UNIVERSIDADE DE SÃO PAULO

INSTITUTO DE FÍSICA

ATRATORES E EXPOENTES DE LYAPUNOV

Cesar Rogério de Oliveira

Orientadora: Profa. Dra. Coraci P. Malta

SBI-IFUSP

305M810T1071

Tese de Doutoramento

Apresentada no Instituto

de Física da Universidade

de São Paulo

UNIVERSIDADE DE SÃO PAULO
SERVIÇO DE BIBLIOTECA E INFORMAÇÃO
1071
e.1
INSTITUTO DE FÍSICA

Defesa em 21/12/87

São Paulo

-1987-

530.13
O48a
D
e.1

FICHA CATALOGRÁFICA
Preparada pelo Serviço de Biblioteca e Informação
do Instituto de Física da Universidade de São Paulo

Oliveira, Cesar Rogério de
Atratores e expoentes de Lyapunov. São Paulo, 1987.

Tese (Doutorado) - Universidade de São Paulo. Instituto de Física. Departamento de Física Nuclear.
Área de Concentração: Física Nuclear
Orientador: Profa. Dra. Coraci Pereira Malta

Unitermos: 1.Atratores; 2.Medidas invariantes; 3.Caos; 4.Expoentes de Lyapunov; 5.Teoria ergódica.

USP/IF/SBI - 46/87

A você leitor,

na esperança que obtenha

algum proveito deste trabalho.

AGRADECIMENTOS

-) à Profa. Dra. Coraci P. Malta, particularmente por sua preocupação com minha pessoa nos últimos anos.

-) a todos os amigos do DM-UFSCar.

-) a todos que de alguma forma colaboraram com este trabalho.

RESUMO

Com base em exemplos, nos fundamentos da Mecânica estatística e na teoria ergódica, é dada uma definição de atrator como uma medida invariante. Vários resultados que corroboram esta definição são demonstrados. Caos é relacionado à presença de um atrator com entropia métrica maior que zero.

O papel dos expoentes de Lyapunov é analizado e é provado que um atrator caótico possui expoentes de Lyapunov positivos em quase todo ponto, e também que, se um atrator possui todos expoentes de Lyapunov estritamente negativos num conjunto de medida atratora maior que zero, então seu suporte é uma órbita periódica assintoticamente estável.

ABSTRACT

Here, a definition of an attractor as an invariant measure is given based on Ergodic Theory, foundations of Statistical Mechanics and some examples. Chaos is related to the presence of an attractor with metric entropy grater tham zero.

It is proved that a chaotic attractor has positiv Lyapunov exponents almost everywhere, and that, if an attractor has every Lyapunov exponents less tham zero in a set of nonzero measure then the support set of the attractor is an asymptotic stable periodic orbit.

CONTEÚDO

CAPÍTULO PÁGINA

PRIMEIRO
Introdução e Propósitos 01

SEGUNDO
Alguns Preliminares 09

TERCEIRO
Sobre a Estatística dos Sistemas Dinâmicos - I 15
a) Fundamentos da Mecânica Estatística 15
b) Medidas Invariantes 20
c) Primeiras Propriedades Estatísticas 30

QUARTO
Sobre a Estatística dos Sistemas Dinâmicos - II 38
a) Entropia Topológica 38
b) Entropia Métrica 41
c) Princípio Variacional da Entropia 47

QUINTO
Atratores e Caos 48
a) A Noção de Atrator 48
b) A Definição de Atrator 54
c) Caos 62

SEXTO
Expoentes de Lyapunov Positivos 67

SÉTIMO
Expoentes de Lyapunov Estritamente Negativos 75

OITAVO
Observações Finais e Conclusão 81

Referências 85

For us, there is only the trying.

The rest is not our business.

[E1]

Cobras cegas são notívagas.

O orangotango é profundamente solitário.

Macacos também preferem o isolamento.

Certas árvores só frutificam de 25 em 25 anos.

Andorinhas copulam no vôo.

O mundo não é o que pensamos.

[deA]

CAPÍTULO PRIMEIRO

INTRODUÇÃO E PROPÓSITOS

A motivação maior deste trabalho vem do difícil problema de encontrar maneiras práticas de classificar os comportamentos assintóticos de sistemas dinâmicos. Entre os comportamentos assintóticos que podem ser encontrados destacam-se os comportamentos aleatórios. Características estatísticas em sistemas físicos já são conhecidas há algum tempo, por exemplo: quando um sistema é submetido a forças aleatórias externas, ou quando existe um número muito grande de graus de liberdade, como ocorre normalmente em mecânica estatística (o caso da mecânica quântica é diferente, pois a estatística vem da interpretação que é dada àfunção de onda).

Existe um outro tipo de comportamento aleatório que só vem sendo estudado recentemente, e é possível dizer que inicia-se com o trabalho de Lorenz [Lo], mas Poincaré já o havia percebido quando estudava o famoso problema de três corpos.

O trabalho de Lorenz, hoje muito conhecido, trata da simulação numérica do sistema de três equações diferenciais ordinárias mostrado abaixo,

$$\begin{aligned} \dot{x} &= 10(-x+y) \\ \dot{y} &= -xz+28x-y \\ \dot{z} &= xy-8/3\ z \end{aligned} \qquad \ldots\ldots\ldots\ldots\ldots \quad (1.1)$$

que foi obtido através de uma aproximação grosseira da equação de Navier-Stokes (veja [Lo]). A motivação de Lorenz vinha da meteorologia e o seu modelo era uma tentativa de descrever o movimento do

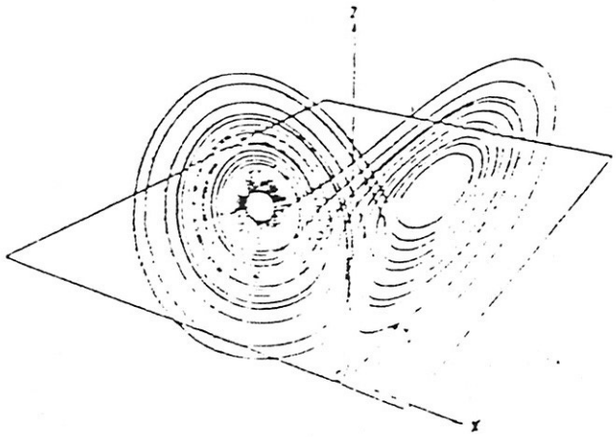

figura 1a

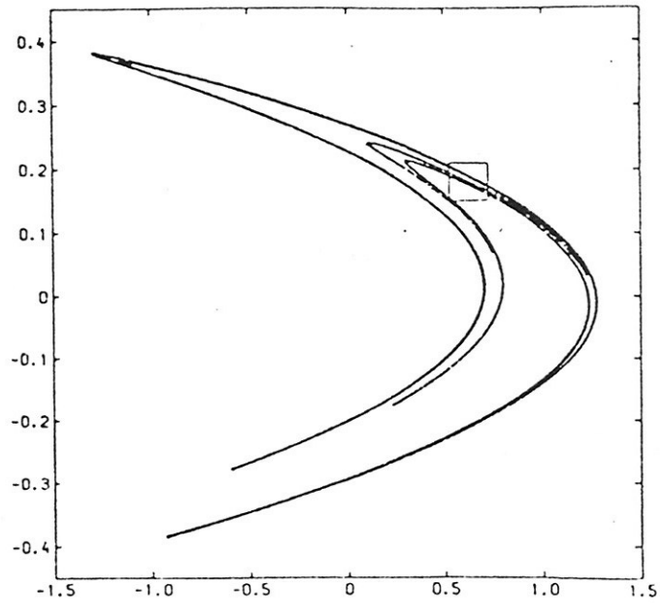

figura 1b

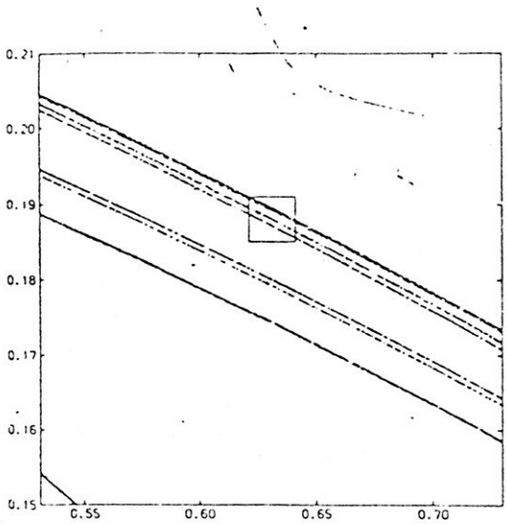

figura 1c

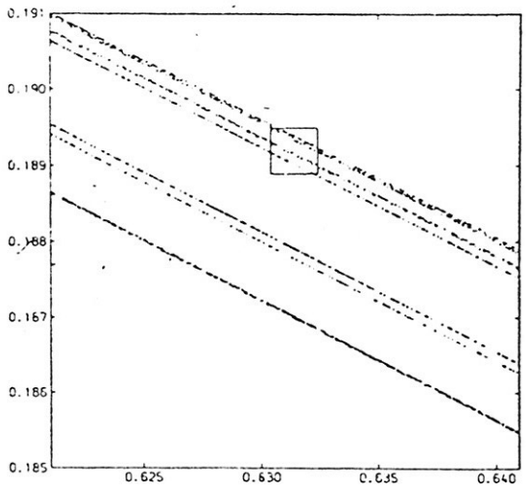

figura 1d

ar na atmosfera.

Uma surpresa foi que as órbitas de (1.1) pareciam não convergir para algo conhecido, mas se enrolavam de maneira a criar belas e complicadas figuras, como a mostrada na figura 1.a.

Nos dias atuais esses tipos de figuras são comuns, e é sabido que equações ou aplicações (tempo discreto) bem simples podem exibir figuras complicadas.

Outra figura que se tornou famosa é a obtida pela aplicação de Hénon [He]

$$x_{n+1} = y_n + 1 - ax_n^2 \qquad \text{.....(1.2)}$$

$$y_{n+1} = bx_n \quad ,$$

onde a=1,4 e b=0,3, reproduzida na figura 1.b

A figura 1.c mostra a ampliação do quadrado indicado em 1.b, ídem para 1.d em relação à 1.c. Localmente estas figuras se parecem com um conjunto de Cantor vezes uma variedade.

Em geral, estas figuras são fractais, ou seja, possuem dimensão fractal não inteira, mas a teoria dos fractais não é estudada aqui, inclusive a relevância física de tal teoria ainda não é clara.

Muito interessante é o fato, já citado acima, de vários sistemas se apresentarem, na prática, imprevisíveis. Esta imprevisibilidade não vem da existência de perturbações externas, nem do grande número de graus de liberdade, mas de não-linearidades. Este comportamento aleatório reflete a ignorância do homem frente as equações diferenciais.

Comportamentos "imprevisíveis", comumente chamados de "caóticos" tem sido observados numa variedade muito grande de sistemas físicos. Nos últimos dez anos, principalmente, apareceram estudos de comportamentos caóticos em fluidos [Lo] [McLM]; lasers [Ha]; física de plasmas [MS], [TM]; reações químicas [TK]; física

do estado sólido [HC]; sistemas estocásticos [AKO] e [Cr]; astrofísica [Co]; transições de fase [AAr]; etc.

Os métodos a serem usados e mesmo os conceitos relevantes para a descrição do "caos" não estão ainda completamente compreendidos.

Neste trabalho investiga-se que conclusões podem ser obtidas a partir dos expoentes de Lyapunov (EL), que são vastamente utilizados como "medida do caos", e são feitas considerações sobre o papel da entropia métrica nos sistemas com características imprevisíveis. É necessário o uso de parte da teoria ergódica, pois somente com esta ferramenta é possível definir os EL e a entropia métrica de forma precisa. Não será dada atenção a modelos físicos particulares.

Os comportamentos "caóticos" não são devidos a transientes, mas ocorrem quando o sistema atinge certo "equilíbrio"; este equilíbrio normalmente se dá pela presença de um "atrator", e é necessária uma definição de atrator que dê significado físico aos expoentes de Lyapunov e à entropia métrica.

Neste trabalho uma teoria estatística dos sistemas dinâmicos é apresentada, tomando como base a teoria ergódica. Estudam-se aplicações contínuas/diferenciáveis $A:X\to X$ onde X é espaço métrico compacto/variedade diferenciável, em particular, o tempo é tomado como um número natural. Em computadores o tempo é sempre discreto, discretizar o tempo pode ser uma forma de fazer aproximações em equações diferenciais, ou ainda, pode-se olhar para alguma seção de Poincaré.

X é visto como o espaço de estados de algum sistema físico, mas para levar em conta os erros experimentais e/ou truncamentos, é mais realista trabalhar com subconjuntos de X e não com pontos individuais; é natural tomar como todos os "estados observá-

veis" uma família de subconjuntos que inclua os abertos da topologia métrica de X.

As considerações probabilísticas são especificadas por uma medida sobre subconjuntos de X, ou seja, a medida de um subconjunto dá a sua probabilidade. A família dos subconjuntos onde faz sentido falar em medida de cada um deles é uma σ-álgebra e, pelo parágrafo anterior, a σ-álgebra importante é aquela que contenha todos os abertos. A menor σ-álgebra que contém os abertos de uma topologia é chamada σ-álgebra de Borel, e os elementos desta σ-álgebra são chamados boreleanos. Quando a frase "o conjunto B de medida..." é usada, está sendo assumido que B é boreleano.

Entre as medidas que se pode definir, aquelas que são preservadas pela aplicação A, $\mu(A^{-1}B)=\mu(B)$, são mais importantes para este trabalho, pois elas indicam uma situação de equilíbrio. Medida invariante é sinônimo de medida preservada.

Medidas de probabilidade são aquelas tais que $\mu(X)=1$, e neste trabalho são consideradas apenas tais medidas, excluindo a medida de Lebesgue nos casos onde é usada como "medida de volume". Se uma propriedade R vale num conjunto de medida μ igual a 1, diz-se que R vale em quase todo ponto e abrevia-se μ-qtp, ou então, que R vale num conjunto de medida μ total.

É possível resumir este trabalho da seguinte forma: Inicialmente é dada uma motivação para o uso da teoria ergódica no estudo do comportamento assintótico dos sistemas dinâmicos; alguns resultados da teoria ergódica relevantes para o trabalho são descritos, e algumas vezes é incluída a motivação destes.

Um atrator é definido como uma medida invariante associada a médias temporais de funções contínuas num conjunto de medida de Lebesgue maior que zero, e caos é relacionado à presença de um atrator com fortes características probabilísticas (veja abai-

xo). Vários resultados que dão suporte a essa definição de atrator são demonstrados.

Uma boa maneira de classificar os atratores seria pelos expoentes de Lyapunov (EL), mas é sabido que tal classificação não é simples. Demonstra-se que EL<0 num conjunto grande (veja abaixo) implica que o suporte do atrator é uma órbita periódica assintoticamente estável. Como corolário deste resultado vem que um atrator caótico possui EL≥0 em quase todo ponto'.

Abaixo são apresentados maiores detalhes do conteúdo deste trabalho.

No segundo capítulo são apresentados uma miscelânea de resultados matemáticos preliminares.

Na seção 3A alguns dos fundamentos da mecânica estatística são discutidos para motivar o uso de medidas invariantes e da teoria ergódica; essa seção é bastante informal. Na seção 3B são estudadas as medidas de probabilidade de aplicações contínuas em espaços métricos compactos, sendo demonstrado, por exemplo, que sempre existe uma medida invariante nestes casos; os conceitos de ergodicidade e mixing também são definidos. Essa seção é a base matemática para os capítulos seguintes. A seção 3C destina-se a mostrar o que é entendido por propriedades estatísticas (das aplicações) diretamente ligadas às medidas invariantes, ergodicidade e mixing. A maior parte das discussões dessa seção é normalmente omitida na literatura.

No quarto capítulo o estudo das propriedades estatísticas é prosseguido, usando os conceitos de entropia topológica e métrica que se baseiam na contagem do número de "órbitas diferentes de um sistema". Na seção 4A é discutida a entropia topológica, que apesar de praticamente não ser utilizada no trabalho, possui grande apelo intuitivo e está relacionada com a entropia métrica, ampla-

mente utilizada.

A entropia métrica, discutida na seção 4B, é um número real positivo que mede quantitativamente o grau de imprevisibilidade, ou informação, de um sistema associado a uma medida invariante. Se a entropia métrica for maior que zero o sistema possui características aleatórias. A seção 4C trata rapidamente da relação entre entropia topológica e métrica. O quarto capítulo é um pouco condensado, e se o leitor desejar pode omitir a leitura da maioria das demonstrações apresentadas, e apenas ler a discussão sobre o significado da entropia, que é o importante.

No capítulo quinto é assumido que X é variedade lisa (variedade C^{∞}),no capítulo sexto A é aplicação C^{1}, e no sétimo capítulo A é considerada C^{2}, mas poderia ser Hölder C^{1}. A compacidade de X é sempre suposta.

Com base em exemplos e nos fundamentos da mecânica estatística, os atratores são definidos como medidas "relevantes", ou seja, aquelas que resultam de médias temporais sobre órbitas com condições iniciais num conjunto de medida de Lebesgue maior que zero; maiores detalhes e discussões estão no capítulo quinto. Alguns resultados que corroboram esta definição são demonstrados.

Na seção 5C a existência de caos em uma aplicação é relacionada à existência de um atrator com entropia métrica maior que zero, em outras palavras, à existência de uma medida invariante com características aleatórias que é atratora. Esta definição é apenas uma formalização precisa da definição de caos que aparece em trabalhos mais empíricos.

De posse de uma medida atratora tem sentido fisicamente falar de entropia métrica e de expoentes de Lyapunov (EL), que são números reais que medem a razão exponencial média que órbitas próximas divergem com o tempo; eles estão bem definidos num

CAPÍTULO SEGUNDO

ALGUNS PRELIMINARES

2A) Sistemas Dinâmicos

A teoria dos sistemas dinâmicos estuda, entre outros tópicos, as propriedades assintóticas de aplicações e, neste trabalho em particular, das aplicações contínuas/diferenciáveis $A:X\to X$, onde X é espaço métrico compacto/variedade lisa (C^{∞}).

Dado um ponto $x\in X$ define-se a órbita de x como o conjunto $o(x)=\{A^n(x): n\in N\}$, onde A^n é o n-ésimo iterado da aplicação, identificando $A^0(x)=x$, e assim, x pode ser chamado de condição inicial.

Uma órbita periódica de período $k\in N$ é um conjunto $\{x_0, x_1, \ldots, x_{k-1}\}$ de elementos de X tal que $A(x_i)=x_{i+1}$, $0\leqslant i<(k-1)$ e $A(x_{k-1})=x_0$; é assumido que $x_i\neq x_j$ se $0\leqslant i\neq j\leqslant(k-1)$. Um ponto x que é uma órbita de período zero é dito ser um ponto fixo de A, e a toda órbita de período k estão associados k pontos fixos distintos de A^k. Esta última observação pode, muitas vezes, simplificar uma demonstração.

Dado um ponto x de X o conjunto ω-limite de x, denotado por $\omega(x)$, consiste de todos os pontos limites da órbita de x, ou seja, $\omega(x)=\bigcap_{n\geqslant 0}\overline{o(A^n(x))}$, onde a barra indica a operação de fecho.

Proposição 2.1 Seja $A:X\to X$ uma aplicação contínua de um espaço métrico compacto e $x\in X$, então $\omega(x)\neq\phi$, $\omega(x)$ é fechado e $A(\omega(x))=\omega(x)$.

Além dos conjuntos ω-limites, o conceito de conjunto não errante de A, denotado por $\Omega(A)$, também é importante no estudo de propriedades assintóticas de A. Um ponto x de X é dito errante

se existe vizinhança aberta V de x tal que os conjuntos V, $A^{-1}(V)$, $A^{-2}(V)$, ..., são disjuntos dois a dois; o conjunto não errante de A é a união de todos os pontos que não são errantes.

Proposição 2.2 Se $A:X\to X$ é contínua e X métrico compacto vale:

(i) $\Omega(A)$ é fechado

(ii) $\omega(x)\subset\Omega(A)$ para todo x em X

(iii) $A(\Omega(A))\subset\Omega(A)$

(iv) toda órbita periódica está contida em $\Omega(A)$

As demonstrações das proposições acima podem ser encontradas na maioria dos livros de sistemas dinâmicos.

Um resultado de topologia geral muito usado neste trabalho é o lema de Urysohn, que pode ser enunciado na seguinte forma:

Proposição 2.3 (Lema de Urysohn) Sejam X um espaço métrico e F e G dois subconjuntos fechados não vazios e disjuntos de X. Então, existe uma função contínua $f:X\to[0,1]$ tal que $f(y)=0$ se y pertence a G e $f^{-1}(1)=F$.

2B) Medidas em Espaços Métricos

O objetivo desta seção é apresentar a relação entre partições e σ-álgebras e três resultados importantes sobre medidas de probabilidades em espaços métricos.

Dado um espaço de medida (X,β,μ), onde X é um espaço métrico, β é uma σ-álgebra e μ uma medida de probabilidade sobre β, define-se uma partição de (X,β,μ) (ou de X) como uma coleção finita de elementos não vazios de β dois a dois disjuntos, tais que a união seja X.

Seja $\eta=(C_1,\ldots,C_k)$ uma partição de X, a coleção de todos os elementos de β que são uniões de elementos de η é uma sub-

σ-álgebra finita de β, denotada por $\beta(\eta)$. Se $\alpha=\{D_1,\ldots,D_f\}$ é uma sub-σ-álgebra finita de β então os conjuntos não vazios da forma $E_1\cap\ldots\cap E_f$, onde E_j denota D_j ou $X\setminus D_j$, formam uma partição finita de X, chamada de a partição associada a α. Desta forma, há uma correspondência biunívoca entre partições e sub-σ-álgebras finitas.

Definição 2.4 Uma medida μ sobre a σ-álgebra de Borel A é dita regular se para qualquer $B\in A$ e $\varepsilon>0$ dado, existem um aberto V e um fechado F com $F\subset B\subset V$ e $\mu(V\setminus F)<\varepsilon$.

Proposição 2.5 Toda medida de probabilidade sobre os boreleanos de um espaço métrico X é regular.

Definição 2.6 Duas medidas de probabilidade μ e ν sobre β são ditas mutuamente singulares se existe $M\in\beta$ tal que $\mu(M)=0$ e $\nu(X\setminus M)=0$.

Definição 2.7 μ é dita absolutamente contínua em relação a ν (denota-se por $\mu\ll\nu$), se $\nu(F)=0$ implica que $\mu(F)=0$.

Teorema 2.8 (Teorema de Decomposiçao de Lebesgue) Sejam μ e ν duas medidas de probabilidade sobre β. Existem um número real p em $[0,1]$ e medidas de probabilidade μ_1 e μ_2 sobre β tais que $\mu=p\mu_1+(1-p)\mu_2$, com $\mu_1\ll\nu$, e μ_2 e ν mutuamente singulares. Tanto p como μ_1 e μ_2 são únicos.

Teorema 2.9 (Teorema de Radon-Nikodyn) Sejam μ e ν medidas de probabilidade sobre β. Então $\mu\ll\nu$ sss (se e somemte se) existe uma função positiva f em $L^1(\nu)$ com $\int f\,d\nu=1$, tal que $\mu(B)=\int_B f\,d\nu$ para todo B em β. A função f é denotada por $d\mu/d\nu$ e é unicamente determinada ν-qtp.

Um caso particular é quando X é variedade lisa e deseja-se definir uma "medida de Lebesgue sobre X", denotada por $\int$, tomando como base a medida de Lebesgue em R^n. Dado um boreleano $B\subset X$ tomam-se cartas locais que cobrem B e define-se a medida de Lebes-

gue de B através de uma forma de volume. Todas as medidas obtidas, quando se considera outras formas de volume, são equivalentes, ou seja, se um conjunto tem medida nula para uma delas, isto ocorrerá com todas as outras medidas obtidas desta forma.

Neste trabalho, salvo casos óbvios, a medida de Lebesgue é utilizada para saber se um conjunto tem medida nula ou não, assim, as formas de volume são equivalentes neste sentido; nestes casos ℓ indica alguma destas medidas.

2C) Teoremas Sobre Integração

Os teoremas abaixo são os mais utilizados para comutar a operação de limite com a de integração.

Teorema 2.10 (Teorema da Convergência Monótona) Se $f_n : X \to R$, $n=1,2,\ldots$, é uma seqüência de funções integráveis tais que $f_n(x)$ é monótona crescente μ-qtp, onde μ é uma medida de probabilidade, e $\sup_n \int f_n \, d\mu < \infty$ então a função $f = \lim_{n\to\infty} f_n$ está definida μ-qtp, é integrável e

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

Teorema 2.11 (Teorema da Convergência Dominada) Sejam μ uma medida de probabilidade e $g : X \to R$ integrável. Se $f_n : X \to R$, $n=1,2,\ldots$ é uma seqüência de funções mensuráveis tal que $|f_n| < g$ μ-qtp e, além disso, $f = \lim_{n\to\infty} f_n$ está definida μ-qtp, então

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

2D) O Teorema de Birkhoff

Teorema 2.12 (Teorema Ergódico de Birkhoff) Se $A : X \to X$ é uma aplicação mensurável que preserva a medida de probabilidade μ e $f \in L^1(\mu)$, então a seqüência $f_n(x) = \sum_{i=0}^{n-1} f(A^i x)/n$ converge μ-qtp à $g \in L^1(\mu)$,

$g \circ A = g$ μ-qtp e $\int f \, d\mu = \int g \, d\mu$.

Este teorema é usado freqüentemente no trabalho, e pode ser considerado como "a base matemática da teoria ergódica". A convergência μ-qtp é a parte difícil da demonstração, que pode ser encontrada em quase todos os livros-texto de teoria ergódica. A demonstração original de Birkhoff é de 1931.

Um enunciado mais rico e um amplo estudo deste teorema pode ser encontrado no livro de K. Petersen [Pet].

2E) Sistemas Hamiltonianos Completamente Integráveis

Quando no texto houver referência a sistemas Hamiltonianos pode-se pensar na "aplicação no tempo um" destes sistemas, já que no texto é considerado que o tempo é discreto. Aplicação no tempo um é o seguinte: dada uma equação diferencial cujas soluções estão definidas em R, considera-se $s(x,t)$ como sendo a solução desta equação no instante t cuja condição inicial em $t=0$ é x. Para t fixo define-se $s_t(x)=s(t,x)$ e pode-se mostrar que esta aplicação é contínua. $s_1(\cdot)$ é a aplicação no tempo um da equação diferencial.

Considere um sistema Hamiltoniano com função de Hamilton $H:X \to R$, onde X é compacto. Uma função F é dita uma integral primeira do sistema se o parêntese de Poisson $\{F,H\}=0$ em todos os pontos de X. Duas funções F e G estão em involução se $\{F,G\}=0$ para todo ponto de X. Uma boa referência para esta seção é o livro do Arnold [Ar], parágrafo 49 e apêndice 8.

<u>Teorema 2.13</u>(Liouville) Se num sistema Hamiltoniano com n graus de liberdade existem n integrais primeiras independentes $F_1, \ldots, F_n$ em involução duas a duas, então:

a) o conjunto de nível $M = \{x: F_i(x)=f_i, \ i=1,\ldots,n\}$ é uma variedade diferenciável invariante pelo fluxo.

b) M é difeomorma a um toro de dimensão n (supondo M conexa)

$$T = \{(\theta_1, \ldots, \theta_n) \bmod 2\pi\}$$

e o fluxo em T é periódico ou quase-periódico, dependendo de certas relações entre os ω_i's definidos abaixo.

c) o fluxo em M pode ser escrito $\dot{\theta} = \omega_i$, $1 \leq i \leq n$, onde $\omega_i \in R$, e as equações de Hamilton podem ser integradas por quadratura.

Existem assim, apenas dois tipos gerais de movimento num sistema que satisfaz o teorema 2.12, ou o movimento é periódico ou quase-periódico, sempre restrito ao toro de dimensão n. Tais sistemas são chamados Sistemas Hamiltonianos Completamente Integráveis.

2F) Algumas Notações

A	σ-álgebra de Borel
M(X)	conjunto das medidas de probabilidade sobre (X,A)
$I(A)$	subconjunto de M(X), das medidas invariantes por A
μ-qtp	μ quase todo ponto
$h_\mu(A)$	entropia métrica de $\mu \in I(A)$
$su(\mu)$	suporte da medida μ
A_*	aplicação de $M(X) \to M(X)$ dada por $(A_*\mu)(B) = \mu(A^{-1}B)$
ℓ	medida de Lebesgue
o(x)	órbita do ponto x
C(X)	conjunto das funções reais contínuas sobre X
$(S_n f)(x) = \sum_{i=0}^{n-1} f(A^i x)$	
sss	se e somente se
∂B	fronteira do conjunto B
χ_B	função característica do conjunto B
δ_x	medida concentrada no ponto $x \in X$
G_μ	conjunto de atração de $\mu \in I(A)$
$\int f \, d\mu = \int_X f \, d\mu = \mu(f)$	

CAPÍTULO TERCEIRO

SOBRE A ESTATÍSTICA DOS SISTEMAS DINÂMICOS - I

Baseado nos fundamentos da mecânica estatística, uma proposta para o estudo estatístico dos sistemas dinâmicos é desenvolvida, e a teoria ergódica surge naturalmente; em seguida é feito um estudo do conjunto das medidas invariantes por uma aplicação contínua $A:X\to X$, onde X é um espaço métrico compacto.

Algumas propriedades consideradas estatísticas são discutidas, e os resultados aqui apresentados não foram desenvolvidos por uma só pessoa ou em um curto espaço de tempo, mas são frutos de trabalhos de muitos especialistas e da intuição de diversas pessoas, e isto torna o assunto um "tanto abstrato" numa primeira leitura. Neste capítulo não há resultados originais, embora uma parte da seção 3C não apareça na literatura.

3A) Fundamentos da Mecânica Estatística

A tentativa de descrever o comportamento de um sistema macroscópico a partir de seus constituintes (átomos e moléculas) através da mecânica estatística esbarra em algumas dificuldades, como: não é possível medir o estado inicial de um sistema com precisão infinita e, em geral, há um número muito grande de graus de liberdade.

O efeito do grande número de graus de liberdade é claro no movimento Browniano, cujas equações de movimento poderiam ser, em princípio, resolvidas mas, na prática o sistema é imprevisível, embora a solução seja única se forem dadas condições ini-

ciais (estes sistemas são chamados determinísticos). É possível levar em consideração o tamanho macroscópico destes sistemas através da teoria de probabilidades e/ou do limite termodinâmico. O objetivo do limite termodinâmico em mecânica estatística é demonstrar que quando o número de átomos, ou moléculas, tende ao infinito (mantendo certas razões constantes) as leis da termodinâmica podem ser aplicadas.

"Aviso ao leitor que nesta seção é feita uma exposição informal, servindo de motivação para o desenvolvimento apresentado nas próximas seções deste capítulo, que formam a base matemática de todo o trabalho. Grande parte desta seção é baseada no artigo de O.Penrose [Pe], e para uma descrição mais detalhada das ligações da mecânica estatística com a termodinâmica sugiro a monografia de Martin-Löf [MaL]".

Já foi mencionada a teoria de probabilidades, mas existem duas maneiras tradicionais de introduzi-la em mecânica estatística, uma desenvolvida principalmente por Boltzmann e a outra por Gibbs. O método de Boltzmann toma como unidade estatística a molécula e a estatística vem do grande número de moléculas constituindo o sistema termodinâmico. O método de Gibbs toma o próprio sistema termodinâmico como unidade estatística, e a estatística surge imaginando-se um conjunto de tais sistemas, chamado ensemble.

Acontece que comportamentos aparentemente imprevisíveis ocorrem em sistemas com poucos graus de liberdade, como no clássico problema dos três corpos, e parecem surgir da precisão finita das condições iniciais; veja o capítulo sexto. É bom ressaltar que este tipo de comportamento, se ocorre, é inevitável, pois não se trata de melhorar a precisão da medida das condições iniciais!

Os sistemas com poucos graus de liberdade deste trabalho são as aplicações contínuas sobre espaços métricos compactos,

ou aplicações diferenciáveis em variedades, e é esperado que uma generalização do método de Gibbs seja adequada ao estudo de suas propriedades estatísticas.

Como dito anteriormente, no método de Gibbs a estatística surge de um conjunto de sistemas idênticos, ou melhor, imagina-se um conjunto de sistemas obedecendo as mesmas leis de movimento mas não necessariamente com as mesmas condiçoes iniciais (um ensemble). A probabilidade de um dado resultado ser observado neste ensemble seria a fração de vezes que o resultado correspondente ocorre numa medição sobre este conjunto, realizada num mesmo instante de tempo.

Para facilitar analogias é suposto que o tempo é discreto, ou seja, o tempo t é um número inteiro.

Em mecânica clássica o estado de um sistema é especificado por um ponto x em um espaço X de dimensão (2k-1) (k seria o número de graus de liberdade do sistema e X uma superfície de energia constante), e o estado de um ensemble é descrito por uma função $g:X\times Z\to R$ que dá a densidade de probabilidade com que os sistemas do ensemble estão distribuídos em X no instante t.

No parágrafo acima foi usado o fato de os sistemas em mecânica clássica preservarem uma medida s absolutamente contínua em relação àmedida de Lebesgue $\mathcal{J}$ sobre X, onde $ds = d\mathcal{J}/\|\text{grad } H\|$, H é a Hamiltoniana e $\|\cdot\|$ indica a norma usual em R^{2K}. Veja, por exemplo, [Khi].

Supondo que uma grandeza física seja descrita por uma função contínua $f:X\to R$, o seu valor médio no ensemble no tempo t é

$$g_t(f) = \int f(x)g(x,t)\, ds, \qquad \text{.....(3.1)}$$

e espera-se que no instante t os valores medidos de f estejam próximos de $g_t(f)$.

Qual $g(\cdot,\cdot)$ corresponde a uma situação física dada?

Esta questão é fundamental e nela reside o surgimento da teoria ergódica!

Se no tempo inicial zero um sistema está em x então no instante t ele estará em x_t, dado pelas soluções das equações de Hamilton. Usando o fato dos sistemas preservarem a medida s, veja [Pe], pode-se mostrar que $g(x,t) = g(x_{-t},0)$, e assim

$$g_t(f) = \int f(x)g(x_{-t},0)\, ds \quad \ldots\ldots\ldots\ldots(3.2)$$

É esperado que (3.1) tenha um "limite de equilíbrio" para t tendendo ao infinito, mas este limite não existe em geral, como pode ser observado no famoso exemplo do oscilador harmônico, onde para f é tomada a coordenada de posição, que oscila indefinidamente.

Para evitar quantidades que não tenham limites define-se o valor de equilíbrio de f como a média temporal

$$g(f) = \lim_{n\to\infty} \sum_{i=0}^{n-1} g_i(f)\,/n, \quad \ldots\ldots\ldots\ldots(3.3)$$

o que é intuitivamente razoável, existe $\wp$-qtp para os sistemas de mecânica clássica e coincide com $\lim_{n\to\infty} g_n(f)$, quando este existe.

Desta forma, é suposto que o que seria observado num experimento é $g(f)$; na realidade, se o comportamento dinâmico é muito complicado pensa-se que esta é a informação que pode ser obtida.

A verificação de que (3.3) está bem definido é uma aplicação do teorema de Birkhoff. Substituindo (3.1) em (3.3) vem

$$g(f) = \lim_{n\to\infty} \sum_{i=0}^{n-1} \int f(x)g(x,i)\, ds\,/n = \int f(x)\bar{g}(x)\, ds \quad \ldots\ldots\ldots(3.4)$$

Do teorema de Birkhoff

$$\bar{g}(x) = \lim_{n\to\infty} \sum_{i=0}^{n-1} g(x,i)\,/n \quad \ldots\ldots\ldots\ldots(3.5)$$

está bem definido $\wp$-qtp, e como f é contínua as operações em (3.4) podem ser justificadas. Deste teorema também segue que $\bar{g}$ é constante sobre as órbitas, isto é, $\bar{g}(x)=\bar{g}(x_t)$, implicando que a medida $\bar{g}ds$ também é conservada.

Se para toda função conveniente $g(\cdot,\cdot)$ o limite (3.5) converge para a mesma função $\bar{g}$ o sistema, ou a medida $\bar{g}ds$, é chamado(a) ergódico(a). Assim, para os sistemas ergódicos em mecânica estatística há apenas uma densidade importante, ou melhor, apenas uma medida invariante ($\bar{g}ds$), importante.

A condição de ergodicidade não é suficiente para garantir que $\lim_{n\to\infty} g_n(f) = \bar{g}(f)$, ou seja, que densidades tendam a um limite de equilíbrio sem que haja necessidade de uma média temporal.

Um sistema, ou a medida $\bar{g}ds$, é mixing se

$$\lim_{n\to\infty} g_n(f) = \int f(x)\bar{g}(x)\, ds = g(f) \quad \ldots\ldots\ldots(3.6)$$

independente da $g(\cdot,\cdot)$ tomada. Mixing é uma condição mais forte que ergodicidade, no sentido de que uma medida mixing é ergódica, mas a recíproca não é verdadeira. Pense no exemplo do oscilador harmônico. Na próxima seção estas idéias estão formuladas precisamente.

A definição original de mixing é de Gibbs, e ele mesmo percebeu que a densidade $g(\cdot,\cdot)$ não se aproxima de seu limite $\bar{g}(\cdot)$ pontualmente, mas pelo "limite fraco" indicado em (3.6).

Com base na discussão desta seção é feita uma generalização (na seção 3B) das idéias envolvidas para as aplicações contínuas $A:X\to X$, onde X é métrico compacto. O objetivo remoto seria entender as propriedades estatísticas destas aplicações, das figuras complicadas e bonitas mostradas no primeiro capítulo e da turbulência em fluidos.

Deseja-se estudar os movimentos irregulares "trocando" a aplicação por um ensemble, de maneira análoga à mecânica estatística.

3B) Medidas Invariantes

Sejam $A:X\to X$ uma aplicação contínua onde X é um espaço métrico compacto com métrica d, e A a σ-álgebra de Borel de X. Sendo A contínua, é mensurável.

Para generalizar as idéias da seção anterior deve-se:

(i) pensar num ensemble de aplicações $A:X\to X$

(ii) procurar medidas invariantes por A

(iii) introduzir uma topologia no conjunto das medidas de maneira que a convergência seja análoga àquela de (3.4)

(iv) definir as propriedades de ergodicidade e mixing

Por enquanto, não é possível dizer quais as propriedades que uma medida deve satisfazer para que seja considerada importante neste contexto, e o que é uma medida importante só será apresentado no capítulo quinto, quando for introduzido o conceito de atrator. Desta forma, deve-se estudar todo o conjunto de medidas, em particular o conjunto das medidas invariantes por A, bem como suas propriedades estatísticas de uma maneira geral. Isto é feito neste capítulo.

O conjunto de todas as medidas de probabilidade sobre (X,A) é denotado por M(X), e se $\mu\in M(X)$, a integral $\int f d\mu=\int_X f d\mu$ também é denotada por $\mu(f)$.

Define-se a aplicação $A_*:M(X)\longrightarrow M(X)$ por $(A_*\mu)(E)=\mu(A^{-1}E)$ para $\mu\in M(X)$ e qualquer boreleano E. O conjunto das medidas de probabilidade invariantes por A é o conjunto $I(A)=\{\mu\in M(X):A_*\mu=\mu\}$. Denota-se por C(X) o espaço vetorial das funções contínuas $f:X\to R$ provido da norma $\|f\| = \sup_{x\in X} |f(x)|$.

Surgem duas perguntas: 1) Como determinar as medidas usando funções contínuas (que representariam grandezas físicas) e, 2) será que sempre existe uma medida invariante por A? A relevância

física da pergunta 2) é discutida na seção 3C).

Como todo elemento de M(X) é regular (proposição 2.5) vem que

Proposição 3.1 Se $\mu \in M(X)$ e E é um boreleano, valem:

$$\mu(E)=\inf\{\mu(U): U \text{ é aberto que contém } E\} =$$
$$=\sup\{\mu(F): F \text{ é fechado contido em } E\}.$$

Portanto é suficiente conhecer $\mu(F)$ para todo fechado F de X, ou então $\mu(U)$ para todo aberto U de X, para conhecer μ. Usando isto, é possível obter um resultado importante para o desenvolvimento deste trabalho, ou seja, é possível determinar cada elemento de M(X) integrando funções contínuas, dando assim, uma resposta à primeira questão formulada acima.

Teorema 3.2 Sejam μ, ν elementos de M(X), se $\mu(f)=\nu(f)$ para qualquer f em C(X), então $\mu=\nu$.

Prova Conhecer uma medida μ é conhecer a medida de todo boreleano, mas pela proposição 3.1 basta conhecer a medida de todo subconjunto fechado F. Seja $\varepsilon>0$; como μ é regular existe um aberto U que contém F tal que $\mu(U/F)<\varepsilon$. Pelo lema de Urysohn existe uma função contínua u(x) que assume zero em $X\setminus U$, um em F, e $0\leq u(x)\leq 1$ para qualquer x em X. Portanto

$$\nu(F)\leq\nu(u)=\mu(u)\leq\mu(\chi_U)<\mu(F)+\varepsilon.$$

Como as manipulações acima são simétricas em relação a μ e ν, e ε é arbitrário, vem que $\mu(F)=\nu(F)$. ©

Agora as medidas serão relacionadas com o espaço dual de C(X), denotado por C*(X).

Teorema 3.3 (Teorema de Representação de Riesz) Seja $L:C(X)\to R$ um funcional linear contínuo e positivo tal que $L(1)=1$(normalização).

Então existe $\mu \in M(X)$ tal que $L(f)=\mu(f)$ para qualquer f em C(X).

A prova deste teorema será omitida aqui, mas pode ser encontrada em alguns livros-texto de medida e integração; veja por exemplo [Ta].

Dos teoremas 3.2 e 3.3 conclui-se que há uma bijeção entre M(X) e o subconjunto convexo de C*(X) formado pelos funcionais lineares positivos e normalizados sobre C(X). A convexidade de M(X) é facilmente verificada.

Baseado na seção 3A), será introduzida uma topologia (metrizável) em M(X) tal que, $\mu_n \to \mu$ em M(X) é equivalente a integrais de funções contínuas convergirem no sentido de

$\mu_n(f) \to \mu(f)$, para qualquer f em C(X).

Seja a função $d^*: M(X) \times M(X) \to R+$ dada por

$$d^*(\mu,\nu) = \sum_{i=1}^{\infty} [|\mu(f_i)-\nu(f_i)|/(2^i \|f_i\|)],$$

onde $f_1, f_2, \ldots,$ é um conjunto enumerável denso em C(X). d* é uma métrica em M(X) e a topologia dada por d* é chamada topologia fraca*.

Teorema 3.4 Sejam $\{\mu_n\}$ $n>0$ e μ elementos de M(X). São equivalentes:

(i) $\lim_{n\to\infty} d^*(\mu_n,\mu) = 0$

(ii) $\lim_{n\to\infty} \mu_n(f_i) - \mu(f_i) = 0$ para todo $i>0$ (veja acima)

(iii) $\lim_{n\to\infty} \mu_n(f) - \mu(f) = 0$ para toda f em C(X).

Prova Como $|\mu_n(f_i)-\mu(f_i)| \leq 2^i \|f_i\| d^*(\mu_n,\mu)$ a aplicação $\mu \to \mu(f_i)$ é contínua para todo $i>0$, assim (i)→(ii).

Como $\{f_i\}$ é denso em C(X) segue que $\mu \to \mu(f)$ é contínuo para toda f em C(X), ou seja, (ii)→(iii).

Sejam $j \in N$, $\varepsilon > 0$ e n suficientemente grande tais que

$$\sum_{i=j}^{\infty} 1/2^i \leq \varepsilon/2 \quad e \quad |\mu_n(f_i)-\mu(f_i)| \leq \varepsilon/2 \text{ para } 1 \leq i \leq j,$$

$$d^*(\mu_n,\mu) =$$

$$= \sum_{i=1}^{j} (|\mu_n(f_i)-\mu(f_i)|/(2^i\,||f_i||)) + \sum_{i=j+1}^{\infty} (|\mu_n(f_i)-\mu(f_i)|/(2^i\,||f_i||))$$

$$\leqslant \varepsilon/2 + \sum_{i=j+1}^{\infty} 2||f_i||/(2^i\,||f_i||) \leqslant \varepsilon/2 + \varepsilon/2 = \varepsilon .$$

Como ε é arbitrário, segue que (iii)→(i). ☐

O teorema 3.4 mostra que a topologia fraca* é boa para as interpretações físicas, mas será que se $\mu_n \to \mu$ nesta topologia, então $\mu_n(B) \to \mu(B)$ para todo boreleano B?

Proposição 3.5 Sejam $\{\mu_n\}$ e μ elementos de M(X). São equivalentes:

(i) $\mu_n \to \mu$ na topologia fraca*

(ii) $\lim\limits_{n\to\infty} \sup \mu_n(F) \leqslant \mu(F)$ para todo fechado F

(iii) $\lim\limits_{n\to\infty} \inf \mu_n(U) \geqslant \mu(U)$ para todo aberto U

(iv) para todo boreleano V com $\mu(\partial V)=0$, $\mu_n(V) \to \mu(V)$, onde ∂V é a fronteira de V.

Prova (i)→(ii)

Seja F fechado, e para $k \in N$ seja $F_k = \{x \in X \mid d(x,F) < 1/k\}$. Como F_k decresce a F quando $k \to \infty$ vem que $\mu(F_k) \to \mu(F)$. Pelo lema de Urysohn pode-se escolher funções contínuas f_k tais que $f_k = 1$ em F, $f_k = 0$ em $X \setminus F$ e $0 \leqslant f_k \leqslant 1$. Assim, $\lim\limits_{n\to\infty} \sup \mu_n(F) \leqslant \lim\limits_{n\to\infty} \sup \mu_n(f_k) = \mu(f_k) \leqslant \mu(F_k)$. Tomando o limite para $k \to \infty$, vem que $\lim\limits_{n\to\infty} \sup \mu_n(F) \leqslant \mu(F)$.

(ii)→(iii)

Se U é aberto então $X \setminus U$ é fechado. De (ii) vem que $\lim\limits_{n\to\infty} \sup \mu_n(X\setminus U) \leqslant \mu(X\setminus U)$ e, portanto $\lim\limits_{n\to\infty} \inf \mu_n(U) \geqslant \mu(U)$.

(iii) → (ii)

De forma análoga, se F é fechado então $X \setminus F$ é aberto. De (iii) vem que $\lim\limits_{n\to\infty} \inf \mu_n(X\setminus F) \geqslant \mu(X\setminus F)$ e, portanto, $\lim\limits_{n\to\infty} \sup \mu_n(F) \leqslant \mu(F)$. Assim (ii) e (iii) são equivalentes.

(ii) e (iii) $\rightarrow$ (iv)

Seja V tal que $\mu(\partial V)=0$. Então $\mu(\text{int}V)=\mu(V)=\mu(\bar{V})$, onde $\bar{V}$ é o fecho de V e intV seu interior. De (ii) vem $\lim_{n\to\infty}\sup \mu_n(\bar{V})\leq \mu(\bar{V})=\mu(V)$, e de (iii), $\lim_{n\to\infty}\inf \mu_n(\text{int}V)\geq\mu(\text{int}V)=\mu(V)$. Como intV está contido em V vem que $\mu_n(V)\to\mu(V)$.

(iv) (i)

Sejam $f\in C(X)$, $\varepsilon>0$ e números reais $a=a_0<a_1<\ldots<a_k=b$, tais que $a_i-a_{i-1}<\varepsilon/k$ com $a<f<b$. Tome μ_n e μ como no enunciado.

Existem no máximo um número contável de valores v tais que $\mu\{x\in X: f(x)=v\}>0$. Se y é um ponto da fronteira do conjunto

$$G_i=\{x\in X: a_{i-1}<f\leq a_i\},$$

então $f(y)=a_{i-1}$ ou a_i, e vale $\mu(\partial G_i)=0$ para todo i se os a_i's são tomados fora de um conjunto contável.

Tem-se que

$$\sum_{i=1}^{k} a_{i-1}\mu_n(G_i)\leq\int f\,d\mu_n\leq\sum_{i=1}^{k} a_i\mu_n(G_i),$$

que vale inclusive para μ no lugar de μ_n.

Denotando-se Lim para o limite superior ou inferior vem

$$\sum_{i=1}^{k} a_{i-1}\mu(G_i)\leq \text{Lim}_{n\to\infty}\int f\,d\mu_n\leq\sum_{i=1}^{k} a_i\mu(G_i).$$

Como os extremos desta expressão diferem de no máximo ε, vem que o mesmo ocorre para os limites superior e inferior. Como ε e f são arbitrários, o resultado segue. ☐

Teorema 3.6 Se $\mu\in M(X)$ então $\mu\in I(A)$ sss $\int f\circ A\,d\mu=\int f\,d\mu$ para toda f pertencente a C(X).

Prova Se V é um boreleano $(A_*\mu)(\chi_V)=\mu(\chi_V\circ A)$, e assim o resultado vale para qualquer função simples, ou seja, qualquer combinação linear finita de funções características. O resultado vale para qual-

quer função contínua f considerando suas partes positiva e negativa f+ e f-, respectivamente, e escolhendo sequências crescentes de funções simples convergindo pontualmente a f+ e f-. Basta aplicar o teorema da convergência monótona. ☐

Agora será mostrado que a aplicação A_* é contínua na topologia fraca* e que $I(A)$ é não vazio, um resultado importante que responde a segunda pergunta formulada anteriormente.

Proposição 3.7 A aplicação $A_*: M(X) \to M(X)$ é contínua.

Prova Se $f \in C(X)$, considere $\mu_n \to \mu$ na topologia fraca*. Tem-se

$$(A_* \mu_n)(f) = \mu_n(f \circ A) \to \mu(f \circ A) = (A_* \mu)(f),$$

e como f é arbitrária em C(X) vem que $A_* \mu_n \to A_* \mu$. ☐

Teorema 3.8 Seja (ν_n) uma seqüência em M(X), então qualquer ponto limite da seqüência $(\mu_n) = (\sum_{i=0}^{n-1} (A_*^i \nu_n)/n)$ é um elemento de $I(A)$.

Corolário 3.9 $I(A)$ é não vazio.

A demonstração do corolário 3.9, que originalmente é devida a Krylov e Bogolioubov, decorre da demonstração do teorema 3.8 e do lema 3.10, onde é mostrada a existência de pontos limites da seqüência (μ_n).

Prova (Teorema 3.8) Seja μ_{n_j} uma subseqüência convergindo a μ em M(X). A existência de tal subseqüência é garantida pelo lema 3.10. Seja $f \in C(X)$,

$$|(A_* \mu)(f) - \mu(f)| = \lim_{n_j \to \infty} |(A_* \mu_{n_j})(f) - \mu_{n_j}(f)| =$$
$$= \lim_{n_j \to \infty} |\sum_{i=0}^{n_j-1} ((A_*^{i+1} \nu_{n_j})(f) - (A_*^i \nu_{n_j})(f))/n_j| =$$
$$= \lim_{n_j \to \infty} |(A_*^{n_j} \nu_{n_j})(f) - \nu_{n_j}(f)|/n_j$$

$$\langle \lim_{n_j \to \infty} 2\|f\|/n_j = 0.$$

Portanto $A_* \mu = \mu$, já que f é arbitrária. ☐

Lema 3.10 M(X) é compacto na topologia fraca*.

Prova Como M(X) é métrico, será mostrado que qualquer seqüência $\{\mu_n\}$ possui subseqüência convergente, e para isso será usado o fato de existir um subconjunto enumerável $f_1, f_2, \ldots$, denso em C(X). Como antes, $\|f\| = \sup\{|f(x)| : x \in X\}$ para f em C(X).

A seqüência de números reais $\{\mu_n(f_1)\}$ é limitada por $\|f_1\|$, portanto possui subseqüência $\{\mu_n^1(f_1)\}$ convergente. Da mesma forma a seqüência $\{\mu_n^1(f_2)\}$ possui subseqüência $\{\mu_n^2(f_2)\}$ convergente, e inclusive $\{\mu_n^2(f_1)\}$ também é convergente.

Repetindo o raciocínio acima constrói-se uma subseqüência $\{\mu_n^j\}$ de $\{\mu_n\}$ que converge para $f_1, \ldots, f_j$. Desta forma, a seqüência $\{\mu_n^n(f_j)\}$ converge para todo j, e então para todo f em C(X), já que $f_1, f_2, \ldots$, é denso.

Seja $L(f) = \lim_{n \to \infty} \mu_n^n(f)$; como L é um funcional linear positivo limitado e $L(1) = 1$, vem do teorema 3.3 que existe $\mu \in M(X)$ tal que $L(f) = \mu(f)$ para todo f em C(X). ☐

Bem, ainda para generalizar a seção 3A é necessário definir as propriedades de mixing e ergodicidade, tomando como base as equações (3.6) e (3.4), respectivamente.

Definição 3.11 Seja $\mu \in I(A)$. μ é dita ergódica (mixing) se para qualquer $\sigma \in M(X)$ absolutamente contínua em relação a μ, a seqüência $\{\sum_{i=0}^{n-1} A_*^i \sigma / n\}$ ($\{A_*^i \sigma\}$) converge a μ.

Na seção 3A a medida $d\mu$ corresponderia a $\bar{g}ds$. Esta definição torna-se mais natural se for um pouco trabalhada. Por

exemplo, suponha que μ é ergódica, h é uma função positiva em $L^1(\mu)$ e f em C(X). Definindo $\sigma \in M(X)$ por $d\sigma = h\, d\mu/\mu(h)$, vale que

$$\sum_{i=0}^{n-1} (\int (f\circ A^i) h\, d\mu)/n = \sum_{i=0}^{n-1} (\int (f\circ A^i)\, d\sigma)/n \,.\, \mu(h) =$$
$$= \int f\, d((\sum_{i=0}^{n-1} A_*^i \sigma)/n) \,.\, \mu(h) \to \int f\, d\mu \,.\, \int h\, d\mu,$$

ou seja, se h for uma "densidade de probabilidade" ($\int h\, d\mu = 1$) vem o correspondente à equação (3.4) (com $\bar{g}=1$). É interessante notar que vale a volta, pois se $\sum_{i=0}^{n-1} (\int (f\circ A^i) g\, d\mu)/n \to \int f\, d\mu . \int g\, d\mu$ para qualquer f em C(X) e $g \in L^1(\mu)$, então μ é ergódica. Resultados análogos valem para mixing.

Uma outra maneira de caracterizar ergodicidade e mixing é usando funções em $L^2(\mu)$.

Teorema 3.12 Seja $\mu \in I(A)$.

(i) μ é ergódica sss para quaisquer f,g em $L^2(\mu)$

$$\lim_{n\to\infty} (\sum_{i=0}^{n-1} \int (f\circ A^i) g\, d\mu)/n = \int f\, d\mu \,.\, \int g\, d\mu.$$

(ii) μ é mixing sss para quaisquer f,g em $L^2(\mu)$

$$\lim_{n\to\infty} \int (f\circ A^n) g\, d\mu = \int f\, d\mu \,.\, \int g\, d\mu.$$

Existem várias maneiras de se caracterizar ergodicidade e mixing, e dependendo do problema particular envolvido, uma maneira pode ser mais conveniente que as outras. Abaixo são dadas várias caracterizações.

Teorema 3.13 Seja $\mu \in I(A)$. São equivalentes:

(i) μ é ergódica

(ii) para quaisquer B,E em A $\lim_{n\to\infty} (\sum_{i=0}^{n-1} \mu(A^{-i}B \cap E))/n = \mu(B).\mu(E)$

(iii) os únicos elementos B de A com $A^{-1}B=B$ satisfazem $\mu(B)=0$ ou $\mu(B)=1$

(iv) os únicos elementos B de A com $\mu(A^{-1}B \Delta B)=0$ são tais que $\mu(B)=0$ ou $\mu(B)=1$ (Δ indica diferença simétrica)

(v) para todo B em A com $\mu(B)>0$ e $n\in N$ vale $\mu(\bigcup_{j=n}^{\infty} A^{-j}B) = 1$

(vi) para todos B,E em A tais que $\mu(B)>0$ e $\mu(E)>0$ existe $n>0$ com $\mu(A^{-n} B \cap E) > 0$

(vii) sempre que $f\in L^{\infty}(\mu)$ (ou $L^{1}(\mu)$) e $f\circ A=f$ μ-qtp, então f é constante μ-qtp.

Teorema_3.14 Seja $\mu\in I(A)$. São equivalentes:

(i) μ é mixing

(ii) para quaisquer B,E em A $\lim_{n\to\infty}\mu(A^{-n}B \cap E) = \mu(B) \cdot \mu(E)$.

Foi mencionado anteriormente que mixing é uma propriedade mais forte que ergodicidade; o próximo teorema coloca isto de forma precisa.

Teorema_3.15 Seja $\mu\in I(A)$. Se μ é mixing então μ é ergódica.

Prova Aplicando o teorema_3.14(ii) para o boreleano F=B=E tal que $A^{-1}F=F$, vem $\mu(F)=[\mu(F)]^2$, o que implica em $\mu(F)=0$ ou 1. Pelo teorema 3.13(iii) μ é ergódica. ☐

As demonstrações dos teoremas_3.12,3.13 e 3.14 já são bem conhecidas em teoria ergódica e podem ser encontradas na maioria dos livros-texto, por exemplo, em [Wa].

Não se pode deixar de citar alguns exemplos onde os conceitos de ergodicidade e mixing podem ser testados. Se A é a aplicação identidade sobre (X,β,μ), esta será mixing sss todos os elementos da σ-álgebra possuem medida zero ou um.

Se A é a rotação do círculo $A(x)=rx$ (mod 1), a medida de Lebesgue normalizada dx será ergódica sss r não é uma raiz da unidade; além disso, nunca será mixing. Este exemplo ilustra o fato de que ergodicidade não implica necessariamente em mixing.

A verificação destas propriedades e outros exemplos podem ser encontrados nos livros-texto, bem como outras caracterizações de ergodicidade e mixing. Veja, por exemplo, [AA], [Bi] e [Wa]. Aqui a preocupação principal é com os conceitos envolvidos e como a estatística aparece.

Um resultado ligado à mecânica clássica é o seguinte:

Teorema 3.16 Seja $\mu \in I(A)$ tal que $\mu(U)>0$ para todo aberto não vazio U de X. Se μ é ergódica então o conjunto $Z=\{x \in X: o(x)$ é densa em $X\}$ possui medida μ total.

Prova Seja $\{U_n : n>0\}$ uma base da topologia de X. Para que o(x) seja denso em X é necessário e suficiente que $x \in \bigcap_{n\geq 1} \bigcup_{k\geq 0} (A^{-k}U_n)$. Como μ é ergódica e $A^{-1}\bigcup_{k\geq 0} A^{-k}U_n \subset \bigcup_{k\geq 0} A^{-k}U_n$ vale que $\mu(\bigcup_{k\geq 0} A^{-k}U_n) = 0$ ou 1. Como U é aberto e não vazio segue que $\mu(\bigcup_{k\geq 0} A^{-k}U_n)=1$, e portanto $\mu(Z)=1$.

$\square$

Como em mecânica clássica a medida s (veja a seção 3A) dá medida não nula para todo aberto não vazio, $s \ll \ell$ e $\ell \ll s$, vem que se s é ergódica então quase todo ponto em relação à medida de Lebesgue possui órbita densa. Historicamente este resultado é conhecido como quase-ergodicidade.

3C) Primeiras Propriedades Estatísticas

Empiricamente é comum definir-se que uma aplicação, ou um sistema dinâmico em geral, comporta-se estocasticamente quando os dados experimentais, ou numéricos, parecem não possuir uma distribuição assintótica. A maneira que estes comportamentos estocásticos podem surgir será discutida em capítulos posteriores.

O interesse é na análise matemática deste fenômeno, e nesta seção são discutidas algumas propriedades consideradas estatísticas. A primeira pessoa a descobrir como leis determinísticas do movimento podem ser tratadas estatísticamente foi Boltzmann, que introduziu a palavra "ergódico" em 1887. O objetivo principal de Boltzmann era, como ele próprio escreveu:

"O propósito deste artigo é dar uma prova puramente analítica e completamente geral da segunda lei da termodinâmica, assim como descobrir o teorema correspondente em mecânica".

O que é entendido por propriedades estatísticas? Esta seção é destinada a responder parcialmente esta pergunta. A resposta está ligada ao conceito de ensemble e das propriedades desenvolvidas nas seções anteriores deste capítulo. O quarto capítulo é uma continuação da resposta àpergunta acima, por isso o título desta seção é 'Primeiras Propriedades Estatísticas'.

Como antes, $A: X \to X$ é uma aplicação contínua do espaço métrico compacto X, e $I(A)$ é o conjunto das medidas de probabilidade sobre os boreleanos invariantes pela ação de A.

É esperado que uma grandeza física representada por uma função contínua $f: X \to R$ seja observada experimentalmente através de sua média sobre um ensemble. Pelo teorema de Birkhoff é sabido que $\mu(f)=\mu(\bar{f})$ para toda $\mu \in I(A)$, onde $\bar{f}(x) = \lim_{n\to\infty} (\sum_{i=0}^{n-1} f(A^i x))/n$ é a média temporal de f no ponto x. Para facilitar a notação, seja

$(S_n f)(x) = \sum_{i=0}^{n-1} f(A^i x)$. Como $\bar{f}(Ax)=\bar{f}(x)$ μ-qtp vem do teorema 3.13 que, se μ é ergódica, então $\bar{f}(x)=\mu(f)$ μ-qtp. De uma forma mais geral:

Teorema 3.17 Se $\mu \in I(A)$ é ergódica, existe $Z \subset X$ tal que $\mu(Z)=1$ e $\bar{f}(x) = \lim_{n\to\infty} (S_n f)(x)/n = \mu(f)$, para quaisquer $f \in C(X)$ e $x \in Z$.

Prova O resultado segue se for demonstrado num subconjunto denso. Como X é métrico compacto existe um subconjunto denso enumerável $\{u_i\}$ de C(X). Pelo teorema de Birkhoff $\bar{u}_i(x) = \mu(u_i)$ μ-qtp. Como a intersecção enumerável de conjuntos de medida 1 também possui medida 1, o resultado vale no subconjunto denso acima. ∎

Desta maneira, os elementos estatísticos aparecem não só pela idéia de ensemble devida a Gibbs, mas também através de médias temporais ao longo de órbitas, e esta será uma idéia central em todo o desenvolvimento deste trabalho.

Cabe ressaltar que dado um ponto x em X este define unicamente sua órbita, e a estatística só aparece quando são considerados objetos relacionados com a aplicação A, como por exemplo a aplicação A_x, e não através da evolução de pontos isolados.

Definição 3.18 O ponto y de X, ou a órbita de y (o(y)), é dita estatisticamente regular se existe a média $\bar{f}(y)$ para toda função contínua f.

Em geral, a demonstração da existência de pontos estatisticamente regulares não é direta, mas se existir uma medida invariante μ o teorema de Birkhoff garante a existência de médias temporais de funções contínuas μ-qtp, e com uma modificação bem simples do teorema 3.17 é fácil mostrar que existe um conjunto $Z \subset X$

tal que $\mu(Z)=1$ e todo ponto de Z é estatisticamente regular.

Proposição 3.19 Para cada ponto estatisticamente regular $y \in X$ existe uma única medida $\tau_y \in I(A)$ tal que $\bar{f}(y)=\tau_y(f)$ para qualquer $f \in C(X)$.

Prova Seja $L_y : C(X) \to R$ o funcional linear $L_y(f)=\bar{f}(y)$. É simples a verificação de que L_y satisfaz as condições do teorema 3.3 e que $L_y(f) = L_{A^y}(f)$ para qualquer f em C(X). Portanto existe $\tau_y \in I(A)$ tal que $L_y(f)=\tau_y(f)$ para qualquer f em C(X). A unicidade segue do teorema 3.2. ©

A cada medida invariante está associado um conjunto de pontos estatisticamente regulares e, pela proposição 3.19, a cada ponto estatisticamente regular está associada uma medida invariante, assim o problema da existência de medidas invariantes pode ser visto naturalmente como o problema da existência de médias temporais. Desta forma, a existência de medidas invariantes já é uma propriedade estatística.

Existem teoremas sobre a existência de medidas invariantes, sendo o mais famoso o de Krylov e Bogolioubov (Corolário 3.9). Este é um campo ativo de pesquisa, pois além do interesse puramente matemático existe a motivação exposta acima.

Se B é um boreleano é possível interpretar $\mu(B)$, para μ em $I(A)$, como a probabilidade de se encontrar $A^n(x)$ em B para n tomado aleatoriamente em N. $\mu(A^{-1}B)=\mu(B)$ é um tipo de "equação de continuidade", pois há a conservação da probabilidade: a probabilidade de encontrar $A(x)$ em B é igual à soma das probabilidades dos conjuntos disjuntos que são levados em B por A.

Outro resultado interessante que usa apenas o fato da existência de uma medida de probabilidade invariante é o já bem conhecido teorema de recorrência de Poincaré, que pode ser considera-

do o primeiro teorema da teoria ergódica! Este teorema é famoso devido a suas implicações físicas e filosóficas, que não são discutidas aqui.

Teorema 3.20 (Teorema de Recorrência de Poincaré) Sejam B um boreleano e $\mu \in I(A)$. Se P é o conjunto dos pontos $x \in B$ tais que $A^i(x) \in B$ para infinitos valores de $i > 0$, então $\mu(B) = \mu(P)$.

Prova O conjunto P é igual a $B - \bigcup_{i \geq 1} B_i$, onde $B_i = \{x \in B : A^j(x) \notin B$ para $j \geq i\}$.

A idéia é mostrar que todo B_i possui medida μ nula, implicando que $\mu(B) = \mu(P)$.

Bem, $B_i = B - \bigcup_{j \geq i} A^{-j}(B) \subset \bigcup_{j \geq 0} A^{-j}(B) - \bigcup_{j \geq i} A^{-j}(B) =$

$$= \bigcup_{j \geq 0} A^{-j}(B) - A^{-i} \bigcup_{j \geq 0} A^{-j}(B).$$

Como

$$\bigcup_{j \geq i} A^{-j}(B) \subset \bigcup_{j \geq 0} (B),$$

vem que

$$\mu(B_i) = \mu(\bigcup_{j \geq 0} A^{-j}(B)) - \mu(A^{-i} \bigcup_{j \geq 0} A^{-j}(B)) = 0. \ \square$$

Este teorema é uma generalização, para qualquer medida invariante, do que ocorre com uma medida concentrada numa órbita periódica, onde os pontos desta órbita realmente retornam neles mesmos pela aplicação sucessiva de A.

Corolário 3.21 Qualquer elemento de $I(A)$ dá medida 1 ao conjunto não errante de A.

Prova Seja $\{U_n\}$ base da topologia de X. O complementar do conjunto não errante é a união dos U_n tais que $U_n, A^{-1}U_n, A^{-2}U_n, \ldots,$ são dois a dois disjuntos; pelo teorema 3.20 $\mu(U_n) = 0$ para qualquer $\mu \in I(A)$. $\square$

Este corolário mostra que, do ponto de vista de medi-

da, a parte interessante da dinâmica ocorre no conjunto não errante. Se uma medida μ é ergódica, então as médias temporais de grandezas físicas representadas por funções contínuas coincidem μ-qtp, e isto é muito forte, pois com probabilidade um as médias temporais da função contínua f independem da condição inicial. Além disso, do teorema 3.12 é possível calcular estas médias, $\bar{f}(x)=\mu(f)$ - μ-qtp. Historicamente este resultado é formulado como "médias temporais coincidem com médias espaciais".

Supondo ainda a ergodicidade de μ é fácil calcular, para x num conjunto de medida μ total, o tempo médio de visita da órbita de x num boreleano E, definido como

$$\lim_{n\to\infty} \text{cardinalidade}\{j\in N: A^j(x)\in E,\ j<n\}/n = \lim_{n\to\infty} (S_n\chi_E)(x)/n,$$

que pelo teorema de Birkhoff, é igual a $\mu(\chi_E)=\mu(E)$. Assim, se μ é ergódica, a órbita de x (para x num conjunto de medida μ total) visita qualquer subconjunto mensurável de X com freqüência igual à medida deste subconjunto.

A condição de mixing pode ser colocada na forma

$$\lim_{n\to\infty} \mu(A^{-n}E\cap F)/\mu(F) = \mu(E),$$

para todos boreleanos E,F (é claro que $\mu(F)\neq 0$). Isto diz que, do ponto de vista de μ, a aplicação A distribui todo boreleano E uniformemente em X, no sentido de que a fração de elementos de E em F será igual a medida de E, independente de F! Há uma "diluição uniforme" de E em X, como uma gota de vinho se dilui num copo d'água após ser misturada com uma colher. É desta analogia que vem o nome mixing.

Também é possível inferir a natureza estatística dos sistemas analisando o comportamento das funções de correlação. Existência de correlação entre duas grandezas indica uma relação causal entre estas variáveis, mas a condição de mixing garante que a função de correlação

$$\lim_{n\to\infty} [\mu((f\circ A^n).h) - \mu(f).\mu(h)]$$

é zero para todas f,h em C(X). Veja o teorema 3.12.

Considerando (A,μ) como a rotação do círculo com a medida de Lebesgue μ, foi afirmado anteriormente que, se a rotação for racional (órbitas periódicas) então μ não é ergódica nem mixing, mas se a rotação for irracional μ é ergódica mas não é mixing. A dinâmica no caso de rotação irracional é muito simples para ser mixing, mas complicada o suficiente para ser ergódica.

Seja A a aplicação identidade de forma que todo elemento da σ-álgebra de Borel possui medida zero ou um, então esta medida é mixing.

Considerando agora (A,δ_p) onde p é um ponto fixo de A, e δ_p é a medida concentrada em p (vale 1 em todo conjunto que contém p e zero de outra forma), então δ_p é ergódica e, se f,h são elementos de C(X), vale

$$\delta_p((f\circ A^n).h)=f(A^n p).h(p)=f(p).h(p)=\delta_p(f).\delta_p(h),$$

ou seja, vale a propriedade de mixing quando restringe-se às funções contínuas.

Embora tenham um sabor estatístico, os conceitos de ergódico e mixing não são suficientes para dar uma formulação precisa dos comportamentos aleatórios, vistos facilmente em computadores, experiências de laboratório, e na natureza.

O conceito de entropia métrica vem preencher esta lacuna na teoria estatística dos sistemas dinâmicos, e foi introduzido por Kolmogorov há mais de trinta anos, que estava preocupado em encontrar invariantes por conjugação (mod 0).

Para finalizar este capítulo são apresentadas algumas propriedades do conjunto $I(A)$. Se $\mu,\nu \in I(A)$ e $p\in[0,1]$ é fácil ver que $(p\mu+(1-p)\nu)\in I(A)$, ou seja, $I(A)$ é subconjunto convexo de M(X). Seja $\{\mu_n\}$ uma seqüência em $I(A)$ convergindo a μ, como $A_*\mu_n = \mu_n$ e

A_* é contínuo (proposição 3.2), vem que $A_*\mu=\mu$, ou seja, μ também é uma medida invariante.

Teorema 3.22 O subconjunto $I(A)\subset M(X)$ tem as seguintes propriedades:

(i) é convexo

(ii) é compacto

(iii) os membros ergódicos são exatamente os pontos extremos de $I(A)$

(iv) se $\mu,\nu\in I(A)$, $\nu\ll\mu$ e μ é ergódica, então $\mu=\nu$.

Prova As demonstrações de (i) e (ii) estão dadas acima.

(iii)

Se $\mu\in I(A)$ não é ergódica então existe $B\in A$ tal que $A^{-1}B=B$ e $\mu(B)$ não é 0 nem um. Sejam μ_1, μ_2 elementos de $I(A)$ dados por

$$\mu_1(E)=\mu(B\cap E)/\mu(B) \text{ e } \mu_2(E)=\mu((X\setminus B)\cap E)/\mu(X\setminus B).$$

Então $\mu(E)=\mu(B)\mu_1(E)+(1-\mu(B))\mu_2(E)$, ou seja, μ não é ponto extremo.

Agora suponha que μ é ergódica e $\mu=p\mu_1+(1-p)\mu_2$, onde $\mu_1,\mu_2\in I(A)$ e $p\in[0,1]$. Para provar (iii) é necessário mostrar que $p=0$ ou 1. Vê-se que $\mu_1\ll\mu$, então seja $g\geq 0$ em $L^1(\mu)$, a derivada de Radon-Nikodyn $d\mu_1/d\mu$.

Sejam $G=\{x\in X: g(x)<1\}$ e $F=\{x\in X: g(x)>1\}$; será mostrado que $\mu(G)=\mu(F)=0$, concluindo então, que $\mu_1=\mu_2=\mu$.

Bem,

$$\mu_1(G\cap A^{-1}G)+\mu_1(G\setminus A^{-1}G)=\mu_1(G)=$$
$$=\mu_1(A^{-1}G)=\mu_1(A^{-1}G\cap G)+\mu_1(A^{-1}G\setminus G),$$

de onde vem que $\mu_1(G\setminus A^{-1}G)=\mu_1(A^{-1}G\setminus G)$. Como $g<1$ em $G\setminus A^{-1}G$ e $g>1$ em $A^{-1}G\setminus G$ e, desde que

$$\mu(A^{-1}G\setminus G)=\mu(A^{-1}G)-\mu(A^{-1}G\cap G)=\mu(G)-\mu(A^{-1}G\cap G)=\mu(G\setminus A^{-1}G),$$

tem-se que $\mu(G\setminus A^{-1}G)=0=\mu(A^{-1}G\setminus G)$. Portanto $\mu(A^{-1}G\Delta G)=0$, e pela ergodicidade de μ vem que $\mu(G)=0$ ou 1 (veja teorema 3.13(iv)).

Se $\mu(G)=1$ então $\mu_{\perp}(X)=\int_G q\, d\mu < \mu(G)=1$. contradizendo o fato de $\mu_{\perp}(X)=1$; assim, $\mu(G)=0$. De forma análoga vem que $\mu(F)=0$.

(iv)

Segue como corolário da segunda parte da prova de (iii).

∎

Nesta seção foi tentado mostrar um pouco do que está por detrás das expressões matemáticas. As mesmas palavras podem ser repetidas para o conceito de entropia, apresentado no quarto capítulo.

CAPÍTULO QUARTO

SOBRE A ESTATÍSTICA DOS SISTEMAS DINÂMICOS - II (ENTROPIA)

Neste capítulo é prosseguido o estudo das propriedades estatísticas das aplicações contínuas $f:X\to X$, com X métrico compacto. Agora, o grande número de "órbitas diferentes possíveis" da aplicação pode levar à imprevisibilidade de medidas sucessivas de grandezas físicas; a formalização deste conceito é discutida sob alguns pontos de vista, que levam às definições de entropia topológica e métrica.

Um resultado esperado, o Princípio Variacional da Entropia, é rapidamente discutido. A entropia métrica é um número positivo, e se for maior que zero é possível dizer que o sistema possui características imprevisíveis, e isto será muito importante quando for dada a definição de "caos" (veja o capítulo quinto). Este capítulo não traz resultados originais, apenas a maneira de expor o tema é nova, assim várias demonstrações não são apresentadas; o importante é a "noção intuitiva" do conceito de entropia.

4A) Entropia Topológica

Como ressaltado no primeiro capítulo, devido à resolução dos aparelhos de medida, os possíveis "estados observáveis" no espaço métrico compacto X são os boreleanos. Restringindo-se aos abertos e, naturalmente, a uma cobertura de X, pode-se sempre obter pela compacidade de X, uma subcobertura finita.

Seja $C=\{C_1,C_2,\ldots,C_n\}$ uma cobertura de abertos de X.

O objetivo desta seção é fazer uma contagem do número assintótico de órbitas módulo cobertura (a ser definido abaixo) da aplicação contínua $A:X\to X$. Tomando um ponto $x\in X$, se $A^n(x)\in C_{j_n}$, $1\leq j_n\leq k$, define-se uma <u>órbita de x módulo cobertura</u> C como: $C_{j_0}, C_{j_1}, C_{j_2}, \ldots$; isso também significa que $x\in (C_{j_0}\cap A^{-1}C_{j_1}\cap A^{-2}C_{j_2}\cap \ldots)$.

Seja $K(C)$ o número de conjuntos em uma subcobertura da cobertura C de menor cardinalidade. Se B é outra cobertura de X, define-se a junção $C\vee B$ como a cobertura formada pelos conjuntos da forma $C_j\cap B_i$, onde $C_j\in C$ e $B_i\in B$. $\bigvee_{l=1}^{n} Z_l$ é a junção de n coberturas Z_l por abertos de X. Vale que $A^{-1}(C\vee B)= A^{-1}C\vee A^{-1}B$.

Para fazer uma contagem do número assintótico de órbitas módulo cobertura possíveis, deve-se estimar $K(\bigvee_{i=0}^{n-1} A^{-i}C)$ quando n cresce indefinidamente. O seguinte teorema faz tal contagem, onde $H_T(C)$ denota $\log(K(C))$, e para evitar possíveis confusões $\log_e=\log=\ln$.

<u>Teorema 4.1</u> Se C é uma cobertura por abertos de X, então o limite $h(A,C) = \lim_{n\to\infty} H_T(K(\bigvee_{i=0}^{n-1} A^{-i}C)/n$ existe.

$H_T(C)$ é chamado de entropia da cobertura C, e $h(A,C)$ de entropia de A relativa a C. A prova do <u>teorema 4.1</u> seguirá facilmente do lema de subaditividade e de propriedades de $H_T(C)$. Uma cobertura Z_2 é dita um refinamento da cobertura Z_1 ($Z_1<Z_2$), se todo elemento de Z_2 é subconjunto de um elemento de Z_1.

<u>Lema 4.2</u> Se C, B são coberturas por abertos de X, valem:

(i) $H_T(C)\geq 0$

(ii) Se $C<B$ então $H_T(C)\leq H_T(B)$

(iii) $H_T(C\vee B)\leq H_T(C)+H_T(B)$

(iv) Se $P:X\to X$ é uma aplicação contínua, então $H_T(P^{-1}C)\leq H_T(C)$.

Lema_4.3 (subaditividade) Se $\{a_n\}_{n=0}^{\infty}$ é uma seqüência de números reais tais que $a_{n+k} \leq a_n + a_k$ para todo n,k, então existe $\lim_{n\to\infty} a_n/n$, e é igual ao $\inf_n a_n/n$.

Prova Fixando $k>0$, cada n pode ser escrito na forma $n=mk+r$, com $0\leq r<k$. Assim, $a_n/n \leq a_{mk+r}/(mk) \leq a_r/(mk) + ma_k/(mk) =$

$$=a_r/(mk) + a_k/k.$$

Como $n\to\infty$ implica que $m\to\infty$, e $\limsup_{n\to\infty} a_n/n \leq a_k/k$, ou ainda, $\limsup_{n\to\infty} a_n/n \leq \inf_k a_k/k$ mas $\inf_k a_k/k \leq \liminf_{n\to\infty} a_n/n$. ☐

Prova (Teorema 4.1) Chamando $a_n = H_T(\bigvee_{i=0}^{n-1} A^{-i}C)$ e usando o lema_4.2 (iii) e (iv) vem que $a_{n+k} \leq a_n + a_k$. Agora, basta aplicar o lema_4.3.☐

Uma aplicação com um número muito grande de órbitas módulo cobertura diferentes é um bom candidato a ter propriedade de "imprevisível".

Definição_4.4 A entropia topológica de A é o número real $h(A)\geq 0$ dado por $h(A) = \sup_C h(A,C)$, onde o C percorre as coberturas (por abertos) finitas de X.

Desta forma, a entropia topológica dá o número máximo possível (assintótico) de órbitas módulo cobertura.

O próximo resultado mostra que $h(A)$ é uma característica da aplicação $A:X\to X$ independente de, por exemplo, mudanças de coordenadas; a entropia topológica é invariante por conjugação topológica.

Teorema_4.5 Se $A_1:X_1\to X_1$ e $A_2:X_2\to X_2$ são aplicações contínuas de espaços métricos compactos X_1 e X_2, e $\psi:X_1\to X_2$ é um homeomorfismo tal

que $\Psi A_1 \Psi^{-1} = A_2$, então $h(A_1)=h(A_2)$.

Prova Seja C uma cobertura por abertos finita de X_2, então

$$h(A_2,C) = \lim_{n\to\infty} H_T\left(\bigvee_{i=0}^{n-1} A_2^{-i}C\right)/n = \lim_{n\to\infty} H_T\left(\bigvee_{i=0}^{n-1} \Psi^{-1} A_2^{-i} C\right)/n,$$

e pelo lema 4.2(iv)

$$h(A_2,C)=\lim_{n\to\infty} H_T\left(\bigvee_{i=0}^{n-1} A_1^{-i}(\Psi^{-1}C)\right)/n = h(A_1, \Psi^{-1}C).$$

Portanto $h(A_1)\geq h(A_2)$, mas como as manipulações acima são simétricas em relação a A_1 e A_2, vem que $h(A_1)=h(A_2)$. ☐

4B) Entropia Métrica

Nesta seção continua-se a "contagem do número de órbitas diferentes possíveis" iniciado na seção anterior. Bem, não é claro o que deva ser um "número grande de órbitas distintas" para que haja instabilidades, mas o problema pode ser resolvido utilizando-se conceitos da teoria de informação, que medem o grau de incerteza de uma aplicação. Abaixo estas idéias estão mais explícitas.

Pode-se dizer que a teoria de informação começou com um trabalho de Shannon em 1948, e desde então tem sido aplicada a vários ramos da ciência, como à teoria dos sistemas dinâmicos, que é o caso deste trabalho. O que Shannon fez foi dar um significado universal à famosa expressão $-\sum_i p_i \log p_i$, embora o primeiro a utilizá-la foi Boltzmann por volta de 1870. Uma excelente exposição da teoria de informação para os propósitos deste trabalho está em [Kh2].

Suponha que haja n eventos mutuamente exclusivos E_1, ...,E_n com probabilidades de ocorrência p_1,...,p_n, respectivamente. Procura-se uma quantidade $H(p_1,...,p_n)$ que possa medir a incerteza deste esquema. Para ficar mais claro o que está sendo chamado de incerteza, é bom um exemplo:

Pergunta_1 Se um dado D_1 que possui em todas suas faces o número 1 é lançado, que número o leitor arriscaria neste lance?

Pergunta_2 Se um dado D_2 possui o número 1 em cinco faces e o número 2 na sexta face é lançado, que número o leitor arrisca agora?

Pergunta_3 Se um dado D_3 possui o número 1 em duas faces e os números 2, 3, 4 e 5 nas outras, respectivamente, que número o leitor arrisca desta vez?

Última_Pergunta Se o dado D_4 é um dado normal (e honesto), que número o leitor arriscaria num lance com D_4?

Vê-se que a incerteza dos resultados varia com o dado que é lançado. Dois fatos são relevantes: o dado D_1 dá como resultado o número 1 com probabilidade total, portanto não existe incerteza, e a incerteza vai aumentando monotonicamente de D_1 a D_4, sendo que é máxima para D_4, onde todas as possibilidades são igualmente prováveis (todos os palpites são equivalentes). A função $H(\cdot)$ quantifica esta idéia.

Suponha que haja dois esquemas $\eta=\{E_1,\ldots,E_n\}$ e $\xi=\{F_1,\ldots,F_m\}$, e deseja-se saber a incerteza associada se um deles é conhecido. Se é sabido que o resultado F_j ocorreu, e que p_{ji} é a probabilidade de E_i ocorrer nesta condição, a incerteza associada é $H_j(\eta)=H(p_{j1},\ldots,p_{jn})$. Se $p_1,\ldots,p_m$ são as respectivas probabilidades associadas a ξ, é razoável tomar como incerteza condicional (a incerteza de η conhecendo ξ) $H(\eta/\xi)$ o valor esperado de $H_j(\eta)$ em ξ, ou seja, $H(\eta/\xi)=\sum_j p_j H_j(\eta)$.

A realização dos dois esquemas η e ξ induz uma incerteza $H(\eta\vee\xi)=H(\xi)+H(\eta/\xi)$, e se η e ξ são probabilisticamente independentes deve-se ter $H(\eta\vee\xi)=H(\eta)+H(\xi)$.

O seguinte teorema, cuja prova pode ser encontrada em [Kh2], mostra a "unicidade" da "função incerteza" $H(\cdot)$.

Teorema 4.6 Sejam $\Omega_\ell=\{(p_1,\dots,p_\ell)\in R^\ell : p_i\geq 0 \text{ e } \sum_{i=1}^{\ell} p_i=1\}$ e $H:\bigcup_{\ell=1}^{\infty}\Omega_\ell\to R$ com as seguintes propriedades:

(i) $H|_{\Omega_\ell}$ (H restrito a Ω_ℓ) é contínua e simétrica

(ii) $H|_{\Omega_\ell}$ toma seu maior valor para $p_i=1/\ell$, $1\leq i\leq \ell$

(iii) $H(\eta\vee\xi)=H(\xi)+H(\eta/\xi)$

(iv) $H(p_1,\dots,p_\ell,0)=H(p_1,\dots,p_\ell)$

(v) $H(p_1,\dots,p_\ell)=0$ sss algum $p_i=1$.

Então $H(p_1,\dots,p_\ell)=\lambda\sum_{i=1}^{\ell} p_i\log p_i$, com $\lambda<0$ fixo ($0\log 0=0$).

Escolhe-se $\lambda=-1$ e chama-se $H(p_1,\dots,p_\ell)=-\sum_{i=1}^{\ell} p_i\log p_i$, de entropia; esta quantidade mede a incerteza ou a quantidade de informação obtida numa medição. Obviamente, qualquer valor de $\lambda<0$ poderia ter sido escolhido.

A introdução dos conceitos acima em teoria ergódica foi feita por Kolmogorov em 1958, mas seu objetivo era encontrar um invariante por isomorfismo (mod 0), e com isso ele foi capaz de responder algumas questões importantes na época.

Deseja-se definir a "entropia de uma aplicação" $A:X\to X$, que tenha uma interpretação probabilística análoga à interpretação de $H(\cdot)$ discutida acima. Considere uma partição finita $\xi=\{B_1,\dots,B_m\}$ como os "estados observáveis" e suponha que a probabilidade de B_i seja $\mu(B_i)$, para $\mu\in M(X)$.

Como existe uma correspondência biunívoca entre partições e sub-σ-álgebras finitas de A, define-se:

Definição 4.7 A entropia da partição $\xi=\{B_1,\dots,B_m\}$ (ou da sub-σ-álgebra associada $A(\xi)$) relativa a $\mu\in M(X)$, é o número

$$H_\mu(\xi)=H_\mu(A(\xi))=-\sum_{i=1}^{m}\mu(B_i)\log\mu(B_i).$$

$H_\mu(\xi)$ mede a incerteza contida na partição ξ. Se são dadas duas partições finitas $\xi=\{B_1,\dots,B_m\}$ e $\eta=\{C_1,\dots,C_k\}$, a entropia de ξ dada η é

$$H_\mu(\xi/\eta)=-\sum_{j=1}^{k}\mu(C_j)\sum_{i=1}^{m}\mu(B_i\cap C_j)/\mu(C_j)\log[\mu(B_i\cap C_j)/\mu(C_j)=$$
$$=-\sum_{i,j}\mu(B_i\cap C_j)\log[\mu(B_i\cap C_j)/\mu(C_j)],$$

omitindo os termos em que $\mu(C_j)=0$. Repare que $\mu(B_j\cap C_j)/\mu(C_j)$ faz o papel de p_{ji} na discussão anterior.

A realização de duas partições ξ e η é representada pela junção $\xi\vee\eta$, definida por $\xi\vee\eta=\{B_i\cap C_j : 1\leq i\leq m,\ 1\leq j\leq k\}$. η é um refinamento de ξ, $\xi<\eta$, se todo elemento de η está contido num elemento de ξ.

A incerteza da partição ξ quando se considera a dinâmica da aplicação contínua A, é o limite para n crescendo de $H_\mu(\bigvee_{i=0}^{n-1}A^{-i}\xi)$.

Teorema 4.8 Se $\mu\in I(A)$ e $\xi=\{B_1,\dots,B_m\}$ é uma partição de X, então o limite $h_\mu(A,\xi)=\lim_{n\to\infty}H_\mu(\bigvee_{i=0}^{n-1}A^{-i}\xi)/n$ existe.

Lema 4.9 As seguintes propriedades são válidas, onde ξ e η são partições de X.

(i) $H_\mu(\eta\vee\xi)=H_\mu(\eta)+H_\mu(\xi/\eta)$

(ii) $\xi<\eta$ então $H_\mu(\xi)\leq H_\mu(\eta)$

(iii) $H_\mu(\xi)\geq H_\mu(\xi/\eta)$

(iv) $H_\mu(\xi\vee\eta)\leq H_\mu(\xi)+H_\mu(\eta)$

(v) Se $\mu\in I(A)$ então $H_\mu(A^{-1}\xi/A^{-1}\eta)=H_\mu(\xi/\eta)$ e $H_\mu(A^{-1}\xi)=H_\mu(\xi)$.

A demonstração deste lema segue das propriedades das partições e da convexidade da função $\phi(x)=x\log x$. As demonstrações destas propriedades, e de outras, podem ser encontradas em [Wa].

Prova (Teorema 4.8) Seja $a_n = H_\mu(\bigvee_{i=0}^{n} A^{-i}\xi)$. Usando o lema 4.9(iv) e (v) vem que $a_{n+k} \leq a_n + a_k$. O resultado segue aplicando o lema 4.3. ©

Definição 4.10 Chama-se entropia métrica de A associada a $\mu \in I(A)$ (ou de (A,μ)) o número

$$h_\mu(A) = \sup\{h_\mu(A,\xi) : \xi \text{ é partição finita de } X\}.$$

Se $h_\mu(A) > 0$ pode-se dizer que o sistema é imprevisível, no sentido dos jogos de dados discutido anteriormente; novas medidas proporcionam novas informações, e são necessárias para atualizar o conhecimento do observador.

Onde está a contagem do número de órbitas? Agora será feita uma contagem do número de órbitas distintas módulo partição (definidas de maneira análoga às órbitas modulo cobertura), levando-se em conta a probabilidade de cada tipo de órbita (módulo partição).

Sejam η uma partição finita de X e $\mu \in I(A)$. A órbita módulo partição η do ponto $x \in X$, do tempo 0 até (n-1), pode ser representada pelo elemento da partição $\bigvee_{i=0}^{n-1} A^{-i}\eta$ que contém x; seja $F_n(x)$ tal elemento. A probabilidade desta órbita módulo η (de 0 até (n-1)) é $\mu(F_n(x))$, e se $\mu(F_n(x)) \to 0$ rapidamente para $n \to \infty$ e para x num conjunto de medida grande, significa que há muitos tipos distintos de órbitas módulo partição η.

Teorema 4.11 (Shannon-McMillan-Breiman) Se $\mu \in I(A)$, o limite $\bar{h}_\mu(A,\eta,x) = \lim_{n\to\infty} \log\mu(F_n(x))/n$ existe μ-qtp e em $L^1(\mu)$.

Seja $\bar{h}_\mu(A,\eta)$ a média ponderada de $\bar{h}_\mu(A,\eta,x)$ sobre todo o espaço X, ou seja, $\bar{h}_\mu(A,\eta) = \int \bar{h}_\mu(A,\eta,x)\, d\mu(x)$. Trabalhando um pouco com esta expressão vem

$$\bar{h}_\mu(A,\eta)=\lim_{n\to\infty}\int \log\mu(F_n(x))\,d\mu/n = \lim_{n\to\infty}\sum_{j=1}^{K(n)}\left[\int \log\mu(F_n^j)\,d\mu\right]/n$$

onde F_n^j representam os elementos da partição $\bigvee_{i=0}^{n-1} A^{-i}\eta$. Conclui-se então que $\bar{h}_\mu(A,\eta)=\lim_{n\to\infty}\sum_{j=1}^{K(n)}[\mu(F_n^j)\log\mu(F_n^j)]/n = h_\mu(A,\eta)$. Veja o teorema 4.8.

Assim, a entropia métrica de (A,μ) também pode ser definida por $h_\mu(A)=\sup\{\bar{h}_\mu(A,\eta): \eta$ é partição finita de $X\}$.

Como $\bar{h}_\mu(A,\eta)$ é o número médio de órbitas módulo η levando-se em conta a probabilidade de cada tipo de órbita, interpreta-se também $h_\mu(A)$ como o número médio máximo de órbitas módulo partição, com a média ponderada pela probabilidade de cada tipo destas órbitas.

A demonstração do teorema 4.11 pode ser encontrada em [Bi], e não será apresentada aqui.

Teorema 4.12 Se $A_i: X_i \to X_i$, $i=1,2$, são aplicações contínuas dos espaços de probabilidade (X_i, A_i, μ_i), com $\mu_i \in I(A)$, isomórficas (mod 0), então $h_{\mu_1}(A)=h_{\mu_2}(A)$.

A prova segue os mesmos passos da prova do teorema 4.5. As demonstrações das seguintes propriedades podem ser encontradas em [Wa].

Proposição 4.13 Se $\mu \in I(A)$ e η, ξ são partições finitas de X, então:

(i) $h_\mu(A,\eta) \leq H_\mu(\eta)$

(ii) $h_\mu(A,\eta\vee\xi) \leq h_\mu(A,\eta)+h_\mu(A,\xi)$

(iii) se $\eta < \xi$ então $h_\mu(A,\eta) \leq h_\mu(A,\xi)$

(iv) $h_\mu(A,A^{-1}\eta)=h_\mu(A,\eta)$

(v) para $k \in N$, $k \neq 0$, vale $h_\mu(A^k)=k h_\mu(A)$.

4C) O Princípio Variacional da Entropia

Na seção 4B foi definida $h_\mu(A)$ com base em $H(p_1, \ldots, p_n)$, que representava a incerteza de um esquema $E_1, \ldots, E_n$, e é máxima quando todos os eventos são igualmente prováveis.

Suponha que todos os elementos da partição $\eta = \{B_1, \ldots, B_s\}$ possuam mesma probabilidade, ou seja, $\mu(B_i) = 1/s$, $1 \leq i \leq s$. Assim $H_\mu(\eta) = \log(s)$, ou seja, a informação contida na partição η é igual ao logarítmo do número de elementos da partição, mas isto lembra a entropia topológica!

Teorema 4.14 (Princípio Variacional) Se $A: X \to X$ é uma aplicação contínua do espaço métrico compacto X, então $h(A) = \sup\{h_\mu(A) : \mu \in I(A)\}$.

Este resultado foi conjecturado no trabalho original sobre a entropia topológica [AKM]. Sua demonstração foi devida a esforços de diversas pessoas. Uma demonstração bem interessante foi dada por Misiurewicz, e está reproduzida em [Wa].

CAPÍTULO QUINTO

ATRATORES E CAOS

Neste capítulo é apresentada uma definição natural de atrator, que se distingue das outras por levar em conta as características estatísticas dos sistemas dinâmicos em variedades lisas, e não partir de conceitos topológicos, o que é tradicional. As vantagens deste caminho são discutidas, e algumas características topológicas são demonstradas.

A seção 5A mostra alguns exemplos. Na seção 5B a definição de atrator é apresentada, bem como alguns resultados que corroboram tal definição.

Fundindo o conceito de "imprevisível" dado pela entropia métrica e a definição de atrator, a definição de caos fica precisa, e é apresentada na última seção deste capítulo.

5A) A Noção de Atrator

Observou-se que, em geral, as figuras complicadas e bonitas mostradas no capítulo primeiro ocorrem em certas regiões do espaço X , de maneira que é nestas regiões que se localiza a parte interessante da dinâmica. Na tentativa de se definir bem estas regiões foi introduzido o conceito de "atrator". Ao contrário do que pode parecer à primeira vista, este conceito é muito difícil de ser definido de forma satisfatória.

Os atratores são os objetos de interesse na descrição do comportamento assintótico das aplicações, particularmente das aplicações que possuem características probabilísticas. "Um atrator

é algo que atrai as órbitas cujas condições iniciais estão numa certa 'área' ". Esta frase dá a idéia intuitiva de atrator, embora seja muito vaga.

Na realidade, os atratores se apresentam, em geral, como "entes" observados em simulações numéricas, em experimentos de laboratório e na natureza.

Vários autores têm trabalhado na definição de atrator, por exemplo, Smale [Sm1], Auslander, Bhatia e Seibert [ABS], Ruelle [Ru2], Eckmann [Ec], Guckenheimer e Holmes [GH], Cosnard e Demongeot [CD], Milnor [Mil], e outros.

É comum, nas definições de atrator, este ser dado em termos de características geométricas, que possua uma vizinhança aberta de pontos que são "atraídos", e que satisfaça alguma condição de minimalidade.

O objetivo principal deste capítulo é dar uma definição de atrator e, a partir disto, definir "atratores caóticos".

É bom apresentar alguns exemplos, usando o termo atrator apenas num sentido intuitivo.

1) Seja S^1 a circunferência unitária, $A:S^1\to S^1$ dado por

$$A(x)=x+\operatorname{sen}^2(x/2).$$

Perto de $x_0=0$ existem pontos que se afastam de x_0 e outros que se aproximam de x_0, mas x_0 "atrai" todos os pontos de S^1. Veja a figura 5a.

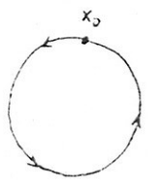

figura 5a

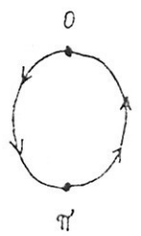

figura 5b

2) Considere $A: S^1 \to S^1$, $A(x)=x+\text{sen}^2 x$. Os pontos 0 e π possuem pontos arbitrariamente próximos deles que são 'atraídos' para o outro. Veja a figura_5b.

3) Seja $A:[-a,a] \to [-a,a]$, onde $a=64/(5\sqrt{5})$, dado por

$$A(x) = 4x(1-x^2)^2 \text{ se } x\in[-1,1], \text{ e}$$

$$A(x)=0 \text{ no complementar de } [-1,1].$$

Esta aplicação é de classe C^1, a origem é um ponto fixo repulsivo, $A'(0)>1$, mas todos os pontos do domínio, excluindo $\sqrt{1/2}$ e $-\sqrt{1/2}$ são 'atraídos' para ela.

4) O problema funcional de Feigenbaum [CE] é o seguinte: encontrar uma função unimodal $A:[-1,1]\to[-1,1]$ que satisfaça

$$A(x) = -1/\alpha \; A(A(-\alpha x)) \;, \alpha=-A'(1)>0.$$

Esta função admite órbitas periódicas instáveis de período 2^n para todo n natural, e um conjunto de Cantor C, tal que todos os pontos num conjunto de medida de Lebesgue total convergem a C pela aplicação sucessiva de A. Dada uma vizinhança aberta deste conjunto de Cantor, pode-se encontrar n tal que uma órbita de período 2^n está nesta vizinhança. Assim, não há vizinhança aberta de pontos que convergem a C. Veja [CE] para maiores detalhes e outras referências.

5) Este exemplo foi retirado de Garrido e Simó [GS]. Seja X o toro bidimensional e $A:X\to X$ dado por

$$A\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} + K\begin{pmatrix} 1 \\ \alpha \end{pmatrix}(\text{sen}^2\pi y + \text{sen}^2\pi x)$$

com $\alpha=(\sqrt{5}-1)/2$ e k suficientemente pequeno. Abaixo são utilizados os conceitos de variedade estável e de variedade instável de um

UNIVERSIDADE DE SÃO PAULO
SERVIÇO DE BIBLIOTECA E INFORMAÇÃO
2071
e.1
INSTITUTO DE FÍSICA

ponto fixo p de um aplicação invertível $B:X\to X$, definidos por: a variedade estável de p é o conjunto dos pontos x de X tais que $d(B^n x,p)\to 0$ quando $n\to\infty$, e a variedade instável de p é o conjunto dos y em X tais que $d(B^{-n} y,p)\to 0$ quando $n\to\infty$.

A origem é o único ponto fixo da aplicação A dada acima, e possui variedades estável e instável densas em X. Foi demonstrado o seguinte resultado: sejam $\varepsilon,\delta>0$ e $B(\varepsilon)$ a bola aberta de raio ε centrada na origem, então existe $T(\varepsilon,\delta)$ tal que para todo $n>T$ vale

$$1/n \text{ cardinalidade}\{j\in N: j<n \text{ e } A^j(q)\in B(\varepsilon)\} > (1-\delta)$$

para todo q em X. Isto mostra que, mesmo sendo a variedade instável da origem densa em X, ela "atrai" todos os pontos do toro.

6) Suponha que $A:X\to X$ é um difeomorfismo C^2 que satisfaz o Axioma-A de Smale [Sm]. Neste caso o conjunto não-errante $\Omega(A)$ pode ser decomposto numa união disjunta finita $\Omega(A)=\Omega_1\cup\ldots\cup\Omega_s$, onde cada Ω_j, $1\leq j\leq s$, é invariante por A $(A(\Omega_j)=\Omega_j)$; os Ω_j's são chamados peças do difeomorfismo A. As demonstrações destes resultados podem ser encontrados na monografia de Bowen [Bo]. Uma peça Ω_j é dita topologicamente mixing se dados dois abertos quaisquer V_1 e V_2 em Ω_j (topologia induzida) existe T natural tal que $A^n(V_1)\cap V_2\neq\emptyset$ para todo $n>T$.

Seja Λ uma peça topológicamente mixing que possua uma vizinhança aberta U tal que $\Lambda=\bigcap_{n\geq 0} A^n(U)$. Neste caso, existe um conjunto Λ_1 tal que $U\setminus\Lambda_1$ possui medida de Lebesgue nula, e ainda $\lim_{n\to\infty}\sum_{i=0}^{n-1}\delta_{A^i x}/n=\sigma\in I(A)$ (topologia fraca*) para todo $x\in\Lambda_1$.

Desta forma, todo ponto de U converge a Λ, e se $U\setminus\Lambda_1$ for "desprezado", obtém-se muito mais informações.

Para difeomorfismos de Anosov a prova original é devida a Sinai [Si2], e para Axioma-A é de Ruelle [Ru3]. Veja também [Bo].

7) Seja $A: X \to X$ a aplicação $A(\theta, r, s) = (2\theta, \varepsilon_1 \cos\theta + \varepsilon_2 r, \varepsilon_1 \operatorname{sen}\theta + \varepsilon_2 s)$, onde X é o toro bidimensional maciço, $\varepsilon_2 < \varepsilon_1 < 1/2$, e r e s são como na figura 5c. A ação de A em X é a seguinte: é como se o toro fosse cortado num θ fixo, transformando-se num cilíndro maciço, e depois fosse esticado longitudinalmente e colado dando uma volta ao redor do centro, como mostra a figura5d.

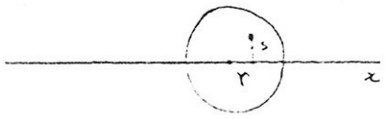

figura 5c

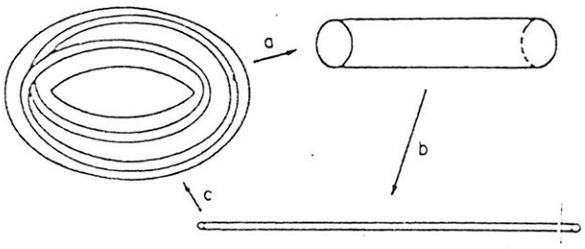

figura 5d

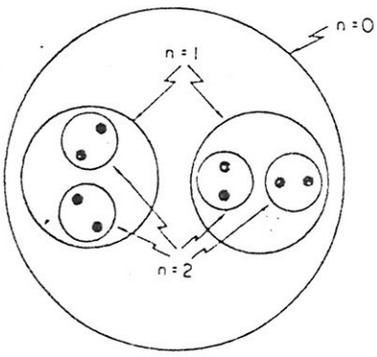

figura 5e

O conjunto $\Lambda = \bigcap_{n \geq 0} A^n(X)$ é localmente o produto de um conjunto de Cantor por um intervalo da reta. Um corte para θ fixo, como mostrado na figura5e, dá uma idéia de Λ.

Existe uma medida $\sigma \in I(A)$ com as propriedades indicadas no exemplo anterior onde X faz o papel do aberto U. Para a demonstração e discussão mais detalhada veja [Sm2] ou [La2]. Uma generalização importante foi obtida em [MR].

8) Nos sistemas Hamiltonianos completamente integráveis com dois ou mais graus de liberdade, todas as órbitas estão restritas a toros de dimensão menor que a dimensão da variedade de energia constante, e os movimentos nos toros são periódicos ou quase periódicos. Neste caso, a intuição diz que não há atratores.

Na próxima seção é dada uma definição de atrator que parte de um ponto de vista diferente do tradicional, ela prioriza

as características estatísticas e não aspectos puramente topológicos, mas características topológicas podem ser associadas de modo natural, como demonstrado na proposição 5.2.

5B) A Definição de Atrator

Seja $A: X \to X$ uma aplicação contínua, onde X é uma variedade lisa compacta. Tendo como motivação a idéia de ergodicidade em mecânica estatística e as discussões das seções 3C e 5A, define-se:

Definição 5.1 Uma medida $\nu \in I(A)$ é dita um atrator para a aplicação A se existe um subconjunto F de X com medida de Lebesgue ℓ estritamente positiva, $\ell(F) > 0$, tal que a seqüência $\{\sum_{i=0}^{n-1} A_*^i \delta_x / n\}$ converge a ν na topologia fraca*, para todo $x \in F$. Se não existe $\nu \in I(A)$ nestas condições diz-se que A não possui atrator.

Definição 5.2 Chama-se conjunto de atração da medida $\mu \in I(A)$ o conjunto G_μ de pontos de X tais que $\{\sum_{i=0}^{n-1} A_*^i \delta_x / n\}$ converge a μ. Diz-se que μ atrai os pontos $x \in G_\mu$, ou que μ atrai as medidas δ_x.

Os resultados abaixo dão suporte à definição de atrator dada acima, como por exemplo: é provado que o conjunto de atração de qualquer medida invariante é mensurável, que se x é atraído por μ então o conjunto ω-limite de x possui medida μ total, e é demonstrada uma maneira de aproximar (na topologia fraca*) um atrator por medidas absolutamente contínuas em relação à medida de Lebesgue.

Proposição 5.3 Para toda $\mu \in I(A)$ G_μ é um boreleano.

Prova Seja $\mu \in I(A)$, e se $f \in C(X)$ sejam

$$\Sigma_f(x) = \limsup_{n\to\infty} (S_n f)(x)/n \quad e \quad \sigma_f(x) = \liminf_{n\to\infty} (S_n f)(x)/n.$$

Como $\{(S_n f)/n\}$ é uma seqüência de funções mensuráveis (contínuas) segue que $\Sigma_f : X \to R$ e $\sigma_f : X \to R$ também são mensuráveis.

Seja Λ_f o conjunto dos pontos $x \in X$ tal que o limite $\lim_{n\to\infty} (S_n f)(x)/n$ existe e é igual a $\mu(f)$.

Como $\Lambda_f = \sum_f^{-1}(\mu(f)) \cap \sigma_f^{-1}(\mu(f))$ vem que Λ_f é um boreleano, já que cada um dos conjuntos na intersecção é boreleano.

Considerando um conjunto enumerável $\{f_1, f_2, f_3, \ldots\}$ e denso em $C(X)$, vem que $G_\mu = \bigcap_{i\geq 1} \Lambda_{f_i}$. Como a intersecção enumerável de boreleanos é um boreleano, a proposição fica provada. ☐

A proposição 5.3 é importante pois garante que o conjunto de atração é mensurável, mesmo que ele não seja um subconjunto aberto de X!

Definição 5.4 Uma órbita periódica P é dita assintoticamente estável se possui vizinhança aberta U, tal que para todo $x \in U$ $d(A^n x, P) \to 0$ quando $n \to \infty$ (d é uma métrica que dá a topologia de X).

As medidas invariantes são uma generalização de órbitas periódicas, e a noção de atrator é uma generalização de órbitas periódicas assintoticamente estáveis. Os resultados abaixo mostram que a definição 5.1 é coerente com isso.

Lema 5.5 [Wa] Sejam $k \in N$ e $x \in X$, então $A^k(x)=x$ sss $\rho_k = \sum_{i=0}^{k-1} A_*^i \delta_x /k$ pertence a $I(A)$ sss $A_*^k(\delta_x) = \delta_x$.

Prova Vale que $A_* \rho_k = \rho_k$, $\sum_{i=0}^{k-1} A_*^{i+1} \delta_x /k = \sum_{i=0}^{k-1} A_*^i \delta_x /k$, o que é equivalente a $A_*^k \delta_x = \delta_x$, ou ainda, $\delta_{A^k x} = \delta_x$, que equivale a $A^k x = x$. ☐

Teorema 5.6 Seja $P=\{p_0, \ldots, p_{k-1}\}$ uma órbita periódica de período k. Se a órbita de x converge a P então $\lim_{n\to\infty} \sum_{i=0}^{n-1} A_*^i \delta_x /n = \rho_k = \sum_{i=0}^{n-1} A_*^i \delta_{p_j} /k$ para todo $0 \leq j \leq (k-1)$.

Prova É claro que ρ_k independe do j escolhido, além disso, pela

proposição 5.5 pode-se supor que $A^n(x)$ converge para o ponto fixo p.

Se $A^n(x)\to p$ e $f\in C(X)$, então dado $\delta>0$ existem $\varepsilon>0$ e T tal que $|f(A^n x)-f(p)|<\delta$ sempre que $d(A^n(x),p)<\varepsilon$, o que ocorre para todo $n>T$. Disto vem

$$\int f\, d\left(\sum_{i=0}^{n-1} A_*^i \delta_x/n\right) = (S_n f)(x)/n.$$

Se $n>T$ vale

$$(S_n f)(x)/n=(S_T f)(x)/n + [(S_n f)(x)-(S_T f)(x)]/n<$$

$$<(S_T f)(x)/n + (n-T)(f(p)+\delta)/n =$$

$$= (S_T f)(x)/n + f(p) + \delta - T(f(p)+\delta)/n,$$

fazendo $n\to\infty$,

$\lim_{n\to\infty} (S_n f)(x)/n\leq f(p)+\delta$. De maneira similar obtém-se

$$f(p)-\delta\leq \lim_{n\to\infty} (S_n f)(x)/n\leq f(p)+\delta,$$

e como δ e f são arbitrários conclui-se que, se $A^n(x)\to p$ então $\lim_{n\to\infty} \sum_{i=0}^{n-1} A_*^i \delta_x/n \to \delta_p$. ☐

O lado físico da definição 5.1 é representado pela maneira que se tem de calcular grandezas fisicamente relevantes dadas por funções contínuas, e também pelo conjunto de atração possuir medida de Lebesgue maior que zero. Aqui está uma suposição fundamental neste trabalho, ou seja, que os atratores (que seriam observados experimentalmente) estão associados a condições iniciais num conjunto de medida de Lebesgue não nula, já que a medida de Lebesgue dá uma boa idéia de amostragem. Deve ser ressaltado que assumir que médias temporais é que serão observadas num experimento de laboratório ou computacional, só pode ser garantido, em cada caso, após o experimento ter sido realizado e corroborado a teoria.

Provar de forma bem geral a existência de médias temporais é um problema antigo e ainda não resolvido. O conhecido teorema de Birkhoff prova a existência de tais médias para pontos

num conjunto de medida invariante total, mas em geral, as medidas invariantes são singulares em relação à medida de Lebesgue!

A existência das médias temporais em $C(X)$ é o lado delicado da definição de atrator, assim, na presente data não é possível descrever precisamente seu domínio de aplicabilidade, mas isto também ocorre com outras definições de atrator. Na prática, as médias temporais parecem existir, embora seja possível imaginar exemplos onde elas não existam.

Como conciliar a visão tradicional de atrator como um subconjunto de X e, como associar as figuras, como aquelas mostradas no primeiro capítulo, com a definição 5.1?

A resposta pode ser dada em termos do suporte da medida atratora e da proposição 5.2. O suporte de uma medida $\mu \in M(X)$ é o conjunto dos pontos $x \in X$ tal que toda vizinhança aberta de x possui medida μ não nula, é onde a medida se concentra. O suporte de μ é denotado por $su(\mu)$, mas conhecer uma medida é um problema mais importante do que conhecer seu suporte.

Uma característica topológica importante associada à definição 5.1 é demonstrada na próxima proposição.

Proposição 5.2 Se $\mu \in I(A)$ e atrai δ_x $(x \in X)$, então $\mu(\omega(x))=1$, onde $\omega(x)$ é o conjunto ω-limite de x.

Prova Sejam $x \in G_\mu$ e $K: X \to R$ uma função em $L^1(\mu)$ tal que $K(A^n x)=K(x)$ para todo $n \in N$, com $K(x) \neq 0$. A existência de tal função é óbvia.

Sejam $\varepsilon > 0$ e $\bar{o}_\varepsilon = \overline{\{y \in X : d(y, \bar{o}(x)) > \varepsilon\}}$, onde $\bar{o}(x)$ é o fecho da órbita de x. Tome a função de Urysohn f_ε que assume 1 em $\bar{o}(x)$, zero em $\bar{o}_\varepsilon$ e $0 \leq f_\varepsilon < 1$ de outro modo.

$$\bar{K}(x) = \lim_{n \to \infty} \sum_{i=0}^{n-1} K(A^i x)/n = \lim_{n \to \infty} \sum_{i=0}^{n-1} [K(A^i x) . f_\varepsilon(A^i x)]/n,$$

o primeiro limite existe pela construção de K, e a igualdade com o segundo membro se dá porque $f_\varepsilon(x) = f_\varepsilon(A(x)) = 1$. Usando as proprieda-

ões de K e f_ε obtém-se

$$\bar{K}(x)=\lim_{n\to\infty}\sum_{i=0}^{n-1}[K(x)f_\varepsilon(A^i x)]/n = K(x)\lim_{n\to\infty}\sum_{i=0}^{n-1} f_\varepsilon(A^i x)/n = K(x)\int f_\varepsilon\, d\mu.$$

Como $K(x)=K(A(x))$ vem que $K(x)=\bar{K}(x)$, ou ainda, $\int f_\varepsilon\, d\mu=1$, ou seja, f_ε é igual a 1 num conjunto de medida total. Segue que a medida $\mu(\bar{o}(x))=1$. De maneira análoga vem que $\mu(\bar{o}(A^n x))=1$ para todo n em N.

Como $\omega(x) = \bigcap_{n\geq 0} \bar{o}(A^n x)$, vem que $\mu(\omega(x))=1$, já que a intersecção enumerável de conjuntos de medida total também possui medida total. ☐

Não necessariamente todo o conjunto ω-limite de um ponto $x\in X$ tem um papel importante na descrição assintótica de sua órbita, mas apenas um subconjunto dele, e isto é enxergado pela medida invariante obtida de $\lim_{n\to\infty}\sum_{i=0}^{n-1} A_x^i\,\delta_x/n$. Esta é uma vantagem em relação a algumas definições onde o atrator é visto como um conjunto.

A definição de atrator de Ruelle [Ru2] também difere bastante das tradicionais, e se baseia no papel desempenhado por pequenas perturbações estocásticas sobre o sistema estudado.

As manipulações com a medida de Lebesgue são, em geral, mais simples que as manipulações com os atratores, assim seria interessante se existisse uma maneira de aproximar um atrator por medidas absolutamente contínuas em relação à medida de Lebesgue. Normalmente não há fórmulas explícitas para as medidas atratoras.

<u>Proposição 5.8</u> Para toda $h\in L^1(\ell)$ positiva, tal que $Y=\int_{G_\nu} h\, d\ell>0$, onde ν é um atrator para $A:X\to X$, vale que a medida $\mu\in M(X)$ dada por

$$d\mu = \chi_{G_\nu}\, h\, d\ell/Y$$

é atraída por ν.

Prova Seja $f\in C(X)$,

$$\sum_{i=0}^{n-1} [\int f(A^i x)\chi_{G_\nu}(x)h(x)\, d\delta/Y]/n = \int \sum_{i=0}^{n-1} [f(A^i x)\chi_{G_\nu}(x)\, d\delta/Y]/n =$$
$$= \int [\sum_{i=0}^{n-1} f(A^i x)/n]\, d\mu .$$

Como $\mu(G_\nu)=1$ vem que

$$\lim_{n\to\infty} \sum_{i=0}^{n-1} [\int f(A^i x)\chi_{G_\nu}(x)h(x)/Y\, d\delta]/n =$$

(pelo teorema da convergência dominada)

$$= \int_{G_\nu} [\lim_{n\to\infty} \sum_{i=0}^{n-1} f(A^i x)/n]\chi_{G_\nu}(x)h(x)/Y\, d\delta =$$
$$\int_{G_\nu} (\int f\, d\nu)\, d\mu = \int f\, d\nu.$$

Como f é arbitrária o resultado segue. ∎

Esta proposição permite aproximar o atrator de várias maneiras, basta variar a função h. Mas, se um resultado é obtido aproximando um atrator por medidas absolutamente contínuas em relação à de Lebesgue e truncando, não se conhece exatamente qual o significado dessa aproximação.

Como um conjunto de medida estritamente positiva pode ter no máximo um número contável de subconjuntos disjuntos de medida maior que zero, segue o seguinte resultado (que "limita" a quantidade de atratores segundo a definição_5.1):

Proposição_5.2 Uma aplicação contínua $A:X\to X$ possui no máximo um número infinito contável de atratores.

Voltando a alguns dos exemplos da seção anterior:

Em 1) existe apenas um atrator δ_0 e $G_{\delta_0} = S^1$. Existem dois atratores para 2), δ_0 e δ_π, sendo $G_{\delta_0} = (x)\pi') \cup \{0\}$ e G_{δ_π} = complementar de G_{δ_0}. Em 3) só δ_0 é atrator, e G_{δ_0} = todo o domínio excluindo $-\sqrt{1/2}$ e $\sqrt{1/2}$.

O exemplo 4) apresenta fortes indicações de que há apenas um atrator, que seria $\nu=\delta_0$ e G_ν = todo o toro.

No caso de difeomorfismos C^2 Axioma-A, há apenas um

atrator em cada peça básica como em 6), e σ é o atrator. Idem para o exemplo 7).

Nos sistemas Hamiltonianos completamente integráveis com dois ou mais graus de liberdade, claramente não existem atratores.

É importante ressaltar que para as aplicações de Hénon e Lorenz, mostradas no capítulo primeiro, não há resultados rigorosos, mas várias simulações numéricas indicam a presença de um único atrator em cada caso.

Definição 5.10 Uma aplicação contínua A é dita unicamente ergódica se $I(A)$ é um conjunto unitário.

Proposição 5.11 [Wa] A aplicação contínua $A:X\to X$ é unicamente ergódica sss para todo $f\in C(X)$ e todo $x\in X$ existe $\mu\in I(A)$ tal que

$$\lim_{n\to\infty} (S_n f)(x)/n = \mu(f).$$

Prova Suponha que exista o limite acima para toda $f\in C(X)$ e todo $x\in X$. Integrando em relação a $\sigma\in I(A)$ e usando o teorema da convergência dominada, vem que

$$\lim_{n\to\infty} \sum_{i=0}^{n-1}\int f(A^i x)\, d\sigma/n = \sigma(f) = \int \mu(f)\, d\sigma = \mu(f).$$

Portanto $\sigma=\mu$ pelo teorema 3.2, e A é unicamente ergódica.

Para demonstrar a outra implicação, suponha que seja válido o seguinte fato: se A é unicamente ergódica, então dada $f\in C(X)$ o limite $\lim_{n\to\infty} (S_n f)(x)/n$ converge uniformemente a uma constante.

Pelo teorema de representação de Riesz, teorema 3.3, o operador $L(f)=\lim_{n\to\infty} (S_n f)(x)/n$ pode ser escrito como $L(f)=\mu(f)$, e facilmente se verifica que $\mu\in I(A)$. Como A é unicamente ergódica, pode-se usar o fato acima, e fica demonstrado o resultado.

Para provar o fato acima, suponha que ele seja falso.

Se fosse verdadeiro a constante deveria ser $\mu(f)$, onde μ é o único elemento de $I(A)$. Se é falsa, existem $h \in C(X)$ e $\varepsilon > 0$ tais que

$$\left| \sum_{i=0}^{n-1} [h(A^i x_n)/n] - \mu(h) \right| \geq \varepsilon \text{ para } n > T(\varepsilon) \text{ e } x_n \in X.$$

Escrevendo $\mu_n = \sum_{i=0}^{n-1} A_i^i \delta_{x_n}/n$, vem que $|\mu_n(h) - \mu(h)| \geq \varepsilon$.

Tomando uma subseqüência $\{\mu_{n_j}\}$ convergente a σ (é claro pelos teorema_3.8 e lema_3.10 , que $\sigma \in I(A)$), vem que

$$|\sigma(h) - \mu(h)| \geq \varepsilon,$$

o que contradiz a suposição de A ser unicamente ergódica. ©

Se $I(A) = \{\mu\}$, então, μ é atrator para A, atrai todo x em X, ou seja, $G_\mu = X$, e pelo teorema_3.22(iii), μ é ergódica.

Deve ser notado que em muitos sistemas reais não se observa um único atrator, há vários atratores coexistindo, muitas vezes "entrelaçados", alguns com fortes características estocásticas (veja a próxima seção) e outros não.

5C) Caos

"Lorenz tells an anecdote concerning his discovery of chaos that illustrates what deterministic chaos entails in practical terms. Since his computer was printing out a list of the values of x, y, z, he could stop the computer at any intermediate time, examine the series of numbers, and then restant it by entering the intermediate set of values. In doing this, he saw that the results very quickly began to differ from the original run. This difference increased as the simulation proceeded. Since the values were not printed out to full machine accuracy, there were very slight differences between the initial conditions in the two runs".

Esta espécie de "comportamento complicado" é o que se tenta compreender e colocar sobre bases conceituais fundamentais. O trecho acima foi retirado de "Caos em Mecânica", J. Koiler, VII Escola de Matemática Aplicada - LNCC (1985) que, por sua vez, foi extraído de D. Campbell et al., Comm. A.C.M. 28, 4(1985) 374-384.

Nesta seção é dada uma alternativa conceitual para estes comportamentos, mas só a prática pode dar uma resposta positiva a qualquer alternativa, assim, tem-se mais um modelo que uma teoria! Lembre-se que a natureza é mais fascinante que os modelos que o homem constrói!

Estes comportamentos são comumente chamados de "caos", e neste trabalho esta nomenclatura é seguida. É necessário distinguir "caos" de órbitas periódicas de período muito longo, de comportamentos simplesmente não periódicos e de transientes.

Os comportamentos complicados vistos por medições realizadas em intervalos de tempo determinados possuem características imprevisíveis e geram informação. A propriedade estatística

que quantifica estas noções é a entropia, como discutido no capítulo quarto.

Definição 5.12 Uma medida $\nu \in I(A)$ é dita caótica se a entropia métrica $h_\nu(A)$ é maior que zero. Uma aplicação A possui caos se existe atrator caótico para A.

É comum chamar os atratores com propriedades estocásticas, como os atratores caóticos, de "atratores estranhos", apelido introduzido por Ruelle e Takens [RT] em 1971.

Da proposição 5.9 é imediato que uma aplicação contínua $A:X \to X$ pode ter no máximo um número contável de atratores caóticos.

Para aplicações que possuem atratores caóticos a entropia métrica deste atrator é maior que zero e novas medições são necessárias para atualizarem o conhecimento do observador, existe um grau de imprevisibilidade nas medições. Esta é uma maneira de explicar o que ocorreu com a simulação de Lorenz.

O automorfismo de toro $A(x,y)=(ax+by,cx+dy)$ (mod 1) com $ad-cb=1$ possui a medida de Lebesgue ℓ como único atrator, e $h_\ell(A) = \log|\lambda|$, onde λ é o autovalor da matriz $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ de maior módulo, assim, se $|\lambda|>1$ a medida de Lebesgue é atrator caótico para A. Um caso particular é a famosa aplicação do gato de Arnold, onde $a=b=c=1$ e $d=2$. Veja [AA].

Inspirado na aplicação de Hénon, R. Lozi introduziu a aplicação no plano $A(x,y)=(1+y-a|x|,\ 1/2\ x)$, onde $a=1,7$. Esta aplicação é muito mais simples de se manipular que a de Hénon, e foi provada a existência de um atrator ergódico e caótico. Veja [Le] e referências lá citadas para maiores detalhes.

Como já dito, a aplicação de Hénon parece possuir apenas um atrator, como indicam simulações numéricas. Estudos numéricos realizados por J. H. Curry [Cu] apontam fortemente que este atrator é caótico; este trabalho de Curry é também interessante por mostrar algumas das dificuldades no cálculo numérico da entropia, assim, torna-se importante estudar as relações entre entropia e expoentes de Lyapunov, definidos no capítulo sexto.

No exemplo da equação funcional de Feigenbaum [CE] ocorre o fenômeno de acúmulação de uma sequência infinita de bifurcações. Para $n=0, 1, 2, \ldots$ um atrator periódico de período 2^n deixa de ser atrator e nasce outro atrator (quando um parâmetro é variado), concentrado numa órbita periódica de período 2^{n+1}. A solução da equação funcional de Feigenbaum é uma aplicação que surge quando $n \to \infty$. Misiurewicz [Mi] mostrou que esta aplicação possui entropia topológica nula, assim, pelo princípio variacional da entropia, teorema 4.14, a aplicação de Feigenbaum não possui atrator caótico.

Tanto a entropia topológica como a entropia métrica medem o grau de imprevisibilidade de um sistema, embora a entropia métrica é, sem dúvida, mais importante; pode-se concluir isto a partir do seguinte raciocínio: se a entropia topológica é maior que zero, então, pelo pricípio variacional, há medidas invariantes com entropia métrica estritamente positiva, mas poderia ocorrer que entre estas medidas não haja atratores, e neste caso os comportamentos aleatórios não seriam observados na prática.

Para concluir este capítulo, algumas observações (informais) referentes às definições de atrator e caos são apresentadas.

Devido principalmente a dificuldades matemáticas, ainda não se conhece profundamente os "atratores" de Hénon e Lorenz, que se tornaram protótipos da "teoria do caos".

Várias explicações (alternativas) possíveis para estes fenômenos aleatórios estão surgindo, como por exemplo:

A existência de um conjunto invariante pela dinâmica que contenha um conjunto enumerável de órbitas periódicas, de períodos arbitrariamente longos, um conjunto não enumerável de órbitas não periódicas e uma órbita densa; em muitos casos a presença de um ponto homoclínico garante a existência de tal conjunto. Outra possibilidade seria a presença de um conjunto invariante que seja um fractal. Muitos autores dizem que há caos na presença de um conjunto com alguma destas duas características.

Uma objeção para as explicações acima é que tais conjuntos poderiam "atrair poucos pontos". Uma análise detalhada da estrutura das órbitas da aplicação $A:[-1,1] \to [-1,1]$, $A(x)=1 - ax^2$, com $a=1,76$, mostra a presença de um conjunto de Cantor invariante, órbitas periódicas de todos os períodos e órbita densa neste conjunto de Cantor; tudo isso parece complicado, mas, excluindo um conjunto de medida de Lebesgue nula, todos os pontos do domínio são atraídos por uma medida ergódica concentrada numa órbita periódica de período três. É óbvio que, pela definição de atrator caótico deste trabalho, tal exemplo não possui caos. Referência: [Lai].

Ainda outra explicação possível para os fenômenos aleatórios em sistemas determinísticos seria que órbitas num conjunto invariante e limitado fossem tais que condições iniciais arbitrariamente próximas divergiriam exponencialmente com o tempo, de forma que pequenas perturbações rapidamente tornar-se-iam macroscopicamente perceptíveis (detalhes estão expostos no capítulo sexto).

Esta última tentativa de explicar os comportamentos complicados se encaixa muito bem na ilustração de Lorenz, apresentada no início desta seção. A noção de que condições iniciais próximas divergem exponencialmente com o tempo é quantificada pelos

expoentes de Lyapunov, que requerem uma medida invariante para serem devidamente difinidos (a partir do poderoso teorema de Oseledec), e estão intimamente relacionados com a entropia métrica e a definição de caos apresentada neste capítulo. O capítulo sexto trata destas questões.

CAPÍTULO SEXTO

EXPOENTES DE LYAPUNOV POSITIVOS

Neste capítulo, o importante conceito de expoentes de Lyapunov (EL) é introduzido e sua ligação com atratores discutida.

Da mesma forma que sem falar em medidas invariantes é impossível falar em entropia métrica, também não é possível falar em EL sem medidas invariantes, pois este conceito é definido com base no teorema de Oseledec [Os], [Ra] (teorema 6.1).

Intuitivamente, os EL medem de maneira quantitativa a divergência de órbitas próximas, noção muito importante quando na abordagem do caos. A presença de EL estritamente positivos é necessária, mas não suficiente para que atratores sejam caóticos.

Os EL podem ser calculados numericamente ou a partir de dados experimentais, assim, é natural tentar usá-los para classificar os atratores mas, tendo como base alguns trabalhos na literatura especializada é mostrado neste capítulo que tal classificação não é simples.

O comportamento complicado descrito por Lorenz (veja a seção 5C) poderia ser explicado argumentando-se que condições iniciais próximas se "afastam" rapidamente uma da outra, tornando a aplicação imprevisível na prática.

Como surgiria esta imprevisibilidade? Ela resultaria do fato de não ser possível os dados experimentais terem precisão

infinita, e se condições iniciais próximas se afastam rapidamente não se poderia saber qual é a "verdadeira órbita"!

Algumas observações são pertinentes: primeiro, este fenômeno não pode ser evitado apenas melhorando a precisão dos dados, segundo, ele não aparece nos chamados exemplos tipo livro-texto, e terceiro, não está relacionado com um grande número de graus de liberdade.

O que se deseja é quantificar a noção de "órbitas que se afastam rapidamente" e relacioná-la com a entropia métrica. Isso é feito usando o teorema de Oseledec (cuja versão dada abaixo já define os EL) e a desigualdade de Margulis-Ruelle.

A história dos EL começa, obviamente, com Lyapunov, que por volta de 1890 procurava um critério para analisar a estabilidade da origem de

$$\dot{x} = M(t)x \ , \ x \in R^k \text{ e } M: R \to \{\text{matrizes } k \times k\},$$

e a partir daí obter informações sobre a estabilidade da origem em

$$\dot{x} = M(t)x + f(t,x), \text{ com } |f(t,x)| \leq c|x|^{1+\alpha} \ , \ c, \alpha > 0.$$

A idéia de Lyapunov foi olhar para a média dos autovalores de $M(t)$ com o tempo, e assim ele demonstrou várias propriedades que foram amplamente generalizadas por Oseledec e publicadas em 1968.

Neste capítulo é suposto que $A: X \to X$ é uma aplicação de classe C^1 e X é variedade compacta lisa. Sejam $D_x A$ a matriz Jacobiana de A no ponto $x \in X$, e $T_x X$ o espaço tangente a $x \in X$.

<u>Teorema 6.1</u> (Oseledec) Para toda $\mu \in I(A)$ existe um boreleano $B \subset X$ com $A(B) \supseteq B$ e $\mu(B) = 1$, tal que, se $x \in B$ valem:

(i) existem números $\lambda^1(x) > \lambda^2(x) > \ldots > \lambda^{s(x)}(x)$, onde $\lambda^{s(x)}(x)$ pode ser $-\infty$; s é mensurável e $s \circ A = s$.

(ii) existem subespaços $T_x X = V^1(x) \supset V^2(x) \supset \ldots \supset V^{s(x)}(x) \supset \{0\} = V^{s(x)+1}(x)$

tais que

$$\lim_{n\to\infty} \log\|(D_x A)v\|/n = \lambda^i(x),$$

para todo $v \in V^i(x)\setminus V^{i+1}(x)$, $1 \leq i \leq s(x)$.

(iii) $\lambda^i(\cdot)$ é mensurável sobre $\{x \in B: s(x) \leq i\}$ e $\lambda^i \circ A = \lambda^i$.

(iv) $(D_x A)V^i(x) \subset V^i(A(x))$ se $i \leq s(x)$.

Definição 6.2 Os números $\lambda^i(x)$ são chamados expoentes de Lyapunov (EL) da aplicação A em x, e $m^i(x) = \dim[V^i(x)] - \dim[V^{i+1}(x)]$ é chamado de multiplicidade de $\lambda^i(x)$ (repare que existem $\dim[T_x X]$ EL).

O ítem (ii) acima diz que o comportamento de

$$(D_x A^n) = (D_{A^n x} A) \circ (D_{A^{n-1} x} A) \circ \ldots \circ (D_x A)$$

é determinado, para n grande e todo $v \in T_x X$, pelos EL. Para uma medida μ ergódica vem de (iii) e (i) que $\lambda^i(\cdot)$ e $s(\cdot)$ são constantes μ-qtp.

A versão original do teorema 6.1 [Os], bem como a demonstração de Raghunathan [Ra], que não supõe invertibilidade de A, é bem mais geral que a versão aqui apresentada.

Se $\lambda^i(x) > 0$ há uma "expansão média" no subespaço $V^i(x)$ numa razão exponencial; grosseiramente, se x_0 é uma condição inicial e Δx_0 é o erro em x_0, então

$$\Delta A^n(x_0) = A^n(x_0 + \Delta x_0) - A^n(x_0) \sim (D_{x_0} A^n)\Delta x_0,$$

e se há EL estritamente positivos $\Delta A^n(x_0)$ pode crescer exponencialmente com o tempo n.

De (ii) vem que para $v \in T_x X$ o EL

$$\lim_{n\to\infty} \log\|(D_x A^n)v\|/n \neq \lambda^1(x),$$

se e somente se $v \in V^2(x)$, que possui codimensão maior ou igual a um; assim, se v é escolhido aleatoriamente em $T_x X$, calcula-se $\lambda^1(x)$ com probabilidade total; por esta razão, o maior EL é mais fácil de ser computado numericamente.

Com as interpretações acima, os EL tornam-se muito importantes, e esta importância também se deve ao fato, já citado acima, deles serem acessíveis numericamente ([BGGS], [SN]). Uma área de intensa pesquisa atualmente é o desenvolvimento de técnicas para o cálculo dos EL a partir de dados experimentais de laboratório, veja [SS], [WSSV], [EKRC] (agradeço ao Prof. J.F. Perez por ter me enviado este último trabalho).

Um estudo numérico detalhado dos EL para a aplicação de Hénon foi feito por Féit [Fe], encontrando para os valores dos parâmetros correspondentes à figura 1.2 $\lambda^1=0,419$ e $\lambda^2=-0,719$, sendo que estes valores parecem independer do ponto $x\in X$, corroborando a conjectura de que existe apenas um atrator para esta aplicação.

Definição 6.3 Uma medida $\mu\in I(A)$ é dita possuir dependência sensível às condições iniciais (dsci) se existe um boreleano B com $\mu(B)>0$ tal que $\lambda^1(x)>0$ para todo $x\in B$.

Será que um atrator caótico da aplicação A possui dsci? E se um atrator possui dsci é um atrator caótico?

A desigualdade de Margulis-Ruelle responde uma destas questões. Seja $\lambda^+(x)$ a soma dos EL positivos no ponto x, incluindo multiplicidade; $\lambda^+(x)=0$ se todos EL são negativos.

Teorema 6.4 (Desigualdade de Margulis-Ruelle) Para toda $\mu\in I(A)$ vale

$$h_\mu(A)\leq \int \lambda^+ \, d\mu.$$

A demonstração do teorema acima pode ser encontrada em [Rui].

Corolário 6.5 Um atrator caótico μ para A possui dsci.

A recíproca do corolário 6.5 não é verdadeira. Seja a

aplicação do exemplo 3) na seção 5b, δ_0 é o único atrator, e vê-se facilmente que $\lambda(0)=\lambda^{+}=\log 4$, e $h_{\delta_0}(A)=0$.

Outro exemplo pode ser dado pela figura ∞ de Bowen e Katok [Ka],[ER], considerando a aplicação no tempo um do fluxo indicado na figura 6.1, com um ponto fixo no infinito.

Esta aplicação possui apenas quatro medidas invariantes ergódicas, e estão concentradas nos pontos fixos p_1, p_2, p_∞ e p, mas o único atrator é δ_p, que atrai um conjunto de medida de Lebesgue total, possui $\lambda^{+}>0$ e $h_{\delta_p}=0$.

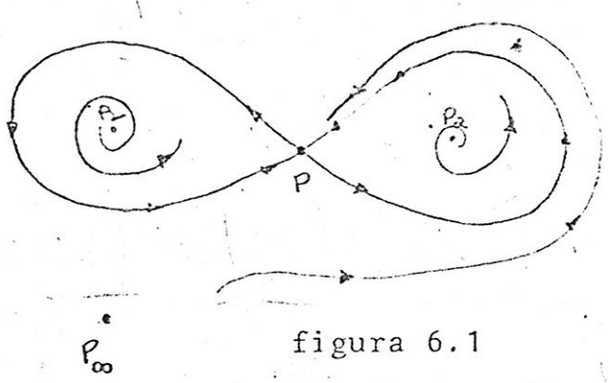

figura 6.1

Estes resultados mostram que dsci para um atrator não é conclusivo quanto a presença de atratores caóticos.

Para atratores caóticos há poucos resultados, apenas algumas classes de sistemas estão bem compreendidas como, por exemplo, os difeomorfismos de Anosov [Si2] e Axioma-A [Ru3], [Bo]. Veja também os resultados de Ledrappier e Young [LY].

Abaixo é apresentado um resultado que vale para atratores, caóticos ou não:

Definição 6.6 Um ponto $y\in X$ segue $x\in X$ topologicamente se $d(A^n x, A^n y)$ tende a zero quando n tende ao infinito.

Definição 6.7 Um ponto y segue x estatisticamente se

$$\lim_{n\to\infty}\sum_{i=0}^{n-1} A_*^i\delta_y/n = \lim_{n\to\infty}\sum_{i=0}^{n-1} A_*^i\delta_x/n,$$

supondo a existência do limite da direita.

Proposição 6.8 Se $\lim_{n\to\infty}\sum_{i=0}^{n-1} A_*^i \delta_x /n$ existe e y segue x topologicamente, então, y segue x estatisticamente.

Prova Note que a proposição garante a existência do $\lim_{n\to\infty}\sum_{i=0}^{n-1} A_*^i \delta_y /n$. Sejam $f\in C(X)$, $\varepsilon,\delta>0$ tais que $|f(w)-f(z)|<\varepsilon$ se $d(w,z)<\delta$, e $d(A^n x, A^n y)<\delta$ se $n\geq T(\delta)$, já que y segue x topologicamente. Vale que

$$\left|\lim_{n\to\infty}\sum_{i=0}^{n-1}[f(A^i x)-f(A^i y)]/n\right| \leq \lim_{n\to\infty}\sum_{i=0}^{n-1}|f(A^i x)-f(A^i y)|/n=$$

$$=\lim_{n\to\infty}\sum_{i=0}^{T-1}|f(A^i x)-f(A^i y)|/n + \lim_{n\to\infty}\sum_{i=T}^{n-1}|f(A^i x)-f(A^i y)|/n \leq$$

$\leq 0 + \lim_{n\to\infty}\varepsilon(n-T(\delta))/n = \varepsilon$. Como ε é arbitrário e o resultado vale para toda $f\in C(X)$, a demonstração fica completa. ©

Uma maneira ingênua de construir um atrator caótico seria pela união de um número arbitrário de órbitas periódicas, todas elas com algum EL>0. Abaixo está demonstrado que um atrator com estas características não é caótico. A idéia de dsci surgiu para quantificar a separação de órbitas próximas, e este resultado indica que se um atrator com dsci não é caótico, é porque "os pontos que se separam estão fora do atrator".

Para demonstrar o resultado indicado acima será utilizado o importante teorema da decomposição ergódica de uma medida invariante. Para a demonstração e maiores detalhes deste teorema veja a boa exposição do tema encontrada no livro de Mañé [Ma]. A versão aqui apresentada não é a mais geral, mas é adequada para os interesses deste trabalho.

Seja i_0 o conjunto de pontos estatisticamente regulares (veja a definição 3.18) de $A:X\to X$. Pela proposição 3.19 existe uma única medida invariante μ_x tal que $\lim_{n\to\infty}(S_n f)(x)/n=\int f\, d\mu_x$, para toda $f\in C(X)$ e $x\in i_0$ dado. Seja i o subconjunto de i_0 tal que μ_x é ergódica.

Proposição 6.9 Para toda $\mu \in I(A)$ vale que $\mu(i)=1$.

Teorema 6.10 (Decomposição Ergódica) Seja $\mu \in I(A)$. Toda $g \in L^1(\mu)$ é μ_x integrável para x num subconjunto de i de medida μ total, e

$$\int \left(\int g \, d\mu_x \right) d\mu(x) = \int g \, d\mu.$$

Por exemplo, se $g=\chi_B$, a função característica do boreleano B, então, pelo teorema 6.10, $\mu(B)=\int \mu_x(B) \, d\mu(x)$, para toda $\mu \in I(A)$. Esquematicamente pode-se escrever $\mu = \int \mu_x \, d\mu(x)$, e diz-se que o lado direito é a decomposição ergódica de μ.

Jacobs demonstrou que um resultado análogo vale para a entropia métrica vista como uma função da medida invariante, ou seja, $h_\mu(A) = \int h_{\mu_x}(A) \, d\mu(x)$. Veja [Ja], [Wa].

Teorema 6.11 Se o atrator $\nu \in I(A)$ é tal que existe um subconjunto $i_\perp \subset i$ com $\nu(i_\perp)=1$ e o suporte de μ_x é uma órbita periódica (instável) para todo $x \in i_\perp$, então ν não é atrator caótico.

Prova Para ν ser caótico $h_\nu(A) = \int h_{\mu_x}(A) \, d\mu(x) > 0$, ou ainda $h_{\mu_x}(A) > 0$ para x num conjunto de medida ν maior que zero. Como a medida μ_x está suportada numa órbita periódica (instável) vem que $h_{\mu_x}(A)=0$ para todo $x \in i_\perp$. Assim, $h_\nu(A)=0$. ☐

O teorema 6.11 mostra que não existem atratores caóticos "concentrados" apenas em órbitas periódicas, não importando quantas sejam e/ou quão instáveis sejam!

O próximo passo é quando os EL são menores ou iguais a zero (negativos) com o maior deles nulo, ou melhor, $\lambda^1(x)=0$ para x num comjunto de medida $\mu \in I(A)$ maior que zero, e não possui dsci. Neste caso várias possibilidades ocorrem, por exemplo, pode haver um ponto fixo instável ou assintoticamente estável. Outra

possibilidade é mostrada pela rotação $A: S^1 \to S^1$ $A(x)=x+\omega \pmod{2\pi}$ e ω é um irracional; a medida de Lebesgue μ é o único atrator e $\lambda^1(x)=0$ μ-qtp.

Um comportamento interessante foi observado por [GOPY], que analisaram nemericamente algumas aplicações, como

$$A(x,\theta) = (3 \tanh x \cos\theta,\ (\theta+2\omega)) \pmod{2\pi}),$$

com $\omega=(\sqrt{5}-1)/2$. As análises indicam a presença de um único atrator ν e com EL independentes da condição inicial, sendo $\lambda^1=0$ e $\lambda^2<0$. As simulações também indicam que o suporte do atrator é um conjunto fractal, veja a figura 6.2. Outros exemplos com estas características são conhecidos.

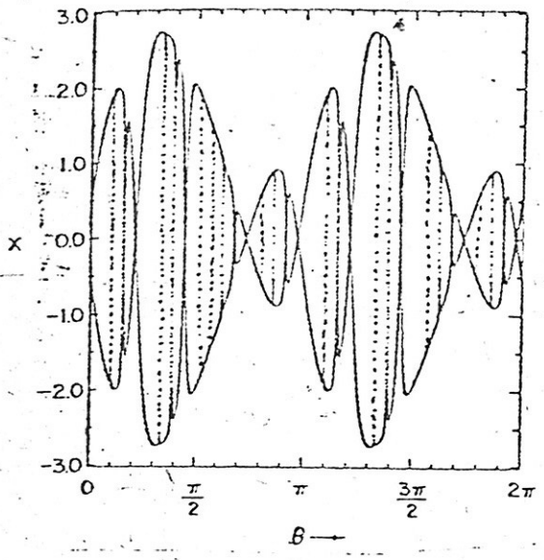

figura 6.2

Bem, olhar para os EL não é decisivo quanto ao comportamento assintótico de uma aplicação, pelo menos quando $\lambda^1 \geq 0$ em quase todo ponto, embora se não há dsci sabe-se que a entropia métrica é nula, e se a referida medida é um atrator, não é atrator caótico.

No próximo capítulo será dada uma classificação dos atratores com EL estritamente negativos num conjunto de medida atratora maior que zero.

CAPÍTULO SÉTIMO

EXPOENTES DE LYAPUNOV ESTRITAMENTE NEGATIVOS

Seria muito útil classificar os atratores a partir dos EL, mas tal classificação não é um problema simples.

O objetivo deste capítulo é demonstrar que os EL resolvem o problema de classificar um atrator se este é uma medida invariante que possui EL<0 num conjunto de medida maior que zero: seu suporte é uma órbita periódica assintoticamente estável.

No sexto capítulo foi discutido o problema de classificar os atratores através dos EL, e o problema é bem complicado.

Neste capítulo é feita uma classificação dos atratores com todos EL estritamente negativos, e para isso serão usados o teorema de decomposição ergódica de uma medida invariante (proposição 6.9 e teorema 6.10), e um resultado devido a Ruelle [Ru4] que classifica as medidas ergódicas que possuem EL estritamente negativos num conjunto de medida total.

A existência de medidas invariantes tais que todos EL são menores que zero num conjunto de medida um, e que não são ergódicas é óbvia, salvo casos particulares. Considere, por exemplo, a medida invariante $\mu=(\mu_1+\mu_2)/2$, onde μ_1 e μ_2 são medidas ergódicas concentradas em órbitas periódicas assintoticamente estáveis distintas, cada uma com todos os EL estritamente negativos.

Teorema_Z.1 [Ru4] Sejam $A:X\to X$ uma aplicação C^2, onde X é uma variedade lisa compacta e $\tau\in I(A)$ ergódica. Se existe um boreleano B tal que $\tau(B)=1$ e todos os EL são estritamente negativos para todo $x\in B$, então o $su(\tau)$ é uma órbita periódica assintoticamente estável.

No teorema_Z.1 a aplicação poderia ser Hölder C^1; a demonstração pode ser encontrada na referência acima.

Do teorema_6.4 vem que $h_\mu(A)\leq\int\lambda^+d\mu$ para toda $\mu\in I(A)$. Se $\lambda^\perp<0$ μ-qtp e μ é ergódica, então (teorema_Z.1) $h_\mu(A)=0$ e o suporte de μ é uma órbita periódica assintoticamente estável. Se μ não é ergódica e/ou $\lambda^\perp(y)\geq 0$ para y num conjunto de medida maior que zero não é possível, até o momento, fazer qualquer afirmação. É bom lembrar que verificar estas propriedades na prática é, em geral, extremamente difícil, assim, o corolário_Z.3 é importante por não supor a ergodicidade do atrator!

Vale que $h_\mu(A)=\int h_{\mu_x}(A)\,d\mu(x)$, onde $\mu=\int\mu_x\,d\mu(x)$ é a decomposição ergódica de μ. Se existe um conjunto F com $\mu(F)>0$ tal que $\lambda^\perp(x)<0$ para todo $x\in F$, poderia ocorrer que $\lambda^\perp>0$ no complementar de F, e assim $h_\mu(A)>0$, sendo μ um candidato a atrator caótico. O corolário_Z.3 mostra que este nunca será o caso.

Teorema_Z.2 Sejam $A:X\to X$ como no teorema_Z.1 e $\nu\in I(A)$. Se $G_\nu\neq\phi$ e se existe um boreleano F com $\nu(F)>0$ tal que todos os EL são menores que zero para todo $x\in F$, então ν é ergódica e seu suporte é uma órbita periódica assintoticamente estável.

Corolário_Z.3 Se $\nu\in I(A)$ é um atrator e existe um boreleano F como acima, então ν é ergódica e seu suporte é uma órbita periódica assintoticamente estável. Em particular, $h_\nu(A)=0$.

Prova (teorema 7.2) Por hipótese $\int \lambda^{\perp} d\nu < 0$, e usando a decomposição ergódica de ν vem que

$$\int_F \lambda^{\perp} d\nu = \int \chi_F \cdot \lambda^{\perp} d\nu = \int [\int_F \lambda^{\perp} d\mu_y] \, d\nu(y) < 0,$$

assim existe um boreleano t, com $\nu(t) > 0$ e $\int_F \lambda^{\perp} d\mu_y < 0$ para todo y em t (se $y \in t$ μ_y é ergódica (teorema 6.10)).

Lema 7.4 A cardinalidade de t é no máximo enumerável, e se $y \in t$ a medida μ_y possui suporte numa órbita periódica assintoticamente estável.

Prova Se $y \in t$ então $\lambda^{\perp}(z) < 0$ para todo z num conjunto de medida μ_y maior que zero, já que $\int_F \lambda^{\perp} d\mu_y < 0$ para todo $y \in t$.

Como $\lambda^{\perp}(\cdot)$ é invariante nas órbitas de A conclui-se, da ergodicidade de μ_y, que $\lambda^{\perp}(z) < 0$ para z num conjunto de medida μ_y total, para todo $y \in t$.

Desta forma, ficam satisfeitas as condições do teorema 7.1 para $\tau = \mu_y$; assim, μ_y possui suporte numa órbita periódica assintoticamente estável, sempre que $y \in t$.

Como cada órbita periódica assintoticamente estável possui vizinhança aberta tal que todo ponto dessa vizinhança converge à órbita periódica pela aplicação sucessiva de A, e essa vizinhança possui medida de Lebesgue maior que zero, segue que a cardinalidade de t é no máximo enumerável, já que duas destas vizinhanças são disjuntas. ©

Usando o lema 7.4 pode-se escrever a medida ν na forma

$$\nu = \sum_{y \in t}{}' q_y \mu_y + q\sigma \qquad \ldots\ldots\ldots (7.1)$$

onde $\sum'$ indica soma sem repetição em μ_y ($y \in t$), e q_y e q são números reais positivos tais que $\sum_{y \in t}{}' q_y > 0$ e $q + \sum_{y \in t}{}' q_y = 1$, no sentido que

$$\nu(f)=\sum_{y\in t} q_y\,\mu_y(f)+q\nu(f)$$

para toda f em C(X).

A decomposição (7.1) pode ser comprovada integrando-se uma função contínua $f\in C(X)$ arbitrária,

$$\nu(f) = \int f\, d\nu = \int \mu_y(f)\, d\nu(y) =$$

$$= \int_t \mu_y(f)\, d\nu(y) + \int_{X\setminus t} \mu_y(f)\, d\nu(y);$$

como t é enumerável, o primeiro termo da soma fica

$$\sum_{y\in t}\mu_y(f)\, d\nu(y) = \sum_{y\in t}{}' \nu(su(\mu_y))\ \mu_y(f),$$

já que o suporte de μ_y é uma órbita periódica.

Agora, se $\nu(X\setminus t)\neq 0$ seja

$$L(g) = [\int \chi_{X\setminus t}(y)\ \mu_y(g)\ d\nu(y)]/\nu(X\setminus t)$$

para $g\in C(X)$. Como L é um funcional linear contínuo positivo e $L(1)=1$ vem, pelo teorema 3.3, que existe uma medida $\xi\in M(X)$ tal que $L(g)=\xi(g)$ para toda $g\in C(X)$. Pela definição do funcional L é imediato que a medida ξ é um elemento de $I(A)$.

A partir das considerações acima, vem que a decomposição (7.1) é válida fazendo a seguinte correspondência:

$$q_y=\nu(su(\mu_y));$$

se $\nu(X\setminus t)\neq 0$ então $\xi=\sigma$ e $q=\nu(X\setminus t)$, e

se $\nu(X\setminus t)=0$ então $q=0$.

Como $\nu(t) = \sum_{y\in t}\nu(y) = \sum_{y\in t}{}'\nu(su(\mu_y)) > 0$, e t é no máximo enumerável, vem que $\nu(su(\mu_w))>0$ para algum $w\in t$, ou seja, existe $w\in t$ tal que $q_w=r>0$. Sem perda de generalidade pode-se supor que $\mu_w=\delta_p$, para algum p em X.

Como p é um ponto fixo assintoticamente estável, veja o lema 7.4, vem que G_{δ_p} é não vazio e contém uma vizinhança

aberta U de p.

Para terminar a demonstração, será provado que todo elemento de G_ν visita o aberto U.

Lema Z.5 $G_\nu \subset G_{\delta_p}$.

Prova Se $G_\nu = \phi$ o lema é trivial. Suponha que $G_\nu \neq \phi$ e que $x \in G_\nu$.

Usando simples manipulações pode-se escrever

$$\nu = r\delta_p + (1-r)\mu,$$

onde $r>0$ e $\mu \in I(A)$, no sentido de que

$$\nu(f) = r\delta_p(f) + (1-r)\mu(f)$$

para toda $f \in C(X)$.

Seja $\varepsilon>0$ tal que $\bar{B}(p,2\varepsilon) \subset U$, onde $\bar{B}(p,2\varepsilon)$ é o fecho da bola aberta de centro p e raio 2ε. Pelo lema de Urysohn existe uma função contínua $u_\varepsilon(s)$ tal que

$$u_\varepsilon(s) = \begin{cases} 1 \text{ se } s \in \bar{B}(p,\varepsilon) \\ 0 \text{ se } s \in X \setminus B(p,2\varepsilon) \\ 0 \leq u_\varepsilon \leq 1 \text{ de outra maneira.} \end{cases}$$

Vale que

$$\liminf_{n\to\infty} (S_n \chi_{B(p,2\varepsilon)})(x)/n \geq \lim_{n\to\infty} (S_n u_\varepsilon)(x) = \int u_\varepsilon \, d\nu,$$

já que x é um elemento de G_ν. Mas,

$$\int u_\varepsilon \, d\nu = r\int u_\varepsilon \, d\delta_p + (1-r)\int u_\varepsilon \, d\nu \geq r\int u_\varepsilon \, d\delta_p = r > 0.$$

Portanto

$$\liminf_{n\to\infty} (S_n \chi_{B(p,2\varepsilon)})(x)/n > 0,$$

assim existe $n \in N$ tal que $A^n(x) \in B(p,2\varepsilon) \subset U \subset G_{\delta_p}$; na realidade, existem infinitos valores de n que satisfazem esta condição. ©

Do lema Z.5 segue que, se $G_\nu \neq \phi$ então $r=1$, e fica provado o teorema Z.2. ©

O corolário 7.3 dá uma classificação completa dos atratores com todos os EL estritamente negativos. Um atrator não pode possuir um conjunto de medida não nula com todos EL estritamente negativos, e ainda ser um atrator caótico. Mais especificamente, um atrator possui EL não negativos num conjunto de medida zero ou um; no primeiro caso o atrator é ergódico e seu suporte é uma órbita periódica assintoticamente estável.

Corolário 7.6 Se $\mu \in I(A)$ é atrator caótico para A, então $\lambda^{\mu}(x) \geq 0$ para x num conjunto de medida μ total.

Uma órbita periódica assintoticamente estável de período muito longo pode, por exemplo, ter a aparência de um fractal, e é esperado que os resultados acima possam, em alguns casos, mostrar que o conjunto observado num experimento é uma órbita periódica, e não um fractal, desde que este será o caso se os EL forem estritamente negativos.

O corolário 7.3 mostra, também, uma maneira prática de procurar órbitas periódicas em um dado sistema físico, basta encontrar uma órbita estatísticamente regular com todos EL menores que zero.

CAPÍTULO OITAVO

OBSERVAÇÕES FINAIS E CONCLUSÃO

Os expoentes de Lyapunov são amplamente utilizados para caracterizar comportamentos assintóticos de sistemas físicos estudados em simulações numéricas e em laboratório, e é comum aceitar (de forma ingênua) que a presença de EL>0 implica em caos, conjecturando-se que a entropia métrica é igual à soma dos EL positivos (veja o teorema 6.4).

A primeira dificuldade encontrada neste trabalho foi a necessidade de uma definição de atrator que compatibilizasse o uso dos EL e da entropia métrica nos estudos dos comportamentos assintóticos de sistemas físicos. A segunda dificuldade surgiu quando, partindo de trabalhos na literatura especializada, concluiu-se que conhecer apenas os EL não é suficiente para se conhecer um atrator.

A definição de atrator dada no capítulo quinto é corroborada por vários resultados demonstrados no trabalho, embora seja difícil descrever seu domínio de aplicabilidade devido, principalmente, à suposição de que médias temporais de funções contínuas existem num conjunto de medida de Lebesgue maior que zero.

Se não se associar uma medida invariante à idéia intuitiva de atrator não se pode falar em expoentes de Lyapunov, tampouco de entropia métrica, e assim, não é possível utilizar os conceitos acima no estudo de comportamentos assintóticos, e aleatórios em particular.

As principais contribuições deste trabalho podem ser

resumidas da seguinte forma:

"Na seção 5B foi dada uma definição de atrator (definição 5.1) para aplicações contínuas $A: X \to X$, onde X é variedade compacta lisa. Os resultados conhecidos nos quais tal definição está baseada foram apresentados nas seções 3A, 3C e 5A.

Esta definição de atrator é fortalecida, principalmente, pelas proposições 5.3, 5.7, 5.8 e 5.9, onde foram demonstrados que -segundo a definição 5.1- :

(i) o conjunto de atração (definição 5.2) de qualquer atrator é mensurável, mesmo não sendo necessariamente um subconjunto aberto,

(ii) se o ponto x de X é atraído pelo atrator $\mu \in I(A)$ então $\mu(\omega(x))=1$ (onde $\omega(x)$ é o conjunto ω-limite de x),

(iii) é possível aproximar (na topologia fraca*) um atrator, de várias maneiras diferentes, por medidas absolutamente contínuas em relação à medida de Lebesgue, e

(iv) existem, no máximo, um número infinito contável de atratores para cada aplicação contínua.

Quanto a utilidade da definição 5.1 foi demonstrado, no sétimo capítulo, que um atrator para uma aplicação diferenciável C^2 com todos EL<0 num conjunto de medida atratora maior que zero é ergódico e possui suporte numa órbita periódica assintoticamente estável. Como corolário deste resultado segue que um atrator caótico, ou seja, um atrator com entropia métrica não nula, possui EL $\geqslant 0$ num conjunto de medida atratora total".

Em geral, resultados são obtidos para medidas invariantes ergódicas, e os resultados do capítulo sétimo são uns dos poucos onde não é suposta a ergodicidade do atrator. Se existe um subconjunto de medida maior que zero onde todos os EL são estritamente negativos, foi mostrado que, se esta medida é um atrator, a ergodicidade é a única possibilidade e, ainda mais, este atrator

possui suporte numa órbita periódica assintoticamente estável.

Usar os EL para classificar os atratores seria um caminho útil, principalmente porque eles podem ser calculados tanto numericamente quanto a partir de dados experimentais, mas quando há EL positivos, como mencionado no capítulo sexto, tal classificação não é simples.

Uma classificação geométrica das medidas que satisfazem a igualdade no teorema 6.1 para difeomorfismos C^2, foi dada por [LY], mas verificar essas condições numa dada aplicação é, geralmente, muito difícil; os difeomorfismos de Anosov e Axioma-A verificam essas condições.

A influência dos resultados da teoria ergódica dos difeomorfismos C^2 de Anosov [Si] e Axioma-A [Ru3] em todo este trabalho é evidente, basta observar o exemplo 6 da seção 5A. O atrator σ referido neste exemplo é ergódico, assim, se todos os EL<0 vale o teorema 7.1 (observe que G_σ contém um conjunto aberto!). Infelizmente, as construções matematicas de atratores, como os Axioma-A, não tem sido descobertos em sistemas reais.

Uma área muito ativa no momento é o estudo dos "fenômenos caóticos" em mecânica quântica, sendo este um dos ramos fundamentais da física atual; esses estudos são delicados. Por exemplo, é comum o uso de variáveis de ângulo-ação, mas vários problemas podem surgir na introdução destas variáveis em mecânica quântica, veja [dOM1]. Pelo que o autor saiba, não existem trabalhos sobre este tema na literatura matemática.

Existem relações interessantes entre EL, entropia métrica e dimensão (fractal, de Hausdorff, ...) do suporte de uma medida invariante; alguns resultados podem ser encontrados em [Le], veja também [LY].

Generalizações para o caso de turbulência em fluidos

em três dimensões são necessárias, inclusive estes estudos podem envolver elementos novos, de forma que a ênfase seja em dificuldades técnicas e na imaginação.

A longo prazo, estes problemas deverão formar um capítulo dos livros-texto de equações diferenciais.

REFERÊNCIAS

[AA] V. I. Arnold, A. Avez, Ergodic Problems of Classical Mechanics Benjamin, N.Y., (1968).

[AAr] F. Argoul, A. Arneodo, "Lyapunov Exponents and Phase Transitions", LNM 1186 (Springer) (1984).

[ABS] J. Auslander, N. P. Bhatia, P. Seibert, "Attractors in Dynamical Systems". Bol. Soc. Mat. Mex. 9(1964)55.

[AKM] R. L. Adler, A. G. Konheim, M. H. McAndrew, "Topological Entropy". Trans. Am. Math. Soc. 114(1965)309.

[AKO] L. Arnold, W. KKliemann, E. Oeljeklaus, "Lyapunov Exponents of Linear Stochastic Systems", LNM 1186 (Springer) (1984).

[Bi] P. Billingsley, Ergodic Theory and Information, Wiley, N.Y., (1965).

[BGGS] G. Benettin, L. Galgani, A. Giorgilli, J.-M. Strelcyn, "Lyapunov Characteristic Exponents ..." Part 1 and 2, Meccanica (1980)9.

[Bo] R. Bowen, Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms, LNM 470 (Springer, Berlin) (1975).

[CE] P. Collet, J.-P. Eckmann, Iterated Maps on The Interval as Dynamical Systems, Birkhauser, Cambridge, MA, (1980).

[CD] M. Cosnard, J. Demongeot, "Attracteurs: une approche deterministe". C. R. Acad. Sc. Paris 300, 15(1985)551.

[Co] G. Contopoulos, "Nonlinear Problems in Stellar Dynamics", Springer Proceedings in Physics 3, (Springer-Verlag) (1985).

[Cr] H. Crauel, "Lyapunov Exponents and invariant measures os stochastic systems", LNM 1186 (Springer) (1984).

[Cu] J. H. Curry, "On Computing the Entropy of the Hénon Attractor" J. Stat. Phys. 26(1981)683.

[deA] C. D. de Andrade, Corpo, Ed. Record (São Paulo).

[deOM1] C. R. de Oliveira, C. P. Malta, "Action-Angle Variables for the Harmonic Oscilator: Ambiguity Spin x Duplication Spin". Annals of Physics 155, 2(1984)447.

[deOM2] C. R. de Oliveira, C. P. Malta, "Bifurcations in a Class of Time Delay Equations", a aparecer em Physical Review A.

[Ec] J.-P. Eckmann, "Roads to Turbulence in Dissipative Dynamical Systems". Rev. Mod. Phys. 53(1981)643.

[EKRC] J.-P. Eckmann, S. O. Kamphorst, D. Ruelle, S. Ciliberto, "Liapunov Exponents from Time Series", priprint IHES (1986).

[El] T. S. Eliot, Poesia, Ed. Nova Fronteira (São Paulo).

[Fe] S. D. Feit, "Characteristic Exponents and Strange Attactors", Comm. Math. Phys. 61(1978)249.

[GH] J. Guckenheimer, P. Holmes, Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields, (Springer) (1983).

[GOPY] C. Grebogi, E. Ott, S. Pelikan, J. A. Yorke, "Strange Attractors That are not Chaotic", Physica 13D, (1984)261.

[GS] L. Garrido, C. Simó, "Some Ideas About Strange Attractors", LNP 179 (Springer), (1983).

[Ha] H. Haken, Physics Letters A 53(1975)77.

[HC] B. A. Huberman, J. Cruchfield, Phys. Rev. Lett. 43(1979)1743.

[He] M. Hénon, "A two-dimensional mapping with a strange attractor" Comm. Math. Phys. 50(1976)69.

[Ja] K. Jacobs, em Ergodic Theory, ed. F. B. Wright, Academic Press 173.

[Ka] A. Katok,, "Lyapunov Exponents, Entropy and Periodic Orbits for Diffeomorphisms". Publ. Math. IHES 51(1980)137.

[Khi] A. I. Khinchin, Statistical Mechanics, Dover, N. Y., (1949).

[Kh2] A. I. Khinchin, Information Theory, Dover, N. Y., (1957).

[La1] O. E. Lanford, "Smooth Transformations on Intervals", Sem. Bourbaki 563, (1980/81).

[La2] O. E. Lanford, "Introduction to the Mathematical Theory of Dynamical Systems", Les Houches-Session XXXVI, (1983).

[Le] Y. Levy, "Ergodic Properties of the Lozi Map", LNM 1109 (Springer) (1983).

[Lo] E. N. Lorenz, "Deterministic Nonperiodic Flow", J. Atmos. Sci. 20(1963)130.

[LY] F. Ledrappier, L.-S. Young, "The Metric Entropy of Diffeomorphisms". Part 1 and 2. Ann. of Math. 122(1985)509.

[Ma] R. Mañé, Teoria Ergódica, Projeto Euclides CNPq, (1983).

[MaL] A. Martin-Lof, Statistical Mechanics and the Foundations of Thermodynamics, LNP 101, (Springer-Berlin,Heidelberg), (1979).

[McML] J. B. McLaughlin, P. C. Martin, "Transition to turbulence in a statically stressed fluid system", Phys. Rev. A 12(1975)186.

[Mi] M. Misiurewicz, "Structure of Mappings of an Interval with Zero Entropy". Publ. Math. IHES 53(1981)5.

[Mil] J. Milnor, "On the Concept of Attractor", Comm. Math. Phys. 99(1985)177. Comm. Math. Phys. 102(1985)517.

[MR] D. Mayer, G. Roepstorff, "Strange Attractors and Assynptotic Measures of Discrete-Time Dissipative Systems", J. Stat. Phys. 31, 2(1983)309.

[Os] V. I. Oseledec, "A Multiplicative Ergodic Theorem. Lyapunov Characteristic Numbers for Dynamical Systems", Moscow Math. Soc. 19(1968)197.

[Pe] O. Penrose, "Foundations of Statistical Mechanics", Rep. Prog. Phys. 42(1979)129.

[Pet] K. Petersen, Ergodic Theory, Cambridge Univ. Press, (1983).

[Ra] M. S. Raghunathan, "A Proof of Oseledec's Multiplicative Ergodic Theorem". Israel J. Math. 32(1979)365.

[RT] D. Ruelle, F. Takens, "On the Nature of Turbulence". Comm. Math. Phys. 20(1971)167.

[Ru1] D. Ruelle, "An Inequality for the Entropy of Differentiable Maps". Bol. Soc. Bras. Mat. 9(1978)83.

[Ru2] D. Ruelle, "Small Randon Perturbation of Dynamical Systems and the Definition of Attractors". Comm. Math. Phys. 82(1981)137.

[Ru3] D. Ruelle, "A Measure Associated with Axiom A Attractors". Am. J. Math. 98(1976)619.

[Ru4] D. Ruelle, "Ergodic Theory of Differentiable Dynamical

Systems". Publ. Math. IHES 50(1979)275.

[Si] Ya. G. Sinai, "Gibbs Measure in Ergodic Theory". Russ. Math. Surveys 27, 4(1972)21.

[Sm1] S. Smale, "Differentiable Dynamical Systems", Bull. Am. Math. Soc. 73(1967)747.

[Sm2] S. Smale, "Dynamical Systems and Turbulence", LNM 615 (Springer)

[SN] I. Shimada, T. Nagashima, "Numerical Approach to Ergodic Problem of Dissipative Dynamical Systems". Prog. Theor. Phys. 61(1979)1605.

[SS] M. Sano, Y. Sawada, "Measurement of the Lyapunov Spectrum from a Chaotic Time Series". Phys. Rev. Lett. 55, 10(1985)1082.

[Ta] S. J. Taylor, Introduction to Measure and Integration, Cambridge University Press (1973).

[Th] R. Thom, Modele Mathématiques de la Morphogénese, Bourgois Ed. Paris, (1980).

[TK] K. Tomita, T. Kai, J. Statistical Physics 21(1979)65.

[TM] Y. Treve, O. P. Manley, "Intrinsic Stochasticity in Plasmas", ed. G. Laval and D. Gresillon (Courtaboeuf-Orsay) (1980)393.

[Wa] P Walters, Introduction to Ergodic Theory, GTM 79, Springer,

(1982).

[WSSV] A. Wolf, J. B. Swift, H. L. Swinney, J. A. Vastano, 'Determining Lyapunov Exponents from a Time Series' Physica 16D, (1985)285.

[Za] G. M. Zaslavskii, Physics Letters A 69 (1978)145.