Universidade Federal de Uberlândia - UFU
Faculdade de Matemática - FAMAT
Coordenação dos Cursos de Bacharelado e Licenciatura em Matemática

Trabalho de Conclusão de Curso

Sistemas Dinâmicos: Alguns Teoremas Clássicos da Teoria Ergódica

MATEUS RIBEIRO DE SOUZA MARRA

Orientado pelo Professor: Dr. Thiago Catalan

Sumário

1 RESUMO

Neste trabalho, faremos inicalmente um passeio pela história da Teoria Ergódica e, em seguida, uma introdução à Teoria da Medida e Integração, no qual trabalharemos com espaços mensuráveis e espaços de medida. Logo após, trabalharemos o Teorema da Recorrência de Poincaré e veremos o conceito de Medidas Invariantes. Por fim, demonstraremos importantes resultados da Teoria Ergódica como por exemplo o Teorema Ergódico de Birkhoff e o Teorema da Decomposição Ergódica.

2 INTRODUÇÃO

Em termos simples, a Teoria Ergódica é a disciplina matemática que estuda sistemas dinâmicos munidos de medidas invariantes. Começaremos por dar as definições precisas destas noções e algumas das principais motivações para o seu estudo.

2.1 Sistemas dinâmicos

Há várias definições, mais ou menos gerais, do que é um sistema dinâmico. Nós nos restringiremos a dois modelos principais. O primeiro deles, ao qual nos referiremos na maior parte do tempo, são as transformações $f : M \to M$ em algum espaço métrico ou topológico M. Heuristicamente, pensamos em f como associando a cada estado $x \in M$ do sistema o estado $f(x) \in M$ em que o sistema se encontrará uma unidade de tempo depois. Trata-se portanto de um modelo de dinâmica com tempo discreto.

Também consideraremos fluxos, que são modelos de sistemas dinâmicos a tempo contínuo. Lembre que um fluxo em M é uma família $f^t : M \to M,\ t \in \mathbb{R}$ de transformações satisfazendo

$$f^0 = \text{identidade e } f^t \circ f^s = f^{t+s} \text{ para todo } t, s \in \mathbb{R}.$$

Fluxos aparecem, por exemplo, associados a equações diferenciais: tome como f^t a transformação que associa a cada ponto x o valor no tempo t da solução da equação que passa por x no tempo zero. Num caso e no outro, sempre iremos supor que o sistema dinâmico é pelo menos mensurável: na maior parte dos casos será até contínuo, ou mesmo diferenciável.

2.2 Teoria Ergódica

A palavra *ergódico* é a concatenação de duas palavras gregas, *ergos* = trabalho e *odos* = caminho, e foi introduzida por Boltzmann, no século XIX, no seu trabalho sobre a teoria cinética dos gases. Os sistemas em que Boltzmann, Maxwell, Gibbs, os principais fundadores da teoria cinética, estavam interessados são descritos por um fluxo hamiltoniano, ou seja, uma equação diferencial da forma

$$\left(\frac{dq_1}{dt}, ..., \frac{dq_n}{dt}, \frac{dp_1}{dt}, ..., \frac{dp_n}{dt}\right) = \left(\frac{\partial H}{\partial p_1}, ..., \frac{\partial H}{\partial p_n}, -\frac{\partial H}{\partial q_1}, ..., -\frac{\partial H}{\partial q_n}\right).$$

Boltzmann acreditava que as órbitas típicas do fluxo preenchem toda a superfície de energia $H^{-1}(c)$ que as contém. A partir desta hipótese ergódica, ele deduzia que as médias temporais de grandezas observáveis (funções) ao longo de órbitas típicas coincidem com as respectivas médias espaciais na superfície de energia, um fato crucial para a sua formulação da teoria cinética.

De fato, esta hipótese é claramente falsa e, com o tempo, tornou-se usual chamar hipótese ergódica a sua consequência de igualdade das médias temporais e espaciais. Sistemas para os quais esta igualdade vale foram chamados ergódicos. E pode dizer-se que uma boa parte da Teoria Ergódica, tal como ela se desenvolveu ao longo do século XX, foi motivada pelo problema de decidir se a maioria dos sistemas hamiltonianos, especialmente aqueles que aparecem na teoria cinética dos gases, são ergódicos ou não.

3 UMA INTRODUÇÃO À TEORIA DA MEDIDA

Neste capítulo inicial recordamos algumas noções e resultados básicos da Teoria da Medida e Integração que serão úteis para o decorrer deste trabalho.

3.1 Espaços mensuráveis

Começamos por introduzir as noções de álgebra e $\sigma-$álgebra de subconjuntos. Considere M um conjunto qualquer.

Uma álgebra de subconjuntos de M é uma família $\mathcal{B}$ de subconjuntos que contém M e é fechada para as operações elementares de conjuntos:

- $A \in \mathcal{B}$ implica $A^c = M \backslash A \in \mathcal{B}$,
- $A \in \mathcal{B}$ e $B \in \mathcal{B}$ implica $A \cup B \in \mathcal{B}$.

Observemos que $A \cap B = (A^c \cup B^c)^c$ e $A \backslash B = A \cap B^c$ também estão em $\mathcal{B}$, quaisquer que sejam $A, B \in \mathcal{B}$. Além disso, por associatividade, a união e a interseção de qualquer número finito de elementos de $\mathcal{B}$ também estão em $\mathcal{B}$.

Uma álgebra diz-se uma $\sigma-$álgebra de subconjuntos de M se também for fechada para uniões enumeráveis:

- $A_j \in \mathcal{B}$ para $j = 1, 2, ..., n, ...$ implica $\cup_{j=1}^{\infty} A_j \in \mathcal{B}$.

Um espaço mensurável é uma par ordenado $(M, \mathcal{B})$ onde M é um conjunto e $\mathcal{B}$ é uma $\sigma-$álgebra de subconjuntos de M. Os elementos de $\mathcal{B}$ são chamados conjuntos mensuráveis. Chamamos de $\sigma-$álgebra gerada por uma família $\mathcal{E}$ de subconjuntos de M a menor $\sigma-$álgebra que contém a família $\mathcal{E}$.

Dado (M, τ) um espaço topológico, isto é, M um conjunto e τ a família dos subconjuntos abertos de M. Então a $\sigma-$álgebra de Borel de M é a σ-álgebra gerada por τ, ou seja, a menor

σ-álgebra que contém todos os subconjuntos abertos.

3.2 Espaços de medida

Agora, vamos introduzir o conceito de medida e analisamos algumas das suas propriedades fundamentais.

Definição 1. *Uma medida num espaço mensurável* $(M, \mathcal{B})$ *é uma função* $\mu : B \to [0, \infty]$ *que satisfaz:*

1. $\mu(\emptyset) = 0$;
2. $\mu(\bigcup_{j=1}^{\infty} A_j) = \sum_{j=1}^{\infty} A_j$ *para quaisquer* $A_j \in \mathcal{B}$ *disjuntos dois-a-dois.*

A tripla $(M, \mathcal{B}, \mu)$ é chamada espaço de medida. Quando $\mu(M) = 1$ dizemos que μ é uma medida de probabilidade e $(M, \mathcal{B}, \mu)$ é um espaço de probabilidade. Dizemos que $A \subset M$ é um conjunto de medida total se $\mu(A) = \mu(M)$, por outro lado se $\mu(A) = 0$ dizemos que A é um conjunto nulo.

3.3 Integração em espaços de medida

Nesta seção definimos a noção de integral de uma função em relação a uma medida e apresentamos os teoremas fundamentais da Teoria da Medida. Para tanto, introduziremos algumas classes de funções. Ao longo desta seção $(M, \mathcal{B}, \mu)$ será sempre um espaço de medida.

Definição 2. *Seja* $\mathcal{B}(\mathbb{R})$ *a* $\sigma-$*álgebra de Borel de* $\mathbb{R}$. *Uma função* $f : M \to \mathbb{R}$ *diz-se mensurável se* $f^{-1}(D) \in \mathcal{B}$ *para todo* $D \in \mathcal{B}(\mathbb{R})$.

O espaço das funções mensuráveis satisfaz diversas propriedades muito úteis. Vamos enunciá-las como proposição:

Proposição 3. *Sejam* f_1, f_2 *funções mensuráveis e* $c_1, c_2 \in \mathbb{R}$. *Então também são mensuráveis as seguintes funções:*

1. $(c_1 f_1 + c_2 f_2)(x) = c_1 f_1(x) + c_2 f_2(x)$,
2. $(f_1 \cdot f_2)(x) = f_1(x) \cdot f_2(x)$,
3. $\max\{f_1, f_2\}(x) = \max\{f_1(x), f_2(x)\}$.

Dizemos que uma função $s : M \to \mathbb{R}$ é simples se existem constantes $\alpha_1, ..., \alpha_k \in \mathbb{R}$ e conjuntos $A_1, ..., A_k \in \mathcal{B}$ disjuntos dois-a-dois tais que

$$s = \sum_{j=1}^{k} \alpha_j \chi_{A_k},$$

onde χ_A é a função característica do conjunto A, isto é, $\chi_A(x)$ é igual a 1 se $x \in A$ e zero caso contrário.

Introduzimos agora a noção de integral. Para tal começamos por definir integral de uma função simples.

Definição 4. *Seja $s = \sum_{j=1}^{k} \alpha_j \chi_{A_k}$ função simples. Definimos a integral de s em relação a μ por:*

$$\int s d\mu := \sum_{j=1}^{k} \alpha_k \mu(A_k)$$

É importante verificar que esta definição é coerente. Isto é, se duas combinações lineares de funções características definem uma mesma função simples, os valores das integrais obtidos a partir das duas combinações devem coincidir, o que de fato ocorre e pode ser facilmente verificado. O próximo passo é definir integral de uma função mensurável qualquer. Para isso, trataremos primeiro do caso da função ser não-negativa. Antes, porém, vejamos o seguinte resultado, que nos diz que qualquer função mensurável é o limite de uma sequência de funções simples mensuráveis:

Teorema 5. *Seja $f : M \to [-\infty, \infty]$ uma função mensurável. Então existe uma sequência $s_1,\ s_2, ...$ de funções simples mensuráveis tais que*

$$\lim_{k\to\infty} s_k(x) = f(x),$$

para todo $x \in M$. Se $f \geq 0$ então a sequência pode ser escolhida de modo que $0 \leq s_1 \leq s_2 \leq ....$

Demonstração. Considere o caso em que $f \geq 0$. Para $k = 0, 1, 2, ...$ e $0 \leq n \leq 2^{2k} - 1$, seja

$$E_k^n = f^{-1}((n2^{-k}, (n+1)2^{-k}]) \text{ e } F_k = f^{-1}((2^k, \infty]),$$

e defina

$$s_k = \sum_{n=0}^{2^{2k}-1} n2^{-k} \chi_{E_k^n} + 2^k \chi_{F_k}.$$

Desta maneira $s_n \leq s_{n+1}$ para todo n, e $0 \leq f - s_n \leq 2^{-n}$ no conjunto de pontos onde $f(x) \leq 2^n$. Portanto $s_n \to f$.

Para $f \leq 0$ a demonstração é análoga e escrevendo $f = f^+ - f^-$ onde $f^+(x) = \max\{f(x), 0\}$ e $f^-(x) = \max\{-f(x), 0\}$ são não-negativas, mais ainda, f^+ e f^- são mensuráveis se e só se, f é mensurável. Logo existe (s_k^+) tendendo a f^+ e (s_k^-) tendendo a f^-, portanto tomando a sequência (s_k), onde $s_k = s_k^+ - s_k^-$, temos que s_k converge para f. □

Definição 6. *Seja $f : M \to [0, \infty]$ uma função mensurável não-negativa. Então, definimos*

$$\int f d\mu = \lim_{n\to\infty} \int s_n d\mu,$$

onde $s_1 \leq s_2 \leq ...$ é uma sequência de funções simples crescentes convergindo para f.

É fácil verificar que o valor da integral não depende da escolha da sequência de funções simples, e portanto esta definição é coerente. Mais geral, dada uma função $f:\ M \to \mathbb{R}$ como podemos escrever $f = f^+ - f^-$, temos o seguinte:

Definição 7. *Seja $f:\ M \to [0,\infty]$ uma função mensurável. Então*

$$\int f d\mu = \int f^+ d\mu - \int f^- d\mu,$$

desde que alguma das integrais do lado direito seja finita.

Definição 8. *Dizemos que uma função é integrável se for mensurável e tiver integral finita. Denotamos o conjunto das funções integráveis por $L_1(M,\mathcal{B},\mu)$ ou, mais simplesmente, por $L_1(M,\mu)$.*

Dada uma função mensurável $f:\ M \to \mathbb{R}$ e um conjunto mensurável E definimos a integral de f sobre E por

$$\int_E f d\mu = \int f \chi_E d\mu,$$

onde χ_E é a função característica do conjunto E.

3.4 Medidas invariantes

Sempre consideraremos medidas μ definida na σ-álgebra de Borel do espaço M. Dizemos que μ é uma probabilidade se $\mu(M) = 1$. Na maior parte dos casos trataremos com medidas finitas, isto é $\mu(M) < \infty$. Neste caso sempre podemos transformar μ numa probabilidade ν: para isso basta definir

$$\nu(E) = \frac{\mu(E)}{\mu(M)} \text{ para cada conjunto mensurável } E \subset M.$$

Uma medida μ diz-se invariante pela transformção $f : M \to M$ mensurável, se

$$\mu(E) = \mu(f^{-1}(E)) \text{ para todo conjunto mensurável } E \subset M. \tag{1}$$

A grosso modo, isto significa que a probabilidade de um ponto estar num dado conjunto e a probabilidade de que a sua imagem esteja nesse conjunto são iguais. Note que a definição (1) faz sentido, uma vez que a pré-imagem de um conjunto mensurável por uma transformção mensurável ainda é um conjunto mensurável. No caso de fluxos, substituímos (1) por

$$\mu(E) = \mu(f^{-t}(E)) \text{ para todo mensurável } E \subset M \text{ e todo } t \in \mathbb{R}. \tag{2}$$

Mas porque medidas invariantes?

Como em todo ramo da Matemática, parte importante da motivação é intrínseca e estética: estas estruturas matemáticas têm propriedades profundas e surpreendentes que conduzem à demonstração de belíssimos teoremas. Igualmente fascinante, idéias e resultados da Teoria Ergódica se aplicam em outras áreas da Matemática que a priori nada têm de probabilístico, por exemplo a Combinatória e a Teoria dos Números.

Outra razão é que muitos fenômenos importantes na Natureza e nas ciências experimentais são modelados por sistemas dinâmicos que deixam invariante alguma medida interessante.

Ainda outra motivação fundamental para que nos interessemos por medidas invariantes é que o seu estudo pode conduzir a informação importante sobre o comportamento dinâmico do sistema, que dificilmente poderia ser obtida de outro modo. O Teorema de Recorrência de Poincaré e o Teorema Ergódico de Birkhoff são excelentes ilustrações do que acabamos de dizer.

Exibiremos no decorrer deste trabalho, a medida de Lebesgue, onde os conjuntos mensuráveis pertencem à $\sigma-$álgebra de Borel gerada por intervalos meio-abertos da reta, ou seja, da forma $[a, b)$ ou $(a, b]$. Quando for o caso, substituiremos a notação μ por m.

4 DESENVOLVIMENTO

Definição 9. *Um ponto $x \in M$ diz-se recorrente se a sua trajetória pelo sistema dinâmico $f : M \to M$ volta arbitrariamente perto de x quando o tempo vai para infinito.*

A dinâmica no conjunto dos pontos não-recorrentes é, em certo sentido, sempre a mesma, independentemente do sistema dinâmico. Por isso, é fundamental compreender o conjunto dos pontos recorrentes, já que ele contém toda a dinâmica interessante do sistema.

4.1 Teorema de Recorrência de Poincaré (versão mensurável)

Teorema 10. *Seja $f : M \to M$ uma transformação mensurável e μ uma medida invariante finita. Seja $E \subset M$ qualquer conjunto mensurável com $\mu(E) > 0$. Então, $\mu-$quase todo ponto $x \in E$ tem algum iterado $f^n(x)$, $n \geq 1$, que também está em E.*

Em outras palavras, o teorema afirma que quase todo ponto de E regressa a E no futuro.

Demonstração. Seja E^0 o conjunto dos pontos $x \in E$ que nunca regressam a E. O nosso objetivo é provar que E^0 tem medida nula. Para isso, começamos por afirmar que as suas pré-imagens $f^{-n}(E^0)$ são disjuntas duas-a-duas. De fato, suponhamos que existem $m > n \geq 1$ tais que $f^{-m}(E^0)$ intersecta $f^{-n}(E^0)$. Seja x um ponto na interseção e seja $y = f^n(x)$. Então $y \in E^0$ e $f^{m-n}(y) = f^m(x) \in E^0$, que está contido em E. Isto quer dizer que y volta pelo menos uma vez a E, o que contradiz a definição de E^0. Esta contradição, prova que as pré-imagens são disjuntas duas-a-duas, como afirmamos.

Isto implica que

$$\mu(\bigcup_{n=0}^{\infty} f^{-n}(E^0)) = \sum_{n=0}^{\infty} \mu(f^{-n}(E^0)) = \sum_{n=0}^{\infty} \mu(E^0).$$

Na última igualdade usamos a hipótese de que μ é invariante, que implica que $\mu(f^{-n}(E^0)) = \mu(E^0)$ para todo $n \geq 1$. Como supomos que a medida é finita, a expressão do lado esquerdo é finita. Por outro lado, à direita temos uma soma de infinitos termos, todos iguais. O único jeito desta soma ser finita é que as parcelas sejam nulas. Portanto, devemos ter $\mu(E^0) = 0$, tal como foi afirmado. $\square$

Corolário 11. *Nas condições do Teorema 10, para $\mu-$quase todo ponto $x \in E$ existem infinitos valores de $n \geq 1$ tais que $f^n(x)$ está em E.*

Demonstração. Para cada $k \geq 1$ vamos representar por E^k o conjunto dos pontos $x \in E$ que regressam a E exatamente k vezes: existem exatamente k valores de $n \geq 1$ tais que $f^n(x) \in E$. Observe que o conjunto dos pontos que regressam a E apenas um número finito de vezes é precisamente

$$\bigcup_{k=1}^{\infty} E^k.$$

Portanto, para provar o corolário, basta mostrar que $\mu(E^k) = 0$ para todo $k \geq 1$. Suponhamos que $\mu(E^k) > 0$ para algum $k \geq 1$. Então, aplicando o Teorema 10 com este E^k no lugar de E, obtemos que quase todo ponto $x \in E^k$ tem algum iterado $f^n(x)$ que está em E^k. Fixemos um tal x e denotemos $y = f^n(x)$. Por definição, y tem exatamente k iterados futuros que estão em E. Como y é um iterado de x, isso implica que x tem $k+1$ iterados futuros em E. Mas isso contradiz o fato de que $x \in E^k$. Esta contradição prova que E^k tem medida nula, relativamente a μ, e portanto o corolário está demonstrado. $\square$

4.2 Teorema de Recorrência de Poincaré (versão topológica)

Definição 12. *Dizemos que um ponto $x \in M$ é recorrente para uma transformação $f: M \to M$ se, para toda vizinhança U de x, existe algum iterado $f^n(x)$ que está em U.*

A definição para fluxos é análoga, apenas nesse caso o tempo n é um número real.

Na formulação topológica do teorema de recorrência supomos que o espaço M admite uma base enumerável de abertos, ou seja, uma família enumerável $\{U_k;\ k \in \mathbb{N}\}$ de abertos tal que todo aberto de M pode ser escrito como união de elementos U_k dessa família.

Teorema 13. *Suponhamos que M admite uma base enumerável de abertos. Seja $f: M \to M$ uma transformação mensurável e μ uma medida invariante finita. Então, $\mu-$quase todo ponto $x \in M$ é recorrente para f.*

Demonstração. Para cada k representamos por U_k^0 o conjunto dos pontos $x \in U^k$ que nunca regressam a U_k. De acordo com o Teorema 10, todo U_k^0 tem medida nula. Consequentemente, a união enumerável

$$\tilde{U} = \bigcup_{k \in \mathbb{N}} U_k^0$$

tem medida nula. Portanto, para demonstrar o teorema será suficiente mostrarmos que todo ponto x que não está em $\tilde{U}$ é recorrente.

Seja $x \in M \backslash \tilde{U}$ e seja U uma vizinhança qualquer de x. A definição de base de abertos implica que existe algum $k \in \mathbb{N}$ tal que $x \in U_k$ e $U_k \subset U$. Como x não está em $\tilde{U}$ também $x \notin U_k^0$. Em outras palavras, x tem algum iterado $f^n(x)$, $n \geq 1$ que está em U_k. Em particular, $f^n(x)$ também está em U. Como a vizinhança U é arbitrária, isto prova que x é um ponto recorrente, como havíamos afirmado. □

4.3 Recorrência para medidas infinitas

As conclusões dos Teoremas 10 e 13 não são verdadeiras, em geral, se omitirmos a hipótese de que a medida μ é finita. O exemplo mais simples é o seguinte:

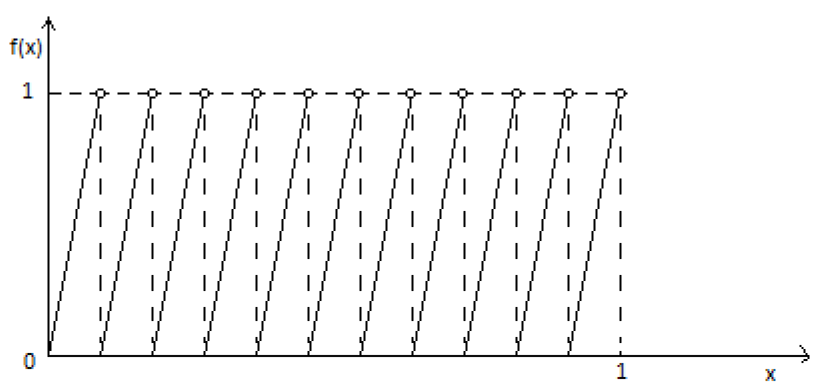
Exemplo. Seja $f: \mathbb{R} \to \mathbb{R}$ a translação de 1 unidade, isto é, $f(x) = x + 1$ para todo $x \in \mathbb{R}$. É fácil verificar que f deixa invariante a medida de Lebesgue em $\mathbb{R}$ (que é infinita). Por outro lado nenhum ponto é recorrente para f.

4.4 Um exemplo de medida invariante (Expansão Decimal)

Agora veremos um exemplo de medida invariante. Seja

$$\begin{aligned} f:\ [0,1] &\to [0,1] \\ x &\mapsto 10x - [10x] \end{aligned}$$

onde $[10x]$ representa o maior inteiro menor ou igual a $10x$. Em outras palavras, f associa a cada $x \in [0,1]$ a parte fracionária de $10x$. O gráfico da transformação f está descrito na seguinte figura:



Afirmamos que a medida de Lebesgue m no intervalo é invariante pela transformação f, isto é, satisfaz a condição (1). Comecemos por supor que E é um intervalo. Então, como ilustrado na figura acima, a pré-imagem $f^{-1}(E)$ consiste de dez intervalos, cada um deles dez vezes mais curto do que E. Logo $m(f^{-1}(E)) = m(E)$, e portanto (1) é satisfeita no caso de intervalos. Por outro lado, a família dos intervalos gera a $\sigma-$álgebra de Borel de $[0,1]$. Portanto, para concluir a demonstração basta usar o seguinte fato geral:

Lema 14. *Seja $f : M \to M$ uma transformação mensurável e μ uma medida finita em M. Suponha que existe uma sub-álgebra geradora $\mathcal{I}$ da $\sigma-$álgebra de M tal que $\mu(E) = \mu(f^{-1}(E))$ para todo $E \in \mathcal{I}$. Então o mesmo vale para todo conjunto mensurável E, isto é, a medida μ é invariante por f.*

Demonstração. Vejamos primeiramente que a família de todos os conjuntos E tais que $\mu(E) = \mu(f^{-1}(E))$ é uma $\sigma-$álgebra. Desde que f é bem-definida, então suponha $\{A_i\}_{i\in\mathbb{N}} \in \mathcal{I}$ família de conjuntos disjuntos, logo os $f^{-1}(A_i)$ também são conjuntos disjuntos. Sabemos que

$$\sum_{i=1}^{\infty} \mu(A_i) = \sum_{i=1}^{\infty} \mu(f^{-1}(A_i)).$$

Mas desde que os $A_i's$ são disjuntos e os $f^{-1}(A_i)$ também são, mais ainda $f^{-1}(\cup A_i) = \cup f^{-1}(A_i)$, logo:

$$\mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i) = \sum_{i=1}^{\infty} \mu(f^{-1}(A_i)) = \mu(\bigcup_{i=1}^{\infty} f^{-1}(A_i)) = \mu(f^{-1}(\bigcup_{i=1}^{\infty} A_i)).$$

Segue que $\bigcup\limits_{i=1}^{\infty} A_i \in \mathcal{I}$. Agora note que se não tivessemos os $A_i's$ disjuntos dois-a-dois, podemos encontrar B_i^s tal que $\cup_{i\in\mathbb{N}} A_i = \cup_{j\in\mathbb{N}} B_i$, onde os $B_i's$ são disjuntos dois-a-dois. Basta pegar $B_1 = A_1,\ B_2 = A_2 \backslash A_1,\ B_3 = A_3 \backslash (A_1 \cup A_2)$, e assim sucessivamente. Portanto temos que $\mathcal{I}$ é uma $\sigma-$álgebra geradora da $\sigma-$álgebra de M, logo para todo E desta $\sigma-$álgebra vale que $\mu(E) = \mu(f^{-1}(E))$. ☐

Agora, vamos explicar como, a partir do fato de que a medida de Lebesgue é invariante pela transformação f, podemos obter conclusões interessantes e não-triviais usando o Teorema de Recorrência de Poincaré (Teorema 10).

Note que f nada mais é do que uma translação das casas decimais de $x \in [0,1]$ para a esquerda. Isto é, considerando $x = 0, a_0a_1a_2...$ com $a_i \in \{0,1,2,3,4,5,6,7,8,9\}$ temos que $f(x) = 0, a_1a_2a_3....$ Com isso, fica muito fácil escrever a expressão do iterado $n-$ésimo, para qualquer $n \geq 1$:

$$f^n(x) = 0, a_n a_{n+1} a_{n+2}...$$

Agora, seja E o subconjunto dos $x \in [0,1]$ cuja expansão decimal começa com o dígito 7, ou seja, tais que $a_0 = 7$. De acordo com o Corolário 11, quase todo elemento de E tem infinitos iterados que também estão em E. Isto quer dizer que existem infinitos valores de n tais que $a_n = 7$. Portanto, provamos que quase todo número x cuja expansão decimal começa por 7 tem infinitos dígitos iguais a 7.

Claro que no lugar de 7 podemos considerar qualquer outro dígito. Podemos ainda considerar blocos de dígitos mais complicados como, por exemplo, o Cpf de uma pessoa aparece infinitas vezes em quase todo número $x \in [0,1]$ que começa com tal Cpf na expansão decimal.

5 TEOREMA DE REPRESENTAÇÃO DE RIESZ

Nesta seção temos por objetivo enunciar e demonstrar o Teorema de Representação de Riesz, um resultado muito importante da Análise Funcional e Teoria da Medida. Mas antes de enunciá-lo precisaremos de alguns resultados preliminares.

Teorema 15. *Suponha M um espaço Hausdorff, $K \subset M$, K compacto e $p \in K^c$. Então existem conjuntos abertos U e W tal que $p \in U$, $K \subset W$ e $U \cap W = \emptyset$.*

Demonstração. Dado $q \in K$, o axioma da separação de Hausdorff implica na existência de conjuntos abertos disjuntos U_q e V_q, tais que $p \in U_q$ e $q \in V_q$. Como K é compacto, existem pontos $q_1, ..., q_n \in K$ tais que

$$K \subset V_{q_1} \cup \cdots \cup V_{q_n}.$$

Portanto, a demonstração termina tomando:

$$U = U_{q_1} \cap \cdots \cap U_{q_n} \quad \text{e} \quad W = V_{q_1} \cup \cdots \cup V_{q_n}.$$

□

Teorema 16. *Se $\{K_\alpha\}$ é uma coleção de subconjuntos compactos de um espaço Hausdorff e se $\cap_\alpha K_\alpha = \emptyset$, então alguma subcoleção finita de $\{K_\alpha\}$ também tem interseção vazia.*

Demonstração. Pegue $V_\alpha = K_\alpha^c$. Fixe um membro K_1 de $\{K_\alpha\}$. Já que nenhum ponto de K_1 pertence a todo K_α, $\{V_\alpha\}$ é uma cobertura aberta de K_1. Portanto $K_1 \subset V_{\alpha_1} \cup \cdots \cup V_{\alpha_n}$ para alguma coleção finita $\{V_{\alpha_i}\}$. Isto implica que

$$K_1 \cap K_{\alpha_1} \cap \cdots \cap K_{\alpha_n} = \emptyset.$$

□

Teorema 17. *Suponha U aberto em um espaço localmente compacto Hausdorff M, $K \subset U$, e K é compacto. Então existe um conjunto aberto V com fecho compacto tal que*

$$K \subset V \subset \overline{V} \subset U.$$

Demonstração. Desde que todo ponto de K possui uma vizinhança com fecho compacto, e desde que K é coberto pela união finita de algumas dessas vizinhanças, então K está contido em um conjunto G com fecho compacto. Se $U = M$, pegue $V = G$.

Por outro lado, seja C o complementar de U. Pelo Teorema anterior, para cada ponto $p \in C$ existe um conjunto aberto correspondente W_p tal que $K \subset W_p$ e $p \notin \overline{W}_p$. Portanto $\{C \cap \overline{G} \cap \overline{W}_p\}$

é uma coleção de conjuntos compactos com interseção vazia. Pelo Teorema acima existem pontos $p_1, ..., p_n \in C$ tais que

$$C \cap \overline{G} \cap \overline{W}_{p_1} \cap \cdots \overline{W}_{p_n} = \emptyset.$$

O conjunto

$$V = \overline{G} \cap W_{p_1} \cap \cdots \cap W_{p_n}$$

então possui as propriedades necessárias, pois

$$\overline{V} \subset \overline{G} \cap \overline{W}_{p_1} \cap \cdots \cap \overline{W}_{p_n}.$$

□

Definição 18. *O suporte de uma função f definida em um espaço topológico M é o fecho do conjunto*

$$\{x;\ f(x) \neq 0\}.$$

A coleção de todas a funções contínuas em M cujo suporte é compacto é denotado por $C_c(M)$.

Notação: A notação

$$K \prec f$$

significa que K é um subconjunto compacto de M, que $f \in C_c(M)$, que $0 \leq f(x) \leq 1$ para todo $x \in M$ e que $f(x) = 1$ para todo $x \in K$. A notação

$$f \prec V$$

significa que V é aberto, $f \in C_c(M)$, $0 \leq f \leq 1$ e o suporte de f está em V.

Lema 19. *(Lema de Urysohn) Suponha M um espaço localmente compacto e Hausdorff, V um aberto em M, $K \subset V$, K compacto. Então existe $f \in C_c(M)$, tal que*

$$K \prec f \prec V. \tag{3}$$

Em termos da função característica, a conclusão acima assegura a existência de uma função contínua que satisfaz $\chi_K \leq f \leq \chi_V$. Note que é fácil encontrar funções semicontínuas que satisfazem isto, por exemplo, χ_K e χ_V.

Demonstração. Pegue $r_1 = 0,\ r_2 = 1$ e seja $r_3,\ r_4,\ r_5,\ ...$ uma enumeração de raízes em $(0, 1)$. Pelo Teorema 17, nós podemos achar conjuntos abertos V_0 e V_1 tais que $\overline{V}_0$ é compacto e

$$K \subset V_1 \subset \overline{V}_1 \subset V_0 \subset \overline{V}_0 \subset V. \tag{4}$$

Suponha $n \geq 2$ e V_{r_1}, ..., V_{r_n} sejam escolhidos de maneira que $r_i < r_j$ implique que $\overline{V}_{r_j} \subset V_{r_i}$. Então um dos números r_1, ..., r_n, digamos r_i, será o maior deles que é menor que r_{n+1}, e outro, digamos r_j, será o menor deles que é maior que r_{n+1}. Usando o Teorema 17 novamente, nós podemos encontrar $V_{r_{n+1}}$ tal que

$$\overline{V}_{r_j} \subset V_{r_{n+1}} \subset \overline{V}_{r_{n+1}} \subset V_{r_i}.$$

Continuando, nós obtemos uma coleção $\{V_r\}$ de conjuntos abertos, um para cada $r \in [0, 1]$, com a seguinte propriedade: $K \subset V_1$, $\overline{V}_0 \subset V$, cada $\overline{V}_r$ é compacto, e

$$s > r \Rightarrow \overline{V}_s \subset V_r. \tag{5}$$

Defina

$$f_r(x) = \begin{cases} r & \text{se } \ x \in \overline{V}_r \\ 0 & \text{caso contrário} \end{cases} \qquad \text{e} \qquad g_s(x) = \begin{cases} 1 & \text{se } \ x \in \overline{V}_s \\ s & \text{caso contrário} \end{cases}, \tag{6}$$

e ainda,

$$f = \sup_r f_r \qquad g = \inf_s g_s \tag{7}$$

É claro que $0 \leq f \leq 1$, sendo $f(x) = 1$ se $x \in K$, e que f tem o suporte em $\overline{V}_0$. A prova será completada mostrando que $f = g$.

A inequação $f_r(x) > g_s(x)$ é possível apenas se $r > s$, $x \in V_r$, e $x \notin \overline{V}_s$. Mas $r > s$ implica $V_r \subset V_s$. Portanto $f_r \leq g_s$ para todo r e s, então $f \leq g$.

Suponha $f(x) < g(x)$ para algum x. Então existem racionais r e s tais que $f(x) < r < s < g(x)$. Já que $f(x) < r$, nós temos $x \notin V_r$; já que $g(x) > s$, nós temos $x \in \overline{V}_s$. Por (5) vemos que isto é uma contradição. Logo $f = g$. □

Teorema 20. *Suponha que $V_1, ..., V_n$ são subconjuntos de um espaço localmente compacto e Hausdorff M, K é compacto e*

$$K \subset V_1 \cup \cdots V_n.$$

Então existem funções $h_i \prec V_i$ $(i = 1, ..., n)$ tais que

$$h_1(x) + \cdots + h_n(x) = 1 \qquad (x \in K). \tag{8}$$

Por causa de (8), a coleção $\{h_1, ..., h_n\}$ é chamada de partição da unidade em K subordinada à cobertura $\{V_1, ..., V_n\}$.

Demonstração. Cada $x \in K$ possui um W_x com fecho compacto $\overline{W}_x \subset V_i$ para algum i. Existem pontos $x_1, ..., x_m$ tais que $W_{x_1} \cup \cdots \cup W_{x_m} \supset K$. Se $1 \leq i \leq n$, seja H_i a união dos $\overline{W}_{x_j}$ que está em V_i. Pelo lema de Urysohn, existem funções g_i tais que $H_i \prec g_i \prec V_i$. Defina

$$h_1 = g_1$$
$$h_2 = (1 - g_1)g_2$$
$$\dots$$
$$h_n = (1 - g_1)(1 - g_2) \cdots (1 - g_{n-1})g_n.$$

Então $h_i \prec V_i$. Por indução, temos que vale

$$h_1 + \cdots + h_n = 1 - (1 - g_1) \cdots (1 - g_n). \tag{9}$$

Desde que $K \subset H_1 \cup \cdots \cup H_n$, então pelo menos um $g_i(x) = 1$ em cada ponto de K. Portanto (9) mostra que (8) vale. □

Agora, finalmente temos ferramentas suficientes para demonstrar o nosso objetivo:

Teorema 21. *(Teorema de Representação de Riesz) Seja M um espaço localmente compacto e Hausdorff, e seja $\Lambda : C_c(M) \to \mathbb{R}$ um funcional linear positivo. Existe uma única medida boreliana μ em M que representa Λ, isto é*

$$\Lambda(f) = \int_M f d\mu \quad \textit{para toda} \quad f \in C_c(M).$$

Chamemos de $\mathcal{M}_F$ a classe dos $E \subset M$ que satisfazem duas condições: $\mu(E) < \infty$, e

$$\mu(E) = \sup\{\mu(K);\ K \subset E,\ K \text{ compacto}\}. \tag{10}$$

Para a demonstração do Teorema 21 precisaremos ver um outro resultado antes:

Lema 22. *Se K é compacto, então $K \in \mathcal{M}_F$, e*

$$\mu(K) = \inf\{\Lambda f;\ K \prec f\}. \tag{11}$$

Demonstração. Se $K \prec f$ e $0 < \alpha < 1$, seja $V_\alpha = \{x;\ f(x) > \alpha\}$. Então $K \subset V_\alpha$, e $\alpha g \leq f$ sempre que $g \prec V_\alpha$. Portanto

$$\mu(K) \leq \mu(V_\alpha) = \sup\{\Lambda g;\ g \prec V_\alpha\} \leq \alpha^{-1}\Lambda f.$$

Fazendo $\alpha \to 1$, concluimos que

$$\mu(K) \leq \Lambda f. \tag{12}$$

Então $\mu(K) < \infty$. Desde que K evidentemente satisfaça (10), $K \in \mathcal{M}_F$.

Se $\varepsilon > 0$, existe $V \supset K$ com $\mu(V) < \mu(K) + \varepsilon$. Pelo Lema de Urysohn, $K \prec f \prec V$ para algum f. Então

$$\Lambda(f) \leq \mu(V) < \mu(K) + \varepsilon$$

tendo isto e (12), temos então que (11) vale. □

Demonstração. (Teorema 21) Claramente, é suficiente provar que

$$\Lambda(f) \leq \int_M f d\mu \tag{13}$$

para todo função $f \in C_c(M)$. Já que, pela linearidade de Λ :

$$-\Lambda(f) = \Lambda(-f) \leq \int_M (-f) d\mu = -\int_M f d\mu,$$

e juntamente com (10) mostramos que a igualdade vale.

Seja K o suporte de $f \in C_c(M)$, e $[a, b]$ um intervalo que contém o domínio de f. Dado $\varepsilon > 0$, escolha y_i, para $i = 0, 1, ..., n$, tal que $y_i - y_{i-1} < \varepsilon$ e

$$y_0 < a < y_1 < \cdots < y_n = b. \tag{14}$$

Pegue

$$E_i = \{x;\ y_{i-1} < f(x) \leq y_i\} \subset K \qquad (i = 1, ..., n). \tag{15}$$

Desde que f é contínua, f é mensurável, e os conjuntos E_i são disjuntos cuja reunião é igual a K. Então existem conjuntos abertos $V_i \supset E_i$ tais que

$$\mu(V_i) < \mu(E_i) + \frac{\varepsilon}{n} \qquad (i = 1, ..., n) \tag{16}$$

e ainda vale $f(x) < y_i + \varepsilon$ para todo $x \in V_i$. Pelo Teorema 20, existem funções $h_i \prec V_i$ tais que $\sum h_i = 1$ em K. Portanto $f = \sum h_i f$, e pelo Lema 22, temos que

$$\mu(K) \leq \Lambda(\sum h_i) = \sum \Lambda h_i.$$

Desde que $h_i f \leq (y_i + \varepsilon) h_i$, e desde que $y_i - \varepsilon < f(x)$ em E_i, nós temos

$$
\begin{aligned}
\Lambda(f) &= \sum_{i=1}^{n} \Lambda(h_i f) \leq \sum_{i=1}^{n} (y_i + \varepsilon)\Lambda(h_i) \\
&= \sum_{i=1}^{n} (|a| + y_i + \varepsilon)\Lambda(h_i) - |a| \sum_{i=1} \Lambda(h_i) \\
&\leq \sum_{i=1}^{n} (|a| + y_i + \varepsilon)[\mu(E_i) + \varepsilon/n] - |a|\mu(K) \\
&= \sum_{i=1}^{n} (y_i - \varepsilon)\mu(E_i) + 2\varepsilon\mu(K) + \frac{\varepsilon}{n} \sum_{i=1}^{n} (|a| + y_i + \varepsilon) \\
&\leq \int_M f d\mu + \varepsilon[2\mu(K) + |a| + b + \varepsilon].
\end{aligned}
$$

Como ε é arbitrário, (13) está demonstrado, e a prova do Teorema 21 está completa. □

Além da importância do Teorema de Representação de Riesz citada no início deste capítulo, tal teorema também nos garante a existência de medidas invariantes para funções, como veremos a seguir.

6 EXISTÊNCIA DE MEDIDAS INVARIANTES

Teorema 23. *(Teorema de Existência) Seja $f : M \to M$ uma transformação contínua num espaço métrico compacto. Então existe pelo menos uma probabilidade invariante por f. O mesmo resultado vale para fluxos.*

Note que nenhuma das hipóteses do Teorema, continuidade ou compacidade, pode ser ignorada. Considere $f : (0,1] \to (0,1]$ dada por $f(x) = x/2$. Suponha que f admite alguma probabilidade invariante. Pelo Teorema 10, relativamente a essa probabilidade quase todo ponto de $(0,1]$ é recorrente. Mas é imediato que não existe nenhum ponto recorrente: a órbita de qualquer $x \in (0,1]$ converge para zero e, em particular, não acumula no ponto inicial x. Isto mostra que f é um exemplo de transformação contínua num espaço não compacto que não admite nenhuma medida probabilidade invariante.

Modificando um pouco o exemplo anterior, podemos mostrar que o mesmo fenômeno pode ocorrer em espaços compactos, se a transformação não é contínua. Considere $f : [0,1] \to [0,1]$ dada por $f(x) = x/2$ se $x \neq 0$ e $f(0) = 1$. Pela mesma razão que antes, nenhum ponto $x \in (0,1]$ é recorrente. Portanto, se existe alguma probabilidade invariante μ ela tem dar peso total ao único ponto recorrente que é $x = 0$. Em outras palavras, μ precisa ser a medida de Dirac δ_0 suportada em zero, que é definida por

$$\delta_0(E) = 1 \text{ se } 0 \in E \quad e \quad \delta_0(E) = 0 \; se \; 0 \notin E.$$

Mas a medida δ_0 não é invariante por f: tomando $E = \{0\}$ temos que E tem medida 1 mas a sua pré-imagem $f^{-1}(E)$ é o conjunto vazio, que tem medida nula. Portanto, esta transformação também não tem nenhuma probabilidade invariante.

6.1 Topologia fraca*

Nesta seção vamos introduzir uma topologia importante no conjunto $\mathcal{M}_1(M)$ das probabilidades borelianas do espaço M, chamada topologia fraca* , que será usada para provar o Teorema 23. A idéia da definição é a seguinte: duas medidas estão próximas se dão integrais próximas para muitas funções contínuas. Mais precisamente, dada uma medida $\mu \in \mathcal{M}_1(M)$, um conjunto finito $F = \{\phi_1, ..., \phi_N\}$ de funções contínuas $\phi_j : M \to \mathbb{R}$, e um número $\varepsilon > 0$, definimos

$$V(\mu, F, \varepsilon) = \left\{\eta \in \mathcal{M}_1(M);\ \left|\int \phi_j\ d\eta - \int \phi_j\ d\mu\right| < \varepsilon \text{ para todo } \phi_j \in F\right\}.$$

Então a topologia fraca* é definida pelos conjuntos $V(\mu, F, \varepsilon)$, com F e ε variável, formando uma base de vizinhanças da medida μ. O seguinte lema ajuda a compreender o significado desta topologia:

Lema 24. *Uma sequência $(\mu_n)_{n\in\mathbb{N}}$ em $\mathcal{M}_1(M)$ converge para uma medida $\mu \in \mathcal{M}_1(M)$ na topologia fraca* se, e somente se*

$$\int \phi\ d\mu_n \to \int \phi\ d\mu \textit{ para toda função contínua } \phi : M \to \mathbb{R}.$$

Demonstração. Para provar a parte "somente se", considere qualquer função contínua ϕ e tome o conjunto $F = \{\phi\}$. Como $\mu_n \to \mu$, temos que dado qualquer $\varepsilon > 0$ existe uma ordem a partir $\overline{n}$ da qual μ_n está na vizinhança $V(\mu, F, \varepsilon)$. Mas isto significa, precisamente, que

$$\left|\int \phi\ d\mu_n - \int \phi\ d\mu\right| < \varepsilon$$

para todo $n \geq \overline{n}$. Em outras palavras, a sequência $\int \phi\ d\mu_n$ converge para $\int \phi\ d\mu$. A recíproca afirma que se $\int \phi\ d\mu_n$ converge para $\int \phi\ d\mu$, para toda função contínua, então dado qualquer F e ε existe uma ordem a partir da qual $\mu_n \in V(\mu, F, \varepsilon)$. Para vermos isso, escrevemos $F = \{\phi_1, ..., \phi_N\}$. A hipótese garante que para cada $1 \leq j \leq N$ existe $\overline{n}_j$ tal que

$$\left|\int \phi_j\ d\mu_n - \int \phi\ d\mu\right| < \varepsilon \text{ para todo } n \geq \overline{n}_j.$$

Tomando $\overline{n} = \max\{\overline{n}_1, ..., \overline{n}_N\}$, temos $\mu_n \in V(\mu, F, \varepsilon)$ para $n \geq \overline{n}$. $\square$

As principais propriedades desta topologia estão dadas no seguinte teorema:

Teorema 25. *$\mathcal{M}_1(M)$ munido da topologia fraca* é metrizável e compacto.*

Vamos começar por demonstrar a metrizabilidade, isto é, que existe uma métrica d que gera a topologia fraca* em $\mathcal{M}_1(M)$. Para isso vamos usar a Proposição 26. Como é usual, denotamos por $C^0(M)$ o espaço das funções contínuas $\phi: M \to \mathbb{R}$, munido da norma da convergência uniforme:

$$||\phi_1 - \phi_2|| = \sup\{|\phi_1(x) - \phi_2(x)|;\ x \in M\}.$$

É claro que $C_c(M) \subset C^0(M)$ e que essas duas classes coincidem se M é compacto.

Proposição 26. *Se M é um espaço métrico compacto então $C^0(M)$ tem subconjuntos enumeráveis densos, ou seja é separável.*

A demonstração desse fato requer um pouco de domínio da Análise Funcional, e se encontra na referência [Rud87].

Agora podemos escolher um subconjunto enumerável $\mathcal{F} = \{\phi_n;\ n \in \mathbb{N}\}$ denso na bola unitária do espaço $C^0(M)$. Feito isso, definimos:

$$d(\mu_1, \mu_2) = \sum_{n=1}^{\infty} \frac{1}{2^n} \left| \int \phi_n \, d\mu_1 - \int \phi_n \, d\mu_2 \right|, \tag{17}$$

para qualquer par de medidas μ_1 e μ_2.

Proposição 27. *A expressão d está bem definida, é uma métrica, e gera a topologia fraca* em $\mathcal{M}_1(M)$.*

Demonstração. Como as funções ϕ estão na bola unitária de $C^0(M)$, ou seja, $\sup|\phi| \leq 1$, e as medidas μ_i são probabilidades, o termo geral da soma é limitado por $2 \cdot 2^{-n}$. Isto garante que a série em (17) converge. O único passo não trivial na prova de que d é uma métrica é mostrar que

$$d(\mu_1, \mu_2) = 0 \Rightarrow \mu_1 = \mu_2.$$

A hipótese $d(\mu_1, \mu_2) = 0$ significa que $\int \phi_j \mu_1 = \int \phi_j d\mu_2$ para toda $\phi_j \in \mathcal{F}$. Agora, dada qualquer ϕ na bola unitária de $C^0(M)$ podemos encontrar uma sequência de elementos de $\mathcal{F}$ convergindo uniformemente para ϕ. Como consequência, a igualdade continua valendo para ϕ:

$$\int \phi \, d\mu_1 = \int \phi \, d\mu_2 \tag{18}$$

para toda ϕ na bola unitária de $C^0(M)$. Como todo elemento de $C^0(M)$ tem algum múltiplo na bola unitária, isto implica que a igualdade (18) é verdadeira para toda função contínua ϕ. Isto quer dizer que $\mu_1 = \mu_2$, como pretendíamos mostrar.

Para provar que d gera a topologia, devemos mostrar que toda bola

$$B(\mu, \delta) = \{\eta \in \mathcal{M}_1(M);\ d(\mu, \eta) < \delta\}$$

contém alguma vizinhança $V(\mu, F, \epsilon)$ e reciprocamente. Dado $\delta > 0$ fixemos $N \geq 1$ suficientemente grande para que

$$\sum_{n=N}^{\infty} 2^{-n} < \frac{\delta}{2}$$

e consideremos $F = \{\phi_1, ..., \phi_N\}$ formado pelos primeiros N elementos do subconjunto enumerável denso. Além disso, consideremos $\epsilon = \delta/2$. Afirmamos que $V(\mu, F, \epsilon) \subset B(\mu, \delta)$. De fato,

$$\nu \in V(\mu, F, \epsilon) \Rightarrow \left| \int \phi_n \, d\mu - \phi_n \, d\nu \right| < \epsilon \qquad \text{para todo} \;\; 1 \leq n \leq N$$

$$\Rightarrow \sum_{n=1}^{\infty} 2^{-n} \left| \int \phi_n \, d\mu - \int \phi_n \, d\nu \right| < \sum_{n=1}^{N} 2^{-n}\epsilon + \sum_{n=N+1}^{\infty} 2 \cdot 2^{-n} < \delta,$$

o que prova a nossa afirmação.

Reciprocamente, dado $F = \{\psi_1, ..., \psi_N\}$ e $\epsilon > 0$, selecionemos elementos $\phi_{n_1}, ..., \phi_{n_N}$ distintos de $\mathcal{F}$ tais que

$$||\phi_{n_j} - \psi_j|| < \frac{\varepsilon}{4} \qquad \text{para todo} \;\; 1 \leq j \leq N.$$

Fixemos $\delta > 0$ suficientemente pequeno para que $2^{n_j}\delta < \varepsilon/4$ para todo $1 \leq j \leq N$.

Afirmamos que $B(\mu, \delta) \subset V(\mu, F, \varepsilon)$. De fato

$$\nu \in B(\mu, \delta) \Rightarrow \sum_{n=1}^{\infty} 2^{-n} \left| \int \phi_n \, d\mu - \int \phi_n \, d\nu \right| < \delta$$

$$\Rightarrow \left| \int \phi_{n_j} \, d\mu - \int \phi_{n_j} \, d\nu \right| < 2^{n_j}\delta \qquad \text{para todo} \;\; 1 \leq j \leq N$$

$$\Rightarrow \left| \int \psi_j \, d\mu - \int \psi_j \, d\nu \right| < 2^{n_j}\delta + \frac{\varepsilon}{2} < \varepsilon \qquad \text{para todo} \;\; 1 \leq n \leq N,$$

e isto finaliza nossa demonstração. □

Resta provar que $(\mathcal{M}_1, \text{fraca}^*)$ é um espaço compacto. Na demonstração vamos utilizar o Teorema de Representação 21 que, interpretado a grosso modo, diz que as integrais são os únicos operadores lineares positivos no espaço das funções contínuas.

Já sabemos que o espaço é metrizável, basta provar que toda sequência $(\mu_k)_{k\in\mathbb{N}}$ em $\mathcal{M}_1(M)$ admite alguma sub-sequência que é convergente na topologia fraca*.

Considere então $F = \{\phi_n; \; n \in \mathbb{N}\}$ um subconjunto enumerável denso na bola unitária de $C^0(M)$. Para cada $n \in \mathbb{N}$, a sequência de números reais $\int \phi_n d\mu_k$, $k \in \mathbb{N}$ é limitada por 1. Portanto, para cada $n \in \mathbb{N}$ existe uma sequência $(k_j^n)_{j\in\mathbb{N}}$ tal que

$$\int \phi_n d\mu_{k_j^n} \;\; \text{converge para algum número} \;\; \Lambda_n \in \mathbb{R} \;\; \text{quando} \;\; j \to \infty.$$

Além disso, cada sequência $(k_j^{n+1})_{n\in\mathbb{N}}$ pode ser escolhida como subsequência da anterior $(k_j^n)_{n\in\mathbb{N}}$ Definamos $l_j = k_j^j$ para cada $j \in \mathbb{N}$. Por construção, a menos de um número finito de termos, $(l_j)_{j\in\mathbb{N}}$ é uma subsequência de cada uma das $(k_j^n)_{j\in\mathbb{N}}$. Logo

$$\int \phi_n \, d\mu_{l_j}^n \to \Lambda_n \quad \text{para todo } \; n \in \mathbb{N}.$$

Daqui se deduz facilmente que

$$\Lambda\varphi = \lim_j \int \varphi \, d\mu_{l_j} \;\; \text{existe, para toda função } \; \varphi \in C^0(M). \tag{19}$$

De fato, suponha primeiro que φ está na bola unitária de $C^0(M)$. Dado qualquer $\varepsilon > 0$ podemos encontrar $\phi_n \in \mathcal{F}$ tal que $||\varphi - \phi_n|| < \varepsilon$. Então,

$$\left| \int \varphi \, d\mu_{l_j} - \int \phi_n \, d\mu_{l_j} \right| \leq \varepsilon$$

para todo j. Como $\int \phi_n \, d\mu_{l_j}$ converge (para Λ_n), segue que

$$\limsup_j \int \varphi \, d\mu_{l_j} - \liminf_j \int \varphi \, d\mu_{l_j} \leq 2\varepsilon.$$

Como ε é arbitrário, concluímos que $\lim_j \int \varphi d\mu_{l_j}$ existe. Isto prova (19) quando a função está na bola unitária. O caso geral reduz-se imediatamente a esse, substituindo φ por $\varphi/||\varphi||$. Assim, completamos a prova de (19).

Finalmente, é claro que o operador $\Lambda : \; C^0(M) \to \mathbb{R}$ definido por (19) é linear e positivo: $\Lambda(\varphi) \geq \min \varphi > 0$ para todo função $\varphi \in C^0(M)$ positiva em todo ponto. Além disso, $\Lambda(1) = 1$. Logo, pelo Teorema de Representação de Riesz, existe alguma probabilidade boreliana μ em M tal que $\Lambda(\varphi) = \int \varphi d\mu$ para toda função contínua φ. Agora a igualdade em (19) pode ser reescrita

$$\int \varphi \, d\mu = \lim_j \int \varphi \, d\mu_{l_j} \quad \text{para toda } \; \varphi \in C^0(M).$$

De acordo com o Lema 24, isto quer dizer que a subsequência $(\mu_{l_j})_{j\in\mathbb{N}}$ converge para μ na topologia fraca*. Isto completa a demonstração do Teorema 25.

□

6.2 Demonstração do Teorema de Existência

Comecemos por introduzir uma notação útil. Dado $f : \; M \to M$ e qualquer medida η em M denota-se por $f_*\eta$ e chama-se imagem de η por f a medida definida por

$$f_*\nu(E) = \nu(f^{-1}(E)) \quad \text{para cada conjunto mensurável } \; E \subset M.$$

Note que η é invariante por f se e somente se $f_*\eta = \eta$.

Lema 28. *A aplicação $f_* : \mathcal{M}_1(M) \to \mathcal{M}_1(M)$ é contínua relativamente à topologia fraca*.*

Demonstração. Para mostrarmos o lema acima, basta mostrar que se μ_n converge para μ na topologia fraca*, então para toda função contínua ϕ temos que

$$\lim_{n\to\infty} \int \phi \; df_*\mu_n = \int \phi \; df_*\mu.$$

De fato, seja η uma medida qualquer, afirmamos que

$$\int \phi \; df_*\eta = \int \phi \circ f \; d\eta.$$

Com efeito, podemos aproximar ϕ por uma sequência de funções simples ϕ_n com $||\phi_n|| \leq ||\phi||$. Observe que isso implica, em particular, que $||\phi_n \circ f|| \leq ||\phi \circ f||$. Observe que se χ_A é função característica, então

$$\int \chi_A \; df_*\eta = \eta(f^{-1}(A)) = \int \chi_A \circ f \; d\eta.$$

Por linearidade, a igualdade acima se estende para as funções simples ϕ_n. Para finalizar, temos que pelo Teorema da Convergência Dominada,

$$\int \phi \; df_*\eta = \lim_{n\to\infty} \int \phi_n \; df_*\eta = \lim_{n\to\infty} \int \phi_n \circ f \; d\eta = \int \phi \circ f \; d\eta$$

o que termina a prova da afirmação.

Para completar a prova do Lema, basta observar que a função $\phi \circ f$ também é contínua, uma vez que f é contínua. Assim,

$$\lim_{n\to\infty} \int \phi \; df_*\mu_n = \lim_{n\to\infty} \int \phi \circ f \; d\mu_n = \int \phi \circ f \; d\mu = \int \phi \; df_*\mu,$$

como queríamos demonstrar. □

Voltando a prova do Teorema de Existência, considere qualquer probabilidade ν em M: por exemplo, a medida de Dirac em um ponto qualquer. Agora tomemos a seguinte sequência de probabilidades

$$\mu_n = \frac{1}{n} \sum_{j=0}^{n-1} f_*^j \nu \tag{20}$$

onde $f_*^j\nu$ é a imagem de ν pelo iterado f^j . Pelo Teorema 25, esta sequência tem algum ponto de acumulação: existe alguma subsequência $(n_j)_{j\in\mathbb{N}}$ e alguma probabilidade $\mu \in \mathcal{M}_1(M)$ tais que

$$\mu = \lim_k \mu_{n_k} = \lim_k \frac{1}{n_k} \sum_{j=0}^{n_k-1} f_*^j \nu. \tag{21}$$

Agora é suficiente provar o seguinte:

Lema 29. *Todo ponto de acumulação de uma sequência $(\mu_n)_{n\in\mathbb{N}}$ é uma probabilidade invariante por f.*

Demonstração. A partir (21), e usando o Lema 28, obtemos que

$$f_*\mu = f_* \left(\lim_k \frac{1}{n_k} \sum_{j=0}^{n_k-1} f_*^j \nu \right) = \lim_k f_* \left(\frac{1}{n_k} \sum_{j=0}^{n_k-1} f_*^j \nu \right) = \lim_k \frac{1}{n_k} \sum_{j=0}^{n_k-1} f_*^{j+1} \nu.$$

A expressão do lado direito pode ser reescrita como:

$$\lim_k \left(\frac{1}{n_k} \sum_{j=0}^{n_k-1} f_*^j \nu - \nu + f_*^{n_k} \nu \right).$$

Afirmamos que $\lim_k \frac{1}{n_k} \nu = 0$ e $\lim_k \frac{1}{n_k} f_*^{n_k} \nu = 0$. A primeira afirmação é óbvia, e para a segunda basta observar que

$$\frac{1}{n_k} f_*^{n_k} \nu(E) = \frac{1}{n_k} \nu(f^{-n_k}(E)) \leq \frac{1}{n_k},$$

para todo conjunto mensurável $E \subset F$. Deste modo obtemos que

$$f_*\mu = \lim_k \frac{1}{n_k} \sum_{j=0}^{n_k-1} f_*^j \nu = \mu,$$

e portanto μ é invariante por f. Isto completa a demonstração do Teorema de Existência. □

Corolário 30. *(Teorema de Recorrência de Birkhoff) Se $f : M \to M$ é uma transformação contínua num espaço métrico compacto então f tem algum ponto recorrente.*

Demonstração. Pelo Teorema de Existência, existe alguma probabilidade $f-$invariante μ. Por outro lado, sabemos que todo espaço métrico compacto admite uma base enumerável de abertos. Portanto, podemos aplicar o Teorema de Recorrência de Poincaré, para concluir que $\mu-$quase todo ponto é recorrente. Em particular, o conjunto dos pontos recorrentes é não vazio, conforme foi afirmado. □

6.3 Teorema Ergódico de Birkhoff

Comecemos por explicar o que entendemos por tempo médio de permanência de uma órbita num conjunto. Dado M espaço métrico qualquer, chamemos por tempo médio de permanência de x em $E \subset M$ a função $\tau: M \to [0, \infty)$ dada por:

$$\tau(E, x) = \lim_{n \to \infty} \tau_n(E, x).$$

onde

$$\tau_n(E, x) = \frac{1}{n} \sum_{j=0}^{n-1} \chi_E(f^j(x)).$$

Note que se $\tau(E, x) > 0$ então temos infinitos iterados de x em E. Mas claro, precisamos mostrar que este limite existe.

Teorema 31. *(Teorema Ergódico de Birkhoff) Seja $f: M \to M$ uma transformação mensurável qualquer e μ uma probabilidade invariante por f. Dado qualquer conjunto $E \subset M$, o tempo médio de permanência $\tau(E, x)$ existe em $\mu-$quase todo todo ponto $x \in M$. Além disso,*

$$\int \tau(E, x) d\mu(x) = \mu(E).$$

Note ainda que,

$$\begin{aligned}
\tau(E, f(x)) &= \lim_{n \to \infty} \frac{1}{n} \sum_{j=1}^{n} \chi_E(f^j(x)) \\
&= \lim_{n \to \infty} \frac{1}{n} \sum_{j=0}^{n-1} \chi_E(f^j(x)) - \frac{1}{n}[\chi_E(x) - \chi_E(f^n(x))] \\
&= \tau(E, x) - \lim_{n \to \infty} \frac{1}{n}[\chi_E(x) - \chi_E(f^n(x))].
\end{aligned}$$

Como a função característica é limitada, o último limite é igual a zero e portanto

$$\tau(E, f(x)) = \tau(E, x).$$

O Teorema Ergódico de Birkhoff pode ser enunciado de modo um pouco mais geral:

Teorema 32. *Seja $f: M \to M$ uma transformação mensurável e μ uma probabilidade invariante por f. Dada qualquer função integrável $\varphi: M \to \mathbb{R}$ o limite*

$$\tilde{\varphi}(x) = \lim_{n \to \infty} \frac{1}{n} \sum_{j=0}^{n} \varphi(f^j(x))$$

existe em $\mu-$quase todo ponto $x \in M$. Além disso,

$$\int \tilde{\varphi}(x)d\mu(x) = \int \varphi(x)d\mu(x).$$

Note que o Teorema 31 é um caso particular em que $\varphi = \chi_E$. Veremos a seguir a demonstração deste caso particular, o caso geral pode ser encontrado na referência [Oli14].

Demonstração. (Teorema 31) A estratégia da prova é a seguinte. Seja $E \subset M$ um conjunto mensurável qualquer. Para cada $x \in M$, definimos

$$\overline{\tau}(E, x) = \limsup_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \chi_E(f^j(x)),$$

$$\underline{\tau}(E, x) = \liminf_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \chi_E(f^j(x)).$$

Analogamente como foi feito para τ temos, para todo $x \in M$ que

$$\overline{\tau}(E, f(x)) = \overline{\tau}(E, x) \text{ e } \underline{\tau}(E, f(x)) = \underline{\tau}(E, x). \tag{22}$$

O principal passo da demonstração consiste em mostrar que

$$\overline{\tau}(E, x) = \underline{\tau}(E, x) \quad \text{para } \ \mu-\text{quase todo ponto } x \in M. \tag{23}$$

É claro que $\overline{\tau}(E, x)$ é sempre maior ou igual que $\underline{\tau}(E, x)$. Portanto, para mostrar (23) será suficiente provar que

$$\int \overline{\tau}(E, x)d\mu \leq \mu(E) \leq \int \underline{\tau}(E, x)d\mu. \tag{24}$$

Vamos provar a primeira desigualdade em (24). A segunda segue de um argumento inteiramente análogo (a segunda desigualdade pode ser deduzida da primeira, aplicada ao complementar E^c, observando que $\mu(E) = 1 - \mu(E^c)$ e $\underline{\tau}(E, x) = 1 - \overline{\tau}(E^c, x)$.

Fixe $\varepsilon > 0$ qualquer. Por definição de $\limsup$, para cada $x \in M$ existem inteiros $t \geq 1$ tais que

$$\frac{1}{n}\#\{j \in \{0, 1, ..., t-1\}; f^j(x) \in E\} \geq \overline{\tau}(E, x) - \varepsilon.$$

Representaremos por $t(x)$ o menor inteiro com esta propriedade. Para tornar a demonstração mais transparente, consideraremos primeiro o caso particular em que a função $x \mapsto t(x)$ é limitada, isto é,

Caso particular: Existe $T \in \mathbb{N}$ tal que $t(x) \leq T$ para todo $x \in M$.

Dado qualquer $x \in M$, definimos uma sequência $x_0,\ x_1, ...,\ x_s$ de pontos em M e uma sequência $t_0,\ t_1, ...,\ t_s$ de números naturais, do seguinte modo:

1. Primeiramente, tomamos $x_0 = x$.
2. Supondo que x_i já foi definido, tomamos $t_i = t(x_i)$ e $x_{i+1} = f^{t_i}(x_i)$.
3. Terminamos quando encontramos x_s tal que $t_0 + t_1 + ... + t_s \geq n$.

Note que todo x_i é iterado do ponto x : de fato $x_i = f^{t_0+...+t_{i-1}}(x)$. Aplicando (22) concluímos que $\tau(E, x_i) = \tau(E, x)$ para todo i. A definição de $t(x_i)$ implica que, dos t_i primeiros iterados de x_i, pelo menos

$$t_i(\overline{\tau}(E, x_i) - \varepsilon) = t_i(\overline{\tau}(E, x) - \varepsilon) \tag{25}$$

estão em E. Isto vale para cada $i = 0, 1, ..., s-1$. Portanto, pelo menos

$$(t_0 + ... + t_{s-1})(\overline{\tau}(E, x) - \varepsilon)$$

dos n primeiros iterados de x, estão em E. Além disso, a última regra na definição das nossas sequências implica que

$$t_0 + t_1 + ... + t_{s-1} \geq n - t_s \geq n - T.$$

Deste modo, mostramos que pelo menos $(n - T)(\tau(E, x) - \varepsilon)$ dos n primeiros iterados de x estão em E. Em outras palavras,

$$\sum_{j=0}^{n-1} \chi_E(f^j(x)) \geq (n - T)(\overline{\tau}(E, x) - \varepsilon) \tag{26}$$

para todo $x \in M$ e todo $n \geq 1$. Integrando a relação (26), obtemos que

$$\sum_{j=0}^{n-1} \int \chi_E(f^j(x)) d\mu \geq (n - T) \int \overline{\tau}(E, x) d\mu - (n - T)\varepsilon.$$

Todas as parcelas no membro da esquerda são iguais a $\mu(E)$, uma vez que a probabilidade μ é invariante por f. Portanto, esta desigualdade pode ser escrita como

$$n\mu(E) \geq (n - T) \int \overline{\tau}(E, x) d\mu - (n - T)\varepsilon.$$

Dividindo os dois termos por n e fazendo n ir para infinito, concluímos que

$$\mu(E) \geq \int \overline{\tau}(E, x) d\mu - \varepsilon.$$

Como $\varepsilon > 0$ é qualquer, isto implica a primeira desigualdade em (24), o que termina a demonstração neste caso.

Caso geral: Vamos indicar as modificações que devem ser feitas relativamente ao caso particular.

Dado $\varepsilon > 0$, começamos por fixar $T \geq 1$ suficientemente grande, de modo que a medida do

$$B = \{y \in M; t(y) > T\}$$

seja menor do que ϵ. Em seguida, na definição das sequências substituímos a regra 2 por

2a. Se $t(x_i) \leq T$, tomamos $t_i = t(x_i)$ e $x_{i+1} = f^{t_i}(x_i)$.

2b. Se $t(x_i) > T$, tomamos $t_i = 1$ e $x_{i+1} = f(x_i)$.

As regras 1 e 3 permanecem inalteradas. A estimativa referente a (25) continua válida, para os valores de i aos quais se aplica a regra 2a, isto é:

$$\sum_{j=0}^{t_i-1} \chi_E(f^j(x_i)) \geq t_i(\overline{\tau}(E,x) - \varepsilon).$$

É claro que esta desigualdade implica a seguinte:

$$\sum_{j=0}^{t_i-1} \chi_E(f^j(x_i)) \geq t_i(\overline{\tau}(E,x) - \varepsilon) - \sum_{j=0}^{t_i-1} \chi_B(f^j(x_i)). \tag{27}$$

A vantagem é que (27) é válida também para os valores de i aos quais se aplica a regra 2b. De fato, nesse caso tem-se $t_i = 1$, o membro da esquerda é maior ou igual que zero e o membro da direita é menor que zero, uma vez que $\tau(E,x)$ é sempre menor ou igual que 1. Isso significa que, no lugar de (26), tem-se

$$\sum_{j=0}^{n-1} \chi_E(f^j(x)) \geq (n-T)(\overline{\tau}(E,x) - \varepsilon) - \sum_{j=0}^{n-1} \chi_B(f^j(x)).$$

Integrando, como fizemos anteriormente, obtemos

$$n\mu(E) \geq (n-T)\int \overline{\tau}(E,x)d\mu - (n-T)\varepsilon - n\mu(B).$$

Dividindo por n e fazendo $n \to \infty$, deduzimos que (lembre que $\mu(B) < \epsilon$)

$$\mu(E) \geq \int \overline{\tau}(E,x)d\mu - \varepsilon - \mu(B) \geq \int \overline{\tau}(E,x)d\mu - 2\varepsilon.$$

Como $\varepsilon > 0$ é arbitrário, segue que

$$\mu(E) \geq \int \overline{\tau}(E,x)d\mu.$$

Isto termina a demonstração. $\square$

Agora que conhecemos algumas das ferramentas da Teoria Ergódica, como o Teorema de Recorrência de Poincaré e o Teorema Ergódico de Birkhoff, já podemos introduzir tópicos mais avançados da Ergodicidade. Com o estudo de Sistemas Ergódicos podemos tirar conclusões mais interessantes do nosso problema de Expansão Decimal visto anteriormente.

7 ERGODICIDADE

Uma transformação $f: M \to M$ diz-se ergódica para uma probabilidade invariante μ (também dizemos que a medida μ é ergódica para f, ou que o sistema (f, μ) é ergódico) se as médias temporais dadas pelo Teorema Ergódico de Birkhoff coincidem em quase todo ponto com as respectivas médias espaciais:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \varphi(f^j(x)) = \int \varphi \, d\mu,$$

para toda função μ-integrável $\varphi: M \to \mathbb{R}$ e $\mu-$quase todo $x \in M$.

Na próxima proposição vamos reescrever esta condição de várias maneiras equivalentes, para ajudar a entender o seu significado. Um conjunto mensurável $A \subset M$ diz-se invariante se $f^{-1}(A) = A$. Uma função mensurável $\psi: M \to \mathbb{R}$ diz-se invariante se $\psi \circ f = \psi$.

Proposição 33. *Seja μ uma probabilidade invariante de uma transformação $f: M \to M$ mensurável. As seguintes condições são equivalentes:*

1. *O sistema (f, μ) é ergódico.*
2. *Para todo subconjunto invariante A tem-se $\mu(A) = 0$ ou $\mu(A) = 1$*
3. *Toda função invariante ψ é constante num conjunto de medida total.*

Demonstração. (1) $\Rightarrow$ (2). Por um lado, a hipótese (1) significa que

$$\tilde{\varphi}(x) = \int \varphi \, d\mu = \mu(A)$$

para quase todo $x \in M$. Por outro lado, como A é invariante, temos que $x \in A$ se e somente se $f(x) \in A$. Isto implica que $\varphi(f^j(x)) = \varphi(x)$ para todo $j \geq 0$ e para todo x. Portanto,

$$\tilde{\varphi}(x) = \varphi(x) = \chi_A(x)$$

para todo $x \in M$. Como a função característica só toma os valores 0 e 1, estas duas igualdades implicam que $\mu(A) = 0$ ou $\mu(A) = 1$, como é afirmado em (2).

(2) $\Rightarrow$ (3). Seja ψ uma função invariante qualquer. Então, a pré-imagem $\psi^{-1}(I)$ de qualquer intervalo $I \subset \mathbb{R}$ é um conjunto invariante. Portanto, pela hipótese (2), essa pré-imagem tem medida zero ou um. Como o intervalo I é qualquer, isto prova que ψ é constante num conjunto com probabilidade μ total.

(3) $\Rightarrow$ (1). Seja φ uma função integrável qualquer. Podemos verificar, para qualquer função integrável φ, que a média temporal $\tilde{\varphi}$ satisfaz $\varphi \circ f = \tilde{\varphi}$ em $\mu-$quase todo ponto. Logo a média temporal $\tilde{\varphi}$ é uma função invariante. Logo, pela hipótese (3), $\tilde{\varphi}$ é constante em quase todo ponto. Então, usando o teorema ergódico,

$$\tilde{\varphi}(x) = \int \tilde{\varphi}\, d\mu = \int \varphi\, d\mu$$

em quase todo ponto. Isto é, o sistema é ergódico. □

7.1 Exemplos e Aplicações

Nesta seção veremos exemplos de Sistemas Ergódicos. Pegue como exemplo inicial o problema de expansão decimal que vimos anteriormente.

Expansão decimal

Considere a transformação mensurável $f: [0,1] \to [0,1]$ dada por $f(x) := 10x - [10x]$, onde $[10x]$ é a parte inteira de $10x$. Seja m a medida de Lebesgue, e note que m é invariante por f. De fato, para qualquer $E \subset [0,1]$, $f^{-1}(E)$ é igual a união de dez conjuntos $E_1, ..., E_{10}$ cada um com medida igual a $\frac{m(E)}{10}$, pois cada um dos dez segmentos de reta que compõe o gráfico de f tem inclinação igual a 10. Segue, portanto, que $m(f^{-1}(E)) = m(E_1) + ... + m(E_{10}) = m(E)$. Note ainda que m é uma probabilidade, uma vez que $m([0,1]) = 1$.

Vamos verificar agora que vale o item (2) da Proposição 33 em nosso sistema (f, m), ou seja, que para todo $E \subset [0,1]$ invariante temos ou $m(E) = 0$ ou $m(E) = 1$. Para tal, suponha que $m(E) > 0$ e fixe $a \in E$ ponto de densidade. Dado $A \subset \mathbb{R}$, dizemos que $p \in A$ é *ponto de densidade* de A se este conjunto preenche a maior parte de qualquer pequena vizinhaça de p, i.e.

$$\liminf_{\epsilon \to 0} \left\{ \frac{m(I \cap A)}{m(I)}; I \subset (p - \epsilon, p + \epsilon), \quad a \in I \right\} = 1. \tag{28}$$

Considere a sequência de intervalos contendo a,

$$I_k = (\frac{m_k}{10^k}, \frac{m_k + 1}{10^k}), \quad m_k \in \mathbb{Z},\ k \in \mathbb{N}.$$

Sendo a ponto de densidade de E, então por (28) temos que quando k tende ao infinito:

$$\frac{m(I_k \cap E)}{m(I_k)} \to 1.$$

Observe que f^k é uma bijeção de I_k afim sobre o intervalo $(0,1)$, de fato, seu gráfico consiste de segmentos de reta de inclinação 10^k. Isso tem a seguinte consequência, que é crucial para o nosso argumento:

$$\frac{m(f^k(J_1))}{m(f^k(J_2))} = \frac{m(J_1)}{m(J_2)}$$

para quaisquer subconjuntos mensuráveis J_1 e J_2 de I_k. Aplicando este fato a $J_1 = I_k \cap E$ e $J_2 = I_k$ obtemos que

$$\frac{m(f^k(I_k \cap E))}{m((0,1))} = \frac{m(I_k \cap E)}{m(I_k)}.$$

Claro que $m((0,1)) = 1$, além disso, como estamos supondo que E é invariante, $f^k(I_k \cap E)$ está contido em E. Desta forma obtemos que

$$m(E) \geq \frac{m(I_k \cap E)}{m(I_k)}.$$

Como o lado direito converge para 1, segue que $m(E) = 1$. Logo (f, m) cumpre o item (2), e portanto (f, m) é ergódico.

Agora vejamos uma aplicação na Teoria dos Números: pegue s o número formado com todos os CPF's do Brasil postos em sequência. Considere o intervalo $E = (0.s,\ 0.s999...)$, obviamente s tem finitos dígitos, e portanto $m(E) > 0$. Logo, pelo Teorema Ergódico de Birkhoff temos que para $\mu-$quase todo ponto $x \in [0,1]$, $\tau(E, x)$ existe. Como nosso sistema (f, m) cumpre o item (1) da Proposição 33 e já foi visto que τ é invariante por f, segue que $\tau(E, x)$ é constante em quase todo ponto de $[0,1]$, logo:

$$m(E) = \int \tau(E,x)dm(x) = \tau(E,x) \int dm(x) = \tau(E,x) \cdot 1.$$

Portanto $\tau(E, x) > 0$, o que implica termos infinitos iterados de x em E para quase todo ponto $x \in [0,1]$, ou seja, infinitos valores de $j \in \mathbb{N}$ com $f^j(x) = 0, s....$ Segue portanto, que o número s aparece infinitas vezes em quase todo ponto de $[0,1]$. Logo, se pegarmos o conjunto A dos pontos em que s aperece finitas vezes, teremos $m(A) = 0$.

Rotação irracional no círculo

Para nós o círculo S^1 será o conjunto dos números complexos com módulo igual a 1. Dado $\alpha \in \mathbb{R}$, a rotação de ângulo α é a multiplicação pelo número complexo $e^{\alpha i}$

$$\begin{aligned} R_\alpha :\ S^1 &\to S^1 \\ z &\mapsto e^{\alpha i z} \end{aligned}$$

É claro que R_α preserva o comprimento dos intervalos (segmentos) de S^1. Além disso, podemos ver facilmente que a medida de Lebesgue (comprimento de arco) é invariante por qualquer R_α . O comportamento dinâmico e ergódico de R_α depende muito da natureza de α, como vamos ver. Dizemos que a rotação é irracional se o número $\alpha/2\pi$ é irracional, e dizemos que a rotação é racional no caso contrário.

Teorema 34. *Se R_α é rotação irracional então R_α é ergódica para a medida de Lebesgue.*

Para a demonstração desta Proposição usaremos fatos simples de Análise de Fourier. Seja μ a medida de Lebesgue no círculo. Chama-se $L^2(\mu)$ o espaço das funções mensuráveis $\psi :\ S^1 \to \mathbb{C}$ cujo quadrado é integrável:

$$\int |\psi|^2 \, d\mu < \infty.$$

É claro que este espaço contém todas as funções mensuráveis limitadas e, em particular, todas as funções características de conjuntos mensuráveis. Outro fato de que necessitamos é que a família de funções $\{\phi_k(z) = z_k : \ k \in \mathbb{Z}\}$ é uma base desse espaço: dada qualquer $\varphi \in L^2(\mu)$ existe uma única sequência $(c_k)_{k\in\mathbb{Z}}$ de números complexos tais que

$$\varphi(z) = \sum_{k\in\mathbb{Z}} c_k z^k$$

para quase todo $z \in S^1$.

Demonstração. (Teorema 34) Pela Proposição 33, basta mostrar que toda função integrável φ que é invariante é constante em $\mu-$quase todo ponto. Observe que se φ é integrável, então automaticamente $\varphi \in L^2(\mu)$.

Usando a expansão de Fourier $\varphi(z) = \sum_{k\in\mathbb{Z}} c_k z^k$, a condição de ser invariante $\varphi \circ R_\alpha = \varphi$ se escreve

$$\sum_{k\in\mathbb{Z}} c_k e^{ki\alpha} z^k = \sum_{k\in\mathbb{Z}} c_k z^k$$

Por unicidade dos coeficientes da expansão em série de Fourier, obtemos que

$$c_k(e^{ki\alpha} - 1) = 0 \quad \text{para todo} \ \ k \in \mathbb{Z}.$$

A hipótese de que a rotação é irracional significa que $e^{ki\alpha} - 1 \neq 0$ para todo $k \neq 0$, e portanto, $c_k = 0$ para todo $k \neq 0$. Ou seja, $\varphi(z) = c_0$ para $\mu-$quase todo $z \in S^1$, como queríamos provar. $\square$

De fato as rotações irracionais satisfazem uma propriedade muito mais forte do que ergodicidade: elas são unicamente ergódicas, o que quer dizer que têm uma única probabilidade invariante (que é a medida de Lebesgue, claro).

Observação. A noção de rotação irracional se estende para dimensões maiores. Dado qualquer $d \geq 1$ chamamos $d-$toro o produto $\mathbb{T}^d = S^1 \times \cdots \times S^1$ do círculo por si mesmo d vezes. A rotação de ângulo $\alpha = (\alpha_1, ..., \alpha_d)$ é a aplicação $R_\alpha : \ \mathbb{T}^d \to \mathbb{T}^d$, $R_\alpha(z_1, ..., z_d) = (e^{i\alpha_1} z_1, ..., e^{i\alpha_d} z_d)$. A rotação é irracional se os números $\alpha_j/2\pi$ são incomensuráveis:

$$m_0 + m_1 \frac{\alpha_1}{2\pi} + \cdots + m_d \frac{\alpha_d}{2\pi} = 0 \Rightarrow m_0 = m_1 = \cdots = m_d = 0,$$

quaisquer que sejam os inteiros $m_0, m_1, ..., m_d$. Usando uma versão multidimensional das idéias anteriores, se prova que uma rotação é ergódica se e somente se ela é irracional.

8 EQUIVALÊNCIA ESPECTRAL

Este capítulo se trata da caracterização das propriedades dinâmicas de uma medida como propriedades de Algebra Linear, relacionadas a um certo operador U_f que iremos definir. Iremos ver, por exemplo, que a invariância da medida μ equivale ao fato do operador preservar distâncias num espaço adequado de funções. Veremos também a ergodicidade como uma propriedade espectral do sistema (f, μ), no sentido de que ela pode ser caracterizada pelos auto-valores do operador U_f.

O espaço $L^2(\mu)$

Antes de definir o operador U_f, precisamos definir o seu domínio, o espaço $L^2(\mu)$:

Definição 35. *O espaço vetorial $L^2(\mu)$ é definido como o conjunto*

$$L^2(\mu) = \{\varphi : \ M \to \mathbb{R}; \int |\varphi|^2 \ d\mu < \infty\}$$

onde duas funções são identificadas se elas coincidem em quase todo ponto com respeito à medida μ.

Podemos munir este espaço vetorial com um produto interno dado por:

$$\langle \varphi, \psi \rangle = \int \varphi \cdot \psi \ d\mu.$$

Segue da desigualdade de Hölder e das propriedades da integral que a expressão acima de fato define um produto interno em $L^2(\mu)$, ou seja, para todo $\varphi, \ \psi, \ \rho \in L^2(\mu)$ e $\alpha \in \mathbb{R}$, vale:

•$\langle \varphi, \varphi \rangle \geq 0$ e $\langle \varphi, \varphi \rangle = 0 \iff \varphi \equiv 0$ em $\mu - q.t.p.$

•$\langle \varphi, \psi \rangle = \langle \psi, \varphi \rangle$

•$\langle \varphi + \rho, \psi \rangle = \langle \varphi, \psi \rangle + \langle \rho, \psi \rangle$ e $\langle \alpha\varphi, \psi \rangle = \alpha \langle \varphi, \psi \rangle = \langle \varphi, \alpha\psi \rangle$

De fato,

$\langle \varphi, \varphi \rangle = \int \varphi^2 \ d\mu \geq 0$ e $\langle \varphi, \varphi \rangle = 0 \iff \int \varphi^2 \ d\mu = 0 \iff \varphi^2 \equiv 0 \ \mu - q.t.p. \iff \varphi \equiv 0 \ \mu - q.t.p.$,

$\langle \varphi, \psi \rangle = \int \varphi \cdot \psi \ d\mu = \int \psi \cdot \varphi \ d\mu = \langle \psi, \varphi \rangle,$

$\langle \varphi + \rho, \psi \rangle = \int (\varphi + \rho)\psi \ d\mu = \int \varphi \cdot \psi + \rho \cdot \psi \ d\mu = \int \varphi \cdot \psi \ d\mu + \int \rho \cdot \varphi \ d\mu = \langle \varphi, \psi \rangle + \langle \rho, \psi \rangle,$
a penúltima igualdade ocorre pois, se φ, ρ e $\psi \in L^2(\mu)$, então da desigualdade de Hölder temos

$$\left| \int \varphi \cdot \psi \ d\mu \right| \leq \left(\int |\varphi|^2 \ d\mu \right)^{\frac{1}{2}} \cdot \left(\int |\psi|^2 \ d\mu \right)^{\frac{1}{2}} < \infty,$$

portanto $\varphi \cdot \psi \in L^2(\mu)$ e da mesma forma $\rho \cdot \psi \in L^2(\mu)$. Por último

$$\langle \alpha\varphi, \psi \rangle = \int \alpha\varphi \cdot \psi \ d\mu = \alpha \int \varphi \cdot \psi \ d\mu = \alpha \langle \varphi, \psi \rangle = \int \varphi \cdot \alpha\psi \ d\mu = \langle \varphi, \alpha\psi \rangle.$$

Logo $\langle \cdot, \cdot \rangle$ é um produto interno.

Naturalmente, uma vez que temos um produto interno em $L^2(\mu)$, podemos definir uma norma por:

$$||\varphi||_2 = \sqrt{\langle \varphi, \varphi \rangle}.$$

Além disso, podemos munir $L^2(\mu)$ com a métrica que provém do produto interno $\langle \cdot, \cdot \rangle$ definida por:

$$d(\varphi, \psi) = ||\varphi - \psi||_2 = \left(\int |\varphi - \psi|^2 d\mu \right)^{\frac{1}{2}}. \tag{29}$$

Da Análise Funcional é conhecido que o espaço $L^2(\mu)$ com a métrica d é um espaço métrico completo.

Defiremos agora a transformação linear que carrega algumas propriedades métricas do sistema (f, μ). Considere $U_f : L^2(\mu) \to L^2(\mu)$ definida simplesmente como:

$$U_f(\varphi) = \varphi \circ f.$$

É imediato verificar que U_f é uma transformação linear. Pois dados $\varphi, \psi \in L^2(\mu)$ e $\lambda \in \mathbb{R}$, note que $U_f(\varphi + \lambda\psi) = (\varphi + \lambda\psi) \circ f = (\varphi) \circ f + (\lambda\psi) \circ f = \varphi \circ f + \lambda(\psi \circ f) = U_f(\varphi) + \lambda U_f(\psi)$.

A próxima proposição traz a primeira tradução entre propriedades de (f, μ) e propriedades da transformação U_f.

Proposição 36. *Seja f uma transformação e μ uma medida em M. f preserva μ se, e só se, para toda $\varphi \in L^2(\mu)$:*

$$||U_f \varphi||_2 = ||\varphi||_2 \tag{30}$$

Um resultado preliminar bastante importante para a demonstração da Proposição 36 é:

Proposição 37. *Seja $f : M \to M$ uma transformação e μ uma medida. Então f preserva μ se, e somente se, para toda função integrável $\phi : M \to \mathbb{R}$ vale:*

$$\int \phi \ d\mu = \int \phi \circ f \ d\mu.$$

Demonstração. Assuma que f preserva μ. Se ϕ é a função característica de algum conjunto, digamos $\phi = \chi_A$, é imediato verificar que $\mu(f^{-1}(A)) = \int \phi \circ f d\mu$, já que $\chi_{f^{-1}(A)} = \phi \circ f$. Assim, fica provado que $\int \phi d\mu = \int \phi \circ f d\mu$, quando ϕ é uma função caracerística. Observe que segue diretamente da linearidade da integral que se ϕ é uma função simples, então a igualdade ainda vale. Finalmente, se ϕ é uma função integrável qualquer, pela definição de integral

$$\int \phi \ d\mu = \lim_{n \to \infty} \int \phi_n \ d\mu,$$

onde (ϕ_n) é uma sequência de funções simples crescendo para ϕ. Por outro lado, $\phi_n \circ f$ é uma sequência de funções crescendo para $\phi \circ f$. Logo,

$$\int \phi \circ f \mu = \lim_{n\to\infty} \int \phi_n \circ f \; d\mu.$$

Como $\int \phi_n d\mu = \int \phi_n \circ f d\mu$, tomando o limite em ambos os lados, vem que

$$\int \phi \; d\mu = \int \phi \circ f \; d\mu.$$

A recíproca é imediata, desde que dado um boreliano A, tomando $\phi = \chi_A$, então

$$\mu(A) = \mu(f^{-1}(A)) \iff \int \phi \; d\mu = \int \phi \circ f \; d\mu.$$

$\square$

Agora podemos demonstrar a Proposição 36.

Demonstração. (Proposição 36) De fato, primeiramente observe que se f preserva a medida μ então para todo boreliano A vale $\mu(A) = \mu(f^{-1}(A))$. Em particular, a igualdade (30) vale se φ for uma função característica de um boreliano. Utilizando a linearidade da integral, podemos estender nossa conclusão para todas funções simples. Para concluir a demonstração da propriedade (30) observe que $L^2(\mu) \subset L^1(\mu)$. Com efeito, se $\phi \in L^2(\mu)$, aplicando a desigualdade de Hölder às funções $||\phi||$ e 1, temos que $\int ||\phi|| d\mu \leq \int ||\phi||^2 d\mu < \infty$. Como toda função ψ em $L^1(\mu)$ pode ser aproximada por uma sequência monótona ψ_n de funções simples, dada $\varphi \in L^2(\mu)$, tomando uma sequência de funções simples s_n convergindo monotonamente para φ e utilizando o Teorema da Convergência Monótona para as sequência s_n e $s_n \circ f$, temos que:

$$\int \varphi \circ f \; d\mu = \lim_{n\to\infty} \int s_n \circ f \; d\mu = \lim_{n\to\infty} \int s_n \; d\mu = \int \varphi \; d\mu.$$

Reciprocamente, provaremos este fato primeiramente para as funções contínuas não-negativas. Se $\varphi \geq 0$ é função contínua, tomando $\psi = \sqrt{\varphi}$ temos que $\psi \in L^2(\mu)$, pois ψ é limitada e μ é uma probabilidade. Como estamos assumindo que $||\psi \circ f||_2 = ||\psi||_2$, temos que:

$$\int ||\psi \circ f||^2 \; d\mu = \int \varphi \circ f \; d\mu = \int \varphi \; d\mu = \int ||\psi||^2 \; d\mu.$$

Para provarmos que a igualdade acima vale quando φ é uma função contínua qualquer, basta observarmos que toda função contínua se escreve como diferença de duas funções positivas limitadas e aplicarmos a igualdade obtida a estas funções. Assim, utilizando a Proposição 37, temos que f preserva μ. $\square$

A igualdade (30) acima significa que U_f é uma *isometria* do espaço $L^2(\mu)$. Em particular, esta propriedade implica diretamente que U_f é uma transformação linear injetiva, pois:

$$||\varphi||_2 > 0 \Rightarrow ||U_f\varphi||_2 = ||\varphi||_2 > 0.$$

Porém, U_f só é sobrejetiva se f for invertível. De fato, dado $\psi \in L^2(\mu)$, existe $\varphi \in L^2(\mu)$ tal que $U_f(\varphi) = \psi$ então U_f é invertível logo existe uma tranformação linear U_g tal que $U_g(\psi) = \varphi$ e

$$\varphi = U_g \circ U_f(\varphi) = U_g \circ \varphi \circ f = \varphi \circ f \circ g \Rightarrow f \circ g = Id \Rightarrow g = f^{-1}.$$

Observe que 1 é sempre um autovalor de U_f e a esse autovalor temos associado um auto-espaço que sempre contém as funções constantes. A Proposição 33 diz que a propriedade da medida μ ser ergódica equivale ao fato de que as únicas funções invariantes por U_f são as funções constantes ou, em outros termos, a dimensão do auto-espaço associado ao auto-valor 1 é igual a um. Isso justifica a afirmação de que a ergodicidade é uma propriedade *espectral*.

Vamos estabelecer quando duas tranformações $f: M \to M$ e $g: N \to N$ preservando medidas μ e ν são equivalentes do ponto de vista espectral. Isto nos permitirá extrair informações sobre as propriedades espectrais do sistema (f,μ) a partir das propriedades espectrais do sistema (g,μ) e vice-versa.

Definição 38. *Sejam $f: M \to M$ e $g: N \to N$ preservando μ e ν, respectivamente. Dizemos que (f,μ) e (g,ν) são espectralmente equivalentes, $(f,\mu) \sim (g,\nu)$, se existir uma transformação linear invertível $A: L^2(\mu) \to L^2(\nu)$ tal que:*

i) $\langle A\varphi, A\psi\rangle = \langle\varphi,\psi\rangle$;

ii) $AU_f = U_gA$.

De fato a relação acima é uma relação de equivalência, pois

a) $(f,\mu) \sim (f,\mu)$:

i) $\langle I\varphi, I\varphi\rangle = \langle\varphi,\varphi\rangle$,

ii) $IU_f = U_fI$,

onde I é a matriz identidade.

b) Se $(f,\mu) \sim (g,\nu) \Rightarrow (g,\nu) \sim (f,\mu)$

i) $\langle A\varphi, A\psi\rangle = \langle\varphi,\psi\rangle \Rightarrow -\langle A\psi, A\varphi\rangle = -\langle\psi,\varphi\rangle \Rightarrow \langle A\psi, A\varphi\rangle = \langle\psi,\varphi\rangle$,

ii) $AU_f = U_gA \Rightarrow U_fA^{-1} = A^{-1}U_g$ (A^{-1} também é transformação linear invertível).

c) Se $(f,\mu) \sim (g,\nu)$ e $(g,\nu) \sim (h,\eta)$, então $(f,\mu) \sim (h,\eta)$

Se $(f,\mu) \sim (g,\nu)^{(*)}$ e $(g,\nu) \sim (h,\eta)^{(**)}$, então existem transformações lineares invertíveis $A: L^2(\mu) \to L^2(\nu)$ e $\tilde{A}: L^2(\nu) \to L^2(\eta)$, tais que $\langle A\varphi, A\psi\rangle = \langle \varphi, \psi\rangle$ e $\langle \tilde{A}\sigma, \tilde{A}\rho\rangle = \langle \sigma, \rho\rangle$. Além disso, $AU_f = U_g A$ e $\tilde{A}U_g = U_h\tilde{A}$, e assim,

$$\langle \tilde{A}.A\varphi, \tilde{A}.A\psi\rangle =^{(**)} \langle A\varphi, A\psi\rangle =^{(*)} \langle \varphi, \psi\rangle,$$

portanto a propriedade (i) se verifica. Por outro lado

$$\tilde{A} \cdot A \cdot U_f =^{(*)} \tilde{A} \cdot U_g \cdot A =^{(**)} U_h \cdot \tilde{A} \cdot A,$$

logo a propriedade (ii) também se verifica. Portanto $\sim$ é uma relação de equivalência.

Observe que, dados (f,μ) e (g,ν) espectralmente equivalentes e A uma equivalência espectral, se φ é auto-função de U_f associada ao auto-valor λ, então $A\varphi$ é autofunção de U_g associada ao auto-valor λ. De fato, $U_f(\varphi) = \lambda\varphi \Rightarrow U_g A(\varphi) = AU_f(\varphi) = \lambda A\varphi$. Assim, o auto-espaço F_1 associado ao auto-valor 1 para operador U_g é simplesmente a imagem por A do auto-espaço E_1 associado ao auto-valor 1 para U_f. Como A é um isomorfismo linear, a dimensão de F_1 é igual dimensão de E_1. Segundo a observação que (f,μ) é ergódica se a dimensão do auto-espaço associado ao auto-valor 1 é igual a 1, podemos afirmar que se (f,μ) e (g,ν) são espectralmente equivalentes, então (f,μ) é um sistema ergódico se, e só se, (g,ν) é ergódico.

8.1 Propriedades de Medidas Ergódicas

Fixemos uma transformação $f: M \to M$ qualquer. Uma medida ν diz-se absolutamente contínua com relação a outra medida μ ($\nu \ll \mu$) se $\mu(E) = 0$ implicar $\nu(E) = 0$. O próximo lema afirma que probabilidades ergódicas são minimais para tal relação:

Lema 39. *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. Seja $\varphi: M \to \mathbb{R}$ uma função mensurável limitada qualquer, e seja $\tilde{\varphi}$ a sua média temporal. Como μ é invariante e ergódica, a média temporal é contante

$$\tilde{\varphi}(x) = \int \varphi \, d\mu$$

para μ-quase todo ponto. Segue portanto, que isto é verdade para ν-quase todo ponto, já que $\nu \ll \mu$. Em particular,

$$\int \tilde{\varphi} \, d\nu = \int \varphi \, d\mu.$$

Por outro lado, pelo Teorema Ergódico 32,

$$\int \tilde{\varphi} \, d\nu = \int \varphi \, d\nu.$$

Portanto, as integrais de φ com relação a μ e em relação a ν coincidem, qualquer que seja a função mensurável limitada φ. Logo, considerando funções características, $\mu = \nu$. $\square$

Naturalmente, se μ_1 e μ_2 são probabilidades invariantes com respeito à f a probabilidade $\mu_1 + t(\mu_1 - \mu_2)$ ainda é invariante. Isso siginfica que o conjunto das probabilidades invariantes é um conjunto *convexo*. Veremos que dentro deste conjunto, as medidas ergódicas desempenham um papel destacado:

Definição 40. *Seja X um conjunto convexo. Um ponto $p \in X$ é dito extremal, se para quaisquer $x, y \in X$ e $t \in [0,1]$, $x + t(y - x) = p$ implica que $t = 0$ ou $t = 1$.*

O lema a seguir afirma que uma probabilidade invariante é ergódica se, e somente se, a probabilidade é ponto extremal no conjunto das probabilidades invariantes:

Lema 41. *Uma probabilidade invariante μ é ergódica se e somente se não é possível escrevê-la na forma*

$$\mu = c_1\mu_1 + c_2\mu_2,$$

com c_1 e c_2 maiores que zero, e μ_1 e μ_2 probabilidades invariantes distintas.

Demonstração. Para provar a parte "se", suponha que μ não seja ergódica. Então existe algum conjunto invariante A com $0 < \mu(A) < 1$. Defina μ_1 e μ_2 como sendo as restrições normalizadas de μ a A e ao seu complementar, respectivamente:

$$\mu_1(E) = \frac{\mu(E \cap A)}{\mu(A)}, \qquad \mu_2(E) = \frac{\mu(E \cap A^c)}{\mu(A^c)}.$$

Como A e A^c são conjuntos invariantes e μ é medida invariante, μ_1 e μ_2 são também probabilidades invariantes. Além disso, $\mu = c_1\mu_1 + c_2\mu_2$, com $c_1,\ c_2 > 0$ e portanto μ não é extremal.

Para provar a recíproca, suponha que μ é ergódica e temos $\mu = c_1\mu_1 + c_2\mu_2$ com $c_1, c_2 > 0$. É claro que $\mu(E) = 0$ implica $\mu_1(E) = \mu_2(E) = 0$, ou seja, μ_1 e μ_2 são absolutamente contínuas com relação a μ. Logo, pelo Lema 39, $\mu_1 = \mu = \mu_2$. Isto prova que μ é extremal. ☐

Em seguida veremos que medidas ergódicas distintas "vivem"em subconjuntos disjuntos do espaço M:

Lema 42. *Sejam $\mu_1, ..., \mu_N$ probabilidades invariantes e ergódicas, todas distintas. Então existem subconjuntos mensuráveis $P_1, ..., P_N$ invariantes disjuntos tais que*

$$\mu_j(P_k) = \begin{cases} 1 & se \quad j = k, \\ 0 & se \quad j \neq k. \end{cases}$$

Demonstração. Fixe qualquer par (j, k) de números distintos em $\{1, ..., N\}$. Pelo Lema 39, a medida μ_j não pode ser absolutamente contínua em relação a μ_k. Em outras palavras, existe algum subconjunto mensurável E tal que $\mu_j(E) > 0$ mas $\mu_k(E) = 0$. Então,

$$\mu_j\left(\bigcup_{j=0}^{\infty} f^{-j}(E)\right) \geq \mu_j(E) > 0, \qquad \text{e} \qquad \mu_k\left(\bigcup_{j=0}^{\infty} f^{-j}(E)\right) = 0.$$

Defina $P_{j,k} = \cap_{m=0}^{\infty} \cup_{j=m}^{\infty} f^{-j}(E)$. Como a sequência de conjuntos na interseção é decrescente com m,

$$\mu_j(P_{j,k}) = \lim_{m\to\infty} \mu_j \left(\bigcup_{j=m}^{\infty} f^{-j}(E) \right). \tag{31}$$

Analogamente podemos verificar para μ_k. Como as medidas μ_j e μ_k são invariantes, e

$$\bigcup_{j=m}^{\infty} f^{-1}(E) = f^{-m} \left(\bigcup_{j=0}^{\infty} f^{-j}(E) \right),$$

a sequência no lado direito de (31) é constante. Concluímos, portanto, que

$$\mu_j(P_{j,k}) = \mu_j \left(\bigcup_{j=0}^{\infty} f^{-j}(E) \right) > 0 \quad \text{e} \quad \mu_k(P_{j,k}) = \mu_k \left(\bigcup_{j=0}^{\infty} f^{-j}(E) \right) = 0.$$

Além disso, $P_{j,k}$ é um conjunto invariante por f. Portanto $\mu_j(P_{j,k}) = 1$, uma vez que μ_j é ergódica. Agora, defina

$$\tilde{P}_j = \bigcap_{k\neq j} P_{j,k} \quad \text{e} \quad P_j = \mathrm{P}_j \bigcup_{k\neq j} \tilde{P}_k.$$

Primeiramente, $\mu_j(\tilde{P}_j) = 1$ e $\mu_k(\tilde{P}_j) = 0$ para todo $k \neq j$. Segue que $\mu_j(P_j) = 1$, e $\mu_k(P_j) = 0$ para todo $k \neq j$. Além disso, os P_j são disjuntos dois-a-dois. □

9 TEOREMA DA DECOMPOSIÇÃO ERGÓDICA

Na linha dos resultados da seção anterior, é natural perguntar se toda medida invariante é uma combinação linear de medidas ergódicas. O teorema que vamos enunciar nesta seção diz, em um certo sentido, que a resposta é afirmativa, exceto que o número de "parcelas"nesta combinação não é necessariamente finito, nem mesmo enumerável, em geral.

Teorema 43. *Seja $f: M \to M$ uma transformação contínua num espaço compacto. Então existe um conjunto mensurável $M_0 \subset M$, uma partição P de M_0 e uma família de probabilidades $\{\nu_P : P \in \mathcal{P}\}$ satisfazendo:*

a) $\nu_P(P) = 1$ para todo elemento P de $\mathcal{P}$;

b) a aplicação $P \to \nu_P$ é mensurável;

c) toda ν_P é invariante e ergódica para f;

tais que, dada qualquer probabilidade $f-$invariante μ, o conjunto M_0 satisfaz $\mu(M_0) = 1$ e, além disso,

d) $\mu(E) = \int \nu_P(E) d\hat{\mu}(P)$ para todo conjunto mensurável $E \subset M$

onde $\hat{\mu}$ é a medida projeção de P em $\mathcal{P}$.

A relação (d) significa que μ é uma combinação convexa das várias probabilidades ergódicas ν_P, em que cada ν_P entra com "coeficiente"igual a $\hat{\mu}(P)$. Dada qualquer partição P de M fica definida a projeção natural $\pi : M \to P$ que associa a cada ponto $x \in M$ o elemento $P(x)$ da partição que o contém. Isto permite definir o que é um subconjunto mensurável da partição: $\mathcal{Q} \subset \mathcal{P}$ é mensurável se, e somente se,

$$\pi^{-1}(\mathcal{Q}) = \text{união dos } P \in \mathcal{Q}$$

é um subconjunto mensurável de M. E fácil ver que esta definição está coerente. Isto é, a família dos subconjuntos mensuráveis é uma $\sigma-$álgebra em P. A medida projeção de μ está definida nesta $\sigma-$álgebra, por

$$\hat{\mu}(\mathcal{Q}) = \mu(\pi^{-1}(\mathcal{Q})).$$

Corolário 44. *Uma transformação f é unicamente ergódica se, e somente se, admite exatamente uma medida invariante ergódica.*

Demonstração. Já sabemos que, se f é unicamente ergódica então a sua probabilidade invariante é ergódica. Isto prova a parte "somente se"do enunciado. Por outro lado, o Teorema 43 mostra que a recíproca também é verdadeira: se f admite apenas uma probabilidade invariante ergódica, então essa é a única probabilidade invariante. □

9.1 Prova do Teorema da Decomposição Ergódica

Seja Z um espaço métrico compacto, μ uma probabilidade em Z, e P uma partição de Z em subconjuntos mensuráveis. Seja $\pi : Z \to P$ a aplicação que associa a cada $z \in Z$ o átomo $P \in \mathcal{P}$ que o contém. Por definição, Q é um subconjunto mensurável de P se e somente se $\pi^{-1}(Q)$ é um subconjunto mensurável de Z. Seja $\hat{\mu}$ o iterado de μ por π, ou seja, $\hat{\mu}$ é a probabilidade em P definida por $\hat{\mu}(Q) = \mu(\pi^{-1}(Q))$ para cada conjunto mensurável $Q \subset P$.

Definição 45. *Uma família de medidas condicionais de μ relativamente a P é uma família $(\mu_P)_{P\in\mathcal{P}}$ de probabilidades em Z tais que*

1. $\mu_P(P) = 1$ para $\hat{\mu}$-quase todo $P \in \mathcal{P}$;

2. dada qualquer função contínua $\varphi : Z \to \mathbb{R}$, a função $\mathcal{P} \ni P \mapsto \int \varphi d\mu_P$ é mensurável e tem-se $\int \varphi d\mu = \int \left(\int \varphi d\mu_P\right) d\hat{\mu}(P)$.

Lema 46. *Se $(\mu_P)_{P\in\mathcal{P}}$ é uma família de medidas condicionais de μ relativamente à partição $\mathcal{P}$, então $\mathcal{P} \ni P \mapsto \int \psi d\mu_P$ é mensurável e*

$$\int \psi d\mu = \int \left(\int \psi d\mu_P\right) d\hat{\mu}(P),$$

para toda função mensurável limitada $\psi : Z \to \mathbb{R}$.

Demonstração. Seja $\mathcal{C}$ a classe das funções mensuráveis limitadas que satisfazem a conclusão do lema. Pela definição de medidas condicionais, essa classe contém todas as funções contínuas. Além disso, suponha que $\varphi_n : Z \to \mathbb{R}$, $n \in \mathbb{N}$, seja uma sequência de funções em $\mathcal{C}$ convergindo pontualmente para alguma função $\varphi : Z \to \mathbb{R}$. Suponha ainda que essa sequência é uniformemente limitada, isto é, existe $K > 0$ tal que $|\varphi_n(z)| \le K$ para todo $z \in Z$ e todo $n \in \mathbb{N}$. Afirmamos que a função limite φ também está em $\mathcal{C}$. Admita, por um instante, que esta afirmação é verdadeira.

Em particular, $P \to \mu_P(E)$ é mensurável, e $\mu(E) = \int \mu_P(E) d\hat{\mu}(P)$, para qualquer conjunto mesurável $E \subset Z$. □

Medidas condicionais, quando existem, são únicas em quase todo ponto:

Proposição 47. *Se $(\mu_P)_{P\in\mathcal{P}}$ e $(\nu_P)_{P\in\mathcal{P}}$ são dois sistemas de medidas condicionais de μ com respeito à P, então $\mu_P = \nu_P$ para $\hat{\mu}$-q.t.p $P \in \mathcal{P}$.*

Demonstração. Suponha o contrário, isto é, existe um conjunto mensurável $\mathcal{Q}_0 \subset \mathcal{P}$ com $\hat{\mu}(\mathcal{Q}_0) > 0$ tal que $\mu_P \neq \nu_P$ para todo $P \in \mathcal{Q}_0$. Seja $\{\varphi_k;\ k \in \mathbb{N}\}$ um subconjunto enumerável e denso de $C^0(Z, \mathbb{R})$, e defina

$$A_k = \{P \in \mathcal{Q}_0; \int \varphi_k \, d\mu_p \neq \int \varphi_k \, d\nu_p\}.$$

Notando que $\cup_k A_k = \mathcal{Q}_0$, existe $\varphi \in C^0(Z, \mathbb{R})$ e um subconjunto $\mathcal{Q}$ de $\mathcal{Q}_0$ tal que $\hat{\mu}(\mathcal{Q}) > 0$ e (trocando os papéis de μ_P e ν_P, se necessário) $\int \varphi d\mu_p > \int \varphi d\nu_p$ para todo $P \in \mathcal{Q}$. Então

$$\int_{\mathcal{Q}} \left(\int \varphi \, d\mu_P \right) \, d\hat{\mu}(P) > \int_{\mathcal{Q}} \left(\int \varphi \, d\nu_P \right) \, d\hat{\mu}(P). \tag{32}$$

De outro modo, pelo Lema 46,

$$\int (\varphi \chi_{\pi^{-1}(\mathcal{Q})}) \, d\mu = \int \left(\int (\varphi \chi_{\pi^{-1}(\mathcal{Q})}) \, d\mu_P \right) \, d\hat{\mu}(P).$$

Por hipótese $\mu_P(P) = 1$ para $\hat{\mu}$ quase todo $P \in \mathcal{P}$. Para cada P, temos que

$$\int (\varphi \chi_{\pi^{-1}(\mathcal{Q})}) \, d\mu = \int \left(\chi_{\mathcal{Q}}(P) \int \varphi \, d\mu_P \right) \, d\hat{\mu}(P) = \int_{\mathcal{Q}} \left(\int \varphi \, d\mu_P \right) \, d\hat{\mu}(P).$$

Analogamente, temos

$$\int (\varphi \chi_{\pi^{-1}(\mathcal{Q})}) \, d\mu = \int_{\mathcal{Q}} \left(\int \varphi \, d\nu_P \right) \, d\hat{\mu}(P)$$

Essas duas últimas igualdades contradizem (32). Logo, $\mu_P = \nu_P$ para $\hat{\mu}-$ quase todo P, como desejávamos. □

Definição 48. *$\mathcal{P}$ é uma partição mensurável se existem conjuntos mensuráveis $E_1,\ E_2, ..., E_n, ...$ de Z tais que existe um conjunto de medida μ total $F_0 \subset Z$ tal que, dado qualquer átomo P de $\mathcal{P}$ podemos escrever*

$$P \cap F_0 = E_1^* \cap E_2^* \cap ... \cap E_n^* \cap ... \cap F_0 \tag{33}$$

onde $E_j^* = E_j$ ou $Z \backslash E_j$ para todo $j \in \mathbb{N}$.

Para demonstrarmos o Teorema da Decomposição Ergódica vamos usar o seguinte resultado, cuja demonstração se encontra na sequência.

Teorema 49. *(Desintegração) Se P é partição mensurável então a probabilidade μ admite alguma família de medidas condicionais relativamente a P.*

Demonstração. (Teorema da Decomposição Ergódica) Seja $f:\ Z \to Z$ contínua num espaço métrico compacto Z, e seja B_f o subconjunto dos pontos $z \in Z$ tais que as médias temporais estão bem definidas na órbita de z : dada qualquer função contínua $\varphi :\ Z \to \mathbb{R}$, a sequência

$$\frac{1}{n}\sum_{j=0}^{n-1} \varphi(f^j(z))$$

converge para algum $\tilde{\varphi}(z) \in \mathbb{R}$ quando $n \to \infty$, ou seja $B_f = \{z \in Z;\ \exists \tilde{\varphi}(z)\}$. Pelo Teorema Ergódico de Birkhoff (Teorema 31) B_f é mensurável e $\mu(B_f) = 1$. Seja $\mathcal{P}$ a partição de Z cujos elementos são

1. $Z \backslash B_f \in \mathcal{P}$ e

2. classes de equivalência em B_f pela relação $z_1 \sim z_2 \iff \tilde{\varphi}(z_1) = \tilde{\varphi}(z_2)$ para toda φ contínua.

Note que $Z \backslash B_f$ é mensurável, e dado $P \in \mathcal{P}$, $P \neq Z \backslash B_f$, para toda φ contínua temos $\tilde{\varphi}(z) = cte$ para todo $z \in P$. Logo, de 1 e 2 temos que $\mathcal{P}$ é mensurável com respeito a qualquer probabilidade μ. Pelo Teorema de Desintegração existe $(\mu_P)_{P \in \mathcal{P}}$ família de medidas condicionais a μ.

Como μ é $f-$invariante ($\mu(Z \backslash B_f) = 0$), qualquer família de medidas condicionais $(\mu_P)_P$ a μ relativamente a $\mathcal{P}$ é tal que μ_P é $f-$invariante e ergódica para $\hat{\mu}-$quase todo $P \in \mathcal{P}$. De fato, $\tilde{\varphi}(f(z)) = \tilde{\varphi}(z),\ \forall z \in P \Rightarrow f(z) \in P \Rightarrow f(P) = P$. $\square$

9.2 Teorema de Desintegração

No intuito de provar o Teorema 49, podemos trocar o espaço Z por qualquer conjunto de medida nula. Logo, não há perda de generalidade supor que F_0 em (33) coincide exatamente com Z, e iremos assumir isso no que se segue. Seja ψ qualquer função mensurável limitada em Z. Para cada $n \geq 1$, e seja $\mathcal{P}_n$ a partição de Z cujos átomos são $E_1^* \cap \cdots E_n^*$. Defina $\tilde{\psi}_n :\ Z \to \mathbb{R}$ como se segue. Se o átomo $P_n(z)$ de $\mathcal{P}_n$ que contém z tem medida $\mu(P_n(z))$ positiva, então

$$\tilde{\psi}_n(z) = \frac{1}{\mu(P_n(z))} \int_{P_n(z)} \psi\ d\mu, \tag{34}$$

caso contrário, $\tilde{\psi}_n(z) = 0$. Claramente, o segundo caso na definição de $\tilde{\psi}_n$ se aplica somente num conjunto de pontos de medida μ igual a zero.

Lema 50. *Dada qualquer função mensurável limitada $\psi : Z \to \mathbb{R}$, existe um subconjunto de medida μ total $F = F(\psi)$ de Z tal que $\tilde{\psi}_n(z)$, $n \geq 1$, converge para algum número real $\tilde{\psi}(z)$, para todo $z \in F$.*

Demonstração. Observe que sempre podemos escrever $\psi = \psi^+ - \psi^-$, onde $\psi^\pm$ são mensuráveis, limitadas e não-negativas: por exemplo, $\psi^\pm = (|\psi| \pm \psi)/2$. Então $\tilde{\psi}_n = \tilde{\psi}_n^+ \tilde{\psi}_n^-$ para $n \geq 1$, e então a conclusão é verdadeira para ψ se ela vale para ψ^+ e ψ^-. Isto mostra que não há restrição em assumir que ψ é não-negativa. Iremos assumir isso de agora em diante.

Para todo $\alpha < \beta$, seja $S(\alpha, \beta)$ o conjunto dos pontos $z \in Z$ tais que

$$\liminf \ \tilde{\psi}_n(z) < \alpha < \beta < \limsup \ \tilde{\psi}_n(z).$$

É claro que dado $z \in Z$, a sequência $\tilde{\psi}_n(z)$ diverge se, e só se, $z \in S(\alpha, \beta)$ para algum par de números racionais α e β. Logo, o lema segue se mostrarmos que $S = S(\alpha, \beta)$ tem medida μ igual a zero para todo α e β.

Para cada $z \in S$, fixe uma sequência de inteiros $1 \leq a_1^z < b_1^z < \cdots < a_i^z < b_i^z < \cdots$ tais que

$$\tilde{\psi}_{a_i^z} < \alpha \ \text{ e } \ \tilde{\psi}_{b_i^z}(z) > \beta \ \text{ para todo } i \geq 1.$$

Defina A_i como sendo a união dos elementos $P_{a_i^z}(z)$ e B_i como sendo a união dos elementos $P_{a_i^z}$ obtidos deste modo, para todos os pontos $z \in S$. Por construção,

$$S \subset A_{i+1} \subset B_i \subset A_i \ \text{ para todo } i \geq 1.$$

Em particular, S está contido no conjunto

$$\tilde{S} = \cap_{i=1}^\infty B_i = \cap_{i=1}^\infty A_i.$$

Dados dois entre os conjuntos $P_{a_i^z}(z)$ que formam A_i, ou eles são disjuntos ou coincidem. isto porque $\mathcal{P}_n$, $n \geq 1$, é uma sequência não-decrescente de partições. Consequentemente, A_i pode ser escrito como uma união de conjuntos $P_{a_i^z}(z)$ dois-a-dois disjuntos. Assim,

$$\int_{A_i} \psi \, d\mu = \sum_{P_{a_i^z}(z)} \int_{P_{a_i^z}} \psi \, d\mu < \sum_{P_{a_i^z}(z)} \alpha\mu(P_{a_i^z}) = \alpha\mu(A_i),$$

para qualquer $i \geq 1$ (as somas são sobre uniões disjuntas). Analogamente,

$$\int_{B_i} \psi \, d\mu = \sum_{P_{b_i^z}(z)} \int_{P_{b_i^z}} \psi \, d\mu > \sum_{P_{b_i^z}(z)} \alpha\mu(P_{b_i^z}) = \beta\mu(B_i).$$

Já que $A_i \supset B_i$ e observando que estamos assumindo que $\psi \geq 0$, segue que

$$\alpha\mu(A_i) > \int_{A_i} \psi \, d\mu \geq \int_{B_i} \psi \, d\mu > \beta\mu(B_i),$$

para todo $i \geq 1$. Tomando o limite quando $i \to \infty$, temos que

$$\alpha\mu(\tilde{S}) \geq \beta\mu(\tilde{S}).$$

Isto implica que $\mu(\tilde{S}) = 0$, e logo $S \subset \tilde{S}$ também é um conjunto de medida nula. $\square$

Dada qualquer função limitada mensurável $\psi : Z \to \mathbb{R}$, iremos representar por $e_n(\psi)$, $e(\psi)$, respectivamente, as funções $\tilde{\psi}_n$, $\tilde{\psi}$ definidas por (34) e o Lema 50.

Seja $\{\varphi_k : k \in \mathbb{N}\}$ algum conjunto enumerável e denso de $C^0(Z, \mathbb{R})$, e seja

$$F_* = \bigcap_{k=1}^{\infty} F(\varphi_k),$$

onde $F(\varphi_k)$ é como dado pelo Lema 50.

Lema 51. *Dada qualquer função contínua $\varphi : Z \to \mathbb{R}$, a sequência $e_n(\varphi)(z)$ converge para $e(\varphi)(z)$ quando $n \to \infty$, para todo $z \in F_*$.*

Demonstração. Fixe $z \in F_*$. É claro que $\psi \mapsto e_n(\psi)(z)$ é um funcional linear limitado em $C^0(Z, \mathbb{R})$, com norma 1, e o mesmo é verdade para $\psi \mapsto e(\psi)(z)$. Para todo $\varepsilon > 0$, escolha k tal que $||\varphi - \varphi_k||_0 < \varepsilon/3$. Então se n é grande o suficiente,

$$\begin{aligned} |e_n(\varphi)(z) - e(\varphi)(z)| &\leq |e_n(\varphi)(z) - e_n(\varphi_k)(z)| + |e_n(\varphi_k)(z) - e(\varphi_k(z)| + |e(\varphi_k)(z) - e(\varphi)(z)| \\ &\leq 2||\varphi - \varphi_k||_0 + \varepsilon/3 < 2\varepsilon/3 + \varepsilon/3 = \varepsilon. \end{aligned}$$

$\square$

Seja $\varphi : Z \to \mathbb{R}$ contínua. Por construção, $e_n(\varphi)$ é constante em cada $P_n \in \mathcal{P}_n$, e logo é constante também em cada átomo P de $\mathcal{P}$, para todo $n \geq 1$. Assim, $e(\varphi)$ é constante em $P \cap F_*$ para cada $P \in \mathcal{P}$. Seja $e_n(\varphi)(P_n)$ o valor de $e_n(\varphi)$ em cada $P_n \in \mathcal{P}_n$. Similarmente, $e(\varphi)(P)$ representa o valor de $e(\varphi)$ em $P \cap F_*$ desde que o último conjunto seja não-vazio. Então, desde que (5.14) define $e_n(\varphi)$ em um subconjunto de medida μ total de Z,

$$\int \varphi \, d\mu = \sum_{\mu(P_n)>0} \int \varphi \, d\mu = \sum_{\mu(P_n)>0} \mu(P_n) e_n(\varphi)(P_n) = \int e_n(\varphi) \, d\mu.$$

Observe também que $|e_n(\varphi)| \leq \sup |\varphi| < \infty$ para cada $n \geq 1$. Assim, nós podemos usar o Teorema da Convergência Dominada para concluir que

$$\int \varphi \, d\mu = \int e(\varphi) \, d\mu. \tag{35}$$

Agora, estamos em condições de construir um sistema de medidas condicionais para μ. Seja P qualquer átomo de P tal que $P \cap F_*$ é não-vazio. É fácil ver que

$$C^0(Z, \mathbb{R}) \ni \varphi \to e(\varphi)(P) \in \mathbb{R}$$

é um funcional não-negativo de $C^0(Z, \mathbb{R})$. De $e_n(1)(P) = 1$ e do Teorema de Representação de Riez (Teorema 21), existe uma única medida de probabilidade μ_P em Z tal que

$$\int \varphi \; d\mu_P = e(\varphi)(P). \tag{36}$$

Devemos definir μ_P mesmo quando P não intersecta F_*. Neste caso tomaremos μ_P como qualquer probabilidade em Z: desde que o conjunto de todos esses átomos P tem medida $\hat{\mu}$ igual a zero em P (em outras palavras, sua união tem medida μ igual a zero em Z), a escolha não é relevante. De acordo com essas definições, (35) pode ser reescrito como

$$\int \varphi \; d\mu = \int \left(\int \varphi \; d\mu_P \right) \; d\hat{\mu}(P),$$

o fato de que $\mathcal{P} \ni P \mapsto \int \varphi \; d\mu_P$ é uma função mensurável é uma consequência direta de (36). Assim, para concluir que $(\mu_P)_{P \in \mathcal{P}}$ forma um sistema de medidas condicionais de μ com respeito a P resta-nos provar que

Lema 52. *$\mu_P(P) = 1$ para $\hat{\mu}-$quase todo $P \in \mathcal{P}$.*

Para isto, usaremos o seguinte resultado auxiliar:

Lema 53. *dada uma função mensurável limitada $\psi : \; Z \to \mathbb{R}$ existe um conjunto de medida $\hat{\mu}$ total $\mathcal{F}(\psi) \subset \mathcal{P}$ tal que o conjunto $P \cap F_*$ é não vazio então $\int \psi d\mu_P = e(\psi)(P)$, para qualquer $P \in \mathcal{F}(\psi)$.*

Podemos provar agora o Lema 53:

Defina $\mathcal{F}_* = \cap_{k, P_k} \mathcal{F}(\chi_{P_k})$, onde a interseção é tomada sobre o conjunto de todos os átomos $P_k \in \mathcal{P}_k$, e todo $k \geq 1$. Desde que esse conjunto é enumerável, F_* tem medida $\hat{\mu}$ total. Afirmamos que a conclusão do lema vale para todo $P \in \mathcal{F}_*$. De fato, seja $k \geq 1$ e P_k elemento de $\mathcal{P}_k$ que contém P. Pela definição de $\mathcal{F}_*$

$$\mu_P(P_k) = \int \chi_{P_k} \; d\mu_P = e(\chi_{P_k})(P). \tag{37}$$

Para cada $n \geq 1$, seja P_n o átomo que contém P. Dado qualquer $z \in P \cap F_*$,

$$e_n(\chi_{P_k})(z) = \frac{1}{\mu(P_n)} \int_{P_n} \chi_{P_k} \; d\mu.$$

Agora, para cada $n \geq k$ temos que $P_n \subset P_k$, e então o último termo é igual a 1. Assim,

$$e(\chi_{P_k})(P) = e(\chi_{P_k})(z) = \lim_{n\to\infty} e_n(\chi_{P_k})(z) = 1.$$

Substituindo isso em (37) obtemos que $\mu_P(P_k) = 1$ para todo $k \geq 1$. Finalmente,

$$\mu_P(P) = \lim_{k\to\infty} \mu_P(P_k) = 1,$$

pois $P_k,\ k \geq 1$, é uma sequência decrescente cuja interseção é P.

Concluímos, desta maneira, a demonstração do Teorema de Desintegração.

□

10 CONCLUSÃO

De acordo com que foi proposto no trabalho, conseguimos unir várias áreas distintas da matemática e usá-las para resolver nossos problemas de dinâmica. Acabamos de ver algumas das mais importantes ferramentas da Teoria Ergódica, uma área relativamente nova e em grande desenvolvimento na pesquisa científica.

11 REFERÊNCIAS BIBLIOGRÁFICAS

[Fol99] FOLLAND, Gerald B. Real Analysis, Modern Techniques and Their Applications. Second Edition. Wiley-Interscience, 1999.

[Oli14] OLIVEIRA, Krerley; VIANA, Marcelo. Fundamentos da Teoria Ergódica. 1ª Edição. SBM, 2014.

[Rud87] RUDIN, Walter. Real and complex analysis. 3 edition. McGraw-Hill, 1987.