(x))} 1_{\cup_{i=1}^m A_i}(a^n\sigma^n(x))\, dp(a^n) \\ &= \lim_{m\to\infty} \int_{K^n} e^{S_n f(a^n\sigma^n(x))} \sum_{i=1}^m 1_{A_i}(a^n\sigma^n(x))\, dp(a^n)\end{aligned}$$

$$
\begin{aligned}
&= \int_{K^n} e^{S_n f(a^n\sigma^n(x))} \sum_{i=1}^{\infty} 1_{A_i}(a^n\sigma^n(x))\, dp(a^n) \\
&= \sum_{i=1}^{\infty} \int_{K^n} e^{S_n f(a^n\sigma^n(x))} 1_{A_i}(a^n\sigma^n(x))\, dp(a^n) \\
&= \sum_{i=1}^{\infty} \mathscr{L}_f^n(1_{A_i})(\sigma^n(x)).
\end{aligned}
$$

Onde usamos o Teorema 1.47 para a função crescente não-negativa $h_m(y) \equiv e^{S_n f(y)} \sum_{i=1}^{m} 1_{A_i}(y)$ e o Teorema 1.51 para a função não-negativa $h(y) \equiv e^{S_n f(y)} 1_{A_i}(y)$. Falta mostrar a mensurabilidade. Fixe $A \in \bar{\mathscr{F}}$. Para todo $x \in \Omega$, temos:

$$
\begin{aligned}
\gamma_n(A|x) &= \int_{K^n} e^{S_n f(a^n\sigma^n(x))} 1_A(a^n\sigma^n(x))\, dp(a^n) \\
&= \int_{K^{\Lambda_n}} \int_{K^{\Lambda_n^c}} e^{S_n f(a^n\sigma^n(x))} 1_A(a^n\sigma^n(x)) \left(dp(a^n) \times \prod_{i\in\Lambda_n^c} \delta_{x_i} \right).
\end{aligned}
$$

Temos $g(a,x) \equiv e^{S_n f(a^n\sigma^n(x))} 1_A(a^n\sigma^n(x)) \geq 0$ $\bar{\mathscr{F}}$-mensurável. Pelo Teorema 1.51, $g^x = \gamma_n(A|x)$ é $\bar{\mathscr{F}}$-mensurável. $\square$

Note que definimos os núcleos de probabilidade apenas nos subconjuntos cofinais de $\mathbb{N}$. Isto ocorre pelo fato das medidas de Gibbs serem caracterizadas nestes conjuntos, que é um fato da Proposição 2.14.

Voltemos ao contexto geral. Suponha que existam um campo aleatório μ e uma especificação γ em (2.5). Então para $n \leq m \in \mathbb{N}$, $A \in \bar{\mathscr{F}}$ e $B \in \mathcal{T}_{\Lambda_n}$ dados, pelas propriedades de esperança condicional temos:

$$
\begin{aligned}
\gamma_{\Lambda_n}(A \cap B|\cdot) &= \mu(A \cap B|\mathcal{T}_{\Lambda_n}) \\
&= \mu(A|\mathcal{T}_{\Lambda_n}) 1_B = \gamma_{\Lambda_n}(A|\cdot) 1_B \;\; \mu - q.c.
\end{aligned}
$$

e

$$
\begin{aligned}
\gamma_{\Lambda_m}\gamma_{\Lambda_n}(A|\cdot) &= \mu(\mu(A|\mathcal{T}_{\Lambda_n})|\mathcal{T}_{\Lambda_m}) \\
&\mu(A|\mathcal{T}_{\Lambda_m}) = \gamma_{\Lambda_m}(A|\cdot) \;\; \mu - q.c.
\end{aligned}
$$

Ou seja, γ_{Λ_n} é próprio μ-q.c. e γ_{Λ_n} e γ_{Λ_m} são consistentes μ-q.c.. Mas queremos estas propriedades sem depender do suporte dos campos aleatórios μ que satisfazem (2.5).

Os conjuntos cofinais, que são os conjuntos de nosso interesse, em $\mathbb{N}$ são da forma $\Lambda_n = \{1, \ldots, n\}$ e são onde as especificações são bem definidas em nosso caso.

Definição 2.13 (Especificação)**.** *Uma especificação com conjunto de parâmetros* $\mathbb{N}$ *e estado de espaços* $(\overline{\Sigma}, \bar{\mathscr{F}})$ *é uma família* $\gamma = (\gamma_{\Lambda_n})_{n\in\mathbb{N}}$ *de núcleos de probabilidade próprios* $\gamma_{\Lambda_n} \equiv \gamma_n$ *de* $\mathcal{T}_{\Lambda_n} \equiv \mathcal{T}_n$ *a* $\bar{\mathscr{F}}$ *que satisfazem a condição de consistência* $\gamma_{n_1}\gamma_{n_0} = \gamma_{n_1}$ *quando* $n_0 \leq n_1$. *Os campos aleatórios no conjunto*

$$\mathcal{G}^{DLR}(\gamma) \equiv \{\mu \in \mathcal{P}(\overline{\Sigma}, \bar{\mathscr{F}}) : \mu(A|\mathcal{T}_n) = \gamma_n(A|\cdot)\,\mu - q.c.\,\forall\, A \in \mathscr{F} \;\; e \;\; n \in \mathbb{N}\}$$

são ditos especificados por γ. *As medidas* μ *que estão neste conjunto são ditas medidas de Gibbs, no sentido de satisfazer a equação DLR, que é dada pelo item (b) da Proposição 2.14.*

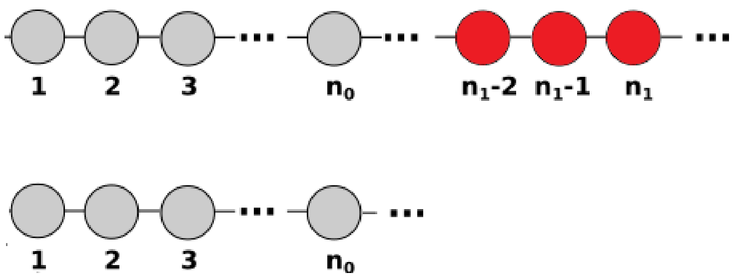

Figura 2.2: *Os sítios em cinza são onde as medidas agem, enquanto os sítios em vermelho denotam os sítios fixados pela condição de fronteira, pois neste exemplo temos* $n_0 < n_1$.

Nosso último resultado sobre especificações é uma caracterização de $\mathcal{G}^{DLR}(\gamma)$.

Proposição 2.14. *Suponha* γ *uma especificação e* μ *um campo aleatório. Então as seguintes afirmações são equivalentes:*

(a) $\mu \in \mathcal{G}^{DLR}(\gamma)$;

(b) $\mu\gamma_{\Lambda_n} = \mu$ *para todo* $\Lambda_n = \{1, \ldots, n\}$.

Demonstração. Basta observar que os itens (a) e (b) são equivalentes como consequência direta da Proposição (2.11). □

Nosso foco agora será nesta família de núcleos do Exemplo 2.12. Novamente, $E = K$ e $\overline{\Sigma} = \Omega$. A condição de consistência, por exemplo, para tal família é a seguinte. Dados $n, r \in \mathbb{N}$ e qualquer função $\bar{\mathscr{F}}$-mensurável limitada $g : \Omega \to \mathbb{R}$ têm-se

$$\gamma_{n+r}(g|x) = \int_{\Omega} \gamma_n(g|\cdot) d\gamma_{n+r}(\cdot|x) \equiv \gamma_{n+r}(\gamma_n(g|\cdot)|x).$$

Mostraremos que a equação acima é válida para os núcleos de probabilidade definidos no Exemplo 2.12, quando o potencial f é normalizado.

Proposição 2.15. *Sejam $f \in C(\Omega)$ um potencial normalizado, g uma função $\bar{\mathscr{F}}$-mensurável e $\{\gamma_n\}_{n\in\mathbb{N}}$ a família de núcleos de probabilidade definida no Exemplo 2.12. Temos que*

$$\mathscr{L}_f^{n+r}(\mathscr{L}_f^n(g)(\cdot))(\sigma^n(y)) = \mathscr{L}_f^{n+r}(g)(\sigma^{n+r}(y)), \forall y \in \Omega. \tag{2.6}$$

Demonstração. Quando o potencial f é normalizado, para cada n e cada g $\bar{\mathscr{F}}$-mensurável, γ_n é dada por

$$\begin{aligned}\gamma_n(g|y) &= \mathscr{L}_f^n(g)(\sigma^n(y)) \\ &= \int_{K^n} e^{S_n f(a^n \sigma^n(y))} g(a^n \sigma^n(y))\, dp^n(a^n).\end{aligned}$$

Mostraremos a igualdade (2.6). Considere $n, r \in \mathbb{N}$ e $\bar{y} = \sigma^{n+r}(y)$.

$$\begin{aligned}
&\mathscr{L}_f^{n+r}(\mathscr{L}_f^n(g)(\cdot))(\bar{y}) \\
&= \int_{K^{n+r}} e^{S_{n+r} f(a^n z^r \bar{y})} \mathscr{L}_f^n(g)(\sigma^n(a^n z^r \bar{y}))\, dp^{n+r}(a^n z^r) \\
&= \int_{K^{n+r}} e^{S_{n+r} f(a^n z^r \bar{y})} \left(\int_{K^n} e^{S_n f(b^n z^r \bar{y})} g(b^n z^r \bar{y})\, dp^n(b^n) \right) dp^{n+r}(a^n z^r) \\
&= \int_{K^{n+r}} \int_{K^n} e^{S_{n+r} f(a^n z^r \bar{y}) + S_n f(b^n z^r \bar{y})} g(b^n z^r \bar{y})\, dp^n(b^n)\, dp^n(a^n)\, dp^r(z^r) \\
&= \int_{K^{n+r}} \int_{K^n} e^{S_{n+r} f(b^n z^r \bar{y}) + S_n f(a^n z^r \bar{y})} g(b^n z^r \bar{y})\, dp^n(a^n)\, dp^n(b^n)\, dp^r(z^r) \\
&= \int_{K^{n+r}} e^{S_{n+r} f(b^n z^r \bar{y})} g(b^n z^r \bar{y}) \left(\int_{K^n} e^{S_n f(a^n z^r \bar{y})}\, dp^n(a^n) \right) dp^n(b^n)\, dp^r(z^r) \\
&= \int_{K^{n+r}} e^{S_{n+r} f(b^n z^r \bar{y})} g(b^n z^r \bar{y})\, dp^{n+r}(b^n z^r) \\
&= \mathscr{L}_f^{n+r}(g)(\bar{y}).
\end{aligned}$$

Onde usamos, nas igualdades, a definição da medida produto, o Teorema de Fubini para a função $e^{S_{n+r}(f)+S_n(f)}g$, pois g é limitada e

$$e^{S_{n+r}(f)+S_n(f)} \leq (2n+r)\|f\|_\infty, \forall\, r, n \in \mathbb{N}.$$

Além disso, $\mathscr{L}_f^n(1)(\sigma^n(y)) = 1$ para quaisquer $n \in \mathbb{N}, y \in \Omega$ e, por fim, dados $a, b \in K$, temos

$$S_{n+r}f(b^n z^r \bar{y}) + S_n f(a^n z^r \bar{y}) = S_{n+r}f(a^n z^r \bar{y}) + S_n f(b^n z^r \bar{y}).$$

□

A Definição 2.13 neste contexto será exibida abaixo.

Definição 2.16. *Dizemos que $\mu \in \mathscr{M}_1(\Omega)$ é uma medida Gibbs DLR para o potencial contínuo f se para qualquer n e qualquer função contínua $f : \Omega \to \mathbb{R}$ tivermos para x μ-q.c. que*

$$E^\mu[f|\sigma^n(\mathscr{F})](x) \equiv \mu(f|\sigma^n(\mathscr{F})) = \int f(y)\, d\gamma_n(y|x).$$

O conjunto de todas as medidas Gibbs DLR para f é denotado por $\mathcal{G}^{DLR}(f)$.

Como consequência da definição acima, temos que os dois itens a seguir são equivalentes:

(a) $\mu \in \mathcal{G}^{DLR}(f)$;

(b) para qualquer $n \in \mathbb{N}$ e $E \in \mathscr{F}$ temos que $\mu(E) = \int \gamma_n(E|\cdot)\, d\mu$.

Temos assim, as duas medidas de Gibbs bem definidas. Felizmente, existe uma relação entre os dois conjuntos, que é a igualdade nas medidas de probabilidade que são automedidas do Operador de Ruelle e as medidas que satisfazem a equação DLR, sendo o conteúdo do teorema a seguir.

Teorema 2.17. *Seja $f \in C(\Omega)$ um potencial e $(\gamma_n)_{n\in\mathbb{N}}$ a especificação definida como no Exemplo 2.12. Então vale*

$$\mathcal{G}^{DLR}(f) = \mathcal{G}^*(f) = \{\mu \in \mathscr{M}_1(\Omega) : \mathscr{L}_f^* \mu = \lambda_f \mu\}.$$

Demonstração. Ver Teorema 2 em [CL16]. □

2.2.2 Funções Locais e Quasilocalidade

Definição 2.18 (Função Local)**.** *Uma função $\varphi : \overline{\Sigma} \to \mathbb{R}$ é dita ser local se φ é $\bar{\mathscr{F}}_n-$mensurável para algum $n \in \mathbb{N}$. Para cada $n \in \mathbb{N}$, denotamos por $\mathcal{L}_n$ o espaço de todas as funções locais limitadas $\bar{\mathscr{F}}_n$-mensuráveis e $\mathcal{L} = \cup_{n\in\mathbb{N}}\mathcal{L}_n$ o conjunto de todas as funções locais limitadas.*

Note que se K é finito, então qualquer função local é contínua. De fato, qualquer elemento $\varphi \in \mathcal{L}_n$ é uma função dependendo apenas das primeiras n coordenadas.

Definição 2.19 (Função quasilocal)**.** *Uma função $\varphi : \overline{\Sigma} \to \mathbb{R}$ é dita ser quasilocal se existe uma sequência $(\varphi_n)_{n\in\mathbb{N}} \in \mathcal{L}$ tal que $\|\varphi - \varphi_n\|_\infty \to 0$, quando $n \to \infty$. Escrevemos $\bar{\mathcal{L}}$ para denotar o espaço de todas as funções quasilocais.*

Definição 2.20. *Dizemos que a especificação $(\gamma_n)_{n\in\mathbb{N}}$ é quasilocal, se para cada $n \in \mathbb{N}$ e $\varphi \in \bar{\mathcal{L}}$ a aplicação $x \mapsto \gamma_n(\varphi|x)$ é quasilocal.*

2.3 Entropias

Nesta subseção serão definidos os conceitos mais importantes deste texto, as entropias variacional e específica. Será mostrada nesta subseção que as duas entropias têm bastante propriedades em comum, o que nos faz desconfiar que, em alguma configuração adequada, estas sejam equivalentes, o que será provado no próximo capítulo. Considere $(K, \mathcal{K})$ um espaço métrico compacto mensurável.

2.3.1 Entropia Variacional

Esta entropia, relacionada à Teoria do Formalismo Termodinâmico, cujo contexto foi definido em (2.1) será trabalhada a partir de uma medida de probabilidade *a priori* p. Tal medida deve ter suporte completo sobre K e, além disso, a entropia será definida sobre o conjunto das medidas invariantes por *shift*, que é o conjunto $\mathscr{M}_\sigma(\Omega)$.

Definição 2.21. *Seja p uma medida de probabilidade fixada a priori sobre K. Dado um potencial $g \in C^\alpha(\Omega)$ normalizado, denotamos por $\mathcal{G}^*(g) = \mathcal{G}^*_p(g)$ o conjunto das medidas de Gibbs, que denota o conjunto das $\mu \in \mathscr{M}_\sigma(\Omega)$, tais que, $\mathscr{L}^*_g(\mu) = \mu$. Para tais potenciais, definimos a entropia variacional de μ por*

$$\mathtt{h}^{\mathtt{v}}(\mu) = \mathtt{h}^{\mathtt{v},p}(\mu) = -\int g \, d\mu.$$

Queremos estender a definição acima para todas as medidas σ-invariantes, para isso, vamos usar o seguinte resultado.

Lema 2.22. *Fixe um potencial α-Hölder contínuo f e uma medida $\mu \in \mathcal{G}^*(\bar{f})$. Chamamos por $C(\Omega)^+$ o espaço das funções contínuas positivas sobre Ω. Temos*

$$\mathtt{h}^{\mathtt{v}}(\mu) + \int_\Omega f\, d\mu = \inf_{u \in C(\Omega)^+} \left\{ \int_\Omega \log \left(\frac{\mathscr{L}_f u}{u} \right) d\mu \right\}. \tag{2.7}$$

Demonstração. Ver Lema 2 em [LMMS15]. □

A equação 2.7 pode ser vista como uma versão estendida da entropia variacional apresentada em [LO09], pois os autores, neste trabalho, trabalham com espaços da forma $[0,1] \times \{1, \ldots, n\}^{\mathbb{N}}$ e, além disso, a entropia não depende da medida *a priori*. Ela é definida por

$$\mathtt{h}(\mu) = \inf_{z \in B([0,1])^+} \int \log \left(\frac{\mathscr{L}_g(z)}{gz} \right) d\mu\,,\ \forall g \in B([0,1])^+,$$

onde $B([0,1])^+$ denota o conjunto de todas as funções de Borel positivas sobre $[0,1]$. Em particular, tomando $g \equiv 1$ e $z = e^f u$, onde $f \in C^\alpha([0,1])$ e $u \in C([0,1])^+$, pode-se mostrar que

$$\mathtt{h}(\mu) + \int_\Omega f\, d\mu \le \inf_{u \in C([0,1])^+} \left\{ \int_\Omega \log \left(\frac{\mathscr{L}_f u}{u} \right) d\mu \right\}. \tag{2.8}$$

Voltemos a equação (2.7). tomando $u = h_f$, onde h_f é a autofunção dada pelo Teorema 2.6, temos

$$\begin{aligned} \mathtt{h}^{\mathtt{v}}(\mu) + \int_\Omega f\, d\mu &\le \int_\Omega \log \left(\frac{\mathscr{L}_f h_f}{h_f} \right) d\mu \\ &= \int_\Omega \log \left(\frac{\lambda_f h_f}{h_f} \right) d\mu \\ &= \log \lambda_f. \end{aligned}$$

Em particular, se μ é medida de equilíbrio para o potencial f, temos que a desigualdade em (2.8) é uma igualdade. Disto, estendemos a definição da entropia variacional para todas as medidas σ-invariantes pela expressão abaixo.

Definição 2.23. *Seja $\mu \in \mathscr{M}_\sigma(\Omega)$. Definimos a entropia de μ por*

$$\mathtt{h}^{\mathtt{v}}(\mu) = \mathtt{h}^{\mathtt{v},p}(\mu) = \inf_{f \in C^\alpha(\Omega)} \left\{ -\int_\Omega f\, d\mu + \log \lambda_f \right\},$$

onde λ_f é o autovalor maximal de $\mathscr{L}_f$, dado pelo Teorema 2.6.

Note que a dependência da medida *a priori* está implícita pelo autovalor λ_f, dado pelo Teorema 2.6, pois o operador de Ruelle depende da medida *a priori* p. Outro conceito importante é o do funcional Pressão e será definido abaixo.

Definição 2.24. *Dado um potencial α-Hölder f, o valor $\mathsf{P}(f)$ será chamado a Pressão (Topológica) de f.*

$$\mathsf{P}(f) = \sup_{\mu \in \mathscr{M}_\sigma(\Omega)} \left\{ \mathtt{h}^{\mathtt{v}}(\mu) + \int_\Omega f \, d\mu \right\}.$$

Uma medida de probabilidade que atinge tal valor maximal é chamado estado de equilíbrio para f.

O último resultado é o princípio variacional da pressão que caracteriza o estado de equilíbrio:

Teorema 2.25 (Princípio variacional)**.** *Seja f potencial α-Hölder contínuo e λ_f o autovalor maximal de $\mathscr{L}_f$, então*

$$\log \lambda_f = \mathsf{P}(f) = \sup_{\mu \in \mathscr{M}_\sigma(\Omega)} \left\{ \mathtt{h}^{\mathtt{v}}(\mu) + \int_\Omega f \, d\mu \right\}.$$

Além disso, o supremo é atingido sobre a medida de Gibbs, ou seja, a medida μ_f que satisfaz $\mathscr{L}^*_{\bar{f}}(\mu_f) = \mu_f$.

Demonstração. Ver Teorema 3 em [LMMS15]. $\square$

Note que além da formulação acima, pela equação (2.8), podemos caracterizar o funcional Pressão por outro método que é enunciado a seguir.

Teorema 2.26 (Pressão como Minimax)**.** *Dado um potencial f α-Hölder contínuo,*

$$\mathsf{P}(f) = \sup_{\mu \in \mathscr{M}_\sigma(\Omega)} \left[\inf_{u \in C(\Omega)^+} \left\{ \int_\Omega \log\left(\frac{\mathscr{L}_f u}{u} \right) d\mu \right\} \right].$$

Demonstração. Ver Lema 2, em [LMMS15], ou a Observação 2, em [LO09]. $\square$

Portanto, os estados de Gibbs e os estados de equilíbrio para f são dados pela mesma medida μ_f, a qual é o único ponto fixo para o dual do operador de Ruelle associado ao potencial normalizado $\bar{f}$.

2.3.2 Entropia Específica

O princípio variacional para medidas de Gibbs em medidas homogêneas é notável. Este assegura que o conjunto $\mathcal{G}^{DLR}(f)$ de todas as medidas de Gibbs homogêneas para um potencial f invariante por translações coincide com a classe de todos os campos aleatórios homogêneos que minimizam a energia específica livre, que é a diferença entre a energia relativa à f e a entropia específica. A minimalidade da energia livre é precisamente o equilíbrio clássico para um sistema físico a uma temperatura constante.

Seja $(\Sigma, \mathscr{F})$ um espaço mensurável arbitrário e μ e ν duas medidas finitas sobre este espaço.

Definição 2.27. *Suponha $\mathcal{A}$ uma sub-sigma-álgebra de $\mathscr{F}$. Defina*

$$\mathscr{H}_{\mathcal{A}}(\mu|\nu) = \begin{cases} \nu(\varphi_{\mathcal{A}} \log \varphi_{\mathcal{A}}), & \textit{se } \mu \ll \nu \textit{ sobre } \mathcal{A} \\ +\infty, & \textit{caso contrário.} \end{cases}$$

Onde $\varphi_{\mathcal{A}}$ é qualquer densidade de Radon-Nikodym de $\mu|_{\mathcal{A}}$ relativa à $\nu|_{\mathcal{A}}$. $\mathscr{H}_{\mathcal{A}}(\mu|\nu)$ é chamada a entropia relativa de μ com respeito à ν.

Usaremos a convenção $0 \log 0 = 0$ e a notação $\varphi_{\mathcal{A}} = \frac{d\mu|_{\mathcal{A}}}{d\nu|_{\mathcal{A}}}$ para simplificar a notação e as contas apresentadas nesta subseção.

Como a função $f(x) = x \log x$ é limitada inferiormente, a integral sempre estará bem definida, talvez $+\infty$. Além disso,

$$\mathscr{H}_{\mathcal{A}}(a\mu|b\nu) = a\mathscr{H}_{\mathcal{A}}(\mu|\nu) + a\mu(\Omega) \log \frac{a}{b}, \quad (a, b > 0). \tag{2.9}$$

De fato:

$$\begin{aligned} \mathscr{H}_{\mathcal{A}}(a\mu|b\nu) &= b\nu\left(\frac{a}{b}\varphi_{\mathcal{A}} \log\left(\frac{a}{b}\varphi_{\mathcal{A}}\right)\right) \\ &= a\left(\int_{\Sigma} \varphi_{\mathcal{A}} \log \frac{a}{b}\, d\nu + \int_{\Sigma} \varphi_{\mathcal{A}} \log \varphi_{\mathcal{A}}\, d\nu\right) \\ &= a\mathscr{H}_{\mathcal{A}}(\mu|\nu) + a\mu(\Omega) \log \frac{a}{b}. \end{aligned}$$

Portanto, sem perda de generalidade, podemos assumir μ e ν medidas de probabilidade.

Tal entropia tem as seguintes propriedades:

Proposição 2.28. *Seja $(\Sigma, \mathscr{F})$ espaço mensurável e $\mathcal{A}$ uma sub-sigma-álgebra de $\mathscr{F}$. Então os seguintes itens são satisfeitos.*

a) $\mathscr{H}_{\mathcal{A}}(\mu|\nu) \geq 0$*;*

b) $\mathscr{H}_{\mathcal{A}}(\mu|\nu) = 0$ *se, e somente se,* $\mu = \nu$ *sobre* $\mathcal{A}$*;*

c) $\mathscr{H}_{\mathcal{A}}(\mu|\nu)$ *é uma função crescente de* $\mathcal{A}$*;*

d) $\mathscr{H}_{\mathcal{A}}(\mu|\nu)$ *é uma função convexa do par* (μ,ν)*.*

Demonstração. Ver Proposição 15.5 em [Geo11]. □

Agora, considere um espaço mensurável da forma $(\overline{\Sigma},\bar{\mathscr{F}}) = (\Sigma^{\mathbb{N}},\mathscr{F}^{\mathbb{N}})$, como na Definição 2.8. Para cada $n \in \mathbb{N}$, denote por $\bar{\mathscr{F}}_{\Lambda_n}$ a sigma-álgebra gerada pelas projeções $\{\pi_m\}_{m\leq n}$ dada por

$$\bar{\mathscr{F}}_{\Lambda_n} = \{(\pi_{\Lambda_n})^{-1}(A) : A \in \mathscr{F}^{\Lambda_n}\}.$$

Assim, dadas quaisquer μ,ν medidas finitas sobre $(\Sigma,\bar{\mathscr{F}}_{\Lambda_n})$, a quantidade

$$\mathscr{H}_{\Lambda_n}(\mu|\nu) \equiv \mathscr{H}_{\bar{\mathscr{F}}_{\Lambda_n}}(\mu|\nu)\,.$$

é a entropia relativa de μ com respeito à ν sobre $\bar{\mathscr{F}}_{\Lambda_n}$.

Assumiremos que $(\overline{\Sigma},\bar{\mathscr{F}})$ é equipada com uma medida finita *a priori* p. p será fixada através do seguinte. Para cada $n \in \mathbb{N}$ tomamos $\pi_n^{-1}(p^{\Lambda_n})$ denotando a medida finita sobre $(\Omega,\bar{\mathscr{F}}_{\Lambda_n})$ que é dada por

$$\pi_n^{-1}(p^{\Lambda_n})(\pi_n \in A) = p^{\Lambda_n}(A), \quad (A \in \mathscr{F}^{\Lambda_n}).$$

Se $p \in \mathscr{M}_1(\Sigma)$, então $\pi_n^{-1} = \boldsymbol{p}\,|_{\bar{\mathscr{F}}_{\Lambda_n}}$, onde $\boldsymbol{p} \equiv \prod_{i=1}^{\infty} p$ denota a medida produto.

Definição 2.29. *Para cada* $\mu \in \mathscr{M}_1(\overline{\Sigma})$ *e* $n \in \mathbb{N}$*, a quantidade*

$$\mathscr{H}_{\Lambda_n}(\mu) \equiv -\mathscr{H}_{\Lambda_n}\left(\mu|\pi_n^{-1}(p^{\Lambda_n})\right) = -\mathscr{H}_{\mathscr{F}^{\Lambda_n}}\left(\pi_n(\mu)|p^{\Lambda_n}\right)$$

é chamada a entropia de μ *em* Λ_n *relativa à* p*.*

Com isto, tiramos a dependência da notação da entropia relativa sobre p. Quando Σ é finito e p é a medida de contagem, pode-se mostrar que a entropia coincide com a entropia de Shannon. Das propriedades da entropia relativa, podemos obter que a entropia específica, definida abaixo, existe e está no intervalo $[-\infty,0]$ para qualquer medida σ-invariante.

Definição 2.30 (Entropia específica)**.** *Para cada* $\mu \in \mathscr{M}_\sigma(\overline{\Sigma})$*, a quantidade* $\mathtt{h}^{\mathtt{s}}(\mu)$ *acima é chamada a Entropia Específica por sítio de* μ *relativa a uma medida a priori* p*.*

$$\mathtt{h}^{\mathtt{s}}(\mu) = \lim_{n\to\infty}\frac{\mathscr{H}_{\Lambda_n}(\mu)}{n}.$$

2.3.3 Propriedades

Novamente, considere $E = K$ e $\Omega = \Sigma$. Um dos nossos resultados originais é mostrar que estas duas entropias são equivalentes. Um dos motivos para tal desconfiança é elas compartilharem várias propriedades similares como provaremos abaixo:

Proposição 2.31. *As entropias* $\mathtt{h}^{\mathtt{s}}$ *e* $\mathtt{h}^{\mathtt{v}}$ *possuem as seguintes propriedades:*

(a) As duas entropias são não-positivas para qualquer medida invariante $\mu \in \mathscr{M}_\sigma(\Omega)$*;*

(b) Se p *é medida a priori sobre* K*, então as duas entropias se anulam em* $\boldsymbol{p} = \prod_{i=1}^{\infty} p$*;*

(c) As duas entropias são semi-contínuas superiormente;

(d) $\mathtt{h}^{\mathtt{v}}$ *é uma função côncava enquanto* $\mathtt{h}^{\mathtt{s}}$ *é afim.*

Observação 2.32. *Apesar de que no item (d), as duas entropias não são afim, como iremos mostrar a igualdade e, portanto,* $\mathtt{h}^{\mathtt{v}}$ *também será uma função afim.*

Demonstração.

(a) Como 0 é um potencial normalizado, segue da definição de $\mathtt{h}^{\mathtt{v}}$ que para qualquer medida invariante μ, $\mathtt{h}^{\mathtt{v}}(\mu) \leq 0$.

Para a entropia específica, basta usar o item (d) da Proposição 2.28 com a Definição 2.29.

(b) No caso da entropia variacional, vamos mostrar que $\boldsymbol{p} = \prod_{i=1}^{\infty} p$ (medida produto) é medida de equilíbrio para $f = 0$. De fato,

$$\begin{aligned}\int_\Omega g \, d\mathscr{L}_0^*(\boldsymbol{p}) = \int_\Omega \mathscr{L}_0 g \, d\boldsymbol{p} &= \int_\Omega \int_K e^0 g(ax) \, dp(a) \, d\boldsymbol{p}(x) \\ &= \int_\Omega g(ax) \, d\boldsymbol{p}(ax).\end{aligned}$$

O caso da entropia específica segue diretamente da definição de entropia relativa.

(c) Fixe $\varepsilon > 0$ e suponha $\mu_n \to \mu$. Pela definição de $\mathtt{h}^{\mathtt{v}}(\mu)$, podemos escolher $f \in C^\alpha(\Omega)$ tal que

$$-\int_\Omega f \, d\mu + \log \lambda_f \leq \mathtt{h}^{\mathtt{v}}(\mu) + \varepsilon.$$

Se n é grande o suficiente, temos $|\int_\Omega f\,d\mu_n - \int_\Omega f\,d\mu| < \varepsilon$. Então,

$$\mathtt{h}^{\mathtt{v}}(\mu_n) \leq -\int_\Omega f\,d\mu_n + \log\lambda_f \leq -\int_\Omega f\,d\mu + \log\lambda_f + \varepsilon \leq \mathtt{h}^{\mathtt{v}}(\mu) + 2\varepsilon,$$

portanto $\limsup_{n\to\infty} \mathtt{h}^{\mathtt{v}}(\mu_n) \leq \mathtt{h}^{\mathtt{v}}(\mu) + 2\varepsilon$.

Como a entropia relativa é dada por um supremo pelo Corolário 15.7, em [Geo11], segue que a entropia relativa é função semicontínua superiormente. Logo a entropia específica que é um limite de funções semicontínuas inferiormente, também possui tal propriedade.

(d) Sejam μ_1 e μ_2 duas medidas de probabilidade σ-invariantes, $\varepsilon \in (0,1)$ e $\mu = \varepsilon\mu_1 + (1-\varepsilon)\mu_2$. Então

$$\begin{aligned}
\mathtt{h}^{\mathtt{v}}(\varepsilon\mu_1 + (1-\varepsilon)\mu_2) = \mathtt{h}^{\mathtt{v}}(\mu) &= \inf_{f\in C^\alpha(\Omega)} \left(-\int_\Omega f\,d\mu + \log\lambda_f\right) \\
&= \inf_{f\in C^\alpha(\Omega)} \left(-\varepsilon\int_\Omega f\,d\mu_1 + (1-\varepsilon)\int_\Omega f\,d\mu_1 + \log\lambda_f\right) \\
&\geq \inf_{f\in C^\alpha(\Omega)} \varepsilon\left(-\int_\Omega f\,d\mu + \log\lambda_f\right) \\
&\quad + \inf_{f\in C^\alpha(\Omega)} (1-\varepsilon)\left(-\int_\Omega f\,d\mu + \log\lambda_f\right) \\
&= \varepsilon\,\mathtt{h}^{\mathtt{v}}(\mu_1) + (1-\varepsilon)\,\mathtt{h}^{\mathtt{v}}(\mu_2).
\end{aligned}$$

Pela Proposição 2.28, temos $\mathtt{h}^{\mathtt{s}}$ função côncava, falta mostrar a convexidade da mesma. Sejam $\mu_1, \mu_2 \in \mathscr{M}_\sigma(\Omega)$, $0 < s < 1$, e $\mu = s\mu_1 + (1-s)\mu_2$. Devemos verificar que $\mathtt{h}^{\mathtt{s}}(\mu) \leq s\mathtt{h}^{\mathtt{s}}(\mu_1) + (1-s)\mathtt{h}^{\mathtt{s}}(\mu_2)$. Sem perda de generalidade, assuma $\mathtt{h}^{\mathtt{s}}(\mu) > -\infty$. Neste caso, $\mu \ll p^{\Lambda_n}$ sobre cada $\bar{\mathscr{F}}_{\Lambda_n}, n \in \mathbb{N}$. Como μ_1 e μ_2 são μ-contínuas, concluímos que para cada $\Lambda_n = \{1, \ldots, n\}$ e $k = 1, 2$ existe uma função $\bar{\mathscr{F}}_{\Lambda_n}$-mensurável $\varphi^k_{\Lambda_n} \geq 0$

tal que $\mu_k = \varphi^k_{\Lambda_n}\boldsymbol{p}$ sobre $\bar{\mathscr{F}}_{\Lambda_n}$. Daí:

$$\begin{aligned}
-\mathscr{H}_{\Lambda_n}(\mu) &= \mu\left(\log(s\varphi^1_{\Lambda_n} + (1-s)\varphi^2_{\Lambda_n})\right)\\
&= (s\mu_1 + (1-s)\mu_2)\left(\log(s\varphi^1_{\Lambda_n} + (1-s)\varphi^2_{\Lambda_n})\right)\\
&= s\mu_1\left(\log(s\varphi^1_{\Lambda_n} + (1-s)\varphi^2_{\Lambda_n})\right)\\
&\qquad + (1-s)\mu_2\left(\log(s\varphi^1_{\Lambda_n} + (1-s)f^2_{\Lambda_n})\right)\\
&\geq s\mu_1\left(\log(s\varphi^1_{\Lambda_n})\right) + (1-s)\mu_2\left(\log((1-s)\varphi^2_{\Lambda_n})\right)\\
&= s\mu_1(\log\varphi^1_{\Lambda_n}) + s\log s\\
&\qquad + (1-s)\log(1-s) + (1-s)\mu_2(\log\varphi^2_{\Lambda_n}).
\end{aligned}$$

Da desigualdade acima, segue que:

$$\mathscr{H}_{\Lambda_n}(\mu) \leq s\mathscr{H}_{\Lambda_n}(\mu_1) + (1-s)\mathscr{H}_{\Lambda_n}(\mu_2) - s\log s - (1-s)\log(1-s).$$

Dividindo a inequação acima por n, com $n \to \infty$, chegamos ao resultado desejado.

$\square$

2.3.4 Kolmogorov-Sinai

A última entropia da qual iremos falar é a Entropia de Kolmogorov-Sinai que tem sua importância, neste texto, quando K é finito. Apesar disso, nesta breve abordagem, iremos considerar $(\Sigma, \mathscr{F}, P)$ um espaço de probabilidade qualquer. Esta entropia depende de uma partição $\mathcal{P}$ do conjunto, que será definida abaixo.

Definição 2.33. *Seja $(\Sigma, \mathscr{F}, P)$ um espaço de probabilidade. Uma partição de M é uma família finita ou enumerável $\mathcal{P}$ de subconjuntos mensuráveis de M disjuntos dois-a-dois cuja união tem medida total. Ou seja, a união destes subconjuntos é Σ a menos de subconjuntos de medida nula.*

Um integrante da partição que contém x é denotado por $\mathcal{P}(x)$. A expressão $\mathcal{P} \vee \mathcal{Q}$ denota a soma de duas partições $\mathcal{P}$ e $\mathcal{Q}$, onde os elementos desta são conjuntos da forma $P \cap \mathcal{Q}$ com $D \in \mathcal{P}$ e $Q \in \mathcal{Q}$. Ou seja,

$$\vee_n \mathcal{P}_n = \{\cap_n D_n : D_n \in \mathcal{P}_n \text{ para cada } n\}.$$

Dada uma partição $\mathcal{P}$, sua função de informação associada é dada por:

$$I_{\mathcal{P}} : \Sigma \to \mathbb{R}, \ I_{\mathcal{P}}(x) = -\log \mu(\mathcal{P}(x)).$$

A função $I_{\mathcal{P}}$ é mensurável e, portanto, denotamos a entropia da partição $\mathcal{P}$ pela quantia

$$H_P(\mathcal{P}) = \int_{\Sigma} I_{\mathcal{P}} \, dP = \sum_{D \in \mathcal{P}} -P(D) \log P(D).$$

As partições $\mathcal{P}$ e $\mathcal{Q}$ são ditas independentes se $P(D \cap Q) = P(D)P(Q)$ para todo $D \in \mathcal{P}$ e todo $Q \in \mathcal{Q}$. Logo, $H_P(\mathcal{P} \vee \mathcal{Q}) = H_P(\mathcal{P}) + H_P(\mathcal{Q})$ e, geralmente, vale a desigualdade $(\leq)$.

Exemplo 2.34. *Considere $\Omega = [0,1]$ munido da medida de Lebesgue. Para cada $n \geq 1$ considere a partição $\mathcal{P}^n$ nos subintervalos $((i-1)/10^n, i/10^n]$ com $1 \leq 1 \leq 10^n$. Então*

$$H_P(\mathcal{P}^n) = \sum_{i=1}^{10^N} -10^{-n} \log 10^{-n} = n \log 10.$$

O exemplo a seguir será de suma importância para o cálculo da entropia da aplicação *shift*.

Exemplo 2.35. *Seja $\Omega = \{1, \ldots, d\}^{\mathbb{N}}$ munido de uma medida produto $P = p^{\mathbb{N}}$. Denotamos $p_i = p(\{i\})$ para cada $i \in \{1, \ldots, d\}$. Para cada $n \geq 1$, seja $\mathcal{P}^n$ a partição de Ω em cilindros $[1; a_1, \ldots, a_n]$ de comprimento n. A entropia de $\mathcal{P}^n$ é*

$$\begin{aligned} H_P(\mathcal{P}^n) &= \sum_{a_1, \ldots, a_n} -p_{a_1} \ldots p_{a_n} \log(p_{a_1} \ldots p_{a_n}) \\ &= \sum_j \sum_{a_1, \ldots, a_n} -p_{a_1} \ldots p_{a_j} \ldots p_{a_n} \log p_{a_j} \\ &= \sum_j \sum_{a_j} -p_{a_j} \log p_{a_j} \sum_{a_i, i \neq j} p_{a_1} \cdots p_{a_{j-1}} p_{a_{j+1}} \cdots p_{a_n}. \end{aligned}$$

A última soma é igual a 1, uma vez que p é medida de probabilidade. Portanto,

$$H_P(\mathcal{P}^n) = \sum_{j=1}^{n} \sum_{a_j=1}^{d} -p_{a_j} \log p_{a_j} = \sum_{j=1}^{n} \sum_{i=1}^{d} -p_i \log p_i = -n \sum_{i=1}^{d} p_i \log p_i.$$

Lema 2.36. *Toda partição finita tem entropia finita. De fato,* $H_P(\mathcal{P}) \leq \log \#\mathcal{P}$ *e vale a igualdade se, e somente se,* P *é a medida de contagem normalizada relativa a esta partição.*

Demonstração. Seja $\mathcal{P} = \{D_1, D_2, \ldots, D_n\}$ e considere os números $t_i = 1/n$ e $x_i = P(D_i)$. Pela desigualdade de Jensen para $\phi(x) = -x \log x$:

$$\frac{1}{n} H_P(\mathcal{P}) = \sum_{i=1}^{n} t_i \phi(x_i) \leq \phi\left(\sum_{i=1}^{n} t_i x_i\right) = \phi\left(\frac{1}{n}\right) = \frac{\log n}{n}.$$

Portanto, $H_P(\mathcal{P}) \leq n$. Além disso, a igualdade ocorre se, e somente se, $P(D_i) = 1/n$ para todo $i = 1 \ldots n$. □

A Entropia condicional de uma partição $\mathcal{P}$ dada uma partição $\mathcal{Q}$ é o número

$$H_P(\mathcal{P}|\mathcal{Q}) = \sum_{D \in \mathcal{P}} \sum_{Q \in \mathcal{Q}} -P(D \cap Q) \log \frac{P(D \cap Q)}{P(Q)}.$$

Logo a entropia condicional mensura a informação acrescentada ao sistema pela partição $\mathcal{P}$, dada a informação da partição $\mathcal{Q}$.

Sejam $\mathcal{P}$ e $\mathcal{Q}$ duas partições de Σ, $\mathcal{P}$ é dita menos fina que $\mathcal{Q}$, denotada por $\mathcal{P} \prec \mathcal{Q}$, se dado um elemento de $\mathcal{Q}$, ele está contido em algum elemento da partição $\mathcal{P}$, a menos de medida nula. A soma $\mathcal{P} \vee \mathcal{Q}$ é a partição menos fina das partições $\mathcal{R}$ satisfazendo $\mathcal{P} \prec \mathcal{R}$ e $\mathcal{Q} \prec \mathcal{R}$. A proposição abaixo nos dá algumas propriedades desta entropia condicional.

Proposição 2.37. *Sejam* $\mathcal{P}, \mathcal{Q}$ *e* $\mathcal{R}$ *partições com entropia finita. Então,*

(a) $H_P(\mathcal{P} \vee \mathcal{Q}|\mathcal{R}) = H_P(\mathcal{P}|\mathcal{R}) + H_P(\mathcal{Q}|\mathcal{P} \vee \mathcal{R})$;

(b) se $\mathcal{P} \prec \mathcal{Q}$ *então* $H_P(\mathcal{P}|\mathcal{R}) \leq H_P(\mathcal{Q}|\mathcal{R})$ *e* $H_P(\mathcal{R}|\mathcal{P}) \geq H_P(\mathcal{R}|\mathcal{Q})$;

(c) $\mathcal{P} \prec \mathcal{Q}$ *se, e somente se,* $H_P(\mathcal{P}|\mathcal{Q}) = 0$.

Demonstração. Ver Lema 9.1.5 em [VO14]. □

Agora considere $T : \Sigma \to \Sigma$ uma transformação mensurável P-invariante. Seja $\mathcal{P}$ uma partição de Σ com entropia finita, denotamos

$$\mathcal{P}^n = \vee_{i=0}^{n-1} T^{-i}(\mathcal{P}) \text{ para cada } n \geq 1.$$

Note que o elemento $\mathcal{P}^n(x)$ contendo $x \in M$ é dado por:

$$\mathcal{P}^n(x) = \mathcal{P}(x) \cap T^{-1}(\mathcal{P}(T(x))) \cap \ldots \cap T^{-n+1}(\mathcal{P}(T^{n-1}(x))).$$

Tal sequência de partições $\mathcal{P}^n$ é não-decrescente, ou seja, $\mathcal{P}^n \prec \mathcal{P}^{n+1}$ para todo n. Por consequente, $H_P(\mathcal{P}^n)$ é uma sequência não-decrescente. Além disso, esta sequência é subaditiva:

Proposição 2.38. *$H_P(\mathcal{P}^{m+n}) \leq H_P(\mathcal{P}^m) + H_P(\mathcal{P}^n)$ para todo $m, n \geq 1$.*

Demonstração. Ver Lema 9.1.7 em [VO14]. □

A entropia de T com respeito à medida de probabilidade P e considerando $\mathcal{P}$ a partição de interesse o limite

$$h_P(T, \mathcal{P}) = \lim_n \frac{1}{n} H_P(\mathcal{P}^n) = \inf_n \frac{1}{n} H_P(\mathcal{P}^n). \tag{2.10}$$

Tal entropia aumenta quanto se a partição for mais fina. De fato, $\mathcal{P} \prec \mathcal{Q}$ implica $\mathcal{P}^n \prec \mathcal{Q}^n$ para todo n. Pela proposição anterior, tem-se $H_P(\mathcal{P}^n) \leq H_P(\mathcal{Q}^n)$ para todo n. Daí,

$$\mathcal{P} \prec \mathcal{Q} \Rightarrow h_P(T, \mathcal{P}) \leq h_P(T, \mathcal{Q}).$$

Por fim, definimos a entropia de T e medida *a priori* p sendo

$$h_\mu(T) = \sup_{\mathcal{P}} h_\mu(T, \mathcal{P}). \tag{2.11}$$

Exemplo 2.39. *Considere o shift $\sigma : \Omega \to \Omega$ em $\Omega = \{1, \ldots, d\}^{\mathbb{N}}$, munido da medida produto $P = p^{\mathbb{N}}$. Seja $\mathcal{P}$ a partição de Σ em cilindros $[a]$ com $a = 1, \ldots, d$. Então $\mathcal{P}^n$ é a partição em cilindros $[1; a_1, \ldots, a_n]$ de comprimento n. Usando o cálculo do Exemplo 2.35, concluímos que*

$$H_P(\sigma, \mathcal{P}) = \lim_n \frac{1}{n} H_P(\mathcal{P}^n) = \sum_{i=1}^{d} -p_i \log p_i.$$

Antes de enunciarmos o Teorema de Kolmogorov-Sinai, o resultado que nos permite calcular a entropia da aplicação *shift* da maneira que queremos é enunciado a seguir.

Proposição 2.40. *Se $\mathcal{P}$ é partição com entropia finita, então $h_P(T, \mathcal{P}) = h_P(T, \mathcal{P}^k)$ para todo $k \geq 1$.*

Demonstração. Ver Lema 9.1.13 em [VO14]. □

Para o cálculo desta entropia, a principal dificuldade está em achar o supremo na expressão (2.11). Mas Kolomogorov e Sinai deixaram a tarefa menos árdua identificando certas partições $\mathcal{P}$ que realizam o supremo, isto é, tais que $h_P(T, \mathcal{P}) = h_P(T)$. O resultado principal é o seguinte:

Teorema 2.41 (Kolmogorov-Sinai)**.** *Seja $\mathcal{P}_1 \prec \ldots \prec \mathcal{P}_n \prec \ldots$ uma sequência não-decrescente de partições com entropia finita tais que $\cup_{n=1}^{\infty} \mathcal{P}_n$ gera a sigma-álgebra dos conjuntos mensuráveis a menos de medida nula. Então*

$$h_P(T) = \lim_n h_P(T, \mathcal{P}_n).$$

Demonstração. Ver Teorema 9.2.1 em [VO14]. □

Corolário 2.42. *Seja $\mathcal{P}$ uma partição com entropia finita tal que a união dos seus iterados $\mathcal{P}^n = \vee_{j=0}^{n-1} T^{-j}(\mathcal{P})$, $n \geq 1$ gera a sigma-álgebra dos conjuntos mensuráveis. Então $h_P(T) = h_P(T, \mathcal{P})$.*

Demonstração. Aplicando o Teorema 2.41 à sequência $\mathcal{P}^n$, lembrando que $h_P(T, \mathcal{P}^n) = h_P(T, \mathcal{P})$ para todo n, de acordo com a Proposição 2.40. □

Com este último resultado, completamos o cálculo da entropia do *shift* e será enunciado como um corolário.

Corolário 2.43. *A entropia do shift $\sigma : \Omega = \{1, \ldots, d\}^{\mathbb{N}} \to \Omega$ é dada pela seguinte igualdade*

$$H(\mu) \equiv h_P(\sigma) = \lim_{n\to\infty} -\frac{1}{n} \sum_{a_1,\ldots,a_n} p_{a_1} \ldots p_{a_n} \log(p_{a_1} \ldots p_{a_n}) = \sum_{i=1}^{d} -p_i \log p_i. \tag{2.12}$$

Prova-se na seção 9.2.2 de [VO14] que esta entropia é semicontínua superiormente e isto é importante, pois provaremos na próxima seção que as entropias específica e variacional são iguais e no caso do espaço K ser finito, nós podemos relacionar as três entropias.

Além disso, se K é finito, as entropias variacional e a de Kolmogorov-Sinai tem a seguinte relação.

Proposição 2.44. *Seja $\bar{f}$ um potencial α-Hölder contínuo normalizado e $\mu_{\bar{f}} \in \mathcal{G}^*(\bar{f})$. Então para qualquer $\mu \in \mathscr{M}_\sigma(\Omega)$, têm-se:*

$$H(\mu) + \int_\Omega \bar{f}\, d\mu \leq 0,$$

com a igualdade se, e somente se, $\mu = \mu_{\bar{f}}$.

Demonstração. Ver Proposição 3.4 de [PP90]. □

Capítulo 3

Principais Resultados

O principal objetivo desta tese é estabelecer um princípio variacional que caracteriza todas as medidas de equilíbrio para um potencial f contínuo que está na classe de Walters. Tal caracterização é obtida de um modo cuidadoso e é o objetivo deste capítulo.

3.1 A Igualdade das Entropias em Espaços de Estados Finitos

Antes de abordarmos o caso geral, investigaremos o caso finito, onde a igualdade entre as entropias variacional e específica é provada usando a Entropia de Kolmogorov-Sinai, quando μ é uma medida de probabilidade de Gibbs, no sentido de [LMMS15].

3.1.1 Medidas de Equilíbrio

Neste caso, $K = \{1, \dots, n\}$, p é uma medida *a priori* com suporte total sobre K e $\Omega = K^{\mathbb{N}}$. A Entropia de Kolmogorov-Sinai para o nosso operador de interesse, o *shift* é dada por (2.12). Disto, iremos trabalhar com propriedades das entropias variacional e específica para chegarmos a igualdade. A entropia variacional, quando K é finito tem a seguinte representação:

Proposição 3.1 (Proposição 7 , [LMMS15])**.** *Seja $K = \{1, \dots, d\}$ e $p = \sum_{i=1}^{d} p_i \delta_i$ uma medida a priori sobre K. Dado um potencial α-Hölder contínuo, normalizado, f e sua medida de Gibbs correspondente $\mu \in \mathcal{G}^*(f)$, temos o seguinte:*

(a) $H(\mu) = \mathtt{h}^{\mathtt{v}}(\mu) - \sum_{i=1}^{d} \log p_i\, \mu([i])$;

(b) $\mathtt{h}^{\mathtt{v}}(\mu) = -\lim_{n\to\infty} \frac{1}{n} \sum_{i_1\ldots i_n} \mu([i_1 \ldots i_n]) \log\left(\frac{\mu([i_1\ldots i_n])}{p_{i_1}\cdots p_{i_n}}\right)$, *onde*

$$[1; i_1 \ldots i_n] \equiv [i_1 \ldots i_n] = \{x \in K^{\mathbb{N}} : x_1 = i_1, \ldots, x_n = i_n\}$$

denota um cilindro sobre Ω.

Demonstração. (a) Seja $\mu \in \mathcal{G}^*(f)$, então

$$\int_K e^{f(ax)} dp(a) = \sum_{i=1}^{d} e^{f(ix)} p_i = 1, \ \forall x \in K^{\mathbb{N}}.$$

Além disso, $\mathscr{L}_f^*(\mu) = \mu$ implica que μ é ponto fixo para o operador de Ruelle normalizado $f + \log(\overline{P})$, onde $\overline{P}$ é a projeção sobre a primeira coordenada dada por $\overline{P}(y_1, y_2, \ldots) = p_{y_1}$. Portanto, pela Proposição 2.44

$$H(\mu) = -\int_\Omega (f + \log(\overline{P}))\, d\mu = \mathtt{h}^{\mathtt{v}}(\mu) - \int_\Omega \log(\overline{P})\, d\mu$$

$$= \mathtt{h}^{\mathtt{v}}(\mu) - \sum_{i=1}^{d} \log(p_i)\mu([i]).$$

(b) Pela equação acima e $\mu \in \mathscr{M}_\sigma(\Omega)$, para qualquer $n \geq 1$.

$$H(\mu) = \mathtt{h}^{\mathtt{v}}(\mu) - \frac{1}{n} \int_\Omega \log(\overline{P}) + \ldots + \log(\overline{P} \circ \sigma^{n-1})\, d\mu$$

$$= \mathtt{h}^{\mathtt{v}}(\mu) - \frac{1}{n} \int_\Omega \log(p_{x_1} \ldots p_{x_n})\, d\mu(x)$$

$$= \mathtt{h}^{\mathtt{v}}(\mu) - \frac{1}{n} \sum_{i_1\ldots i_n} \mu([i_1 \ldots i_n]) \log(p_{i_1} \ldots p_{i_n}).$$

Portanto,

$$\mathtt{h}^{\mathtt{v}}(\mu) = H(\mu) + \frac{1}{n} \sum_{i_1\ldots i_n} \mu([i_1 \ldots i_n]) \log(p_{i_1} \ldots p_{i_n})$$

$$= H(\mu) + \lim_{n\to\infty} \frac{1}{n} \sum_{i_1\ldots i_n} \mu([i_1 \ldots i_n]) \log(p_{i_1} \ldots p_{i_n})$$

$$= -\lim_{n\to\infty} \frac{1}{n} \sum_{i_1\ldots i_n} \mu([i_1 \ldots i_n]) \log\left(\frac{\mu([i_1 \ldots i_n])}{p_{i_1} \ldots p_{i_n}}\right).$$

Onde na última equação, usamos o Corolário 2.43.

□

Queremos que a entropia variacional tenha a mesma expressão. Para isso, relembre a definição de entropia relativa dada em (2.27) e considere a mesma partição por cilindros feita na proposição acima, a seguinte proposição nos dará a expressão de interesse.

Proposição 3.2. *Se $\bar{\mathscr{F}}$ é gerada por uma partição finita $\bar{\mathscr{F}}$-mensurável $\mathcal{P}$ de Ω, então a entropia relativa é dada pela fórmula:*

$$\mathscr{H}_{\mathcal{A}}(\mu|\boldsymbol{p}) = \sum_{D\in\mathcal{P}} \mu(D)\log\left(\frac{\mu(D)}{p(D)}\right). \tag{3.1}$$

Demonstração. De fato, se f é $\bar{\mathscr{F}}_{\mathcal{A}_\mathcal{P}}$-mensurável, então f é função simples, onde $\bar{\mathscr{F}}_{\mathcal{A}_\mathcal{P}} = \sigma(D_1,\ldots,D_n)$. Como nesta partição os geradores são dois a dois disjuntos, $\sigma(D_1,\ldots,D_n) = \{B_1,\ldots,B_p\}$, onde $B_j = \cup_{r=1}^{r_j} D_{m_r}$. Para qualquer função $\bar{\mathscr{F}}_{\mathcal{A}_\mathcal{P}}$-mensurável, temos:

$$\varphi = \sum_{j=1}^{p} d_j 1_{B_j} = \sum_{j=1}^{p} 1_{\cup_{r=1}^{r_j} D_{m_r}} = \sum_{j=1}^{p}\sum_{r=1}^{r_j} d_j 1_{D_{m_r}} \tag{3.2}$$

Note que nesta sigma-álgebra, temos $\binom{n}{1}$ conjuntos com um gerador, $\binom{n}{2}$ elementos com dois geradores (união) e assim sucessivamente. Da mesma forma, fixado D_i, temos $\binom{n-1}{1}$ conjuntos com um gerador, exceto D_i, da mesma forma $\binom{n-1}{2}$ conjuntos com 2 geradores exceto D_i. Ou seja, cada gerador D_i da sigma-álgebra aparece em $\sum_{i=1}^{n}\left[\binom{n}{i} - \binom{n-1}{i}\right] + \binom{n}{n}$ elementos da mesma. Portanto, a equação (3.2) pode ser reescrita por:

$$\sum_{i=1}^{n}\left(\sum_{j=1}^{a} c_j\right) 1_{D_i},\ \textit{com } a = \sum_{i=1}^{n}\left[\binom{n}{i} - \binom{n-1}{i}\right] + \binom{n}{n}$$

e c_j é o fator que acompanha algum $D_i \cup \left(\cup_{j=1}^{m_{j_i}} D_j\right)$. Assim, a função $\frac{d\mu|_{\bar{\mathscr{F}}_{\mathcal{A}_\mathcal{P}}}}{dp|_{\bar{\mathscr{F}}_{\mathcal{A}_\mathcal{P}}}}$ é uma função simples da forma $\sum_{i=1}^{n} c_i 1_{D_i}$. Queremos determinar as constantes c_i, escolha um dos geradores da sigma-álgebra D_j. Daí:

$$\mu(D_j) = \int_{D_j} \frac{d\mu|_{\bar{\mathscr{F}}_{\mathcal{A}_\mathcal{P}}}}{dp|_{\bar{\mathscr{F}}_{\mathcal{A}_\mathcal{P}}}}\, dp|_{\bar{\mathscr{F}}_{\mathcal{A}_\mathcal{P}}}.$$

Como estamos integrando em um conjunto $\mathcal{A}_{\mathcal{P}}$-mensurável, temos que $dp = dp|_{\bar{\mathscr{F}}_{\mathcal{A}_{\mathcal{P}}}}$, sobre $\mathcal{A}_{\mathcal{P}}$. Daí:

$$\begin{aligned}
\mu(D_j) &= \int_{D_j} \frac{d\mu|_{\bar{\mathscr{F}}_{\mathcal{A}_{\mathcal{P}}}}}{dp|_{\bar{\mathscr{F}}_{\mathcal{A}_{\mathcal{P}}}}}\, dp \\
&= \int_{D_j} \sum_{i=1}^{n} c_i 1_{D_i}\, dp \\
&= \sum_{i=1}^{n} c_i \int 1_{D_j} 1_{D_i}\, dp \\
&= c_j \int 1_{D_j}\, dp \\
&= c_j p(D_j).
\end{aligned}$$

Portanto, $c_j = \frac{\mu(D_j)}{p(D_j)}$.

Integrando a função $\log\left(\frac{d\mu|_{\bar{\mathscr{F}}_{\mathcal{A}_{\mathcal{P}}}}}{dp|_{\bar{\mathscr{F}}_{\mathcal{A}_{\mathcal{P}}}}}\right)$, obtemos:

$$\begin{aligned}
\int_{D_j} \left(\frac{d\mu|_{\bar{\mathscr{F}}_{\mathcal{A}_{\mathcal{P}}}}}{dp|_{\bar{\mathscr{F}}_{\mathcal{A}_{\mathcal{P}}}}}\right) \log\left(\sum_{j=1}^{n} \frac{\mu(D_j)}{p(D_j)} 1_{D_j}\right) dp|_{\bar{\mathscr{F}}_{\mathcal{A}_{\mathcal{P}}}} \\
&= \int 1_{D_j} \log\left(\sum_{j=1}^{n} \frac{\mu(D_j)}{p(D_j)} 1_{D_j}\right) d\mu|_{\bar{\mathscr{F}}_{\mathcal{A}_{\mathcal{P}}}} \\
&= \int_{D_j} \sum_{j=1}^{n} \log\left(\frac{\mu(D_j)}{p(D_j)}\right) 1_{D_j}\, d\mu|_{\bar{\mathscr{F}}_{\mathcal{A}_{\mathcal{P}}}} \\
&= \int \log\left(\frac{\mu(D_j)}{p(D_j)}\right) d\mu \\
&= \mu(D_j) \log\left(\frac{\mu(D_j)}{p(D_j)}\right).
\end{aligned}$$

□

Estamos caminhando para a igualdade entre as duas entropias neste caso. O corolário abaixo é o ponto chave entre esta igualdade.

Corolário 3.3. *Dada $\mu \in \mathscr{M}_\sigma(\Omega)$ Usando a mesma partição via cilindros, como no item (b), da Proposição 3.1 para cada $n \in \mathbb{N}$ temos*

$$\mathscr{H}_{\Lambda_n}(\mu|\boldsymbol{p}) = \sum_{i_1 \dots i_n} \mu([i_1 \dots i_n]) \log \left(\frac{\mu([i_1 \dots i_n])}{p_{i_1} \cdots p_{i_n}} \right). \tag{3.3}$$

Demonstração. Para $n \geq 1$ qualquer, basta tomar em (3.1) a partição finita $\bar{\mathscr{F}}$-mensurável gerada por cilindros da forma $[i_1, \dots, i_n]$, com $i_i \in \{1, \dots, d\}, i \in 1, \dots, n$. □

Pela definição da entropia específica, temos que as duas entropias de interesse são iguais. Assim, temos a igualdade no nosso caso particular que será enunciado abaixo.

Teorema 3.4. *Seja K um espaço finito. Dado um potencial f normalizado α-Hölder contínuo e sua medida de Gibbs $\mu_f \in \mathcal{G}^*(f)$ correspondente, temos*

$$\mathtt{h}^{\mathtt{s}}(\mu_f) = \mathtt{h}^{\mathtt{v}}(\mu_f). \tag{3.4}$$

Corolário 3.5. *Nas mesmas condições do Teorema 3.4, se a medida a priori p for a medida de contagem normalizada, então*

$$\mathtt{h}^{\mathtt{s}}(\mu_f) = \mathtt{h}^{\mathtt{v}}(\mu_f) = H(\mu_f) + \log d.$$

Demonstração. Basta substituir $p_i = \frac{1}{d}$ na fórmula (3.3). □

Observação 3.6. *Note que o cálculo da entropia variacional depende apenas da medida a priori e da partição. Como mostraremos no Teorema A que as duas entropias são iguais, podemos obter um teorema de caracterização de quando estas entropias são finitas para cada medida de probabilidade a priori.*

3.1.2 Exemplo e Obstruções

Considere $K = [0, 1]$ e p a medida de probabilidade uniforme sobre K. Esta tem suporte completo sobre K. Para cada partição finita sobre K, existirá uma outra mais fina de modo que para todo $n \in \mathbb{N}$,

$$\mathtt{h}^{\mathtt{s}}(\mu_f) = H(\mu_f) + \log n, \; f \in C^\alpha(\Omega). \tag{3.5}$$

Em particular, para $f \equiv 0$, $\boldsymbol{p} = \prod_{i \in \mathbb{N}} p$ é a própria medida de equilíbrio para 0, mas note que a expressão acima diverge quando $n \to \infty$, pelo Teorema 2.41 e o Corolário 2.43. De fato, para cada $n \in \mathbb{N}$, podemos particionar o intervalo $[0, 1]$ por $0 = x_0, \dots, x_i, x_{i+1}, \dots, x_n = 1$, formando n intervalos da forma $[x_i, x_{i+1}]$ de modo que $\sum_{i=1}^n (x_i - x_{i-1}) = 1$. Mas para cada partição desta forma $\mathcal{P}_n \equiv \cup_{i=0}^{n-1}[x_i, x_{i+1}]$, podemos arrumar uma partição mais fina que a anterior, $\mathcal{P}_{n+1}$, onde $\mathcal{P}_n \prec \mathcal{P}_{n+1}$ e com um elemento a mais. Daí, a expressão (3.5) diverge, pois a primeira parcela vai para 0 enquanto a outra diverge.

Mas, pela definição de entropia específica, temos $\mathtt{h}^{\mathtt{s}}(\mu_f) = \mathtt{h}^{\mathtt{s}}(\mu_0) = \mathtt{h}^{\mathtt{s}}(\boldsymbol{p}) = 0$. Ou seja, a técnica aplicada ao caso finito não funciona para espaços compactos arbitrários. Por isso, a exposição que virá a seguir é necessária para o desenvolvimento da teoria para o caso onde K é um espaço compacto arbitrário.

3.2 A Igualdade das Entropias em Espaços de Estados Compactos

Vamos obter uma igualdade entre as Entropias Específica e Variacional para todas as medidas σ-invariantes sobre Ω quando K é espaço métrico compacto arbitrário.

Relembremos algumas coisas importantes que iremos utilizar agora.

Seja (K, d_K) um espaço métrico arbitrário e considere o espaço simbólico $\Omega = K^{\mathbb{N}}$ equipado com a métrica d_Ω a qual induz a topologia produto. Tal métrica é dada pela expressão a seguir:

$$d_\Omega(x, y) \equiv \sum_{n=1}^{\infty} 2^{-n} \frac{d_K(x_n, y_n)}{(1 + d_K(x_n, y_n))}.$$

Dada uma medida *a priori* p sobre K e um potencial α-Hölder contínuo f, temos que o operador de Ruelle $\mathscr{L}_f : C(\Omega) \to C(\Omega)$ leva φ a $\mathscr{L}_f(\varphi)$, dado pela seguinte expressão:

$$\mathscr{L}_f(\varphi)(x) = \int_K \exp(f(ax))\varphi(ax)\, \mathrm{d}p(a), \qquad \text{onde} \quad ax \equiv (a, x_1, x_2, \ldots).$$

Um potencial f é dito normalizado se $\mathscr{L}_f(1)(x) = 1$ para todo $x \in \Omega$. Usando a generalização do Teorema de Ruelle-Perron-Frobenius fornecida em [LMMS15], podemos associar a qualquer potencial α-Hölder f um potencial normalizado $\bar{f}$ dado por

$$\bar{f} = f + \log h_f - \log(h_f \circ \sigma) - \log \lambda_f, \tag{3.6}$$

onde λ_f é autovalor maximal de $\mathscr{L}_f$ e h_f é uma autofunção α-Hölder estritamente positiva associada a λ_f. Pelos Teoremas 2.6 e 2.17, temos que o conjunto $\mathcal{G}^*(\bar{f}) \equiv \{\mu \in \mathscr{M}_1(\Omega) : \mathscr{L}_{\bar{f}}^* \mu = \mu\}$ é único e está contido em $\mathscr{M}_\sigma(\Omega)$. Esta medida é denotada a medida de Gibbs associada ao potencial f.

3.2.1 Um resultado Auxiliar Essencial

Agora, consideremos a seguinte família de núcleos de probabilidade $(\gamma_n)_{n\geq 1}$, onde para cada $n \geq 1$ o núcleo $\gamma_n : \bar{\mathscr{F}} \times \Omega \to [0,1]$ é dado por $\gamma_n(A|y) \equiv \mathscr{L}_{\bar{f}}^n(1_A)(\sigma^n(y))$, onde $\bar{f} \in W(\Omega)$ é um potencial normalizado. Para qualquer medida de probabilidade ν, sobre Ω, e $y \in \Omega$ definimos uma medida de probabilidade $\nu\gamma_n(\cdot|y) \equiv \int_\Omega \gamma_n(\cdot|y)\, d\nu(y)$. Observamos que se ν é tal que $\mathscr{L}_{\bar{f}}^*\nu = \nu$, então para qualquer $A \in \bar{\mathscr{F}}$ e $y \in \Omega$ temos

$$
\begin{aligned}
\nu\gamma_n(A|y) &= \int_\Omega \mathscr{L}_{\bar{f}}^n(1_A) \circ \sigma^n(y)\ d\nu(y) = \int_\Omega \mathscr{L}_{\bar{f}}^n(1_A)(y)\ d\nu(y) \\
&= \int_\Omega 1_A(y)\, d[(\mathscr{L}_{\bar{f}}^*)^n\nu](y) = \nu(A).
\end{aligned}
$$

Nosso primeiro resultado é inspirado pelo Teorema 15.30 de [Geo11]. As hipóteses deste não são satisfeitas em nosso contexto. Primeiramente, estamos trabalhando sobre um reticulado unidimensional unilateral. Além disso, a família de núcleos considerada não é uma especificação associada a uma interação uniformemente integrável invariante por translações e o conjunto de parâmetros é $\mathbb{N}$ ao invés de $\mathbb{Z}$. A diferença de nossa prova é o uso das propriedades do operador de Ruelle e seu dual.

Mas antes do teorema em si, provaremos uma série de resultados para deixar a prova do Teorema 3.12 inteligível. A seguir, mostraremos as igualdades

$$
\int_\Omega \bar{f}\ d\mu = \lim_{n\to\infty} \frac{1}{n} \int_\Omega S_n(\bar{f})(x_{\Lambda_n} y_{\Lambda_n^c})\ d\mu(x) \tag{3.7}
$$

e

$$
\int_\Omega \left[S_n(f)(x_{\Lambda_n} y_{\Lambda_n^c}) - \log \int_\Omega \exp(S_n(f)(x_{\Lambda_n} w_{\Lambda_n^c}))\, d\nu(w) \right] d\mu(x) = o(n) \tag{3.8}
$$

em forma de lemas. A σ-invariância de μ será fundamental.

Lema 3.7. *Seja f um potencial contínuo em $C(\Omega)$ e $\mu \in \mathscr{M}_\sigma(\Omega)$, onde Ω é espaço métrico compacto. Então a igualdade* (3.7) *é válida.*

Demonstração. Fixe $y \in \Omega$. Queremos mostrar que

$$
\int_\Omega f(x)\ d\mu(x) = \lim_{n\to\infty} \frac{1}{n} \int_\Omega S_n(f)(x_{\Lambda_n} y_{\Lambda_n^c})\ d\mu(x),
$$

onde $x_{\Lambda_n} = (x_1, \ldots x_n)$ e $y_{\Lambda_n}^c = (y_{n+1}, \ldots)$. Primeiramente, provaremos o seguinte limite:

$$
\lim_{n\to\infty} \int_\Omega f(x_{\Lambda_n} y_{\Lambda_n^c})\ d\mu(x) = \int_\Omega f(x)\ d\mu(x). \tag{3.9}
$$

Defina $f_n(x) = f(x_{\Lambda_n} y_{\Lambda_n^c})$. Temos que $f_n(x) \to f(x)$ pontualmente e $f_n(x) \leq \|f\|_\infty$, pois μ é medida de probabilidade e, assim, $\|f\|_\infty$ é μ-integrável. Portanto, o Teorema da convergência dominada de Lebesgue é válido e a igualdade segue.

Vamos provar (3.7). Da σ-invariância de μ, temos

$$\begin{aligned}\int_\Omega S_n(f)(x_{\Lambda_n} y_{\Lambda_n^c})\, d\mu(x) &= \int_\Omega f(x_{\Lambda_n} y_{\Lambda_n^c}) + \ldots + (f \circ \sigma^{n-1})(x_{\Lambda_n} y_{\Lambda_n^c})\, d\mu(x) \\ &= n \int_\Omega f(x_{\Lambda_n} y_{\Lambda_n^c})\, d\mu(x).\end{aligned}$$

Daí, dividindo a igualdade acima por n, com $n \to \infty$,

$$\begin{aligned}\lim_{n\to\infty} \frac{1}{n} \int_\Omega S_n(f)(x_{\Lambda_n} y_{\Lambda_n^c})\, d\mu(x) &= \lim_{n\to\infty} \frac{n}{n} \int_\Omega f(x_{\Lambda_n} y_{\Lambda_n^c})\, d\mu(x) \\ &= \int_\Omega f(x)\, d\mu(x).\end{aligned}$$

Onde a última igualdade é dada por (3.9). □

Por fim, a igualdade (3.8) é o conteúdo do lema a seguir.

Lema 3.8. *Sejam f um potencial contínuo em $C(\Omega)$ e $\mu, \nu \in \mathscr{M}_\sigma(\Omega)$, com Ω espaço métrico compacto. A expressão*

$$\int_\Omega \left[S_n(f)(x_{\Lambda_n} y_{\Lambda_n^c}) - \log \int_\Omega \exp(S_n(f)(x_{\Lambda_n} w_{\Lambda_n^c}))\, d\nu(w) \right] d\mu(x)$$

é de ordem $o(n)$.

Demonstração. Para quaisquer $n \in \mathbb{N}$ e $y \in \Omega$ fixado, temos:

$$\begin{aligned}&\int_\Omega \left[S_n(f)(x_{\Lambda_n} y_{\Lambda_n^c}) - \log \int_\Omega \exp(S_n(f)(x_{\Lambda_n} w_{\Lambda_n^c}))\, d\nu(w) \right] d\mu(x) \\ &= \int_\Omega (S_n(f)(x_{\Lambda_n} y_{\Lambda_n^c})) \\ &\qquad - \left(\log \int_\Omega \exp(S_n(f)(x_{\Lambda_n} w_{\Lambda_n^c}))\, d\nu(w) \right) d\mu(x) \\ &\leq \int_\Omega (S_n(f)(x_{\Lambda_n} y_{\Lambda_n^c})) \\ &\qquad - \left[\int_\Omega \log(\exp(S_n(f)(x_{\Lambda_n} w_{\Lambda_n^c})))\, d\nu(w) \right] d\mu(x)\end{aligned}$$

$$
\begin{aligned}
&= \int_\Omega (S_n(f)(x_{\Lambda_n} y_{\Lambda_n^c})) \\
&\qquad - \left(\int_\Omega S_n(f)(x_{\Lambda_n} w_{\Lambda_n^c})\, d\nu(w) \right) d\mu(x) \\
&= \int_\Omega (S_n(f)(x_{\Lambda_n} y_{\Lambda_n^c})) - n \left[\int_\Omega f(x_{\Lambda_n} w_{\Lambda_n^c})\, d\nu(w) \right] d\mu(x) \\
&= \int_\Omega \int_\Omega n f(x_{\Lambda_n} y_{\Lambda_n^c}) - n f(x_{\Lambda_n} w_{\Lambda_n^c})\, d\nu(w)\, d\mu(x) \\
&= \int_\Omega \left[\int_\Omega n f(x_{\Lambda_n} y_{\Lambda_n^c}) - n f(x_{\Lambda_n} w_{\Lambda_n^c})\, d\mu(x) \right] d\nu(w),
\end{aligned}
$$

onde na primeira desigualdade acima usamos a Desigualdade de Jensen para a função $-\log$, em seguida usamos a σ-invariância de μ e ν e por último o Teorema de Fubini.

Dividindo a expressão acima por n e em seguida tomando o limite quando $n \to \infty$ segue do Teorema da Convergência Dominada e do Lema 3.7 o resultado. □

Em [ACR17], mostra-se que (3.8) do Lema 3.8 é dominada por outra estimativa, quando f é função α-Hölder contínua, e sua desigualdade de interesse é o conteúdo dos próximos lemas.

Lema 3.9. *Seja f um potencial α-Hölder contínuo sobre Ω, então para todos $x, y, w \in \Omega$ e $n \geq 1$,*

$$
|S_n(f)(x_{\Lambda_n} y_{\Lambda_n^c}) - S_n(f)(x_{\Lambda_n} w_{\Lambda_n^c})| \leq \sum_{j=0}^{n-1} 2^{-\alpha j} \mathrm{Hol}_\alpha(f).
$$

Demonstração. Relembre que $S_n f(x) = \sum_{k=0}^{n-1} f(\sigma^k(x))$. Daí,

$$
\begin{aligned}
|S_n(f)(x_{\Lambda_n} y_{\Lambda_n^c}) - S_n(f)(x_{\Lambda_n} w_{\Lambda_n^c})| &= |f(x_{\Lambda_n} y_{\Lambda_n^c}) + \ldots + f \circ \sigma^{n-1}(x_{\Lambda_n} y_{\Lambda_n^c}) \\
&\qquad - f(x_{\Lambda_n} w_{\Lambda_n^c}) - \ldots - f \circ \sigma^{n-1}(x_{\Lambda_n} w_{\Lambda_n^c})| \\
&= |f(x_1 \ldots x_n y_{n+1} \ldots) + \ldots + f(x_n y_{n+1} \ldots) \\
&\qquad - f(x_1 \ldots x_n w_{n+1} \ldots) - \ldots - f(x_n w_{n+1} \ldots)| \\
&\leq |f(x_1 \ldots x_n y_{n+1} \ldots) - f(x_1 \ldots x_n w_{n+1} \ldots)| + \\
&\qquad + |f(x_i \ldots x_n y_{n+1} \ldots) - f(x_i \ldots x_n w_{n+1} \ldots)| +
\end{aligned}
$$

$$+ |f(x_n y_{n+1} \ldots) - f(x_n w_{n+1} \ldots)|$$

$$\leq \mathrm{Hol}_\alpha(f)[(d_\Omega(x_1 \ldots x_n y_{n+1} \ldots, x_1 \ldots x_n w_{n+1} \ldots))^\alpha +$$

$$+ (d_\Omega(x_i \ldots x_n y_{n+1} \ldots, x_i \ldots x_n w_{n+1} \ldots))^\alpha +$$

$$+ (d_\Omega(x_n y_{n+1} \ldots, x_n w_{n+1} \ldots))^\alpha]$$

$$= \mathrm{Hol}_\alpha(f) \left[\left(\sum_{j=1}^{\infty} \left(\frac{1}{2}\right)^j \frac{d_K((x_1 \ldots x_n y_{n+1} \ldots)_j, (x_1 \ldots x_n w_{n+1} \ldots)_j)}{1 + d_K((x_1 \ldots x_n y_{n+1} \ldots)_j, (x_1 \ldots x_n w_{n+1} \ldots)_j)} \right)^\alpha + \right.$$

$$+ \ldots +$$

$$+ \left(\sum_{j=1}^{\infty} \left(\frac{1}{2}\right)^j \frac{d_K((x_i \ldots x_n y_{n+1} \ldots)_j, (x_i \ldots x_n w_{n+1} \ldots)_j)}{1 + d_K((x_i \ldots x_n y_{n+1} \ldots)_j, (x_i \ldots x_n w_{n+1} \ldots)_j)} \right)^\alpha +$$

$$+ \ldots +$$

$$\left. + \left(\sum_{j=1}^{\infty} \left(\frac{1}{2}\right)^j \frac{d_K((x_n y_{n+1} \ldots)_j, (x_n w_{n+1} \ldots)_j)}{1 + d_K((x_n y_{n+1} \ldots)_j, (x_n w_{n+1} \ldots)_j)} \right)^\alpha \right].$$

Note que algumas parcelas destes n somatórios se anulam, pois algumas projeções coincidem. Por exemplo, no i-ésimo somatório, da primeira a $(n-i)$-ésima parcelas são zero. Retirando estes termos nulos, os somatórios acima são iguais a:

$$\mathrm{Hol}_\alpha(f) \left[\left(\sum_{j=n+1}^{\infty} \left(\frac{1}{2}\right)^j \frac{d_K((x_1 \ldots x_n y_{n+1} \ldots)_j, (x_1 \ldots x_n w_{n+1} \ldots)_j)}{1 + d_K((x_1 \ldots x_n y_{n+1} \ldots)_j, (x_1 \ldots x_n w_{n+1} \ldots)_j)} \right)^\alpha + \right.$$

$$+ \ldots +$$

$$+ \left(\sum_{j=n-i+1}^{\infty} \left(\frac{1}{2}\right)^j \frac{d_K((x_i \ldots x_n y_{n+1} \ldots)_j, (x_i \ldots x_n w_{n+1} \ldots)_j)}{1 + d_K((x_i \ldots x_n y_{n+1} \ldots)_j, (x_i \ldots x_n w_{n+1} \ldots)_j)} \right)^\alpha +$$

$$+ \ldots +$$

$$\left. + \left(\sum_{j=1}^{\infty} \left(\frac{1}{2}\right)^j \frac{d_K((x_n y_{n+1} \ldots)_j, (x_n w_{n+1} \ldots)_j)}{1 + d_K((x_n y_{n+1} \ldots)_j, (x_n w_{n+1} \ldots)_j)} \right)^\alpha \right].$$

Como $x/(1+x) \leq 1$, para todo $x \in \mathbb{R}^+$, os somatórios acima são limitados por

$$
\begin{aligned}
\mathrm{Hol}_\alpha(f)&\left[\left(\sum_{j=n+1}^{\infty}\left(\frac{1}{2}\right)^j\right)^\alpha+\ldots+\left(\sum_{j=n-i+1}^{\infty}\left(\frac{1}{2}\right)^j\right)^\alpha+\right.\\
&\left.+\ldots+\left(\sum_{j=1}^{\infty}\left(\frac{1}{2}\right)^j\right)^\alpha\right]\\
&\leq \mathrm{Hol}_\alpha(f)[2^{(-n+1)\alpha}+\ldots+2^{(-n+i+1)\alpha}+\ldots+1]\\
&=\sum_{j=0}^{n-1}2^{-\alpha j}\mathrm{Hol}_\alpha(f).
\end{aligned}
$$

Portanto, temos o que se deseja. □

Desta estimativa, podemos concluir a estimativa de interesse que é o conteúdo do próximo Lema.

Lema 3.10. *Sejam φ e ψ funções contínuas em $C(\Omega)$. Então*

$$\left|\log\int_\Omega \exp(\varphi)d\mu-\log\int_\Omega \exp(\psi)d\mu\right|\leq\|\varphi-\psi\|_\infty.$$

Demonstração.

$$
\begin{aligned}
\left|\log\int_\Omega \exp(\varphi)d\mu-\log\int_\Omega \exp(\psi)d\mu\right| &= \left|\log\frac{\int_\Omega \exp(\varphi)d\mu}{\int_\Omega \exp(\psi)d\mu}\right|\\
&=\left|\log\frac{\int_\Omega \exp(\varphi+\psi-\psi)d\mu}{\int_\Omega \exp(\psi)d\mu}\right|\\
&\leq\left|\log\frac{\int_\Omega \exp(\psi+\|\varphi-\psi\|_\infty)d\mu}{\int_\Omega \exp(\psi)d\mu}\right|\\
&=\|\varphi-\psi\|_\infty.
\end{aligned}
$$

□

Quando o potencial é α-Hölder contínuo, a cota superior provém do Lema 3.9.

Corolário 3.11. *Sejam f um potencial α-Hölder contínuo e $\mu, \nu \in \mathscr{M}_\sigma(\Omega)$, com Ω espaço métrico compacto. A expressão*

$$\int_\Omega \left[S_n(f)(x_{\Lambda_n} y_{\Lambda_n^c}) - \log \int_\Omega \exp(S_n(f)(x_{\Lambda_n} w_{\Lambda_n^c}))\, d\nu(w) \right] d\mu(x)$$

é de ordem $o(n)$.

Demonstração. Para qualquer $n \in \mathbb{N}$, temos:

$$\begin{aligned}
&\int_\Omega [\ S_n(f)(x_{\Lambda_n} y_{\Lambda_n^c}) - \log \int_\Omega \exp(S_n(f)(x_{\Lambda_n} w_{\Lambda_n^c}))\ d\nu(w) \Big]\ d\mu(x) \\
&= \int_\Omega \Big[\log \big[\exp(S_n(f)(x_{\Lambda_n} y_{\Lambda_n^c})) \big] \\
&\qquad - \log \int_\Omega \exp(S_n(f)(x_{\Lambda_n} w_{\Lambda_n^c}))\, d\nu(w) \Big]\ d\mu(x) \\
&= \int_\Omega \left[\log \left[\exp(S_n(f)(x_{\Lambda_n} y_{\Lambda_n^c})) \int_\Omega 1\, d\nu(z) \right] \right. \\
&\qquad \left. - \log \int_\Omega \exp(S_n(f)(x_{\Lambda_n} w_{\Lambda_n^c}))\ d\nu(w) \right] d\mu(x) \\
&= \int_\Omega \left[\log \int_\Omega \big[\exp(S_n(f)(x_{\Lambda_n} y_{\Lambda_n^c}))\, d\nu(z) \big] \right. \\
&\qquad \left. - \log \int_\Omega \exp(S_n(f)(x_{\Lambda_n} w_{\Lambda_n^c}))\ d\nu(w) \right] d\mu(x) \\
&\le \sup_{x,y,w \in \Omega} |S_n(f)(x_{\Lambda_n} y_{\Lambda_n^c}) - S_n(f)(x_{\Lambda_n} w_{\Lambda_n^c})| \\
&\le \sum_{j=0}^{n-1} 2^{-\alpha j} \mathrm{Hol}_\alpha(f).
\end{aligned}$$

Onde na penúltima desigualdade, usamos que μ é medida de probabilidade e o Lema 3.10. Na última igualdade, recorremos ao Lema 3.9. □

Finalmente, temos condições de enunciar e provar o Teorema a seguir, onde a igualdade a ser mostrada nos permite mostrar os resultados principais desta tese. Além disso, o funcional semicontínuo que está no lado esquerdo da igualdade abaixo é o funcional que nos permite enunciar o Princípio Variacional no nosso contexto.

Teorema 3.12. *Seja p uma medida de probabilidade a priori sobre K tendo suporte total e $f \in W(\Omega)$. Então para cada $\mu \in \mathscr{M}_\sigma(\Omega)$ e $\nu \in \mathcal{G}^*(\bar{f})$, o seguinte limite existe:*

$$\lim_{n\to\infty} \frac{1}{n} \mathscr{H}_{\Lambda_n}(\mu|\nu) \equiv \mathtt{h}(\mu|\nu) = \log \lambda_f - \int_\Omega f \, d\mu - \mathtt{h}^{\mathtt{s}}(\mu)$$

$$\therefore \qquad \mathtt{h}^{\mathtt{s}}(\mu) = \log \lambda_f - \int_\Omega f \, d\mu - \mathtt{h}(\mu|\nu).$$

Demonstração. Se para algum $n \in \mathbb{N}$, tivermos $\mathscr{H}_{\Lambda_n}(\mu|\nu) = +\infty$, então $\mathscr{H}_{\Lambda_n}(\mu|\nu\gamma_n(\cdot|y)) = +\infty$ e, portanto, $\mathscr{H}_{\Lambda_n}(\mu|\gamma_n(\cdot|y)) = +\infty$. Assim $\mu|_{\bar{\mathscr{F}}_{\Lambda_n}}$ não é absolutamente contínua com respeito a $\gamma_n(\cdot|y)|_{\bar{\mathscr{F}}_{\Lambda_n}}$. Como f é limitada, segue da definição de $\gamma_n(\cdot|y)$ que $\mu|_{\bar{\mathscr{F}}_{\Lambda_n}}$ não é absolutamente contínua com respeito a $\boldsymbol{p}|_{\bar{\mathscr{F}}_{\Lambda_n}}$. Além disso, a entropia relativa é uma função crescente de $\bar{\mathscr{F}}_{\Lambda_n}$, temos $\mathscr{H}_{\Lambda_j}(\mu|\gamma_j(\cdot|y)) = +\infty$, $\forall\, j \geq n$ e, também, $\mathtt{h}^{\mathtt{s}}(\mu) = -\infty$, o que prova o teorema neste caso. Portanto, podemos assumir $\mathscr{H}_{\Lambda_n}(\mu|\gamma_n(\cdot|y)) < +\infty$, para todo $n \in \mathbb{N}$.

Para simplificar a notação escrevemos $d\boldsymbol{p} \equiv \prod_{i\in\mathbb{N}} dp$ para denotar a medida produto. Seja $\bar{f}$ um potencial normalizado à f e $\gamma_n(A|y) \equiv \mathscr{L}_{\bar{f}}^n(1_A)(\sigma^n(y))$. Se $\mathscr{H}_{\Lambda_n}(\mu|\gamma_n(\cdot|y))$ é finita para toda $n \geq 1$ então

$$\begin{aligned}
\mathscr{H}_{\Lambda_n}(\mu|\gamma_n(\cdot|y)) &= \int_\Omega \frac{d\mu|_{\bar{\mathscr{F}}_{\Lambda_n}}}{d\gamma_n(\cdot|y)|_{\bar{\mathscr{F}}_{\Lambda_n}}} \log \frac{d\mu|_{\bar{\mathscr{F}}_{\Lambda_n}}}{d\gamma_n(\cdot|y)|_{\bar{\mathscr{F}}_{\Lambda_n}}} \, d\gamma_n(\cdot|y) \\
&= \mathscr{L}_{\bar{f}}^n \left(\frac{d\mu|_{\bar{\mathscr{F}}_{\Lambda_n}}}{d\gamma_n(\cdot|y)|_{\bar{\mathscr{F}}_{\Lambda_n}}} \log \frac{d\mu|_{\bar{\mathscr{F}}_{\Lambda_n}}}{d\gamma_n(\cdot|y)|_{\bar{\mathscr{F}}_{\Lambda_n}}} \right) (\sigma^n(y)) \\
&= \int_\Omega \exp(S_n(\bar{f})) \frac{d\mu|_{\bar{\mathscr{F}}_{\Lambda_n}}}{d\gamma_n(\cdot|y)|_{\bar{\mathscr{F}}_{\Lambda_n}}} \log \frac{d\mu|_{\bar{\mathscr{F}}_{\Lambda_n}}}{d\gamma_n(\cdot|y)|_{\bar{\mathscr{F}}_{\Lambda_n}}} \prod_{i\in\Lambda_n} dp \times \prod_{i\in\Lambda_n^c} d\delta_{y_i} \\
&= \int_\Omega \frac{d\gamma_n(\cdot|y)|_{\bar{\mathscr{F}}_{\Lambda_n}}}{d\boldsymbol{p}|_{\bar{\mathscr{F}}_{\Lambda_n}}} \frac{d\mu|_{\bar{\mathscr{F}}_{\Lambda_n}}}{d\gamma_n(\cdot|y)|_{\bar{\mathscr{F}}_{\Lambda_n}}} \log \frac{d\mu|_{\bar{\mathscr{F}}_{\Lambda_n}}}{d\gamma_n(\cdot|y)|_{\bar{\mathscr{F}}_{\Lambda_n}}} \prod_{i\in\Lambda_n} dp \times \prod_{i\in\Lambda_n^c} d\delta_{y_i} \\
&= \int_\Omega \frac{d\mu|_{\bar{\mathscr{F}}_{\Lambda_n}}}{d\boldsymbol{p}|_{\bar{\mathscr{F}}_{\Lambda_n}}} \log \frac{d\mu|_{\bar{\mathscr{F}}_{\Lambda_n}}}{d\gamma_n(\cdot|y)|_{\bar{\mathscr{F}}_{\Lambda_n}}} \prod_{i\in\Lambda_n} dp \times \prod_{i\in\Lambda_n^c} d\delta_{y_i},
\end{aligned}$$

onde $S_n(\bar{f}) \equiv \bar{f} + \bar{f}\circ\sigma + \ldots + \bar{f}\circ\sigma^{n-1}$ e $\exp(S_n(\bar{f})) = \frac{d\gamma_n(\cdot|y)|_{\bar{\mathscr{F}}_{\Lambda_n}}}{d\boldsymbol{p}|_{\bar{\mathscr{F}}_{\Lambda_n}}}$ conforme (2.2). Como o integrando acima é $\bar{\mathscr{F}}_{\Lambda_n}$-mensurável, temos

$$\mathscr{H}_{\Lambda_n}(\mu|\gamma_n(\cdot|y)) = \int_\Omega \frac{d\mu|_{\bar{\mathscr{F}}_{\Lambda_n}}}{d\boldsymbol{p}|_{\bar{\mathscr{F}}_{\Lambda_n}}} \log \frac{d\mu|_{\bar{\mathscr{F}}_{\Lambda_n}}}{d\gamma_n(\cdot|y)|_{\bar{\mathscr{F}}_{\Lambda_n}}} \, d\boldsymbol{p}.$$

Das propriedades da derivada de Radon-Nikodym, temos

$$\begin{aligned}
\mathscr{H}_{\Lambda_n}(\mu|\gamma_n(\cdot|y)) &= \int_\Omega \frac{d\mu|_{\bar{\mathscr{F}}_{\Lambda_n}}}{d\boldsymbol{p}|_{\bar{\mathscr{F}}_{\Lambda_n}}} \log\left(\frac{d\mu|_{\bar{\mathscr{F}}_{\Lambda_n}}}{d\gamma_n(\cdot|y)|_{\bar{\mathscr{F}}_{\Lambda_n}}} \frac{d\gamma_n(\cdot|y)|_{\bar{\mathscr{F}}_{\Lambda_n}}}{d\boldsymbol{p}|_{\bar{\mathscr{F}}_{\Lambda_n}}}\right)\, d\boldsymbol{p} \\
&\quad - \int_\Omega \frac{d\mu|_{\bar{\mathscr{F}}_{\Lambda_n}}}{d\boldsymbol{p}|_{\bar{\mathscr{F}}_{\Lambda_n}}} \log \frac{d\gamma_n(\cdot|y)|_{\bar{\mathscr{F}}_{\Lambda_n}}}{d\boldsymbol{p}|_{\bar{\mathscr{F}}_{\Lambda_n}}}\, d\boldsymbol{p} \\
&= \int_\Omega \frac{d\mu|_{\bar{\mathscr{F}}_{\Lambda_n}}}{d\boldsymbol{p}|_{\bar{\mathscr{F}}_{\Lambda_n}}} \log \frac{d\mu|_{\bar{\mathscr{F}}_{\Lambda_n}}}{d\boldsymbol{p}|_{\bar{\mathscr{F}}_{\Lambda_n}}}\, d\boldsymbol{p} - \int_\Omega S_n(\bar{f})(x_{\Lambda_n} y_{\Lambda_n^c})\, d\mu(x) \\
&= -\mathscr{H}_{\Lambda_n}(\mu) - \int_\Omega S_n(\bar{f})(x_{\Lambda_n} y_{\Lambda_n^c})\, d\mu(x),
\end{aligned}$$

onde $(x_\Lambda y_{\Lambda^c})_i = x_i$, se $i \in \Lambda$ e $(x_\Lambda y_{\Lambda^c})_i = y_i$, caso contrário. Pelo Lema 3.7, temos

$$\int_\Omega \bar{f}\, d\mu = \lim_{n\to\infty} \frac{1}{n} \int_\Omega S_n(\bar{f})(x_{\Lambda_n} y_{\Lambda_n^c})\, d\mu(x).$$

Observe que para qualquer $\mu \in \mathscr{M}_\sigma(\Omega)$, temos $\mathscr{H}_{\Lambda_n}(\mu)/n \to \mathtt{h}^{\mathtt{s}}(\mu)$ quando $n \to \infty$. Esta convergência aliada a (3.7) implica a existência do seguinte limite

$$\lim_{n\to\infty} \frac{\mathscr{H}_{\Lambda_n}(\mu|\gamma_n(\cdot|y))}{n} = -\mathtt{h}^{\mathtt{s}}(\mu) - \int_\Omega \bar{f}\, d\mu = \log \lambda_f - \mathtt{h}^{\mathtt{s}}(\mu) - \int_\Omega f\, d\mu,$$

onde na última igualdade usamos que $\mu \in \mathscr{M}_\sigma(\Omega)$ e a expressão (3.6).

Para terminarmos a prova, basta mostrar que $n^{-1}\mathscr{H}_{\Lambda_n}(\mu|\gamma_n(\cdot|y)) \to \mathtt{h}(\mu|\nu)$, quando $n \to \infty$, para qualquer escolha de $y \in \Omega$.

$$\begin{aligned}
&\mathscr{H}_{\Lambda_n}(\mu|\nu) = \mathscr{H}_{\Lambda_n}(\mu|\nu\gamma_n(\cdot|y)) \\
&\quad = \int_\Omega \frac{d\mu|_{\bar{\mathscr{F}}_{\Lambda_n}}}{d[\nu\gamma_n(\cdot|y)|_{\bar{\mathscr{F}}_{\Lambda_n}}]} \log \frac{d\mu|_{\bar{\mathscr{F}}_{\Lambda_n}}}{d[\nu\gamma_n(\cdot|y)|_{\bar{\mathscr{F}}_{\Lambda_n}}]} d[\nu\gamma_n(\cdot|y)] \\
&\quad = \int_\Omega \frac{d\mu|_{\bar{\mathscr{F}}_{\Lambda_n}}}{d\boldsymbol{p}|_{\bar{\mathscr{F}}_{\Lambda_n}}} \left(\frac{d[\nu\gamma_n(\cdot|y)|_{\bar{\mathscr{F}}_{\Lambda_n}}]}{d\boldsymbol{p}|_{\bar{\mathscr{F}}_{\Lambda_n}}}\right)^{-1} \log \frac{d\mu|_{\bar{\mathscr{F}}_{\Lambda_n}}}{d[\nu\gamma_n(\cdot|y)|_{\bar{\mathscr{F}}_{\Lambda_n}}]} d[\nu\gamma_n(\cdot|y)] \\
&\quad = \int_\Omega \frac{d\mu|_{\bar{\mathscr{F}}_{\Lambda_n}}}{d[\nu\gamma_n(\cdot|y)|_{\bar{\mathscr{F}}_{\Lambda_n}}]} \log \frac{d\mu|_{\bar{\mathscr{F}}_{\Lambda_n}}}{d[\nu\gamma_n(\cdot|y)|_{\bar{\mathscr{F}}_{\Lambda_n}}]} d[\nu\gamma_n(\cdot|y)]
\end{aligned}$$

$$= \int_\Omega \log \frac{d\mu|_{\bar{\mathscr{F}}_{\Lambda_n}}}{d[\nu\gamma_n(\cdot|y)|_{\bar{\mathscr{F}}_{\Lambda_n}}]} \frac{d[\gamma_n(\cdot|y)|_{\bar{\mathscr{F}}_{\Lambda_n}}]}{d\boldsymbol{p}|_{\bar{\mathscr{F}}_{\Lambda_n}}} d\mu - \int_\Omega \log \frac{d[\gamma_n(\cdot|y)|_{\bar{\mathscr{F}}_{\Lambda_n}}]}{d\boldsymbol{p}|_{\bar{\mathscr{F}}_{\Lambda_n}}} d\mu$$

$$= \int_\Omega \log \frac{d\mu|_{\bar{\mathscr{F}}_{\Lambda_n}}}{d[\gamma_n(\cdot|y)|_{\bar{\mathscr{F}}_{\Lambda_n}}]} d\mu + \int_\Omega \log \frac{d[\gamma_n(\cdot|y)|_{\bar{\mathscr{F}}_{\Lambda_n}}]}{d\boldsymbol{p}|_{\bar{\mathscr{F}}_{\Lambda_n}}} \left(\frac{d[\nu\gamma_n(\cdot|y)|_{\bar{\mathscr{F}}_{\Lambda_n}}]}{d\boldsymbol{p}|_{\bar{\mathscr{F}}_{\Lambda_n}}} \right)^{-1} d\mu$$

$$= \mathscr{H}_{\Lambda_n}(\mu|\gamma_n(\cdot|y))$$

$$\quad + \int_\Omega \left[S_n(f)(x_{\Lambda_n} y_{\Lambda_n^c}) - \log \int_\Omega \exp(S_n(f)(x_{\Lambda_n} w_{\Lambda_n^c}))\ d\nu(w) \right] d\mu(x)$$

$$= \mathscr{H}_{\Lambda_n}(\mu|\gamma_n(\cdot|y)) + o(n),$$

onde a última igualdade é obtida usando o Lema 3.8 (Ou o Corolário 3.11 para quando o potencial f é α-Hölder contínuo). □

Pela definição de entropia específica, o corolário a seguir é válido.

Corolário 3.13. *Para toda $f \in W(\Omega)$ e $\mu \in \mathscr{M}_\sigma(\Omega)$ temos $\mathtt{h}(\mu|\mu_{\bar{f}}) \geq 0$. Em particular, se $\mu \in \mathcal{G}^*(\bar{f})$ então $\mathtt{h}(\mu|\mu_{\bar{f}}) = 0$.*

Demonstração. Basta observar que $\mathtt{h}(\mu|\mu_{\bar{f}})$ é limite de valores não-negativos, logo o seu limite também é não-negativo. A segunda conclusão segue diretamente da definição de medida de equilíbrio, pois elas maximizam a pressão. □

Tal resultado será útil no princípio variacional que iremos mostrar no Teorema 3.21, pois este funcional linear semicontínuo, $\mathtt{h}(\mu|\mu_{\bar{f}})$, caracterizará as medidas de equilíbrio para o potencial f.

3.2.2 A Igualdade das Entropias Específica e Variacional

Como estamos comparando duas entropias, é natural se indagar sobre uma relação de ordem entre as entropias. No começo deste capítulo, provamos a igualdade, no caso finito, para as medidas de equilíbrio.

Agora, estenderemos a igualdade para medidas σ-invariantes, no caso quando K é um espaço métrico compacto arbitrário.

Teorema A. *Para toda $\mu \in \mathscr{M}_\sigma(\Omega)$, temos $\mathtt{h}^{\mathtt{s}}(\mu) = \mathtt{h}^{\mathtt{v}}(\mu)$.*

Demonstração. Dada $\mu \in \mathscr{M}_\sigma(\Omega)$ segue do Corolário 3.13 que

$$\mathtt{h}^{\mathtt{s}}(\mu) \leq \log \lambda_g - \int_\Omega g\ d\mu, \qquad \forall\, g \in C^\alpha(\Omega).$$

Portanto, $\mathtt{h}^{\mathtt{s}}(\mu) \leq \mathtt{h}^{\mathtt{v}}(\mu)$.

O resto da prova é feita por contradição. Por razões técnicas, vamos trabalhar sobre $\mathscr{M}_s(\Omega)$, o espaço vetorial de todas as medidas com sinal finitas Borelianas equipadas com a topologia fraca-$*$, pois pelo Teorema 1.67, este espaço é o dual topológico das funções contínuas sobre Ω. As extensões $\mathtt{h}^{\mathtt{s}}, \mathtt{h}^{\mathtt{v}} : \mathscr{M}_s(\Omega) \to [-\infty, 0]$ de $\mathtt{h}^{\mathtt{s}}$ e $\mathtt{h}^{\mathtt{v}}$, respectivamente, serão dadas por

$$\mathtt{h}^{\mathtt{s}}(\mu) = \begin{cases} \mathtt{h}^{\mathtt{s}}(\mu), & \text{se } \mu \in \mathscr{M}_\sigma(\Omega) \\ -\infty, & \text{caso contrário.} \end{cases} \qquad \mathtt{h}^{\mathtt{v}}(\mu) = \begin{cases} \mathtt{h}^{\mathtt{v}}(\mu), & \text{se } \mu \in \mathscr{M}_\sigma(\Omega) \\ -\infty, & \text{caso contrário.} \end{cases}$$

A prova será dividida em dois casos.

Caso 1. Suponha que exista $\nu \in \mathscr{M}_s(\Omega)$ tal que $-\mathtt{h}^{\mathtt{v}}(\nu) < -\mathtt{h}^{\mathtt{s}}(\nu) < +\infty$. Como a extensão da entropia específica multiplicada por -1, $-\mathtt{h}^{\mathtt{s}} : \mathscr{M}_s(\Omega) \to [0, +\infty]$ é convexa, semicontínua inferiormente e limitada inferiormente, temos que seu epigrafo $\text{epi}(-\mathtt{h}^{\mathtt{s}}) \equiv \{(\mu, t) \in \mathscr{M}_\sigma(\Omega) \times \mathbb{R} : -\mathtt{h}^{\mathtt{s}}(\mu) \leq t\}$ é um subconjunto convexo fechado de $\mathscr{M}_s(\Omega) \times \mathbb{R}$ pela Proposição 1.77 e da semicontinuidade inferior de $-\mathtt{h}^{\mathtt{s}}$.

Pela afirmação, temos $(\nu, -\mathtt{h}^{\mathtt{v}}(\nu)) \notin \text{epi}(-\mathtt{h}^{\mathtt{s}})$. Portanto, o Teorema 1.72 nos garante um $c \in \mathbb{R}$ e um funcional linear contínuo $F : \mathscr{M}_s(\Omega) \times \mathbb{R} \to \mathbb{R}$ tais que para todo $(\mu, t) \in \text{epi}(-\mathtt{h}^{\mathtt{s}})$ tenhamos $F(\nu, -\mathtt{h}^{\mathtt{v}}(\nu)) < c < F((\mu, t))$. $\mathscr{M}_s(\Omega)$ é equipada com a topologia fraca-$*$, o Teorema 1.66, portanto, nos garante que o funcional F pode ser representado a seguir $F((\mu, t)) = \int_\Omega \varphi \, d\mu + at$, para algum $\varphi \in C(\Omega)$ e $a \in \mathbb{R}$.

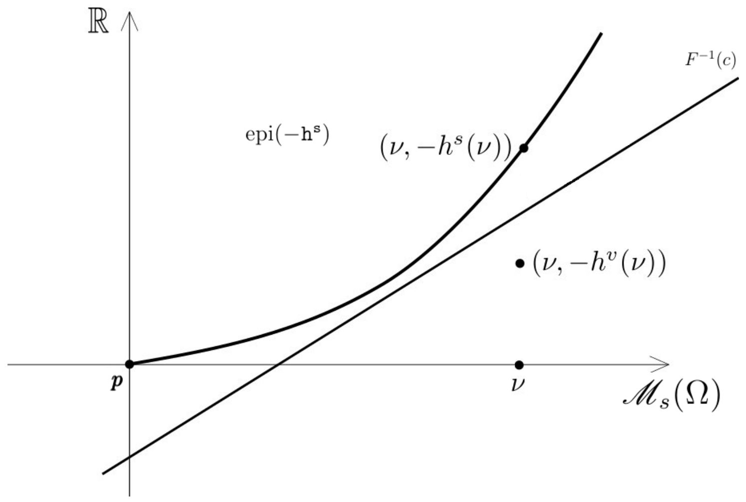

Figura 3.1: $-\mathtt{h}^{\mathtt{v}}(\nu) < -\mathtt{h}^{\mathtt{s}}(\nu) < +\infty$

Da inequação anterior, temos

$$\int_\Omega \varphi \, d\nu - a\mathtt{h}^{\mathtt{v}}(\nu) < c < \int_\Omega \varphi \, d\mu + at, \quad \text{para toda } (\mu, t) \in \text{epi}(-\mathtt{h}^{\mathtt{s}}).$$

Como Ω é compacto, $C^\alpha(\Omega)$ é denso em $C(\Omega)$ com respeito a norma uniforme. Daí, a menos de uma pequena perturbação em c podemos assumir φ na inequação acima como uma função α-Hölder contínua. Desta inequação é fácil deduzir que $a > 0$. De fato, seja $(\nu, -\mathtt{h}^{\mathtt{s}}(\nu)) \in \mathrm{epi}(-\mathtt{h}^{\mathtt{s}})$. Segue que

$$\int_\Omega \varphi \, d\nu - a\mathtt{h}^{\mathtt{v}}(\nu) < \int_\Omega \varphi \, d\nu - a\mathtt{h}^{\mathtt{s}}(\nu),$$

o que nos dá o resultado. Sem perda de generalidade, podemos assumir $a = 1$. Portanto, pelo Teorema 2.25, Corolário 3.13, a definição de $\mathrm{epi}(-\mathtt{h}^{\mathtt{s}})$ e a inequação acima, temos

$$\begin{aligned}\log \lambda_{(-\varphi)} &= \sup_{\mu \in \mathscr{M}_\sigma(\Omega)} \{ \int_\Omega (-\varphi) \, d\mu + \mathtt{h}^{\mathtt{s}}(\mu) \} \\ &\leq -c < \int_\Omega (-\varphi) \, d\nu + \mathtt{h}^{\mathtt{v}}(\nu) \leq \log \lambda_{(-\varphi)}.\end{aligned}$$

Uma contradição. Segue que para toda $\mu \in \mathscr{M}_\sigma(\Omega)$ tal que $\mathtt{h}^{\mathtt{s}}(\mu) > -\infty$ temos $\mathtt{h}^{\mathtt{s}}(\nu) = \mathtt{h}^{\mathtt{v}}(\nu)$.
Caso 2. Agora vamos provar que para toda $\mu \in \mathscr{M}_\sigma(\Omega)$ tal que $-\mathtt{h}^{\mathtt{s}}(\mu) = +\infty$, temos $-\mathtt{h}^{\mathtt{v}}(\mu) = +\infty$. A ideia da prova é reduzir esta ao caso anterior. Suponha que para algum $\nu \in \mathscr{M}_\sigma(\Omega)$ tenhamos $-\mathtt{h}^{\mathtt{v}}(\nu) < -\mathtt{h}^{\mathtt{s}}(\nu) = +\infty$. Como $-\mathtt{h}^{\mathtt{s}}$ é uma função semicontínua inferiormente e convexa de $\mathscr{M}_s(\Omega)$ a $[0, +\infty]$, então é um supremo pontual de uma família $\mathfrak{F}$ de funções contínuas afim de $\mathscr{M}_s(\Omega)$ a $\mathbb{R}$, ver Proposição 1.81. Um elemento genérico de $\mathfrak{F}$ é uma função do tipo $\mu \longmapsto \xi(\mu) + C$, onde $\xi \in \mathscr{M}_s(\Omega)^*$ e como observado antes pode ser representado por $\mu \longmapsto \int_\Omega g \, d\mu + C$, para algum $g \in C(\Omega)$.

Seja $\mathfrak{D}$ denotando a família de todas as funções afim da forma $\mathscr{M}_s(\Omega) \ni \mu \longmapsto \int_\Omega g \; d\mu - \log \lambda_g$, com g variando em $C^\alpha(\Omega)$. Como para toda $\mu \in \mathscr{M}_s(\Omega)$, temos $\mathtt{h}^{\mathtt{s}}(\mu) \leq \mathtt{h}^{\mathtt{v}}(\mu)$, podemos assumir que $\mathfrak{D} \subsetneq \mathfrak{F}$. Dada $M > -\mathtt{h}^{\mathtt{v}}(\nu)$ existe $F \in \mathfrak{F} \setminus \mathfrak{D}$ tal que $F(\nu) > M$. Por outro lado, temos $\mathtt{h}^{\mathtt{s}}(\prod_{i\in\mathbb{N}} p) = 0$, então $F(\prod_{i\in\mathbb{N}} p) \leq 0$. Sem perda de generalidade podemos assumir a última desigualdade estrita e para toda $\mu \in \mathscr{M}_s(\Omega)$, temos $F(\mu) = \int_\Omega \varphi \, d\mu + at$, para algum $\varphi \in C^\alpha(\Omega)$ e $a \in \mathbb{R}$. Note que F é dominada por $-\mathtt{h}^{\mathtt{s}}$ e separa $\mathrm{epi}(-\mathtt{h}^{\mathtt{s}})$ de $(\nu, -\mathtt{h}^{\mathtt{v}}(\nu))$. Portanto, $a \neq 0$ e a prova se reduz ao caso anterior. □

Observação 3.14. *Note que não precisamos trabalhar com as funções Walters contínuas neste teorema, pois Ω é compacto e, pelo Teorema 1.37, temos $\overline{C^\alpha(\Omega)}^{\|\cdot\|_\infty} = C(\Omega)$. Vamos mostrar que*

$$\mathtt{h}^{\mathtt{v}}(\mu) = \inf_{g \in C^\alpha(\Omega)} \{ - \int_\Omega g \, d\mu + \log \lambda_g \} = \inf_{g \in C(\Omega)} \{ - \int g \, d\mu + \log \lambda_g \}.$$

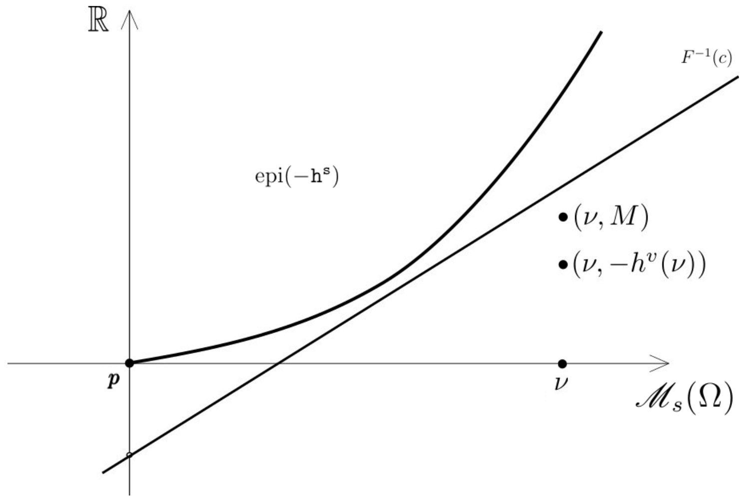

Figura 3.2: $-\mathtt{h}^{\mathtt{v}}(\nu) < -\mathtt{h}^{\mathtt{s}}(\nu) = +\infty$

Para simplificar a notação, chamaremos $\inf_{g\in C(\Omega)}\{-\int g\,d\mu + \log\lambda_g\}$ *de* a. *A expressão está bem definida, pois em [CL16] temos a pressão definida para todos os potenciais contínuos.*

Claramente, $a \leq \mathtt{h}^{\mathtt{v}}(\mu)$, *pois* $C^{\alpha}(\Omega) \subseteq C(\Omega)$. *Dado* $\varepsilon > 0$, *existe* $g \in C(\Omega)$ *tal que* $|-\int g\,d\mu + \log\lambda_g - a| < \varepsilon/2$. *Por densidade, existe uma sequência* $(g_n)_{n\in\mathbb{N}}$ *de funções* α*-Hölder contínuas tal que* g_n *converge à* g*, uniformemente. Além disso,* $\lambda_{g_n} \to \lambda_g$, *pelo Corolário 2 em [CL16], e a diferença entre integrais vai pra zero pela definição da norma da convergência uniforme. Assim, tome* n *suficientemente grande de modo que*

$$\left|-\int g\,d\mu + \log\lambda_g - \left(-\int g_n\,d\mu + \log\lambda_{g_n}\right)\right| < \frac{\varepsilon}{2}.$$

Portanto,

$$\begin{aligned}\left|-\int g_n\,d\mu + \log\lambda_{g_n} - a\right| &\leq \left|-\int g\,d\mu + \log\lambda_g - \left(-\int g_n\,d\mu + \log\lambda_{g_n}\right)\right| \\ &\quad + \left|-\int g\,d\mu + \log\lambda_g - a\right| \\ &< \varepsilon/2 + \varepsilon/2 = \varepsilon.\end{aligned}$$

Segue que a *é uma cota inferior de* $\mathtt{h}^{\mathtt{v}}$. *Daí,* $a \geq \mathtt{h}^{\mathtt{v}}(\mu)$ *e, assim, temos a igualdade de interesse. Como a classe de funções Walters contínuas é maior que as funções* α*-Hölder e todos os resultados que usamos implicitamente neste teorema para as funções* α*-Hölder também são válidos nas funções Walters, a adaptação está concluída.*

Voltemos ao caso onde K é finito. Pelo corolário 3.3, dada uma medida *a priori* p sobre K a entropia específica de μ é

$$-\lim_{n\to\infty}\frac{1}{n}\sum_{i_1\dots i_n}\mu([i_1\dots i_n])\log\left(\frac{\mu([i_1\dots i_n])}{p_{i_1}\dots p_{i_n}}\right).$$

Separando o logaritmo do somatório, obtemos

$$\mathtt{h}^{\mathtt{s}}(\mu) = H(\mu) + \lim_{n\to\infty}\frac{1}{n}\sum_{i_1\dots i_n}\mu([i_1\dots i_n])\log(p_{i_1}\dots p_{i_n})$$

$$= \sum_{i=1}^{d} -p_i\log p_i + \lim_{n\to\infty}\frac{1}{n}\sum_{i_1\dots i_n}\mu([i_1\dots i_n])\log(p_{i_1}\dots p_{i_n}).$$

Como a primeira parcela da soma acima é finita, $\mathtt{h}^{\mathtt{s}}(\mu)$ diverge se, e somente se, $\lim_{n\to\infty}(1/n)\sum_{i_1\dots i_n}\mu([i_1\dots i_n])\log(p_{i_1}\dots p_{i_n})$ diverge. Da igualdade provada acima entre as entropias específica e variacional, tal resultado pode ser estendido para a entropia variacional e o faremos em forma de Corolário a seguir

Corolário 3.15. *Sejam K um conjunto finito,p uma medida de probabilidade a priori sobre K e $\mu \in \mathscr{M}_\sigma(\Omega)$. $\mathtt{h}^{\mathtt{v}}(\mu)(\mathtt{h}^{\mathtt{s}}(\mu))$ é finita se, e somente se, o limite*

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i_1\dots i_n}\mu([i_1\dots i_n])\log(p_{i_1}\dots p_{i_n})$$

é finito.

3.3 Unicidade dos Estados de Equilíbrio e o Princípio Variacional

Finalmente, mostraremos a unicidade dos estados de equilíbrio para potenciais Walters contínuos. É um dos conceitos mais importantes do Formalismo Termodinâmico e da Mecânica Estatística. Além disso, tal unicidade nos possibilita caracterizar as medidas que são valores minimais do funcional semicontínuo $\mathtt{h}(\cdot|\mu_{\bar{f}})$, o que nos possibilita enunciar o Princípio Variacional, neste contexto.

3.3.1 A Unicidade dos Estados de Equilíbrio

Antes de provar este teorema, alguns lemas são necessários.

O primeiro lema a seguir nos garante a quasilocalidade de $\gamma_n(\cdot|y)$, para todo $n \in \mathbb{N}$. Para isto, devemos mostrar que para cada função φ local, $\gamma_n(\varphi|y)$ é quasilocal conforme a Seção 2.2.2.

Lema 3.16. *Seja f um potencial contínuo. As seguintes afirmações são verdadeiras:*

a) Existe uma sequência $(f_n)_{n\in\mathbb{N}}$ tal que esta é contínua, $\bar{\mathscr{F}}_{\Lambda_n}$-mensurável com $\lim_{n\to\infty} \|f_n - f\|_\infty = 0$;

b) O núcleo de probabilidade

$$\gamma_n(\varphi|y) = \mathscr{L}_f^n(\varphi)(\sigma^n(y))$$

é quasilocal, para todo $n \in \mathbb{N}$ e φ função quasilocal.

Demonstração. O item a) é a Proposição 3, em [CL16]. Provaremos o item b). Considere φ uma função quasilocal e $\{\varphi_n\}_{n\in\mathbb{N}}$ a sequência de funções limitadas locais que converge uniformemente a φ. Fixado m natural, defina a função ψ_m por:

$$\psi_m(y) \equiv \int_{K^n} \exp(S_n(\bar{f}_m(a^n y_{\Lambda_n^c})))\varphi(a^n y_{\Lambda_n^c})dp^n(a^n).$$

Onde $\bar{f}_m(x) = f(x_{\Lambda_m} z_{\Lambda_m^c})$, com $z \in \Omega$ fixo, é a função de interesse. Claramente ela satisfaz o item a). Segue que $S_n(\bar{f}_m)$ é local. Como φ_m é local, ϕ_m é local. Além disso, ψ_m é limitada, pois

$$\|\psi_m(y)\|_\infty \leq \exp(n\|\bar{f}\|_\infty)(1 + \|\varphi\|_\infty),$$

para m suficientemente grande. Pelo Teorema da convergência dominada, temos $\psi_m(y) \to \gamma_n(\varphi|y)$ para todo $y \in \Omega$. Afirmamos que a convergência é uniforme. De fato, para qualquer $y \in \Omega$:

$$\begin{aligned}
&|\psi_m(y) - \gamma_n(\varphi|y)| \\
&= \left| \int (\exp(S_n(\bar{f}_m))\varphi_m - \exp(S_n(\bar{f}))\varphi)\, dp^n \right| \\
&\leq \int \left|(\exp(S_n(\bar{f}_m))\varphi_m - \exp(S_n(\bar{f}_m))\varphi + \exp(S_n(\bar{f}_m))\varphi - \exp(S_n(\bar{f}))\varphi\right|\, dp^n \\
&\leq \exp(n\|\bar{f}\|_\infty)\|\varphi_m - \varphi\|_\infty + \|\varphi\|_\infty \|\exp(S_n(\bar{f}_m)) - \exp(S_n(\bar{f}))\|_\infty \\
&\to 0.
\end{aligned}$$

Onde na primeira parcela acima usamos a definição de φ e na segunda a definição de $\bar{f}_m$.

$\square$

Os seguintes lemas técnicos ajudam a nos garantir a unicidade do estado de equilíbrio no caso especial, que é quando a entropia relativa $\mathscr{H}_{\Lambda_n}(\mu|\mu_{\bar{f}})$ é uniformemente limitada.

Lema 3.17. *Suponha que para todas $v, w \in L^2(\Omega, \mu_{\bar{f}})$ com $\int_\Omega w \, d\mu_{\bar{f}} = 0$ temos $\int_\Omega v \circ \sigma^n w \, d\mu_{\bar{f}} \to 0$ quando $n \to +\infty$. Então $\mu_{\bar{f}}$ é misturadora e, em particular, ergódica.*

Demonstração. Ver Proposição 3, em [LMMS15], junto ao Teorema 2.3 em [PP90]. $\square$

Lema 3.18. *Se μ e ν são probabilidades invariantes tais que μ é ergódica e ν é absolutamente contínua com relação a μ, então $\mu = \nu$.*

Demonstração. Ver Lema 4.3.1 em [VO14]. $\square$

Falamos de sistemas ergódico e misturadores (*mixing*) sem uma definição. A observação abaixo é exatamente a definição destes conceitos.

Observação 3.19. *Seja $(\Sigma, \mathscr{F}, \mu)$ um espaço de probabilidade arbitrário. Um sistema (σ, μ) é dito misturador(mixing) se para quaisquer conjuntos mensuráveis $A, B \in \mathscr{F}$:*

$$\lim_{n\to\infty} \mu(\sigma^{-n}(A) \cap B) - \mu(A)\mu(B) = 0.$$

Algumas referências, como [Wal82] consideram esta condição denotando um sistema misturador forte. Já um sistema (σ, μ) é dito ergódico se para todo conjunto E $\mathscr{F}-$mensurável com medida positiva tivermos

$$\mu(E) = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} 1_E(\sigma^j(x)) \, \mu - q.c.$$

Um sistema misturador é ergódico. Para a prova deste fato, basta olhar a Seção 7.1.1 de [VO14].

Teorema B. *Se $f \in W(\Omega)$ então $\mathtt{h}(\mu|\mu_{\bar{f}}) = 0$ se, e somente se $\mu \in \mathcal{G}^*(\bar{f})$. Em particular, o conjunto dos estados de equilíbrio de f é unitário.*

Demonstração. Se $f \in W(\Omega)$, pela Seção 2.1.1, $\mathcal{G}^*(\bar{f})$ é unitário e pelo Corolário 3.13, $\mathtt{h}(\mu|\mu_{\bar{f}}) = 0$.

Reciprocamente, suponha que $\mu \in \mathscr{M}_\sigma(\Omega)$ é tal que $\mathtt{h}(\mu|\mu_{\bar{f}}) = 0$. Para provar que $\mu \in \mathcal{G}^*(\bar{f})$ é suficiente mostrar que para cada $n_0 \in \mathbb{N}$ fixado temos $\mu\gamma_{n_0} = \mu$, ver item b) da Definição 2.16 e o Teorema 2.17. Como assumimos que $\mathtt{h}(\mu|\mu_{\bar{f}}) = 0$, e sabemos que $n \mapsto \mathscr{H}_{\Lambda_n}(\mu|\mu_{\bar{f}})$ é função não decrescente, pela Proposição 2.28, segue que $\mathscr{H}_{\Lambda_n}(\mu|\mu_{\bar{f}}) < \infty$, para qualquer $\Lambda_n \subset \mathbb{N}$.

Primeiramente, provaremos o teorema no caso $\sup\{\mathscr{H}_{\Lambda_n}(\mu|\mu_{\bar{f}}) : n \in \mathbb{N}\} < +\infty$. Sob esta afirmação temos $\mathscr{H}_{\Lambda_n}(\mu|\mu_{\bar{f}}) < +\infty$, para todo $n \in \mathbb{N}$, e, assim, existe a derivada de Radon-Nikodym

$$\varphi_n \equiv \frac{d\mu|_{\bar{\mathscr{F}}_{\Lambda_n}}}{d\mu_{\bar{f}}|_{\bar{\mathscr{F}}_{\Lambda_n}}}. \tag{3.10}$$

Observe que $\varphi_n \geq 0$ e $\bar{\mathscr{F}}_{\Lambda_n}$-mensurável. é fácil ver que $(\varphi_n)_{n\in\mathbb{N}}$ é uma martingale relativa a $\mu_{\bar{f}}$. De fato, como $\bar{\mathscr{F}}_{\Lambda_n}$ define uma filtração crescente, para todo $m > n$ temos $\varphi_n = \mu_{\bar{f}}(\varphi_m|\bar{\mathscr{F}}_{\Lambda_n})$ $\mu_{\bar{f}}$-q.c. pela unicidade da derivada de Radon-Nikodym.

Mostraremos que esta martingale converge na norma $L^1(\mu_{\bar{f}})$ à alguma função $\bar{\mathscr{F}}$-mensurável $\varphi \geq 0$. Para isto, é suficiente provar que esta sequência é uniformemente $\mu_{\bar{f}}$-integrável. De fato, para qualquer $K > 1$ temos

$$\begin{aligned}
\int_\Omega \varphi_n 1_{\{\varphi_n \geq K\}}\, d\mu_{\bar{f}} &\leq \frac{1}{\log K}\int_\Omega [\varphi_n \log\varphi_n]\, 1_{\{\varphi_n\geq K\}}\, d\mu_{\bar{f}} \\
&< \frac{1}{\log K}\int_\Omega [1 + \varphi_n\log\varphi_n]\, 1_{\{\varphi_n\geq K\}}\, d\mu_{\bar{f}} \\
&\leq \frac{1}{\log K}(1 + \sup\{\mathscr{H}_{\Lambda_n}(\mu|\mu_{\bar{f}}) \ : n \in \mathbb{N}\}),
\end{aligned}$$

onde na última inequação usamos que $1 + x\log x \geq 0$, para todo $x \geq 0$. Destas estimativas e da afirmação que $\sup\{\mathscr{H}_{\Lambda_n}(\mu|\mu_{\bar{f}}) \ : n \in \mathbb{N}\} < +\infty$, temos que $(\varphi_n)_{n\in\mathbb{N}}$ é uma martingale uniformemente integrável e, portanto, convergente em $L^1(\mu_{\bar{f}})$. Por conseguinte, existe $\varphi \in L^1(\mu_{\bar{f}})$ tal que

$$\lim_{n\to\infty}\int_\Omega |\varphi_n - \varphi|\, d\mu_{\bar{f}} = 0.$$

Afirmamos que $\varphi = \frac{d\mu}{d\mu_{\bar{f}}}$. De fato, definindo

$$X_n = \mu_{\bar{f}}\left(\frac{d\mu}{d\mu_{\bar{f}}}\middle|\bar{\mathscr{F}}_{\Lambda_n}\right) = \frac{d\mu|_{\bar{\mathscr{F}}_{\Lambda_n}}}{d\mu_{\bar{f}}|_{\bar{\mathscr{F}}_{\Lambda_n}}} \qquad \mu_{\bar{f}} - \text{q.c.}$$

O Teorema 1.63 nos garante que X_n converge à

$$X = \mu_{\bar{f}}\left(\frac{d\mu}{d\mu_{\bar{f}}}\bigg|\bar{\mathscr{F}}_\infty\right) = \mu_{\bar{f}}\left(\frac{d\mu}{d\mu_{\bar{f}}}\bigg|\bar{\mathscr{F}}\right) = \frac{d\mu}{d\mu_{\bar{f}}} \qquad \mu_{\bar{f}} - \text{q.c.}$$

A última igualdade segue do item (E10) do Teorema 1.58 e a sigma-álgebra gerada pela união das $\bar{\mathscr{F}}_{\Lambda_n}$ é $\bar{\mathscr{F}}$, pois $\Lambda_n \uparrow \mathbb{N}$. Além disso, o Lema 3.17 nos garante $\mu_{\bar{f}}$ ergódica. Como μ é medida de probabilidade σ-invariante e $\mu \ll \mu_{\bar{f}}$, segue pelo Lema 3.18 que $\mu = \mu_{\bar{f}}$ e o teorema está provado neste caso.

Agora, o teorema no caso geral. Dado $n_0 \in \mathbb{N}$, provaremos que $\mu\gamma_{n_0} = \mu$. Dividiremos esta prova em três passos.

Passo 1. Para cada $\delta > 0$ e cada $n_1 > n_0$, existe $n_2 \in \mathbb{N}$ tal que $n_0 < n_1 < n_2$ e

$$\mathscr{H}_{\Lambda_{n_2}}(\mu|\mu_{\bar{f}}) - \mathscr{H}_{\Lambda_{n_2}\setminus\Lambda_{n_1}}(\mu|\mu_{\bar{f}}) \leq \delta, \tag{3.11}$$

onde $\mathscr{H}_{\Lambda_0}(\mu|\mu_{\bar{f}}) \equiv 0$. De fato, seja $m_1 \in \mathbb{N}$ tal que se $n \geq \max\{n_1, m_1\}$, então $n^{-1}\mathscr{H}_{\Lambda_n}(\mu|\mu_{\bar{f}}) \leq \delta/n_1$. Tomando $m_2 = \lceil \max\{n_1, m_1\}/n_1 \rceil$ temos

$$\frac{1}{m_2}\sum_{k=1}^{m_2}(\mathscr{H}_{\Lambda_{kn_1}}(\mu|\mu_{\bar{f}}) - \mathscr{H}_{\Lambda_{(k-1)n_1}}(\mu|\mu_{\bar{f}})) \leq \frac{1}{m_2}\mathscr{H}_{\Lambda_{m_2 n_1}}(\mu|\mu_{\bar{f}}) < \delta.$$

Portanto existe algum $n_2 > n_1$ tal que $\mathscr{H}_{\Lambda_{n_2}}(\mu|\mu_{\bar{f}}) - \mathscr{H}_{\Lambda_{n_2-n_1}}(\mu|\mu_{\bar{f}}) \leq \delta$. Como $\mu, \mu_{\bar{f}} \in \mathscr{M}_\sigma(\Omega)$, segue que $\mathscr{H}_{\Lambda_{n_2-n_1}}(\mu|\mu_{\bar{f}}) = \mathscr{H}_{\Lambda_{n_2}\setminus\Lambda_{n_1}}(\mu|\mu_{\bar{f}})$ e a afirmação no passo 1 está provada.

Passo 2. Seja φ_{n_2} o funcional definido como em (3.10). Dado $\varepsilon > 0$, existe $\delta > 0$ tal que

$$\int_\Omega \left|\varphi_{n_2} - \frac{d\mu|_{\bar{\mathscr{F}}_{\Lambda_{n_2}\setminus\Lambda_{n_1}}}}{d\mu_{\bar{f}}|_{\bar{\mathscr{F}}_{\Lambda_{n_2}\setminus\Lambda_{n_1}}}}\right| \, d\mu_{\bar{f}} < \varepsilon$$

toda vez que $n_1 < n_2$ e $\mathscr{H}_{\Lambda_{n_2}}(\mu|\mu_{\bar{f}}) - \mathscr{H}_{\Lambda_{n_2}\setminus\Lambda_{n_1}}(\mu|\mu_{\bar{f}}) < \delta$. De fato, se n_1 e n_2 são dois números inteiros satisfazendo $n_1 < n_2$. Então temos

$$\varphi_{n_2} = 0 \quad \mu_{\bar{f}} - q.c. \text{ sobre o conjunto } \left\{\frac{d\mu|_{\bar{\mathscr{F}}_{\Lambda_{n_2}\setminus\Lambda_{n_1}}}}{d\mu_{\bar{f}}|_{\bar{\mathscr{F}}_{\Lambda_{n_2}\setminus\Lambda_{n_1}}}} = 0\right\},$$

pois a última derivada de Radon-Nikodym é igual a $\mu_{\bar{f}}(\varphi_{n_2}|\bar{\mathscr{F}}_{\Lambda_{n_2}\setminus\Lambda_{n_1}})$. Considere a função $\psi : [0,\infty) \to \mathbb{R}$, dada por $\psi(x) = 1 - x + x\log x$. Para algum

$0 < r < \infty$, temos a seguinte desigualdade $|1 - x| \leq r\psi(x) + \varepsilon/2$, para toda $x \geq 0$. Portanto,

$$\int_\Omega \left| \varphi_{n_2} - \frac{d\mu|_{\bar{\mathscr{F}}_{\Lambda_{n_2} \setminus \Lambda_{n_1}}}}{d\mu_{\bar{f}}|_{\bar{\mathscr{F}}_{\Lambda_{n_2} \setminus \Lambda_{n_1}}}} \right| d\mu_{\bar{f}}$$

$$= \int_\Omega \frac{d\mu|_{\bar{\mathscr{F}}_{\Lambda_{n_2} \setminus \Lambda_{n_1}}}}{d\mu_{\bar{f}}|_{\bar{\mathscr{F}}_{\Lambda_{n_2} \setminus \Lambda_{n_1}}}} \left| 1 - \varphi_{n_2} \left(\frac{d\mu|_{\bar{\mathscr{F}}_{\Lambda_{n_2} \setminus \Lambda_{n_1}}}}{d\mu_{\bar{f}}|_{\bar{\mathscr{F}}_{\Lambda_{n_2} \setminus \Lambda_{n_1}}}} \right)^{-1} \right| d\mu_{\bar{f}}$$

$$\leq r \int_\Omega \frac{d\mu|_{\bar{\mathscr{F}}_{\Lambda_{n_2} \setminus \Lambda_{n_1}}}}{d\mu_{\bar{f}}|_{\bar{\mathscr{F}}_{\Lambda_{n_2} \setminus \Lambda_{n_1}}}} \, \psi\left(\varphi_{n_2} \left(\frac{d\mu|_{\bar{\mathscr{F}}_{\Lambda_{n_2} \setminus \Lambda_{n_1}}}}{d\mu_{\bar{f}}|_{\bar{\mathscr{F}}_{\Lambda_{n_2} \setminus \Lambda_{n_1}}}} \right)^{-1} \right) d\mu_{\bar{f}} + \frac{\varepsilon}{2}$$

$$= r \int_\Omega \varphi_{n_2} \log \left[\varphi_{n_2} \left(\frac{d\mu|_{\bar{\mathscr{F}}_{\Lambda_{n_2} \setminus \Lambda_{n_1}}}}{d\mu_{\bar{f}}|_{\bar{\mathscr{F}}_{\Lambda_{n_2} \setminus \Lambda_{n_1}}}} \right)^{-1} \right] d\mu_{\bar{f}} + \frac{\varepsilon}{2}$$

$$= r(\mathscr{H}_{\Lambda_{n_2}}(\mu|\mu_{\bar{f}}) - \mathscr{H}_{\Lambda_{n_2} \setminus \Lambda_{n_1}}(\mu|\mu_{\bar{f}})) + \frac{\varepsilon}{2}.$$

Tomando $\delta = \varepsilon/2r$, o passo 2 está provado.

Passo 3. Para provar que $\mu\gamma_{n_0} = \mu$, fixamos uma função local ϕ (uma função $\bar{\mathscr{F}}_{\Lambda_n}$-mensurável limitada para algum $n \in \mathbb{N}$) e $\varepsilon > 0$. Como $\bar{f}$ é um potencial Walters contínuo, segue do Lema 3.16 que a família de núcleos de probabilidade $(\gamma_n)_{n \geq 1}$ é quasilocal. Portanto, existe uma função quasilocal $\bar{\mathscr{F}}_{\Lambda^c_{n_0}}$-mensurável $\tilde{\phi}$ tal que

$$\sup_{y \in \Omega} \left| \tilde{\phi}(y) - \int_\Omega \phi(x) d\gamma_{n_0}(x|y) \right| < \varepsilon.$$

De fato,

$$\int_\Omega \phi(x) d\gamma_{n_0}(x|y) = \gamma_{n_0}(\phi|y) = \mathscr{L}^{n_0}_{\bar{f}}(\phi)(\sigma^{n_0}(y))$$

$$= \int_{K^{n_0}} \exp[S_n(\bar{f}(a^n y_{n_0+1} y_{n_0+2} \ldots))]\phi(a^n y_{n_0+1} y_{n_0+2} \ldots) dp^n(a^n).$$

Observe que o núcleo de probabilidade acima, visto como uma função mensurável na segunda entrada não depende das primeiras n coordenadas de y. Ou seja, não depende de $\bar{\mathscr{F}}_{\Lambda_n}$, daí a função quasilocal aproximante pode ser tomada na sigma-álgebra de interesse.

Seja $n_1 > n_0$ tal que ϕ é $\bar{\mathscr{F}}_{\Lambda_{n_1}}$-mensurável e $\tilde{\phi}$ é $\bar{\mathscr{F}}_{\Lambda_{n_1}\setminus\Lambda_{n_0}}$-mensurável. Escolha δ nos termos de ε como no Passo 2, e defina n_2 nos termos de n_1 e δ como no passo 1. Então

$$\begin{aligned}
\left|\int_\Omega \phi\; d\mu\gamma_{n_0} - \int_\Omega \phi\; d\mu\right| &\le \int_\Omega \left|\int_\Omega \phi(x) d\gamma_{n_0}(x|y) - \tilde{\phi}(y)\right|\; d\mu(y) \\
&+ \left|\int_\Omega \tilde{\phi}\; d\mu - \int_\Omega \tilde{\phi} \frac{d\mu|_{\bar{\mathscr{F}}_{\Lambda_{n_2}\setminus\Lambda_{n_1}}}}{d\mu_{\bar{f}}|_{\bar{\mathscr{F}}_{\Lambda_{n_2}\setminus\Lambda_{n_1}}}}\; d\mu_{\bar{f}}\right| \\
&+ \int_\Omega \frac{d\mu|_{\bar{\mathscr{F}}_{\Lambda_{n_2}\setminus\Lambda_{n_1}}}}{d\mu_{\bar{f}}|_{\bar{\mathscr{F}}_{\Lambda_{n_2}\setminus\Lambda_{n_1}}}}(y) \left|\int_\Omega \phi(x) d\gamma_{n_0}(x|y) - \tilde{\phi}(y)\right|\; d\mu_{\bar{f}}(y) \\
&+ \left|\int_\Omega \frac{d\mu|_{\bar{\mathscr{F}}_{\Lambda_{n_2}\setminus\Lambda_{n_1}}}}{d\mu_{\bar{f}}|_{\bar{\mathscr{F}}_{\Lambda_{n_2}\setminus\Lambda_{n_1}}}}(y) \left(\int_\Omega \phi(x) d\gamma_{n_0}(x|y) - \phi(y)\right)\; d\mu_{\bar{f}}(y)\right| \\
&+ \|\phi\|_\infty \int_\Omega \left|\varphi_{n_2} - \frac{d\mu|_{\bar{\mathscr{F}}_{\Lambda_{n_2}\setminus\Lambda_{n_1}}}}{d\mu_{\bar{f}}|_{\bar{\mathscr{F}}_{\Lambda_{n_2}\setminus\Lambda_{n_1}}}}\right|\; d\mu_{\bar{f}} \\
&+ \left|\int_\Omega \varphi_{n_2}\phi\; d\mu_{\bar{f}} - \int_\Omega \phi\; d\mu\right|.
\end{aligned}$$

Como $\tilde{\phi}$ é $\bar{\mathscr{F}}_{\Lambda_{n_2}\setminus\Lambda_{n_1}}$-mensurável e ϕ é $\bar{\mathscr{F}}_{\Lambda_{n_2}}$-mensurável, o segundo e último termo do lado direito da desigualdade são zero. Pela escolha de $\tilde{\phi}$, o primeiro e terceiro termos são limitados por ε. O quinto termo é limitado por $\varepsilon\|\phi\|_\infty$ dada a escolha de n_2. O quarto termo se anula pelas equações DLR. De fato, como $\mu_{\bar{f}} \in \mathcal{G}^*(\bar{f})$, segue do Teorema 2.17 que $\mu_{\bar{f}}\gamma_{n_0} = \mu_{\bar{f}}$, portanto

$$\begin{aligned}
&\left|\int_\Omega \frac{d\mu|_{\bar{\mathscr{F}}_{\Lambda_{n_2}\setminus\Lambda_{n_1}}}}{d\mu_{\bar{f}}|_{\bar{\mathscr{F}}_{\Lambda_{n_2}\setminus\Lambda_{n_1}}}}(y) \left(\int_\Omega \phi(x) d\gamma_{n_0}(x|y) - \phi(y)\right)\; d\mu_{\bar{f}}(y)\right| \\
&= \left|\int_\Omega \int_\Omega \frac{d\mu|_{\bar{\mathscr{F}}_{\Lambda_{n_2}\setminus\Lambda_{n_1}}}}{d\mu_{\bar{f}}|_{\bar{\mathscr{F}}_{\Lambda_{n_2}\setminus\Lambda_{n_1}}}}(x)\phi(x)\, d\gamma_{n_0}(x|y)\, d\mu_{\bar{f}}(y) - \int_\Omega \phi\, \frac{d\mu|_{\bar{\mathscr{F}}_{\Lambda_{n_2}\setminus\Lambda_{n_1}}}}{d\mu_{\bar{f}}|_{\bar{\mathscr{F}}_{\Lambda_{n_2}\setminus\Lambda_{n_1}}}}\; d\mu_{\bar{f}}\right| \\
&= 0,
\end{aligned}$$

onde na primeira igualdade usamos que γ_{n_0} é um núcleo de probabilidade próprio e as equações DLR. $\square$

Observação 3.20. *É fácil verificar que* $|1-x| \le r\psi(x) + \varepsilon/2$ *para algum*

$r > 0$, pois definindo a função

$$f(r) = \begin{cases} r\psi(x) + (1-x) + \varepsilon/2, & \textit{se } x < 1; \\ r\psi(x) + (x-1) + \varepsilon/2, & \textit{caso contrário.} \end{cases}$$

Temos f contínua e $f' = \psi(x) \geq 0$ para todo x. Como a função é crescente, existe algum r de interesse.

3.3.2 Pressão Topológica e a Transformada de Legendre-Fenchel

Como consequência da unicidade dos estados de equilíbrio, da igualdade entre as entropias e do Teorema 3.12, podemos deduzir o Princípio Variacional em nosso contexto:

Teorema 3.21. *Para cada $f \in W(\Omega)$ e $\mu \in \mathscr{M}_\sigma(\Omega)$, $\mathtt{h}(\mu|\mu_{\bar{f}}) \geq 0$. A igualdade ocorre se, e somente se, $\mu \in \mathcal{G}^*(\bar{f})$. Em outras palavras, $\mathcal{G}^*(\bar{f})$ é o conjunto onde o funcional semicontínuo $\mathtt{h}(\cdot|\mu_{\bar{f}}) : \mathscr{M}_\sigma(\Omega) \to [0, \infty]$ atinge seu mínimo zero.*

Note que pelo teorema acima, o funcional Pressão, definido em 2.24, satisfaz

$$\begin{aligned} \mathsf{P}(f) &= \sup_{\mu \in \mathscr{M}_\sigma(\Omega)} \left\{ \int_\Omega f\, d\mu + \mathtt{h}^{\mathtt{v}}(\mu) \right\} \\ &= \sup_{\mu \in \mathscr{M}_s(\Omega)} \left\{ \int_\Omega f\, d\mu + \mathtt{h}^{\mathtt{v}}(\mu) \right\}. \end{aligned} \tag{3.12}$$

Por outro lado, a entropia variacional, explicitada na Observação 3.14, satisfaz

$$\begin{aligned} -\mathtt{h}^{\mathtt{v}}(\mu) &= \sup_{f \in C^\alpha(\Omega)} \left\{ \int_\Omega f\, d\mu - \mathsf{P}(f) \right\} \\ &= \sup_{f \in C(\Omega)} \left\{ \int_\Omega f\, d\mu - \mathsf{P}(f) \right\}. \end{aligned} \tag{3.13}$$

Afirmamos que $-\mathtt{h}^{\mathtt{v}}$ é a Transformada de Legendre-Fenchel (conjugada convexa) de P. Ou seja, elas estão em dualidade. Mostraremos que P é uma Γ-regularização. Das propriedades de supremo, é fácil ver que P é convexa. Além disso, P é Lipschitzianas na topologia forte. Assim,

$$\forall\, c \in \mathbb{R}, \qquad \bar{C} \equiv \{f \in C(\Omega) \,|\, \mathsf{P}(f) \leq c\}$$

é fechado na topologia forte. Como P é convexa, $\bar{C}$ é convexo. Daí, $\bar{C}$ é convexo na topologia fraca e, portanto, P é l.s.c. De (3.12) e (3.13), segue que $\mathsf{P}^* = -\mathtt{h}^{\mathtt{v}}$ e $(-\mathtt{h}^{\mathtt{v}})^* = \mathsf{P}$ e assim, concluímos o resultado.

Observações Finais

Devemos mencionar que nossos resultados sobre a unicidade dos estados de equilíbrio para potenciais no espaço de Walters não podem ser deduzido dos resultados obtidos em [GKLM18]. O primeiro motivo é que neste trabalho os autores assumem que a dinâmica é do tipo *finite-to-one* o que não ocorrerá para o caso de alfabetos infinitos. O segundo motivo é que o operador de Ruelle associado a um potencial Walters não possui, em geral, a propriedade do *gap* espectral, veja [CS16]. Nesses casos, o argumento dado em [GKLM18] não pode ser utilizado. Embora as referências [LMMS15], [CS16], [CL16] já tenham considerado alfabetos compactos e estados de equilíbrio neste contexto, os resultados de unicidade discutidos nestes artigos se referem às automedidas do dual do operador de Ruelle. Como mencionado antes, no caso do alfabeto finito, a unicidade das automedidas implica na unicidade dos estados de equilíbrio, pelo menos para potenciais Hölder e Walters. Para o caso de alfabetos métricos compactos gerais novos resultados são necessários. Estes resultados são obtidos no Capítulo 3.

Em [LMMS15], [CS16], [CL16], o conceito de entropia é considerado no contexto de alfabetos métricos compactos. Elel é definido por um princípio variacional, mas não é reconhecido como sendo menos a transformada de Legendre-Fenchel do funcional pressão. Como subproduto do Teorema A nós provamos que: a entropia considerada nestes trabalhos é, de fato, menos a transformada de Legendre-Fenchel da pressão e, além disso, é igual a entropia específica. Esta igualdade, o conteúdo do Teorema A, possui uma consequência extra por causa da Proposição 15.14 de [Geo11] que é: a entropia considerada nestes trabalhos é uma função afim, quando restrita ao subespaço das medidas Borelianas de probabilidade σ-invariantes. Embora seja um resultado conhecido para sistemas dinâmicos com entropia topológica finita, isto é novo para dinâmica simbólica com alfabetos infinitos.

Como essas entropias dependem de uma medida *a priori*, a pergunta é natural. Existe uma melhor medida *a priori* a se trabalhar? Quando K é finito, os Corolários 3.5 e 3.15 sugerem que a medida de contagem normalizada é uma medida bastante adequada para ser escolhida como medida *a priori*.

Com esta escolha, todas as medidas σ-invariantes tem entropia específica (variacional) finita. Quando K é um grupo compacto não-enumerável, uma escolha bastante útil é a medida de Haar.

É possível utilizar os resultados obtidos nesta tese para estudar alguns problemas onde $\Omega = \mathbb{N}^{\mathbb{N}}$. Note que neste caso o alfabeto é não-compacto e portanto alguma técnica de compactificação deve ser utilizada. Uma maneira de se fazer isto é usar uma ideia explorada em [LMMS15]. A técnica consiste em considerar $K_0 = \{z_i, i \in \mathbb{N}\}$, uma sequência infinita de pontos em $[0,1)$ e $z_\infty := 1 = \lim_{i\to\infty} z_i$. Supomos que $z_1 = 0$. Portanto, cada ponto de K_0 é isolado e temos apenas um ponto de acumulação, $z_\infty = 1$. Considerando a métrica euclideana induzida, então $K = K_0 \cup \{1\}$ é conjunto compacto. O alfabeto K_0 pode ser identificado como $\mathbb{N}$, e K tem um ponto $z_\infty = 1$ fazendo o papel do infinito. Tome $\Omega_0 = K_0^{\mathbb{N}}$ e $\Omega = K^{\mathbb{N}}$. Temos Ω_0 um conjunto não compacto. Pela Proposição 12, em [LMMS15], o Teorema 2.6 é válido para este espaço. Daí, segue que que os resultados aqui apresentados são válidos para os potenciais α-Hölder contínuos em Ω que também são contínuos em sua restrição, Ω_0.

Outra questão que vem a mente é se a classe de potenciais contínuos, onde os Teoremas 3.12, A e B são válidos, pode ser estendida para uma classe mais geral. Certamente, pois todos os resultados para os potenciais provados no Capítulo 3 dependem apenas da continuidade do potencial. A restrição a $W(\Omega)$ se dá pelo Teorema de Ruelle-Perron-Frobenius. Se mostrarmos que este teorema pode ser estendido para uma classe mais ampla (Classe de Bowen, por exemplo) ou, pelo menos, mostrar a existência de uma autofunção mensurável associada a uma automedida ,a extensão desses resultados é intuitiva. Além disso, em [LMST09], mostra-se que as medidas de equilíbrio são aproximadas por medidas empíricas. Espera-se que fazendo as adaptações necessárias, no nosso contexto, as medidas de equilíbrio também tenham esta propriedade.

Apesar de termos restrições técnicas sobre a regularidade de f, dada uma medida de equilíbrio ν para este potencial, o funcional semicontínuo $\mathtt{h}(\cdot|\nu)$ pode ser estendido para todas as funções contínuas, pois este depende apenas da informação espectral maximal do operador de Ruelle, pelo Teorema 3.12. Daí, sendo μ uma medida σ-invariante, $\mathtt{h}(\mu|\nu)$ pode ser definido por

$$\mathtt{h}(\mu|\nu) \equiv \log \rho(\mathscr{L}_f) - \int_\Omega f\, d\mu - \mathtt{h}^{\mathtt{v}}(\mu).$$

Note que a expressão acima é similar a provada no Teorema 3.12, mas a entropia específica $\mathtt{h}^{\mathtt{s}}(\mu)$ é trocada por $\mathtt{h}^{\mathtt{v}}(\mu)$ e podemos ver que $\mathtt{h}(\mu|\nu) \geq 0$. Por outro lado, não sabemos, em geral, quando $n^{-1}\mathscr{H}_{\Lambda_n}(\mu|\gamma_n(\cdot|y))$ converge a $\mathtt{h}(\mu|\nu)$. Contra-exemplos, no reticulado $\mathbb{Z}$, dados na Seção A.5.2 de [vEFS93] (sistemas clássicos) e em [MvE98] (sistemas quânticos) sugerem que esta questão da convergência pode ser um problema bastante delicado.

Outra questão que se coloca é a possibilidade de estender nossos resultados para o caso de *subshifts*. A importância de tal extensão se deve a belíssima conexão entre dinâmicas uniformemente hiperbólicas em variedades compactas e dinâmica simbólica em *subshifts* de $\{1, \ldots, d\}^{\mathbb{N}}$. Esta conexão é feita através das chamadas partições de Markov que permitem obter semi-conjugações topológicas entre tais dinâmicas. Já que as provas dos resultados obtidos nesta tese são baseadas em resultados da teoria geral de Análise Convexa e alguns teoremas de compacidade, acreditamos que eles possam ser generalizados para tal contexto pelo menos quando os *subshifts* forem compactos.

Referências Bibliográficas

[ACR17] D. Aguiar, L. Cioletti, and R. Ruviaro. A variational principle for the specific entropy for symbolic systems with uncountable alphabets. *arXiv preprint arXiv:1707.01762*, 2017.

[AMO01] R. P. Agarwal, M. Meehan, and D. O'Regan. *Fixed point theory and applications*, volume 141 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 2001.

[Ash00] R. B. Ash. *Probability and measure theory*. Harcourt/Academic Press, Burlington, MA, second edition, 2000. With contributions by Catherine Doléans-Dade.

[Bal00] V. Baladi. *Positive transfer operators and decay of correlations*, volume 16 of *Advanced Series in Nonlinear Dynamics*. World Scientific Publishing Co., Inc., River Edge, NJ, 2000.

[Bar95] R. G. Bartle. *The elements of integration and Lebesgue measure*. Wiley Classics Library. John Wiley & Sons, Inc., New York, 1995. Containing a corrected reprint of the 1966 original [ıt The elements of integration, Wiley, New York; MR0200398 (34 #293)], A Wiley-Interscience Publication.

[BCL+11] A. T. Baraviera, L. Cioletti, A. O. Lopes, J. Mohr, and R. R. Souza. On the general one-dimensional XY model: positive and zero temperature, selection and non-selection. *Rev. Math. Phys.*, 23(10):1063–1113, 2011.

[Bre57] L. Breiman. The individual ergodic theorem of information theory. *Ann. Math. Statist.*, 28:809–811, 1957.

[Bre11] H. Brezis. *Functional analysis, Sobolev spaces and partial differential equations*. Universitext. Springer, New York, 2011.

[BW16] R. Bhattacharya and E. C. Waymire. *A basic course in probability theory*. Universitext. Springer, Cham, second edition, 2016.

[CL16] L. Cioletti and A. O. Lopes. Ruelle operator for continuous potentials and DLR-Gibbs measures. *Preprint arXiv:1608.03881*, 2016.

[CS16] L. Cioletti and E. A. Silva. Spectral properties of the Ruelle operator on the Walters class over compact spaces. *Nonlinearity*, 29(8):2253–2278, 2016.

[Dob68a] R. L. Dobrushin. Description of a random field by means of conditional probabilities and conditions for its regularity. *Teor. Verojatnost. i Primenen*, 13:201–229, 1968.

[Dob68b] R. L. Dobrushin. Gibbsian random fields for lattice systems with pairwise interactions. *Funkcional. Anal. i Priložen.*, 2(4):31–43, 1968.

[ET76] I. Ekeland and R. Temam. *Convex analysis and variational problems*. North-Holland Publishing Co., Amsterdam-Oxford; American Elsevier Publishing Co., Inc., New York, 1976. Translated from the French, Studies in Mathematics and its Applications, Vol. 1.

[Fol99] G. B. Folland. *Real analysis*. Pure and Applied Mathematics (New York). John Wiley & Sons, Inc., New York, second edition, 1999. Modern techniques and their applications, A Wiley-Interscience Publication.

[Geo11] H-O. Georgii. *Gibbs measures and phase transitions*, volume 9 of *de Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, second edition, 2011.

[GKLM18] P. Giulietti, B. Kloeckner, A. Lopes, and D. Marcon. The calculus of thermodynamical formalism. *to appear in J. Eur. Math. Soc. (JEMS)*, 2018.

[GMS67] G. Gallavotti and S. Miracle-Sole. Statistical mechanics of lattice systems. *Comm. Math. Phys.*, 5:317–323, 1967.

[KL51] S. Kullback and R. A. Leibler. On information and sufficiency. *Ann. Math. Statistics*, 22:79–86, 1951.

[KW41] H. A. Kramers and G. H. Wannier. Statistics of the two-dimensional ferromagnet. I. *Phys. Rev. (2)*, 60:252–262, 1941.

[Lim05] E. L. Lima. *Espaços métricos*, volume 4 of *Projeto Euclides [Euclid Project]*. Instituto de Matemática Pura e Aplicada, Rio de Janeiro, 2005.

[LMMS15] A. O. Lopes, J. K. Mengue, J. Mohr, and R. R. Souza. Entropy and variational principle for one-dimensional lattice systems with a general *a priori* probability: positive and zero temperature. *Ergodic Theory Dynam. Systems*, 35(6):1925–1961, 2015.

[LMST09] A. O. Lopes, J. Mohr, R. R. Souza, and Ph. Thieullen. Negative entropy, zero temperature and Markov chains on the interval. *Bull. Braz. Math. Soc. (N.S.)*, 40(1):1–52, 2009.

[LO09] A. O. Lopes and E. R. Oliveira. Entropy and variational principles for holonomic probabilities of IFS. *Discrete Contin. Dyn. Syst.*, 23(3):937–955, 2009.

[LR69] O. E. Lanford, III and D. Ruelle. Observables at infinity and states with short range correlations in statistical mechanics. *Comm. Math. Phys.*, 13:194–215, 1969.

[McM53] B. McMillan. The basic theorems of information theory. *Ann. Math. Statistics*, 24:196–219, 1953.

[Mon41] E. W. Montroll. Statistical mechanics of nearest neighbor systems. *The Journal of Chemical Physics*, 9(9):706–721, 1941.

[MvE98] H. Moriya and A. van Enter. On thermodynamic limits of entropy densities. *Lett. Math. Phys.*, 45(4):323–330, 1998.

[NG58] R. E. Nettleton and M. S. Green. Expression in terms of molecular distribution functions for the entropy density in an infinite system. *The Journal of Chemical Physics*, 29(6):1365–1370, 1958.

[Ons44] L. Onsager. Crystal statistics. I. A two-dimensional model with an order-disorder transition. *Phys. Rev. (2)*, 65:117–149, 1944.

[Osb14] M. S. Osborne. *Locally convex spaces*, volume 269 of *Graduate Texts in Mathematics*. Springer, Cham, 2014.

[Per57] A. Perez. Sur la théorie de l'information dans le cas d'un alphabet abstrait. In *Transactions of the first Prague conference on information theory, statistical decision functions, random processes held at Liblice near Prague from November 28 to 30, 1956*, pages 209–243. Publishing House of the Czechslovak Academy of Sciences, Prague, 1957.

[PP90] W. Parry and M. Pollicott. Zeta functions and the periodic orbit structure of hyperbolic dynamics. *Astérisque*, (187-188):268, 1990.

[Rue67] D. Ruelle. A variational formulation of equilibrium statistical mechanics and the Gibbs phase rule. *Comm. Math. Phys.*, 5:324–329, 1967.

[Rue68] D. Ruelle. Statistical mechanics of a one-dimensional lattice gas. *Comm. Math. Phys.*, 9:267–278, 1968.

[Rue69] D. Ruelle. *Statistical mechanics: Rigorous results*. W. A. Benjamin, Inc., New York-Amsterdam, 1969.

[Sch71] H. H. Schaefer. *Topological vector spaces*. Springer-Verlag, New York-Berlin, 1971. Third printing corrected, Graduate Texts in Mathematics, Vol. 3.

[Sha48] C. E. Shannon. A mathematical theory of communication. *Bell System Tech. J.*, 27:379–423, 623–656, 1948.

[vEFS93] A. C. D. van Enter, R. Fernández, and A. D. Sokal. Regularity properties and pathologies of position-space renormalization-group transformations: scope and limitations of Gibbsian theory. *J. Statist. Phys.*, 72(5-6):879–1167, 1993.

[VO14] M. Viana and K. Oliveira. *Fundamentos da Teoria Ergódica*, volume 90. Sociadade Brasileira de Matemática, 2014.

[Wal75] P. Walters. A variational principle for the pressure of continuous transformations. *Amer. J. Math.*, 97(4):937–971, 1975.

[Wal76] Peter Walters. A generalized Ruelle Perron-Frobenius theorem and some applications. pages 183–192. Astérisque, No. 40, 1976.

[Wal82] P. Walters. *An introduction to ergodic theory*, volume 79 of *Graduate Texts in Mathematics*. Springer-Verlag, New York-Berlin, 1982.

[Wal01] P. Walters. Convergence of the Ruelle operator for a function satisfying Bowen's condition. *Trans. Amer. Math. Soc.*, 353(1):327–347, 2001.