ASPECTOS ELEMENTARES DA TEORIA ERGÓDICA

Ricardo Mañé

A teoria Ergódica estuda transformações que preservam medida e como tal seus resultados costumam ser enunciados na linguagem da Teoria da Medida. O propósito deste artigo é dar uma noção dos métodos da Teoria Ergódica através duma série de exemplos de teoremas de Mecânica, Teoria de Números e Geometria, que são essencialmente corolários de teoremas da Teoria Ergódica e que admitem uma formulação elementar ao alcance de estudantes com conhecimentos básicos de Análise em $\mathbb{R}^n$ e integração de Riemann. Estes exemplos serão dados na seção 1. Nas seções 3 e 4 daremos as idéias básicas das provas de alguns deles. A seção 2 será dedicada à introdução informal de algumas noções da Teoria da Medida, imprescindíveis para desenvolver essas idéias.

1. Exemplos de Teoria Ergódica

Antes de entrar nos enunciados dos teoremas que nos servirão de exemplo da metodologia da Teoria Ergódica, é bom lembrar o conceito de conjunto de medida zero. Um conjunto Q em $\mathbb{R}^n$ é um cubo se é um produto cartesiano de intervalos limitados. O volume, vol(Q), do cubo Q define-se como o produto dos comprimentos dos lados. Dizemos que um conjunto $S \subset R^n$ é de medida zero se para todo $\varepsilon > 0$ existem cubos $Q_1, Q_2, \ldots$ tais que $\bigcup_i Q_i \supset S$ e $\sum_i \text{vol}(Q_i) < \varepsilon$. Dizemos que uma propriedade de pontos dum conjunto $K \subset \mathbb{R}^n$ vale para quase todo ponto de K, se o conjunto de pontos em que não vale é de medida zero. Por exemplo, sabemos que uma função limitada $f: Q \to \mathbb{R}$, onde $Q \subset \mathbb{R}^n$ é um cubo, é Riemann inte-

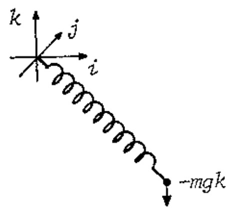
---

(*) Este artigo foi publicado antes no *Noticiário da SBM*, Ano XI, Nº 1, maio de 1980. Reproduzimo-lo aqui, com a permissão do autor, para que os leitores de *Matemática Universitária* tomem contato com uma bela, vigorosa e florescente teoria, apresentada com rara habilidade neste ensaio (N.R.).

grável se e só se é contínua em quase todo ponto de Q.

A seguir introduziremos os exemplos.

Exemplo 1: *O pêndulo-mola.* Seja uma partícula de massa m ligada por uma mola a um ponto fixo. Suponhamos que as únicas forças que agem sobre esta partícula são: a da gravitação, que é $-mg\ k$, onde k é o vetor unitário vertical (ver figura) e g é a constante gravitacional; e a força de restauração da mola. Se lançamos a partícula com condição inicial (x,v) (isto é, do ponto x com vetor velocidade v), a evolução da partícula é descrita por uma função $(x(t), v(t))$, $t > 0$ que diz que no tempo t a partícula encontra-se no ponto $x(t)$ com velocidade $v(t)$. Nosso primeiro teorema estabelece que para quase toda condição inicial a partícula move-se recorrentemente. Informalmente, isto significa que ela retorna infinitas vezes a posições e velocidades arbitrariamente próximas da inicial. Para enunciar rigorosamente esta propriedade comecemos por convir que o ponto fixo ao que está ligada a mola é a origem em $\mathbb{R}^3$. Então a condição inicial (x, v) tem que pertencer a $(\mathbb{R}^3 - \{0\}, \mathbb{R}^3)$. O zero é excluido como possível posição porque isso implicaria reduzir a mola a um único ponto, o que não é fisicamente natural. Dizemos que uma função $x\colon [0,\infty) \ni t \to (x(t), v(t)) \in (\mathbb{R}^3 - \{0\}, \mathbb{R}^3)$ é recorrênte se existem $t_1 < t_2 < \ldots \to +\infty$ tais que



$$\lim_{j\to+\infty} x(t_j) = x(0),$$

$$\lim_{j\to+\infty} v(t_j) = v(0).$$

Teorema 1. *Para quase todo* $(x, v) \in (\mathbb{R} - \{0\}, \mathbb{R})$ *o movimento do pêndulo-mola com condição inicial* (x, v) *é recorrente.*

Naturalmente a prova desta propriedade requer precisar a força que a mola exerce. Denotemos $F(x)$ esta força quando a par tícula está no ponto x. Ao escrever $F(x)$ e não $F(x, v)$ já estamos supondo que a força não depende da velocidade, o que é fisicamente plausível ainda que não seja verdade para velocidades muito grandes. As outras hipóteses serão

I) $F(x)$ é colinear com x;

II) $F(x)$ só depende de x ;

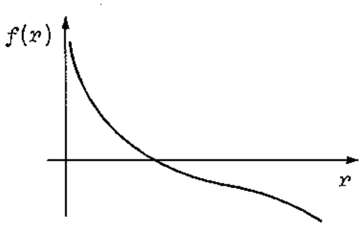
III) Por (I) e (II) podemos escrever:

$$F(x) = f(\|x\|)\frac{x}{\|x\|}$$

onde $f\colon [0, +\infty) \to \mathbb{R}$. Suporemos que para valores grandes de r $f(r)$ é < 0 e monótona decrescente.

IV) $\lim\limits_{r \to 0^+} f(r) = +\infty$

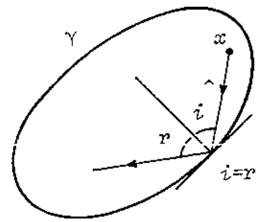
Deixamos ao leitor o exame da credibibilidade das hipóteses I, II e III. A IV merecé algumas explicações. Significa que quando tentamos contrair a mola a um ponto, esta pode desenvolver toda a força que seja necessária para evitá-lo. Fisicamente é acreditável que contrair a mola a um ponto seja impossível, e a hipótese IV é uma forma matematicamente interessante de expressá-lo. Quando na seção 3 demonstrarmos este teorema, usaremos a equação diferencial que descreve o movimento da partícula, e a hipótese IV será fundamental para evitar a existência de soluções chamadas de colisão, que em nosso caso seriam movimentos $(x(t), v(t))$ definidos só para t menor que um certo t^+ e tais que $\lim\limits_{t \to t^+} x(t) = 0$.



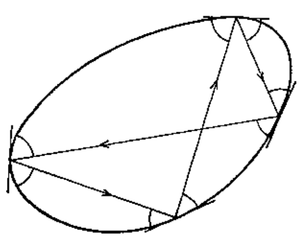
Este fenômeno das soluções de colisão é um problema frequente em questões de Mecânica. Em [9] o leitor pode encontrar uma descrição do interessantíssimo problema das colisões no problema dos corpos, que é onde mais claramente se percebem as complicações decorrentes da existência deste tipo de movimentos.

De fato, o Teorema 1 é só uma das muitas aplicações que tem na Mecânica um teorema demonstrado por Poincaré em 1887, na sua obra *Nouvelles Méthodes de Mécanique Céleste*. Nosso exemplo seguinte também decorre deste teorema.

Exemplo 2: *A mesa de bilhar*. Consideremos uma curva convexa fechada plana diferenciável γ e no seu interior uma partícula (a bola de bilhar) que se move com velocidade constante, exceto quando colide com o bordo, em cujo caso sua velocidade conserva seu valor absoluto mas muda sua direção de forma que o ângulo de incidência é igual ao ângulo de reflexão, ou seja, a normal à curva no ponto do impacto é a bissetriz do ângulo entre os vetores velocidade antes e depois da colisão. Dada a condição inicial (x, v), isto é, se lançamos a partícula do ponto $x \in \Omega$ (onde Ω é o interior da curva) com vetor velocidade $v \neq 0$, esta lei sobre as colisões determina univocamente o movimento $(x(t), v(t))$ para todo $t > 0$. Dizemos que o movimento é periódico se existe $T > 0$ tal que $(x(t+T), v(t+T)) = (x(t), v(t))$ para todo $t > 0$. É equivalente exigir que a trajetória descrita pela partícula seja uma poligonal fechada. Se essa poligonal contém m vértices, dizemos que o período do movimento é m. Na figura vemos um exemplo de movimento periódico com período 4. O seguinte teorema sobre existência de movimentos periódicos é devido a Birkhoff.



Teorema 2. *Para todo $m \geq 2$ existe um movimento periódico de período m.*



Birkhoff provou este teorema com um elegantíssimo argumento totalmente elementar. (Ver [10], "Lecture" 10.) Aqui o demonstraremos a partir de um teorema (também devido a Birkhoff) sobre a existência de pontos fixos de certas transformações de cilindros.

Dizemos que o movimento $(x(t), v(t))$ é recorrente se existem $t_1 < t_2 < \ldots \to +\infty$ tais que

$$\lim_{j\to+\infty} x(t_j) = x(0),$$

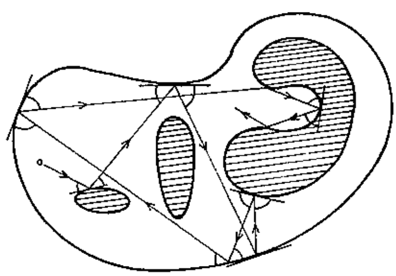
$$\lim_{j\to+\infty} v(t_j) = v(0).$$

O teorema de Poincaré que mencionamos no Exemplo 1 também pode-se aplicar aqui (como veremos na seção 3) e obtemos o seguinte resultado:

Teorema 3. *Para quase toda condição inicial o movimento da bola de bilhar é recorrente.*

De fato, este último teorema vale ainda que a curva γ não seja convexa; e ainda mais, vale se existem obstáculos em Ω, isto é, curvas fechadas contidas em Ω com as quais a bola colide seguindo a mesma lei que descrevemos antes.

A mesa de bilhar continua sendo um tópico de pesquisa de grande interesse. Uma excelente exposição sobre bilhares encontra-se em [10], "Lecture" 10. Mais informação pode ser encontrada em [2]; ou em [8], Cap. II, §6.



Exemplo 3. *Números Decimais.* Seja p um inteiro, $0 \leq p \leq 9$. Tomemos $x \in [0,1)$ e escrevamo-lo em forma decimal:

$$x = 0, x_1 x_2 x_3 \ldots$$

Se $n \geq 1$ definamos $N(n)$ como o número de vezes que p aparece no conjunto $\{x_1,\ldots,x_n\}$.

Teorema 4. *Para quase todo* $x \in [0,1)$ *vale*

$$\lim_{n\to+\infty} \frac{N(n)}{n} = \frac{1}{10}.$$

Ou seja, para quase todo $x \in [0,1)$, cada dígito p ocupa assintoticamente 10% das posições. De certa forma o resultado parece natural ainda que, evidentemente, não seja trivial. Se tivéssemos que advinhar qual é o valor do limite de $N(n)/n$ (supondo que não questionássemos a existência do limite) é claro que diríamos 1/10, só porque intuimos no algorítmo decimal uma certa simetria ou indistinguibilidade entre os dígitos. Muito mais difícil resulta improvisar sobre a mesma questão se o algorítmo decimal é substituido pelo das frações contínuas. Isto é, todo irracional $x\in[0, 1]$ pode-se escrever como:

$$x = \cfrac{1}{n_1 + \cfrac{1}{n_2 + \cfrac{1}{n_3 + \ddots}}}$$

onde $n_1, n_2, \ldots$ são inteiros positivos. Mais precisamente isto significa

$$x = \lim_{j\to+\infty} \cfrac{1}{n_1 + \cfrac{1}{n_2 + \cfrac{1}{\ddots\, n_{j-1} + \cfrac{1}{n_j}}}}.$$

Para detalhes sobre este algorítmo ver [8], Cap. I, §5; ou [3]. Seja então p um inteiro > 0 e $N(m)$ o número de vezes que aparece no conjunto $\{n_1,\ldots,n_m\}$. Qual é o limite de $N(m)/m$ se $m\to+\infty$?

Teorema 5. *Para quase todo* $x \in [0,\ 1)$ *vale*:

$$\lim_{m\to+\infty} \frac{N(n)}{m} = \frac{1}{\log 2} \log \frac{(p+1)^2}{p(p+2)}.$$

Para a prova deste teorema ver [3] ou [8].

Exemplo 4. *Integração*. Neste exemplo estudaremos o problema de calcular a integral

$$\int_0^1 f\, dx$$

de uma função contínua $f\colon [0,\ 1] \to \mathbb{R}$ como médida sobre um subconjunto $x_1, x_2, \ldots$ de pontos de $[0,\ 1]$. O objetivo é descrever sequências tais que para toda função contínua $f\colon [0,\ 1] \to \mathbb{R}$ vale:

$$\lim_{n\to+\infty} \frac{f(x_1) + \ldots + f(x_n)}{n} = \int_0^1 f\, dx. \qquad (*)$$

Estas sequências denominam-se uniformemente distribuídas. Tomemos $x \in (0,\ 1)$ e escrevamo-lo em forma decimal $x = 0,\ t_1\ t_2\ t_3, \ldots$. Definamos:

$$\begin{aligned} x_1 &= 0,\ t_1\ t_2\ t_3\ \ldots \\ x_2 &= 0,\ t_2\ t_3\ \ldots \\ x_3 &= 0,\ t_3\ \ldots \\ &\ \vdots \\ x_n &= 0,\ t_n\ t_{n+1}\ \ldots \\ &\ \vdots \end{aligned}$$

Teorema 6. *Para quase todo* $x \in (0,\ 1)$ *vale a igualdade* $(*)$ *se* $x_1, x_2, \ldots$ *são definidos como acima.*

Outros exemplos de sequências uniformemente distribuídas são a sequência $x_n = x + n\alpha - [x+n\alpha]$, onde α é um irracional e $x \in [0,\ 1]$ ou, se $P(t) = a_n t^n + \ldots + a_1 t + a$ é um polinômio com algum a_j irracional, a sequência $P(n) - [P(n)]$. Provas destes dois últimos resultados encontram-se em [10].

Exemplo 5. *Variedades de Curvatura Negativa.* Até agora, como anunciamos no começo, nossos exemplos têm obedecido a regra de admitir uma formulação numa linguagem matemática elementar. Vamos encerrar nossa série de exemplos com um que foge a essa regra, já que concerne às geodésicas numa variedade compacta M, sem bordo, bidimensional, com curvatura Gaussiana < 0 em todos os seus pontos. A razão pela qual admitimos este exemplo "não elementar" é seu caráter protagonista na teoria ergódica contemporânea. De fato, o teorema que a este respeito mencionaremos é devido a Hedlund e Hopf e data dos anos 30. Mas nos anos 60, Anosov renovou o estudo deste problema, estendendo os teoremas de Hedlund e Hopf a qualquer dimensão e motivando a introdução de um certo tipo de transformação de varieaddes, que Smale batizou de transformações de Anosov, e que tem sido, desde então, um tópico central da teoria ergódica.

Usando a estrutura Riemanniana de M podemos definir a área de abertos $U \subset M$ cujo bordo seja bem comportado; por exemplo, que seja uma união de curvas continuamente diferenciáveis. Denominaremos *simples* este tipo de regiões. Seja $\gamma\colon \mathbb{R} \to M$ uma geodésica. Para todo $T > 0$ o conjunto $\{0 < t < T \mid \gamma(t) \in U\}$ é um aberto em $\mathbb{R}$. Seja $\tau_T(U)$ a soma dos comprimentos das componentes conexas de $\{0 < t < T \mid \gamma(t) \in U\}$.

Teorema 7. *Para quase todo $x \in M$ e quase todo v no espaço tangente a M em p, se $\gamma\colon \mathbb{R} \to M$ é a geodésica com $\gamma(0) = p$ e $\dot{\gamma}(0) = v$, vale*

$$\lim_{T\to+\infty} \frac{1}{T}\,\tau_T(U) = \frac{\text{área}(U)}{\text{área}(M)}$$

para toda região simples U.

Ou seja, que para quase toda condição inicial a geodésica passa em cada região uma percentagem de tempo igual à percentagem da variedade ocupada pela região.

2. Teoria da Medida

Para poder explicar as idéias básicas envolvidas nas provas dos teoremas enunciados na seção anterior precisaremos primeiro introduzir os conceitos fundamentais da teoria da medida. Esta teoria pode ser pensada numa primeira aproximação como um intento de generalização de conceitos como área, volume, massa, etc., que associam a certos conjuntos um número real. Comecemos considerando os conceitos de área ou volume. Se Q é um cubo em $\mathbb{R}^n$, tínhamos definido seu volume, $\mathrm{vol}(Q)$, como o produto dos comprimentos dos lados. Até aqui é fácil concordar com que isto é uma definição de área quando $n = 2$, volume se $n = 3$ e sua generalizações naturais para $n > 3$. O problema que se coloca agora é se é possível estender esta definição (digamos quando $n = 2$) de forma que a cada subconjunto se associe um número que corresponda com nossa noção intuitiva de área. Um primeiro intento é usar a chamada medida exterior μ^* que se define como segue: Se $A \subset \mathbb{R}^n$ definamos $\mu^*(A)$ como o ínfimo das somas $\sum_j \mathrm{vol}(Q_j)$ onde $Q_1, Q_2, \ldots$ é uma família finita ou enumerável de cubos tais que $\bigcup_j Q_j \supset A$. Convencionamos que $\mu^*(A) = \infty$ se a série $\sum_j \mathrm{vol}(Q_j)$ é divergente para toda família de cubos cobrindo A. Esta noção não é satisfatória porque existem pares de conjuntos A e B, limitados e disjuntos, tais que $\mu^*(A \cup B) > \mu^*(A) + \mu^*(B)$. Isto contraria a expectativa de que a área duma união disjunta seja a soma das áreas. Pior ainda é o seguinte fenômeno: existem conjuntos $A_1, A_2, A_3, \ldots$ no disco $D = \{x \in R^2 \mid \|x\| \leq 1\}$ tais que

I) São disjuntos e $\bigcup_{j=1}^{\infty} A_j = D - \{0\}$

II) $\mu^*(A_1) > 0$

III) Para todo n existe uma rotação T_n de R^2 com centro na oriem que transforma A_1 em A_n.

IV) $\mu^*(A_1) = \mu^*(A_2) = \ldots$ (Esta propriedade é de fato um corolário de III).

Este fenômeno mostra claramente que não é bom tomar μ^* como uma definição de área. Não é natural que um conjunto com área finita possa ser dividido em *infinitas* partes disjuntas com a mesma área. Além disso, a existência dos A_j mostra que não existe um conceito de área satisfatório que possa ser aplicado a todo subconjunto de $\mathbb{R}^2$. Se assim fosse teríamos D dividido em infinitas partes com a mesma área (porque as rotações deveriam preservar a área). Isto só poderia acontecer se todos os conjuntos tivessem área zero. Mas uma união enumerável de conjuntos de área zero deveria ter área zero. Ou seja, a área do disco resultaria ser zero.

A solução para estes paradoxos é diminuir nossas pretensões e reconhecer que não faz sentido falar em área de qualquer subconjunto, e que devemos nos restringir aos conjuntos em que faz sentido aplicar o conceito de área. Uma idéia seria usar compactos. Se A e B são compactos disjuntos vale $\mu^*(A \cup B) = \mu^*(A) + \mu^*(B)$. O problema agora é que a classe resulta muito restrita para trabalhar com ela. Resulta insuficiente do mesmo modo que os números racionais são insuficientes para a Análise.

Se tentássemos levar adiante alguma das teorias que se apoiam no conceito de medida (como a teoria da integração ou a teoria ergódica), constantemente encontraríamos o problema de que certas operações entre conjuntos, quando aplicadas a compactos não dão conjuntos compactos; por exemplo, diferenças ou uniões enumeráveis. O problema é então encontrar uma família de subconjuntos de $\mathbb{R}^n$, fechada sob as operações de tomar uniões ou interseções (finitas ou enumeráveis) e sobre a qual a medida exterior tenha a propriedade aditiva: se $A_1, A_2, \ldots$ pertencem à família e são disjuntos então $\mu^*(\bigcup_n A_n) = \bigcup_r \mu^*(A_n)$.

Foi Lebesgue quem, no começo deste século, resolveu este problema. Aqui resolveremos a questão de uma forma ligeiramente diferente, que resulta mais adequada aos propósitos da teoria ergódica. Primeiro introduzimos o conceito de *conjunto boreliano*:

Lema. *Existe uma única família* $\mathcal{B}$ *de subconjuntos de* $\mathbb{R}^n$ *satisfazendo às seguintes condições:*

I) $A \in \mathcal{B} \Longrightarrow$ *seu complementar* $A^c \in \mathcal{B}$;

II) $A_1, A_2, \ldots \in \mathcal{B} \Longrightarrow \bigcup_n A_n \in \mathcal{B}$;

III) *Todo fechado pertence a* $\mathcal{B}$;

IV) *Se* $\mathcal{B}'$ *é outra família de subconjuntos de* $\mathbb{R}^n$ *satisfazendo* (I), (II) e (III) então $\mathcal{B} \subset \mathcal{B}'$.

De (I) e (III) segue que todo aberto pertence a $\mathcal{B}$. Além disso, se $A_1, A_2, \ldots \in \mathcal{B}$ segue de (I) e (II) que $\bigcap_n A_n \in \mathcal{B}$ porque $\bigcap_n A_n = (\bigcup_n A_n^c)^c$.

A prova deste Lema é trivial. É só tomar todas as famílias satisfazendo (I), (II) e (III) (o conjunto de todos os subconjuntos é uma tal família) e intercecioná-las. Os conjuntos de $\mathcal{B}$ denominam-se *borelianos*.

Teorema. *Se* $A_1, A_2, \ldots \in \mathcal{B}$ *vale*

$$\mu^*(\bigcup_n A_n) = \sum_n \mu^*(A_n).$$

O Lema e este Teorema resolvem nosso problema. A solução dada por Lebesgue era diferente. Ele encontrou uma família $\mathcal{B}'$ satisfazendo (I), (II) e (III) (e portanto por (IV) contendo $\mathcal{B}$) onde ainda vale a propriedade do Teorema anterior. Os conjuntos dessa família denominam-se conjuntos mensuráveis à Lebesgue e coincidem com os borelianos a menos de conjuntos de medida exterior zero, isto é, se A é mensurável à Lebesgue pode-se escrever como $A = A_0 \cup A_1 - A_2$, onde A_0 é um boreliano e A_1, A_2 são conjuntos de medida zero.

Introduziremos agora o conceito geral de medida:

Definição. *Seja* $S \subset \mathbb{R}^n$ *um boreliano. Uma medida sobre* S *é uma função* $\mu\colon \mathcal{B}(S) \to [0, \infty]$ *onde* $\mathcal{B}(S)$ *indica a família de todos* *os borelianos contidos em* S, *tal que, se* $A_1, A_2, \ldots \in B(S)$ *são disjuntos, então*

$$\mu\Big(\bigcup_n A_n\Big) = \sum_n \mu(A_n).$$

O teorema anterior diz que μ^* define uma medida em $\mathbb{R}^n$ (que se costuma chamar *medida de Lebesgue*).

3. O Teorema de Recorrência de Poincaré

Nesta seção demonstraremos os Teoremas 1 e 3 da seção 1. Eles decorrem de um teorema geral sobre transformações que preservam medida devido, a Poincaré. Primeiro precisaremos o conceito de transformação que preserva medida. Sejam $S \subset R^n$ um boreliano e μ uma medida em S. Dizemos que uma transformação $T\colon S \circlearrowleft$ *preserva medida* se para todo $A \in B(S)$ vale $T^{-1}(A) \in B$ e $\mu(T^{-1}(A)) = \mu(A)$.

Vejamos alguns exemplos deste tipo de transformações.

Exemplo 1. Se S é aberto e $T\colon S \circlearrowleft$ é uma bijeção continuamente diferenciável, então T preserva a medida de Lebesgue se e só se seu Jacobiano (isto é o determinante da derivada de T) é igual a 1 em todos os pontos.

Exemplo 2. Sejam $U \subset \mathbb{R}^n$ um aberto e $f\colon U \to \mathbb{R}^n$ uma função continuamente derivável. Considere a equação diferencial

$$\dot{x} = f(x).$$

Suponhamos que para todo $p \in U$ exista uma solução $x\colon \mathbb{R} \to U$ desta equação com condição inicial p, isto é, $x(0) = p$. Para todo $t \in \mathbb{R}$ definimos $\phi_t\colon U \circlearrowleft$ como $\phi_t(p) = x(t)$ se x é a solulução da equação com condição inicial p. Seja $\mathrm{div}_p f$ a divergência de f no ponto p, isto é,

$$\mathrm{div}_p f = \frac{\partial f_1}{\partial x_1}(p) + \ldots + \frac{\partial f_n}{\partial x_n}(p)$$

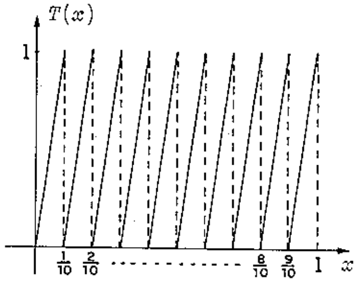
se $f = (f_1,\dots,f_n)$. A fórmula de Liouville estabelece:

$$|\det\,(\phi_t)'(p)| = \int_0^t \operatorname{div}_{\phi_s(p)} f\,ds.$$

Segue que ϕ_t preserva a medida de Lebesgue para todo t se e só se $\operatorname{div} f \equiv 0$.

Exemplo 3. A seguinte transformação será útil na análise dos Exemplos 3 e 4 da seção 1. Seja $\phi\colon [0,1) \circlearrowleft$ definida por

$$\exp \int_0^t \operatorname{div}_{\phi_s(p)} f\,ds.$$



$$\phi(x) = 10\left(x - \frac{n}{10}\right) \quad \text{se} \quad \frac{n}{10} \leq x < \frac{n+1}{10}$$

$$n = 0,1,2,\dots,9$$

Então ϕ preserva a medida de Lebesgue. Isto é claro porque se $A \subset [0,\,1)$ é um boreliano, então $\phi^{-1}(A)$ é igual a $\phi^{-1}(A) \cap [0,\,1/10)$, que é $(1/10)A$ (isto é, o conjunto obtido dividindo por 10 todos os elementos de A) e 9 cópias deste conjunto obtidas por translações. Então $\mu(\phi^{-1}(A)) = 10\mu((1/10)A) =$ $= 10(1/10)\mu(A) = \mu(A)$.

O enunciado do teorema de recorrência de Poincaré é o seguinte:

Teorema. *Sejam* $K \subset \mathbb{R}^n$ *um boreliano e* μ *uma medida em* K *tal que* $\mu(K) < \infty$. *Então, se* $T: K \leftarrow$ *preserva medida, quase todo ponto* x *de* K *é recorrente, isto é, existe uma sequência* $n_1 < n_2 < \ldots \to +\infty$ *tal que*

$$\lim_{j\to+\infty} T^{n_j}(x) = x.$$

O significado de "quase todo ponto" é que existe um boreliano com medida zero, tal que no seu complementar todo ponto é recorrente.

A prova deste teorema é muito simples e o leitor pode encontrá-la em [8], Cap. I, §2; ou em [11].

Apliquemos este teorema às transformações que preservam medida ϕ_t geradas pela equação diferencial do Exemplo 2 desta seção. Tomemos $T = \phi_1$ sendo μ a medida de Lebesgue.

Para poder aplicar o teorema de recorrência temos que supor $\mu(U) < \infty$. Então, para quase todo ponto $p \in U$, podemos encontrar uma sequência $n_j \to +\infty$ tal que $T^{n_j}(p) \to p$, ou seja, $(\phi_1)^{n_j}(p) \to p$.

Por propriedades conhecidas das soluções de equações diferenciais ordinárias, temos $\phi_t\ \phi_s = \phi_{t+s}$ para todo par $t, s \in \mathbb{R}$. Mas $\phi_{n_j}(p)$ é, por definição, $x(n_j)$, onde $x: \mathbb{R} \to U$ é a solução da equação diferencial com condição inicial $x(0) = p$. Concluimos então que para quase toda condição inicial, a solução da equação diferencial é recorrente sempre que $\operatorname{div} f \equiv 0$, $\mu(U) < \infty$ e que todas as soluções estejam definidas sobre todo $\mathbb{R}$.

Vamos aplicar este resultado para provar o Teorema 1 da seção 1. Se $(x(t), v(t))$ descreve o movimento do pêndulo-mola, temos que $\dot{x} = v$ pela definição de velocidade; e pela lei de Newton, o produto da massa m pelo vetor aceleração (definido como $\dot{v}$) deve ser igual à força total que age sobre a partícula que é $F(x) - mgk$. Ou seja

$$\dot{x} = v$$

$$\dot{v} = \frac{1}{m} F(x) - gk$$

Sejam $\Phi: (0, +\infty) \rightleftarrows$ tal que

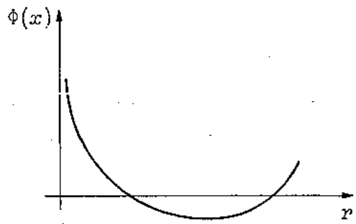
$$\Phi'(r) = -f(r)$$

e $V: \mathbb{R}^3 - \{0\} \to \mathbb{R}$ definida como

$$V(x) = \Phi(\|x\|) + mgx_3.$$

Então

$$\text{grad } V(x) = -(F(x) - mgk).$$



Definamos a função energia $H: (\mathbb{R}^3 - \{0\}) \times \mathbb{R}^3 \to \mathbb{R}$ como

$$H(x, v) = \frac{1}{2} mv^2 + V(x).$$

Se $(x(t), v(t))$ é um movimento do pêndulo-mola temos:

$$\frac{d}{dt} H(x(t), v(t)) = \left\langle \frac{\partial H}{\partial x} (x(t), v(t)), \dot{x}(t) \right\rangle +$$

$$+ \left\langle \frac{\partial H}{\partial v} (x(t), v(t)), \dot{v}(t) \right\rangle =$$

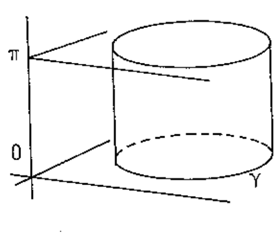
$$= -\langle F(x) - mgk, v \rangle + m\langle v, \dot{v} \rangle = 0.$$

$$= -\langle m\dot{v}, v \rangle + m\langle v, \dot{v} \rangle = 0.$$

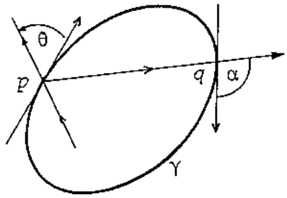
Ou seja, a função energia é constante ao longo de soluções do sistema (*). As hipóteses feitas sobre a mola permitem provar facilmente que $H(x, v) \to +\infty$ se $(x, v) \to \infty$ ou se $x \to 0$. Ou seja, que os conjuntos $U_c = \{(x, v) \mid H(x, v) \leq c\}$ serão compactos contidos em $\mathbb{R}^3-\{0\}\times\mathbb{R}^3$. Isto é fundamental porque permite provar que toda solução do sistema $(*)$ está definida sobre todo $\mathbb{R}$. Se assim não fosse, teríamos soluções $t \longrightarrow (x(t), v(t))$ definidas só em um intervalo $(a, b) \neq \mathbb{R}$ e impossíveis de serem prolongadas a intervalos maiores. Para que isto acontecesse a teoria básica das equações diferenciais estabelece que, se por exemplo $b < +\infty$, dado qualquer compacto K contido no domínio da equação (neste caso $\mathbb{R}^3-\{0\}\times\mathbb{R}^3$) devemos ter $(x(t), v(t)) \notin K$ se t está próximo a b. Mas tomando $K = U_c$ e $c = H(x(0), v(0))$ temos $H(x(t), v(t)) = H(x(0), v(0)) = c$ para todo $t \in (a, b)$, o que significa $(x(t), v(t)) \in U_c = K$ para todo $t \in (a, b)$.

Agora aplicaremos as conclusões que obtivemos sobre a equação $\dot{y} = f(y)$ no caso $y = (x, v)$ e $f(x, v) = (v, m^{-1}F(x)-gk)$ com domínio $U = \mathbb{R}^3-\{0\}\times\mathbb{R}^3$. Como a hipótese $\mu(U) < \infty$ não se cumpre, restrinjamos o domínio a um conjunto $U_c' = \{(x, v) \mid H(x, v) < c\}$. Este conjunto é aberto e o argumento anterior mostra que as soluções da equação estão definidas sobre todo $\mathbb{R}$. A propriedade $\operatorname{div} f \equiv 0$ é de verificação imediata. Então, para quase toda condição inicial em U_c, a solução da equação (que é o movimento do pêndulo mola) é recorrente. Como $\bigcup_{r=1}^{\infty} U_r = U$, isto completa a prova do Teorema 1.

Vejamos agora os teoremas sobre a mesa de bilhar. Para estudar o movimento da partícula consideremos o cilindro $C \subset \mathbb{R}^3$ dado por $C = \{(p, \theta) \mid p \in \gamma, \theta \in [0, \pi]\}$ e definamos uma transformação $T: C \leftarrow$ da seguinte forma: se $(p, \theta) \in C$ e $\theta \in (0, \pi)$, $T(p, \theta)$ será o par (q, α) tal que se a partícula colide com o bordo no ponto p com ângulo de incidência θ, então a colisão seguinte será no ponto q com ângulo de incidência α. (Ver figura.)



Se $\theta = \pi$ ou 0 definimos $T(p, \theta) = (p, \theta)$, qualquer que seja $p \in \gamma$. Birkhoff introduziu uma medida em $\mathcal{C}$ que é preservada por T. Para definir esta medida lembremos primeiro que se $f: C \to \mathbb{R}$ é uma função contínua, então existe uma forma natural de definir a integral de f sobre qualquer subconjunto $S \subset C$ que tenha um bordo bem comportado, digamos uma união de curvas continuamente diferenciáveis. Denotemos essa integral como



$$\int_S f\, d\sigma.$$

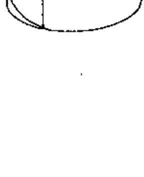
Pode-se provar que se f é sempre ≥ 0 existe uma única medida em C tal que

$$\mu(S) = \int_S f\, d\sigma$$

se C é uma região como as descritas antes.

Seja $\mu = \mu_f$ com f dada por $f(p, \theta) = \operatorname{sen} \theta$. Esta medida é preservada por T. A prova deste fato pode ser encontrada em [10]. Então quase todo ponto $(p, \theta) \in C$ é recorrente. Mas se lançamos a partícula com condição inicial (x, v), tal que sua primeira colisão com o bordo seja em um ponto p com ângulo de incidência θ tais que (p, θ) seja recorrente para T, então o movimento da partícula é recorrente. A partir desta observação é fácil convencer-se (e só um pouco mais difícil demonstrar) que para quase toda condição inicial o movimento da partícula é recorrente.

A existência de movimentos periódicos, que, como mencionamos na seção 1, admite uma prova muito simples e elegante, com recursos totalmente elementares, pode ser deduzida também do teorema seguinte:



Teorema. *Seja* $F: C \circlearrowleft$ *uma aplicação contínua e bijetiva tal que*

1) *A imagem de todo segmento vertical* $r = \{(p, \theta) \mid \theta \in (0, \pi)\}$ *por* F *é uma curva* $F(r)$ *do tipo que indica a figura, isto é, que verticalmente se projeta bijetivamente sobre* $\gamma - \{p\}$.

2) *O conjunto de pontos recorrentes de* F *é denso em* C. *Então, para todo* $m \geq 2$, *existe um ponto* $(p, \theta) \in C$ tal que $F^m(p, \theta) = (p, \theta)$ *e* $F^n(p, \theta) \neq (p, \theta)$ *se* $0 < n < m$.

Apliquemos este teorema a $F = T$. A hipótese (1) pode ser verificada pelo leitor facilmente. A hipótese (2) segue do fato de T preservar a medida μ antes descrita. Se o conjunto de pontos recorrentes não fosse denso seu complementar conteria um aberto U e como todo aberto U contém uma região S do tipo de disco, isto é, com um bordo formado por uma curva fechada simples e continuamente diferenciável. Então

$$\mu(U) \geq \mu(S) = \int_S f\, d\sigma > 0,$$

o que significa que não é verdade que quase todo ponto seja recorrente, contradizendo o teorema de Poincaré.

4. Transformações Ergódicas

Consideremos agora o Teorema 4 da seção 1. Para sua prova usaremos a transformação ϕ definida no Exemplo 3 da seção anterior. Esta transformação tem a propriedade muito fácil de ser verificada, de que se $x = 0, x_1 x_2 x_3 \ldots$ então $x_n = p$ se e só se $\phi^n(x) \in p/10, (p+1)/10)$. Então

$$N(n) = \#\{1 \leq j \leq n \mid \phi^j(x) \in [p/10, (p+1)/10)\}.$$

Ou seja, temos de provar que

$$\lim_{n\to+\infty} \frac{1}{n} \#\{1 \leq j \leq n \mid \phi^j(x) \in [p/10, (p+1)/10)\} = 1/10. \qquad (1)$$

Esta questão é só um caso particular de uma propriedade fundamental na teoria ergódica que se define como segue: Sejam $S \subset \mathbb{R}^n$ um boreliano, μ uma medida em K com $\mu(K) < \infty$ e $T: K \circlearrowleft$ uma transformação que preserva a medida μ. Dizemos que T *é uma transformação ergódica* se para todo boreliano $A \subset K$, vale:

$$\lim_{n\to+\infty} \frac{1}{n} \# \{1 \leq j \leq n \mid T^j(x) \in A\} = \frac{\mu(A)}{\mu(K)}$$

para quase todo x.

Observemos que se T é ergódica e $S \subset K$ é um boreliano tal que $T^{-1}(S) \subset S$ então $\mu(S) = 0$ ou $\mu(K)$. Para provar isto tomamos um boreliano $K_0 \subset K$ com $\mu(K_0) = 0$ e tal que a igualdade anterior valha para $A = S^c$ e todo $x \in K - K_0$. Suponhamos que $\mu(S^c) > 0$. Então não pode valer $S^c \subset K_0$. Tomemos então $x \in S^c - K_0$. Como $x \notin K_0$, vale a igualdade anterior, e como $x \in S^c$ resulta $T^j(x) \in S^c$ para todo $j \geq 1$, pois caso contrário, teríamos $T^j(x) \in S$ o que implicaria $x \in T^{-j}(S)$, que por sua vez está contido em S pela hipótese $T^{-1}(S) \subset S$. Então o limite na igualdade é 1, de onde decorre que $\mu(S^c) = \mu(K)$, ou seja, $\mu(S) = \mu(K - S^c) = \mu(K) - \mu(S^c) = 0$.

Esta condição não só é necessária mas também suficiente.

Teorema (D. Birkhoff, 1931). *T é ergódica se e só se para todo boreliano $S \subset K$ tal que $T^{-1}(S) \subset S$ vale $\mu(S) = 0$ ou $\mu(K)$.*

Em [8] ou [11] podem ser encontradas exposições da prova de Riesz deste teorema que é muito mais simples que a dada originalmente por Birkhoff. Uma exposição interessante comparando várias provas está contida em [3].

Segue deste teorema que se provarmos que ϕ é ergódica, a igualdade (1) estará demonstrada. Para isto usaremos o chamado teorema dos pontos de densidade:

Teorema. *Seja $S \subset \mathbb{R}$ um boreliano, com $\lambda(S) > 0$, λ medida de Lebesgue. Para quase todo ponto $x \in S$ vale a seguinte propriedade:*

Se $I_1 \supset I_2 \supset \ldots$ *são intervalos contendo* x *tais que* $\lim_{n\to+\infty} \mu(I_n)=0$ (*onde* μ *denota a medida de Lebesgue*), *então*

$$\lim_{n\to+\infty} \frac{\mu(S \cap I_n)}{\mu(I_n)} = 1.$$

(Estes pontos chamam-se *pontos de densidade* de S.)

Suponhamos que $S \subset [0, 1]$ é um boreliano que satisfaz $\phi^{-1}(S) \subset S$. Se $\mu(S^c) > 0$, pelo teorema anterior existe um ponto x de densidade de S^c. Tomemos uma sequência de intervalos $I_n = [m_n/10^n, (m_n+1)/10^n]$, onde $0 \leq m_n < 10^n$ é um inteiro, contendo x. Agora observemos que ϕ^n é uma aplicação linear de I_n sobre $[0, 1]$. Então

$$(\phi^n)^{-1}(S) \cap I_n = \frac{1}{10^n} S + \frac{m_n}{10^n}$$

o que implica

$$\mu((\phi^n)^{-1}(S) \cap I_n) = \mu(\frac{1}{10^n} S + \frac{m_n}{10^n}) =$$

$$= \frac{1}{10^n} \mu(S).$$

Mas

$$(\phi^n)^{-1}(S) \subset S$$

e portanto

$$\mu(S \cap I_n) \geq \mu((\phi^n)^{-1}(S) \cap I_n) = \frac{1}{10^n} \mu(S)$$

e

$$\frac{\mu(S \cap I_n)}{\mu(I_n)} = \frac{\mu(S \cap I_n)}{1/10^n} \geq \mu(S).$$

Então

$$\frac{\mu(S^c \cap I_n)}{\mu(I_n)} = \frac{\mu(I_n) - \mu(S \cap I_n)}{\mu(I_n)} \leq 1 - \mu(S).$$

Como o quociente à esquerda converge a 1 quando $n \to +\infty$ (pela definição de ponto de densidade) segue que $1 \leq 1-\mu(S)$, ou seja, $\mu(S) = 0$.

Só resta provar o Teorema 6. Para isso usaremos o seguinte corolário da ergodicidade de ϕ: Para toda f: $[0, 1] \to R$ contínua vale:

$$\lim_{n\to+\infty} \frac{f(\phi(x)) + \ldots + f(\phi^n(x))}{n} = \int_0^1 f\,dx \tag{2}$$

para quase todo x. Primeiro demonstraremos a propriedade quando f é a função característica dum intervalo $J \subset [0, 1]$, isto é, $f(x) = 1$ se $x \in J$ e $f(x) = 0$ se $x \notin J$. Então é fácil ver que:

$$f(\phi(x)) +\ldots+ f(\phi^n(x)) = \#\{1 \leq j \leq n \mid \phi^j(x) \in J\}.$$

Como a integral de f é $\mu(J)$, a fórmula (2) decorre neste caso da ergodicidade de ϕ. No caso em que f é qualquer função contínua, tomemos $\varepsilon > 0$ e sejam $f^- = \sum_{i=1}^m \lambda_i f_i$, $f^+ = \sum_{j=1}^m \lambda_j g_j$ combinações lineares de funções características de intervalos tais que

$$f(x) \leq f^+(x) \leq f(x) + \varepsilon$$

$$f(x) \geq f^-(x) \geq f(x) - \varepsilon$$

para todo x. A propriedade (2) vale para f^+ ou f^- porque vale para funções características de intervalos e os dois lados da igualdade são funções lineares de f. Então

$$\frac{f^-(\phi(x)) + \ldots + f^-(\phi^n(x))}{n} \leq \frac{f(\phi(x)) + \ldots + f(\phi^n(x))}{n} \leq$$

$$\leq \frac{f^+(\phi(x)) + \ldots + f^+(\phi^n(x))}{n}.$$

As expressões das pontas convergem às integrais de f^- e f^+, que diferem em menos de ε da integral de f. Segue que em quase todo ponto os limites superior e inferior da expressão central diferem em menos de ε da integral de f. Como ε é arbitrário, é fácil concluir que o limite dessa expressão existe e coincide com a integral de f em quase todo ponto.

De fato, o argumento anterior só usa a integrabilidade de Riemann de f e não a continuidade. Mais interessante ainda é observar que apenas assumindo que f é a função característica de um boreliano A (não necessariamente um intervalo) vale a fórmula (2) se *definimos*

$$\int_0^1 f\, dx = \mu(A).$$

Suponhamos que *definimos* a integral duma combinação linear $\sum \lambda_i f_i$ de funções características de borelianos A_i como:

$$\int_0^1 (\sum \lambda_i f_i) dx = \sum \lambda_i \mu(A_i).$$

Finalmente, *definimos* uma função f como sendo integrável se para todo $\varepsilon > 0$ existem funções f_ε^+ e f_ε^- que são combinações lineares de funções características de borelianos e tais que $f_\varepsilon^- \leq f \leq f_\varepsilon^- + \varepsilon$, $f_\varepsilon^+ - \varepsilon \leq f \leq f_\varepsilon^+$, e neste caso definimos

$$\int_0^1 f\, dx = \lim_{\varepsilon \to 0} \int_0^1 f_\varepsilon^+\, dx$$

ou, o que é o mesmo,

$$\int_0^1 f\, dx = \lim_{\varepsilon \to 0} \int_0^1 f_\varepsilon^-\, dx.$$

Desta forma obtemos um conceito de integral denominado *integral de Lebesgue*, que estende o de integral de Riemann e para o qual vale (pelo mesmo argumento anterior):

$$\lim_{n \to +\infty} \frac{f(\phi(x)) + \ldots + f(\phi^n(x))}{n} = \int_0^1 f\, dx$$

para quase todo x. De fato, a integral de Lebesgue define-se usualmente usando combinações lineares infinitas de funções características de borelianos disjuntos, isto é, funções da forma $\sum_{i=1}^{\infty} \lambda_i f_i$, o que tem a vantagem de produzir um conceito de integrabilidade que não implica que a função seja limitada (o que acontece com a definição que explicamos antes).

Para a integral de Lebesgue assim definida ainda vale a igualdade (2) em quase todo ponto. Observemos agora que todas estas definições não usam o fato de estarmos trabalhando com funções no intervalo $[0, 1]$. Se o substituimos por um boreliano $K \subset \mathbb{R}^n$ e substituimos a medida de Lebesgue por outra medida μ em K, podemos definir o que é uma função integrável $f: K \to \mathbb{R}$ e qual é sua integral, repetindo a definição anterior sem nenhuma modificação. Se $\phi: K \circlearrowleft$ preserva μ e é ergódica os argumentos anteriores estendem-se sem alterações e concluimos que se $f: K \to R$ é integrável vale

$$\lim_{n \to +\infty} \frac{f(\phi(x)) + \ldots + f(\phi^n(x))}{n} = \frac{1}{\mu(K)} \int_K f \, d\mu \qquad (3)$$

em quase todo x, onde o símbolo à direita denota a integral de f (com respeito à medida μ).

Em muitos casos importantes a transformação T não tem porque ser ergódica. Por exemplo, o caso da mesa de bilhar não dá em geral uma transformação ergódica. Quando a curva γ é uma circunferência, é muito fácil verificar que T preserva o ângulo de incidência, isto é, T é da forma $T(p, \theta) = (F(p, \theta), \theta)$ e portanto todo conjunto da forma $\gamma \times (a, b)$, com $(a, b) \subset [0, \pi]$ é invariante por T. O mesmo vale, por argumentos muito mais complexos, para a maioria das curvas γ. No caso de transformações não ergódicas vale outra versão do teorema de Birkhoff que estabelece que o limite das expressões à direita em (3) existe e define uma função integrável $\tilde{f}(x)$ chamada a média orbital de f.

5. Conclusão

Existe uma extensa e boa literatura sobre Teoria Ergódica. Daremos aqui algumas indicações sobre ela para o leitor interessado em elaborar as idéias esboçadas neste artigo. O livro de Billingsley [3] é excelente. Começando a partir de um nível razoavelmente elementar, desenvolve os conceitos fundamentais da Teoria da Probabilidade, expondo elegantemente suas motivações e aspectos intuitivos até chegar a resultados profundos como o teorema de Birkhoff (do qual, como já mencionamos, apresenta várias provas comparadas) e o teorema de Shannon-McMillan-Breiman, com ajuda do qual introduz a idéia de entropia. Ainda que dedicado essencialmente aos aspectos estatísticos da Teoria Ergódica, contém várias aplicações a outros tópicos (frações contínuas, por exemplo). A relação entre a Teoria Ergódica e a Mecânica Estatística é estudada em [5]. Tratamentos mais atuais (e menos elementares) desta relação encontram-se em [6] e [7]. O livro de Sinai [10] é descrito pelo autor como teoria ergódica "para pedestres" com apenas alguns conhecimentos da Teoria da Medida e variedades. Para a mesma categoria de leitores (quiçá menos "pedestres") foi escrito o livro de Arnold [1]. Ambos são literatura matemática da mais alta categoria. [8] e [11] são livros de texto, isto é, tudo é (ou pretende ser) demonstrado e os conceitos são desenvolvidos com mais detalhes, mas também com mais lentidão. O Capítulo 1 de [8] é o prolongamento natural deste artigo.

Bibliografia

[1] V.I. Arnold, *Méthodes Mathématiques de la Mécanique Classique*, Editions MIR, Moscou, 1976.

[2] V.I. Arnold, A. Avez, *Problèmes Ergodiques de la Mécanique Classique*, Gauthier-Villars, Paris, 1967.

[3] P. Billingsley, *Ergodic Theory and Information*, Wiley Series in Probability and Math. Stat., 1965.

[4] G. Gallavotti, *Lectures on the Billiard*, Dynamical Systems and its Applications, Lecture Notes in Physics 38, Springer Verlag, 1975.

[5] A.I. Khinchin, *Mathematical Foundations of Statistical Mechanics*, Dover, 1949.

[6] O. Lanford, *Time Evolution of Large Classical Systems, Dynamical Systems and its Applications*, Lecture Notes in Physics 38, Springer Verlag, 1975.

[7] G.W. Mackey, *Ergodic Theory and its Significance for Statistical Mechanics and Probability Theory*, Advances in Mathematics 12 (1974) 178-268.

[8] R. Mañe, *Introdução à Teoria Ergódica*, IMPA, Projeto Euclides, Rio de Janeiro, 1983.

[9] D. Saari, *Singularities of Newtonian Gravitational Systems*, Salvador Symposium on Dynamical Systems 479-488, Academic Press, 1973.

[10] Ya. Sinai, *Introduction to Ergodic Theory*, Mathematical Notes, Princeton University Press, 1976.

[11] P. Walters, *Ergodic Theory, Introductory Lectures*. Lecture Notes in Mathematics, 458, Springer Verlag, 1975.

Instituto de Matemática Pura e Aplicada
Estrada Dona Castorina 110 - Jardim Botânico
22.460 Rio de Janeiro, RJ