UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE FÍSICA

CARACTERIZAÇÃO MULTIFRACTAL

MARCOS YAMAGUTI

SBI-IFUSP
305M810T2667

Tese de Doutoramento
submetida ao Instituto de Física
da Universidade de São Paulo

Orientadora: Profa. Carmen Pimentel Cintra do Prado

Banca examinadora:
Profa. Carmen Pimentel Cintra do Prado
Prof. Carlos Eugênio I. Carneiro
Prof. Iberê Luiz Caldas
Prof. Evaldo Mendonça F. Curado
Prof. José Roberto Drugowich de Felício

UNIVERSIDADE DE SÃO PAULO
ex. 1
SERVIÇO DE BIBLIOTECA E INFORMAÇÃO
2667
INSTITUTO DE FÍSICA

Prof. Iberê Luiz Caldas
Presidente da Comissão de Pós-Graduação

São Paulo
1997

530.155352
Y19c
D
ex.1

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

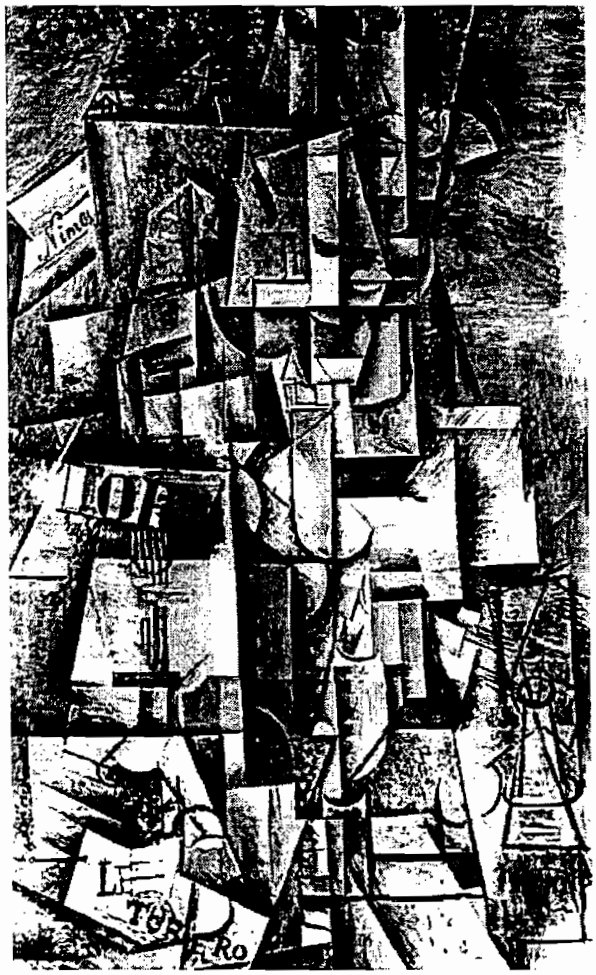
Yamaguti, Marcos
Caracterização Multifractal. São Paulo, 1997.

Tese (Doutoramento) Universidade de São Paulo. Instituto de Física - Departamento de Física Geral.

Área de Concentração: Física do Estado Sólido
Orientador: Profa. Dra. Carmen Pimentel Cintra do Prado

Unitermos: 1. Espectro de Singularidades; 2. Dimensões Generalizadas; 3. Multifractal; 4. Caos Determinístico; 5. Box Counting.

USP/IF/SBI - 27/97



O aficcionado, de Picasso

"... A natureza vê-se reduzida a sólidos geométricos: cubos ou paralelepípedos (as construções), colocados abaixo de pirâmides (os telhados), ao lado de cones, esferas e cilindros (a vegetação)... A ilusão de profundidade é substituida por uma espécie de análise ou decomposicão do objeto... o objeto real se dilui, ficando apenas algumas alusões. De *O Aficcionado* só se pode garantir que tinha bigodes."

Extraído de Os Gênios da Pintura, sobre o Cubismo.

À família,
Às oportunidades,
Aos computadores.

Agradecimentos

Ao Diretor do Instituto de Física da USP, Prof. Gil da Costa Marques,

ao Chefe do Departamento de Física Geral, Prof. Olacio Dietzsch,

e ao Chefe do Grupo de Mecânica Estatística, Prof. Silvio Roberto de Azevedo Salinas pela oportunidade oferecida.

À minha orientadora, Profa. Carmen Pimentel Cintra do Prado, pelo trabalho de orientação, pelos comentários precisos e práticos, e pelo respeito à minha liberdade.

À Kaline Coutinho, pela parceria na administração dos recursos computacionais do grupo, pela amizade e pelo seu alto-astral.

À todos colegas e amigos, pela ajuda, pelas boas conversas, e pelo tempo que passamos juntos. Valeu !

Ao pessoal da secretaria do grupo e do departamento, um obrigado pelo suporte básico.

À FAPESP pelo apoio financeiro.

Ao leitor, pelo interesse.

Um obrigado especial ao Nestor Oiwa, pela idéia que muito contribuiu a este trabalho.

Resumo

A caracterização estática dos sistemas caóticos clássicos dissipativos tem sido realizada através do cálculo das dimensões generalizadas D_q e do espectro de singularidades $f(\alpha)$. Os métodos mais comuns de cálculo numérico dessas funções utilizam algoritmos de contagem de caixa. Porém, esses algoritmos produzem um erro sistemático através de 'caixas espúrias', levando a resultados distorcidos. Por essa razão, estudamos métodos numéricos que não utilizam o algoritmo de contagem de caixa, verificando em que casos eles podem ser aplicados eficazmente e propusemos um novo algoritmo de contagem de caixa que reduz o número de 'caixas espúrias', obtendo melhores resultados.

Abstract

The static caracterization of classical dissipative chaotical systems has been achieved by the calculation of the generalized dimensions D_q and the spectrum of singularities $f(\alpha)$. The most used numerical methods of evaluating these functions are based on box counting algorithms. The results obtained by those methods are distorced by the presence of 'spurious boxes' generated intrinsecally by these algorithms. For this reason, we have studied numerical methods that don't use box counting algorithms, and we have tried to verify in which kind of sets they give best results. We also have proposed a new box counting algorithm that reduces the number of 'spurious boxes', and led to better results.

Índice

Introdução **1**

1 Conceitos e Métodos Numéricos **5**

1.1 Dimensão fractal 5

1.1.1 Dimensão de Hausdorff-Besicovitch 5

1.1.2 Dimensão de caixa 7

1.2 Outras definições de dimensão 7

1.3 Caracterização multifractal completa 8

1.3.1 Dimensões generalizadas 8

1.3.2 Espectro de singularidades 10

1.3.3 Relação entre as dimensões generalizadas e o espectro de singularidades 11

1.4 Classes de métodos para o cálculo numérico de $f(\alpha)$ 13

1.4.1 Contagem de caixa 14

1.4.2 Contagem de distâncias 14

1.5 Métodos numéricos para o cálculo de $f(\alpha)$ 15

1.5.1 Método microcanônico 16

1.5.2 Método canônico 16

1.5.3 Método dos histogramas 17

1.5.4 Método da integral de correlação 18

1.5.5 Método da massa fixa 19

2 Conjuntos multifractais utilizados **21**

2.1 Conjuntos de Cantor generalizados 21

2.2 Atrator do mapa quadrático 22

2.3 Atrator do mapa de Hénon 23

3 O problema das caixas espúrias **25**

4 Métodos de Contagem de Caixa **29**

4.1 Correções Estatísticas 29

4.2 Método Microcanônico 30

4.3 Método Canônico 37

4.4 Método dos Histogramas 44

4.4.1 Parâmetros do método 44

4.4.2 Aplicando o método 45

5 Contagem de Distâncias **53**

5.1 Método da massa fixa 53

5.2 Método da integral de correlação 60

6 Método canônico da integral de correlação **67**

6.1 Teoria 67

6.2 Resultados Numéricos 68

7 Um novo algoritmo de contagem de caixa **77**

7.1 Introdução 77

7.2 Resultados 81

7.2.1 Método Microcanônico 82

7.2.2 Método Canônico 90

7.2.3 Método dos Histogramas 99

Comentários e Conclusões **105**

A Conjuntos de Cantor Generalizados **109**

A.1 Sistema de funções iteradas 109

A.1.1 Algortimo 110

A.2 Método da escada do diabo 111

A.2.1 Algoritmo 112

B Algoritmo de contagem de caixa **115**

B.1 Método numérico 115

C Algoritmo para o Método Microcanônico **119**
C.1 Método numérico 119
C.2 Algoritmo 120

D Algoritmo para o Método Canônico **123**
D.1 Método numérico 123
D.2 Algoritmo 124

E Algoritmo para o Método dos histogramas **127**
E.1 Método numérico 127
E.2 Algoritmo 128

F Algoritmo auxiliado por caixas **131**
F.1 Método numérico 131
F.2 Algoritmo 132

G Algoritmo para o Método da integral de correlação **135**
G.1 Método numérico 135
G.2 Algoritmo 136

H Algoritmo para o Método canônico da integral de correlação **139**
H.1 Método numérico 139
H.2 Algoritmo 140

I Algoritmo para o Método da massa fixa **141**
I.1 Método numérico 141
I.2 Algoritmo 142

J Regresão Linear **145**

K Derivada numérica **147**

L Novo algoritmo de contagem de caixa **151**
L.1 Introdução 151
L.1.1 Variáveis 151
L.2 Primeira parte do algoritmo 152
L.2.1 Algoritmo 153
L.3 Segunda parte do algoritmo 153

M Tempos de Processamento **163**

Introdução

O objetivo deste trabalho é prosseguir com o estudo do problema das chamadas 'caixas espúrias', presente em métodos que realizam o cálculo numérico da função $f(\alpha)$ usando algoritmos de contagem de caixa, problema este detectado em nosso trabalho de mestrado.

A função $f(\alpha)$, segundo o formalismo multifractal, caracteriza completamente um objeto multifractal. Esse formalismo, por sua vez, teve origem nos estudos de sistemas caóticos dissipativos.

Nas últimas décadas ocorreu o desenvolvimento de inúmeros trabalhos sobre sistemas dinâmicos caóticos [3, 34]. Esses sistemas dinâmicos apresentam uma propriedade chamada de sensitividade às condições iniciais, ou seja, trajetórias que iniciam em estados quase iguais vão se tornando exponencialmente diferentes com o desenvolvimento dessas dinâmicas. Dessa forma, após um certo tempo, elas se tornam imprevisíveis. Apesar desse caráter que sugere um mecanismo estocástico, tais dinâmicas são descritas por processos completamente determinísticos. Essas características exigiram novos conceitos e novos métodos, desenvolvidos na tentativa de classificar e entender tais dinâmicas.

Basicamente, os sistemas dinâmicos caóticos estão classificados em conservativos e dissipativos dependendo se a energia é conservada ou não pelo processo dinâmico. Os sistemas dinâmicos caóticos conservativos possuem atualmente um grande instrumental teórico [18, 31], e muito se compreende do mecanismo dinâmico desses sistemas e de suas propriedades. O estudo dos sistemas dinâmicos caóticos dissipativos, porém, tem se desenvolvido de forma mais lenta.

O estudo dos sistemas dinâmicos caóticos dissipativos é baseado na suposição de que o teorema ergódico pode ser aplicado, ou seja, supõe-se que médias em uma trajetória após um tempo infinito e médias espaciais sejam equivalentes em tais sistemas. A distribuição de probabilidades que torna equivalentes essas médias é chamada de medida invariante da dinâmica. Há indicações vindas do cálculo numérico de que o

suporte geométrico da medida invariante desses sistemas é, em uma grande parte dos casos, um objeto fractal. Este tipo de conjunto é conhecido como atrator estranho.

Fractais são objetos geométricos que possuem detalhes em todas as suas escalas de ampliação. Um estudo paralelo feito por matemáticos [26, 27] define e mostra como eles podem ser caracterizados. Um pouco mais tarde, essa definição seria ampliada, com a introdução do conceito de objeto multifractal, e um instrumental próprio seria criado para sua caracterização, definido pela teoria multifractal [32, 9].

A medida invariante associada a um atrator estranho pode ser vista como um objeto multifractal, e ao caracterizar-se dessa forma a medida invariante das diversas dinâmicas caóticas dissipativas, espera-se poder dividi-las em classes de universalidade, o que poderia se constituir em um grande avanço na classificação das mesmas.

A teoria multifractal define duas curvas, conhecidas como dimensões generalizadas D_q e espectro de singularidades $f(\alpha)$, que são capazes de caracterizar completamente um multifractal. Entretanto, o cálculo teórico dessas curvas é muito difícil de ser realizado, existindo pouquíssimos casos onde uma solução analítica é encontrada. Normalmente o cálculo é feito numericamente, e mesmo assim, diversas condições fixadas pela teoria são relaxadas para que ele seja possível. Diversos métodos para o cálculo numérico dessas curvas foram propostos na literatura [1, 4, 10, 28, 33].

Em nosso trabalho de mestrado, estudamos métodos para o cálculo numérico de $f(\alpha)$ baseados em algoritmos de contagem de caixa (BC) [7, 29, 16]. Nesse algoritmo, o conjunto estudado é recoberto por uma grade, que é um conjunto de caixas de igual tamanho justapostas entre si, e conta-se qual a fração de pontos do conjunto há no interior de cada caixa. Essa fração de pontos aproxima a integral da medida invariante no interior da caixa, e é comumente chamada de probabilidade. Realizando o levantamento dessas probabilidades em grades de diferentes tamanhos de caixa, obtém-se a curva $f(\alpha)$.

Observamos que existem distorções na curva $f(\alpha)$ resultantes da aplicação desses métodos, e pudemos mostrar através de um estudo detalhado de Conjuntos de Cantor Generalizados que elas são produzidas pela presença de caixas que são quase completamente preenchidas pelo vazio da estrutura multifractal. Dessa forma, essas caixas são consideradas erroneamente pelos métodos numéricos de contagem de caixa como caixas com baixa probabilidade, distorcendo a curva final. O aparecimento dessas caixas se deve ao fato do algoritmo de contagem de caixa utilizar uma grade como recobrimento do conjunto estudado. Se o recobrimento fosse realizado com todo o rigor

descrito pela teoria, tais caixas não existiriam. Por essa razão, as denominamos de caixas espúrias. O trabalho descrito nesta tese de doutorado é relacionado ao estudo e, principalmente, à busca de soluções para este problema.

Desenvolvemos as seguintes estratégias na busca de soluções para o problema das caixas espúrias:

- Uma vez que essas caixas surgem em todos os métodos que utilizam o algoritmo de contagem de caixa, uma possível solução seria utilizar métodos que não usam esse algoritmo. Estudamos e comparamos outros métodos encontrados na literatura que não se utilizam do algoritmo de contagem de caixa. Esses métodos calculam e classificam distâncias entre pontos do conjunto estudado para o cálculo de $f(\alpha)$. Como alguns dos métodos encontrados na literatura baseados em contagem de distâncias eram métodos indiretos (ou seja, calculam primeiro D_q para depois calcular $f(\alpha)$, o que possui desvantagens), desenvolvemos um novo método [39] que une as vantagens de um método onde não é necessário o uso de um recobrimento para o cálculo das probabilidades com as de um método onde $f(\alpha)$ é calculado diretamente. Chamamos esse método de Método canônico da integral de correlação, e o apresentamos no capítulo 6.

- Criar um algoritmo de contagem de caixa que construísse recobrimentos de melhor qualidade, eliminando ao máximo as caixas espúrias [40]. Os resultados foram bastante satisfatórios e esse algoritmo é apresentado no capítulo 7.

Dessa forma, esta tese se divide em duas partes: Na primeira, que se estende dos capítulos 1 a 5, apresentaremos alguns resultados para métodos encontrados na literatura, que utilizam ou não algoritmos de contagem de caixa, sendo portanto, um trabalho de sistematização e análise comparativa dos mesmos. Na segunda parte, propomos dois novos algoritmos e verificamos quando e em quais condições há alguma melhora no cálculo em relação aos métodos estudados anteriormente. Estes novos métodos são analisados nos capítulos 6 e 7.

Esta tese possui a seguinte estrutura:

- No capítulo 1 descrevemos os principais conceitos da teoria multifractal, introduzindo as curvas dimensões generalizadas D_q e o espectro de singularidades $f(\alpha)$. Também descrevemos suscintamente todos os métodos para o cálculo numérico de $f(\alpha)$ encontrados na literatura que foram analisados neste trabalho.

- No capítulo 2 introduzimos os conjuntos multifractais que escolhemos para testar os métodos citados no capítulo 1, a razão para esta escolha, e a maneira pela qual eles foram simulados numericamente.

- No capítulo 3 o problema das caixas espúrias, desenvolvido por nós no mestrado, é apresentado em detalhes, assim como seus efeitos.

- No capítulo 4 apresentamos os resultados numéricos obtidos a partir dos métodos do cálculo de $f(\alpha)$ baseados em algoritmos de contagem de caixa na sua forma original, como encontrados no literatura, onde pode-se ver o efeito da distorção produzida pelas caixas espúrias.

- No capítulo 5 apresentamos resultados para métodos já existentes na literatura que se utilizam da classificação de distâncias entre pontos do conjunto para o cálculo de $f(\alpha)$. Mostramos em quais casos eles podem ser usados com vantagens quando comparados aos métodos que usam o algoritmo de contagem de caixa. Ressaltamos que nem sempre os métodos de contagem de distâncias apresentam bons resultados.

- No capítulo 6, apresentamos um novo método de contagem de distâncias bem como os resultados obtidos através da sua aplicação. Este método foi inteiramente desenvolvido por nós.

- No capítulo 7, apresentamos um novo algoritmo de contagem de caixa, que constrói recobrimentos melhores que uma grade. Mostramos que os resultados finais obtidos são significantemente melhores do que os calculados através do algoritmo usual, pois este novo algoritmo de fato elimina a maior parte das caixas espúrias.

- Finalmente, nas conclusões, resumimos os nossos resultados.

- Os apêndices trazem detalhes do cálculo numérico para cada um dos métodos estudados nesta tese. A notação utilizada na descrição dos algoritmos nesses apêndices é familiar aos programadores em C.

Capítulo 1

Conceitos e Métodos Numéricos

Neste capítulo vamos descrever as principais definições da chamada caracterização multifractal. Uma delas, o espectro de singularidades, foi o principal objeto de nosso estudo. Assim, após a sua introdução, apresentaremos os métodos numéricos existentes na literatura que têm por objetivo calculá-lo. Classificamos esses métodos em diretos ou indiretos, dependendo da necessidade de um cálculo anterior do expoente de massa, ou ainda métodos de contagem de pontos ou classificação de distâncias, dependendo de como os pontos do conjunto são analisados e de como se obtém uma aproximação para a distribuição de probabilidades de pontos p_i sobre o atrator.

1.1 Dimensão fractal

Dimensão fractal é um nome genérico dado a diferentes definições de dimensão não inteira existentes na literatura. Em um objeto fractal existem estruturas em todas as escalas de ampliação do mesmo. Pode-se interpretar a dimensão fractal como uma quantidade que expressa o aparecimento de estruturas em um conjunto ao ampliá-lo. Essas definições são feitas de tal forma que a dimensão usual é um caso particular das mesmas. A dimensão fractal é o primeiro valor a ser considerado na caracterização fractal.

1.1.1 Dimensão de Hausdorff-Besicovitch

A dimensão de Hausdorff-Besicovitch D_H é a definição formal da dimensão fractal. Antes de apresentarmo-la, porém, é necessário introduzir a definição de um l-recobrimento de um conjunto A. Um l-recobrimento é um conjunto de objetos U_i,

de mesmo formato geométrico e de dimensão D real, com as seguintes propriedades:

- Os tamanhos l_i dos objetos U_i possuem um valor máximo dado por l.
- O conjunto considerado A está contido na união dos objetos U_i.

Define-se a dimensão de Hausdorff D_H do conjunto A [12], considerando-se seus l-recobrimentos:

1. Define-se a quantidade:

$$\Gamma_H(A, D, l, \{l_i\}) = \sum_i l_i^D \tag{1.1}$$

2. escolhe-se, entre todos l-recobrimentos, aquele que reduz essa soma:

$$\Gamma_H(A, D, l) = \inf_{\{l_i\}} \sum_i l_i^D \tag{1.2}$$

3. e finalmente, toma-se o limite de l tendendo a zero.

$$\Gamma_H(A, D) = \lim_{l \to 0} \Gamma_H(A, D, l) \tag{1.3}$$

Pode-se mostrar que a função $\Gamma_H(A, D)$ possui a forma:

$$\Gamma_H(A, D) = \begin{cases} 0 & \text{para} \quad D > D_H \\ \infty & \text{para} \quad D < D_H \end{cases} \tag{1.4}$$

o valor que separa as duas regiões é $D = D_H$, que define a dimensão de Hausdorff.

A função $\Gamma_H(D)$ é a D-medida de Hausdorff. Pode-se interpretá-la como o 'tamanho' (comprimento, área, volume, etc.) de um objeto medido por pequenos objetos de dimensão D. Os casos de dimensão inteira, iguais à definição usual, são mais simples de se visualizar. Por exemplo, no cálculo de Γ_H em uma superfície finita, quando:

- $D = 1$, estamos medindo o 'comprimento' desta superfície com pequenos segmentos de reta, obtendo $\Gamma_H = \infty$, porque uma superfície possui um 'comprimento' infinito.
- $D = 3$, estamos medindo o 'volume' da superfície com pequenos cubos, obtendo $\Gamma_H = 0$, porque esse 'volume' é nulo.
- $D = 2$, o tamanho da superfície está sendo medido usando pequenos quadrados, e temos $\Gamma_H =$ à área da superfície, o único valor de D onde Γ_H é finito.

1.1.2 Dimensão de caixa

O cálculo analítico da dimensão de Hausdorff é possível em apenas poucos casos, de modo que, em geral, o cálculo numérico se faz necessário. Neste, utiliza-se outra definição, chamada *dimensão de caixa*, que na maioria dos casos estudados possui o mesmo valor da Dimensão de Hausdorff [11].

A dimensão de caixa D_B pressupõe um recobrimento do espaço que contém o fractal com uma grade (conjunto de cubos d-dimensionais justapostos de igual formato e tamanho ϵ). Conta-se o número $N(\epsilon)$ de caixas que possuem pelo menos um ponto do fractal. Define-se então:

$$D_B = -\lim_{\epsilon \to 0} \frac{\ln N(\epsilon)}{\ln \epsilon} \tag{1.5}$$

Em uma estimativa numérica não se consegue calcular o limite indicado na eq. 1.5. Então, obtém-se D_B como a inclinação da reta $\ln N(\epsilon) \times \ln \epsilon$, quando ϵ é pequeno. O valor de $N(\epsilon)$ é obtido através de algoritmos conhecidos como *contagem de caixa*, cujo princípio é descrito na sec. 1.4.1.

1.2 Outras definições de dimensão

Usando o conceito de probabilidade, outras definições de dimensão foram criadas. A probabilidade p_E de uma região E de um objeto é a medida do objeto na região E.

$$p_E = \int_E d\mu$$

Quando se realiza um cálculo numérico, a probabilidade p_i é aproximada pela fração de pontos na i-ésima caixa do recobrimento, $p_i = N_i/N$, onde N_i é o número de pontos na i-ésima caixa do recobrimento e N é o número total de pontos do objeto considerado [19]. Tanto a contagem do número de caixas quanto o levantamento das probabilidades são realizados por algoritmos de contagem de caixa.

Note que a introdução do conceito de probabilidade introduz uma modificação na definição do l-recobrimento de um conjunto: É necessário que a intersecção de todos os pares de objetos U_i, U_j com $i \neq j$ do recobrimento não contenha pontos do conjunto, para que a soma de todas as probabilidades do recobrimento seja sempre unitária.

De modo análogo ao utilizado na definição de dimensão de caixa, define-se:

Dimensão de informação σ ([11]):

$$\sigma = \lim_{\epsilon \to 0} \frac{\sum_i p_i \ln p_i}{\ln \epsilon} \tag{1.6}$$

O termo $I = \sum_i p_i \ln p_i$ é a Informação de Shannon. Pode-se interpretar σ como uma quantidade relacionada à taxa do aumento de informação em sucessivas ampliações do objeto considerado.

Dimensão de correlação ν ([14]):

$$\nu = \lim_{\epsilon \to 0} \frac{\ln \sum_i p_i^2}{\ln \epsilon} \tag{1.7}$$

A dimensão de correlação possui este nome devido ao termo p_i^2, que é a probabilidade de dois pontos situarem na mesma caixa de tamanho ϵ.

Numericamente, o limite das expressões acima é sempre aproximado pelo coeficiente angular da reta formada pela curva do numerador do limite contra o denominador $\ln \epsilon$, para valores suficientemente pequenos de ϵ.

Pode-se mostrar que existe uma relação de desigualdade entre as dimensões definidas [11]:

$$\nu \leq \sigma \leq D_H$$

Mais detalhes sobre estas e outras definições de dimensão podem ser obtidas na referência [11].

1.3 Caracterização multifractal completa

Notou-se que apenas alguns números não eram suficientes para uma caracterização completa de um multifractal. Foram definidas duas curvas para a caracterização topológica completa de um multifractal: dimensões generalizadas D_q e espectro de singularidades $f(\alpha)$.

1.3.1 Dimensões generalizadas

Baseando-se no conceito de informação de Renyi, definiu-se as dimensões generalizadas D_q [22]:

$$D_q = \frac{1}{q-1} \lim_{\epsilon \to 0} \frac{\ln \sum_i p_i^q}{\ln \epsilon} \tag{1.8}$$

onde q é real. Note que no caso onde $q = 0$ ou $q = 2$ esta definição reproduz as definições anteriores de dimensão de caixa e dimensão de correlação. No limite de $q \to 1$, seja $\delta q = q - 1$. Para δq tendendo a zero, temos:

$$p_i^q = p_i^{1+\delta q} \approx p_i(1 + \delta q \ln p_i - \ldots)$$

e, como $\sum_i p_i^q = 1$, temos, no mesmo limite :

$$\begin{aligned} \ln \sum_i p_i^q &= \ln(\sum_i p_i^q - 1 + 1) \approx \sum_i p_i^q - 1 \\ &\approx \sum_i p_i + \delta q \sum_i p_i \ln p_i - 1 = \delta q \sum_i p_i \ln p_i \end{aligned}$$

de modo que :

$$D_1 = \frac{1}{\delta q} \delta q \sum_i \frac{p_i \ln p_i}{\ln \epsilon} = \sum_i \frac{p_i \ln p_i}{\ln \epsilon} \tag{1.9}$$

isto é, D_1 reproduz a definição de dimensão de informação.

Define-se também o espectro de massa $\tau(q)$, dado por:

$$\tau(q) = (q-1)D_q \tag{1.10}$$

A curva D_q e q são difíceis de se interpretar, mas pode-se descrever algumas de suas propriedades. Quando q é positivo, as probabilidades maiores do conjunto têm os seus valores aumentados na somatória, e portanto estas determinam D_q com maior influência. Ocorre o mesmo efeito com as probabilidades menores quando q é negativo. Como casos particulares, quando $q \to \infty$, apenas o subconjunto de maior probabilidade é levado em consideração, e quando $q \to -\infty$, apenas o subconjunto de menor probabilidade é considerado. Esses valores são atingidos assintoticamente pela curva D_q.

Pode-se mostrar que a curva D_q é contínua e monotonicamente decrescente [22]. Uma forma típica da curva D_q é apresentada na fig. 1.1.

Análoga à definição de D_H, existe uma definição mais rigorosa de D_q [12]:

1. Definimos uma função

$$\Gamma(A, q, \tau', l, \{l_i\}) = \sum_i \frac{p_i^q}{l_i^{\tau'}} \tag{1.11}$$

onde as probabilidades p_i estão definidas no l-recobrimento do conjunto A, e q e τ' são números reais.

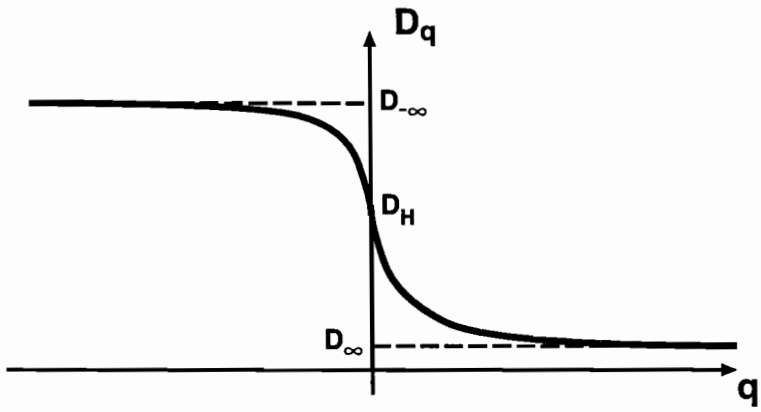


Figura 1.1: Curva D_q típica.

2. A seguir, fazemos:

$$\Gamma(A,q,\tau',l) = \begin{cases} \sup_{\{l_i\}}\{\Gamma(A,q,\tau',l,\{l_i\})\} & \text{para} \quad q > 1 \\ \inf_{\{l_i\}}\{\Gamma(A,q,\tau',l,\{l_i\})\} & \text{para} \quad q < 1 \end{cases} \tag{1.12}$$

3. e depois considera-se o limite $l \to 0$,

$$\Gamma(A,q,\tau') = \lim_{l\to 0} \Gamma(A,q,\tau',l) \tag{1.13}$$

esta curva também possui valor zero ou infinito, exceto para um valor $\tau' = \tau$, onde se faz

$$\Gamma(A,q,\tau) = 1 \tag{1.14}$$

e essa condição forma uma curva $\tau(q) = (q-1)D_q$, determinando D_q para o conjunto.

1.3.2 Espectro de singularidades

A curva D_q não é a única que pode ser definida para caracterizar um multifractal. Lembrando que sempre se procura parâmetros independentes da escala quando se trabalha com fractais, define-se o coeficiente de Lipschitz-Hölder α ([12]) :

$$p_i \sim \epsilon^{\alpha}, \text{ para } \epsilon \to 0 \tag{1.15}$$

Interpreta-se α como uma grandeza relacionada à singularidade da medida (probabilidade) próximo a um ponto do conjunto. É portanto uma propriedade local. Também define-se a função $\rho(\alpha)$, a distribuição de pontos com expoente de Lipschitz-Hölder entre α e $\alpha + d\alpha$.

Pode-se dizer que no limite $\epsilon \to 0$, o número de caixas com α entre α' e $\alpha' + d\alpha'$, é:

$$d\alpha' \rho(\alpha') \epsilon^{-f(\alpha')} \tag{1.16}$$

onde $f(\alpha)$, é uma função contínua em α e é a dimensão fractal do conjunto de pontos com α entre α' e $\alpha' + d\alpha'$. Dessa maneira define-se o espectro de singularidades $f(\alpha)$.

A curva $f(\alpha)$ pode ser interpretada da seguinte forma: Um multifractal típico é visto como um conjunto de infinitos fractais puros interligados, cada um formado pelo conjunto de pontos com mesmo α, cuja dimensão fractal é $f(\alpha)$.

As propriedades da curva $f(\alpha)$ podem ser determinadas estudando-se a sua relação com as dimensões generalizadas.

1.3.3 Relação entre as dimensões generalizadas e o espectro de singularidades

As curvas $f(\alpha)$ e D_q realizam o mesmo tipo de caracterização, pode-se esperar que exista uma relação matemática entre elas.

Reescrevendo a eq. 1.8, através da definição:

$$\chi(q) = \sum_i p_i^q \tag{1.17}$$

temos a expressão:

$$D_q = \frac{1}{q-1} \lim_{\epsilon \to 0} \frac{\ln \chi(q)}{\ln \epsilon} \tag{1.18}$$

Substituindo as expressões que definem $f(\alpha)$ em $\chi(q)$, obtém-se

$$\chi(q) = \int d\alpha' \rho(\alpha') \epsilon^{-f(\alpha')} \, \epsilon^{q\alpha'} \tag{1.19}$$

Como será considerado o limite de $\epsilon \to 0$, o mínimo de $q\alpha' - f(\alpha')$ será o termo significativo na integral. Seja este mínimo o ponto $\alpha' = \alpha(q)$:

$$\left. \frac{\partial}{\partial \alpha} [q\alpha' - f(\alpha')] \right|_{\alpha' = \alpha(q)} = 0 \tag{1.20}$$

$$\left. \frac{\partial^2}{\partial \alpha^2} [q\alpha' - f(\alpha')] \right|_{\alpha' = \alpha(q)} > 0 \tag{1.21}$$

$$\chi(q) = \epsilon^{q\alpha(q) - f(\alpha(q))} \tag{1.22}$$

Das primeiras condições 1.20, 1.21 obtém-se:

$$\left. \frac{\partial f(\alpha')}{\partial \alpha'} \right|_{\alpha' = \alpha(q)} = q \qquad \left. \frac{\partial^2 f(\alpha')}{\partial \alpha'^2} \right|_{\alpha' = \alpha(q)} > 0$$

e substituindo o valor de $\chi(q)$ da eq. 1.22 na eq. 1.18, obtemos finalmente

$$D_q = \frac{1}{q-1}[q\alpha(q) - f(\alpha(q))]$$

que é a relação entre $f(\alpha)$ e D_q.

Resumindo, as funções $f(\alpha)$ e D_q relacionam-se entre si por uma transformação de Legendre:

$$D_q = \frac{1}{q-1}[q\alpha(q) - f(\alpha(q))] \tag{1.23}$$

$$f(\alpha) = \alpha q(\alpha) - (q(\alpha) - 1)D_q(\alpha) \tag{1.24}$$

derivando estas expressões, obtém-se os valores de $q(\alpha)$ e $\alpha(q)$

$$q(\alpha) = \frac{\partial f(\alpha)}{\partial \alpha} \tag{1.25}$$

$$\alpha(q) = \frac{\partial}{\partial q}[(q-1)D_q] \tag{1.26}$$

Descrevemos a seguir mais algumas propriedades conhecidas de $f(\alpha)$. Das equações 1.20 e 1.21, conclui-se que a curva possui concavidade bem definida. Da eq. 1.25 conclui-se que o máximo da curva é em $f(\alpha) = D_B$. Dessa mesma equação, vemos que os extremos da curva $f(\alpha)$ são $\alpha_{min} = D_\infty$ e $\alpha_{max} = D_{-\infty}$. Assim, o ramo correspondente a $q > 0$ é o lado esquerdo da curva, e o ramo correspondente a $q < 0$ é o lado direito da curva $f(\alpha)$. A curva $f(\alpha)$ possui uma localização especial

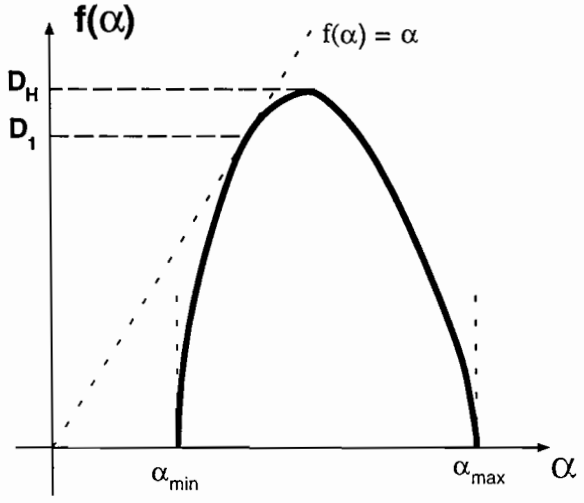


Figura 1.2: Curva $f(\alpha)$ típica.

no plano, que pode ser determinada calculando $[f(\alpha(q)) - \alpha(q)]$, através do seguinte

procedimento: pode-se somar e subtrair α na eq. 1.24, obtendo

$$[f(\alpha(q)) - \alpha(q)] = (q-1)(\alpha(q) - D_q)$$

Da eq. 1.26, resolvendo a derivada, tem-se

$$\alpha(q) = D_q + (q-1)\frac{d}{dq}D_q$$

Substituindo $\alpha(q)$ na expressão anterior, obtém-se finalmente

$$[f(\alpha(q)) - \alpha(q)] = (q-1)^2\frac{d}{dq}D_q$$

Uma vez que a curva D_q é contínua e monotonicamente decrescente, $\frac{d}{dq}D_q \leq 0$, de modo que a curva $f(\alpha)$ se localiza sempre abaixo da bissetriz do primeiro quadrante do plano $\alpha - f(\alpha)$, exceto quando $q = 1$, ponto que pertence obrigatoriamente à mesma bissetriz, ou seja, $f(\alpha(q=1)) = \alpha(q=1) = D_1$.

Da eq. 1.25, os extremos da curva $f(\alpha)$, $\alpha_{min} = D_\infty$ e $\alpha_{max} = D_{-\infty}$ são pontos onde a derivada da curva diverge.

Embora a função $f(\alpha)$ possua uma interpretação mais simples, ela é mais difícil de ser calculada, quando comparada com D_q. A fig. 1.2 mostra uma curva típica de $f(\alpha)$.

1.4 Classes de métodos para o cálculo numérico de $f(\alpha)$

Como a definição das dimensões generalizadas pode ser utilizada diretamente para o seu cálculo numérico, os primeiros métodos numéricos de caracterização multifractal que surgiram calculam essa curva. O espectro de singularidades era calculado em uma segunda etapa, através da transformação de Legendre do expoente de massa $\tau(q)$. Uma vez que o espectro de singularidades possui uma interpretação mais simples, procurou-se criar métodos de cálculo numérico que capazes de calcular diretamente o espectro de singularidades. Dessa forma, os métodos de cálculo do espectro de singularidades se dividem em métodos diretos ou indiretos.

Numericamente, o conjunto estudado geralmente é representado por um conjunto discreto e finito de pontos, e nos diferentes métodos de cálculo numérico, diretos ou indiretos, é necessário obter em algum momento os valores das probabilidades p_i desse

conjunto. Observamos que os métodos encontrados na literatura se dividem em dois grupos: aqueles que se baseiam na contagem dos pontos do conjunto e aqueles que realizam a classificação de distâncias entre pontos do conjunto, para esse mesmo fim. Vamos descrever inicialmente estas duas maneiras de se obter as probabilidades p_i de um conjunto multifractal, e em seguida, os métodos numéricos diretos e indiretos para o cálculo de $f(\alpha)$.

1.4.1 Contagem de caixa

Algoritmos de contagem de caixa (*box counting*) é o nome genérico dado a um conjunto de algoritmos cujo objetivo é realizar numericamente o recobrimento de um conjunto dado. Eles são a tentativa numérica de calcular as probabilidades p_i que mais se aproxima da maneira descrita pela teoria para o cálculo de grandezas da caracterização multifractal.

Como a princípio é muito difícil obter numericamente um recobrimento de caixas de tamanhos e posições diferentes, é feita, nesses algoritmos, uma simplificação: ao invés de considerar um l-recobrimento do conjunto, usa-se uma grade, ou seja, um conjunto de hiper-cubos de mesmo tamanho ϵ justapostos lado a lado, como em um tabuleiro de xadrez, para recobrir o conjunto. A probabilidade p_i da i-ésima caixa da grade é dada pela fração de pontos em seu interior, o que classificaria esses algoritmos como contagem de pontos. O levantamento das probabilidades é realizado em várias grades de tamanho de caixa ϵ diferentes.

Em todos os métodos baseados em contagem de pontos, as grandezas calculadas são sempre obtidas no limite de ϵ tendendo a zero de uma fração cujo denominador é $\ln \epsilon$. Como numericamente não se pode realizar o limite de ϵ tendendo a zero, considera-se como aproximação desse limite o coeficiente angular da reta formada pelo logaritmo do numerador contra o seu respectivo valor de $\ln \epsilon$.

O algoritmo de contagem de caixa que utilizamos está descrito em detalhe no Apendice B.

1.4.2 Contagem de distâncias

Da teoria multifractal, temos que $\nu \leq D_B$, e que a curva D_q é monotônica decrescente. É um fato conhecido que a dimensão de correlação pode ser utilizada como uma aproximação para D_B. A expressão de ν na eq. 1.7 possui um termo p_i^2,

que é a probabilidade de que dois pontos estejam na caixa i, cujo tamanho é ϵ. Ela pode ser aproximada pela probabilidade de dois pontos se situarem a uma distância menor que ϵ. Dessa forma, o numerador da equação acima passa a ser a fração de distâncias entre os pontos que são menores que ϵ, e o método passa a ser considerado como classificação de distâncias.

O algoritmo descrito é conhecido como Algoritmo de Grassberger-Procaccia [14], e é muito utilizado no cálculo da dimensão de correlação e consequentemente de uma aproximação para a dimensão fractal. A vantagem é que ele elimina um problema dos métodos de contagem de caixa, que é a posição relativa da grade ao conjunto. Esse parâmetro altera o número de pontos nas caixas da grade e portanto pode afetar o resultado final. Esta vantagem motivou o surgimento de métodos para o cálculo de $f(\alpha)$ baseados em contagem de distâncias.

Devido à sua simplicidade, não existe um nome especial para o algoritmo de classificação de distâncias. Ele consiste em calcular as distâncias entre os pontos do conjunto e classificá-las, e a melhor maneira de fazê-lo depende do método para o cálculo de $f(\alpha)$ utilizado.

Para maior eficiência numérica do mesmo, costuma-se utilizá-lo associado a um algoritmo auxiliado por caixas, descrito em maior detalhe no Apêndice F.

1.5 Métodos numéricos para o cálculo de $f(\alpha)$

Descrevemos a seguir os métodos numéricos para o cálculo de $f(\alpha)$ que estudamos. Iniciaremos pelos métodos de contagem de caixa, e em seguida vamos introduzir os métodos baseados em contagem de distâncias.

Os métodos de contagem de caixa mais citados na literatura são:

- Método microcanônico: Consiste na aplicação direta da definição de $\tau(q)$ e de sua transformação de Legendre. Como calcula $\tau(q)$ para obter $f(\alpha)$, é um método indireto.

- Método canônico: Também parte da definição de $\tau(q)$, mas é um método direto. A transformação de Legendre fica embutida nas equações do método, como veremos adiante.

- Método dos histogramas: Utiliza a própria definição de $f(\alpha)$, obtendo a curva parametrizada nas colunas de histogramas calculados pelo método, de onde vem

o seu nome. É um método direto.

Os métodos de contagem de distâncias são:

- Método da integral de correlação: Ao invés de usar a probabilidade de pontos serem encontrados em uma caixa da grade, tenta-se calcular a probabilidade de pontos serem encontrados em uma caixa centrada em um ponto do conjunto escolhido aleatoriamente. Essa substituição é feita na definição de $\tau(q)$, portanto este é um método indireto.

- Método da massa fixa: Neste método parte-se de um ponto de vista oposto ao da contagem de caixa: Coleciona-se o tamanho de caixas centradas em pontos do conjunto que possuem uma mesma probabilidade fixada a priori. A curva $f(\alpha)$ é obtida a partir dos histogramas dos tamanhos de caixa encontrados. É um método direto.

Passamos a seguir, a discutir em maior detalhe cada um destes métodos:

1.5.1 Método microcanônico

Este método [1] consiste na aplicação direta da definição de $\tau(q)$ e posteriormente na realização de uma aplicação da transformação de Legendre. Utiliza-se, na sequência, as eqs. 1.10, 1.8, 1.26 e 1.24:

$$\tau(q) = \lim_{\epsilon \to 0} \frac{\ln \sum_i p_i^q(\epsilon)}{\ln \epsilon} \tag{1.27}$$

$$\alpha(q) = \frac{d}{dq}\tau(q) \tag{1.28}$$

$$f(\alpha(q)) = q\alpha(q) - \tau(q) \tag{1.29}$$

O coeficiente $\tau(q)$ é numericamente aproximado pelo coeficiente angular da reta formada pelo numerador da eq. 1.27 contra $\ln \epsilon$. A derivada na eq. 1.28 é obtida através de diferenças discretas, como descrito no apêndice K.

1.5.2 Método canônico

A partir da definição de $\tau(q)$, pode-se obter as seguintes expressões [10] :

$$\mu_i(\epsilon, q) = \frac{p_i^q(\epsilon)}{\sum_j p_j^q(\epsilon)} \tag{1.30}$$

$$\alpha(q) = \lim_{\epsilon \to 0} \frac{\sum_i \mu_i \ln p_i}{\ln \epsilon} \tag{1.31}$$

$$f(\alpha(q)) = \lim_{\epsilon \to 0} \frac{\sum_i \mu_i \ln \mu_i}{\ln \epsilon} \tag{1.32}$$

Cria-se um novo conjunto de probabilidades μ_i para ser utilizado nas expressões que fornecerão α e $f(\alpha)$. Substituindo μ_i nessas equações, verifica-se facilmente que elas originam-se da transformação de Legendre de $\tau(q)$.

O nome dado a estes dois métodos tem origem na sua semelhança formal com a Termodinâmica e Mecânica Estatística. No Método microcanônico, $f(\alpha)$ é calculado como se fosse a energia no ensamble microcanônico, e no Método canônico, a expressão de $f(\alpha)$ lembra o cálculo da média da energia no ensamble canônico.

1.5.3 Método dos histogramas

Neste método [28] utiliza-se diretamente as expressões 1.15 e 1.16 que definem $f(\alpha)$:

$$\alpha = \lim_{\epsilon \to 0} \frac{\ln p_i(\epsilon)}{\ln \epsilon} \tag{1.33}$$

$$N(\epsilon) \approx \rho(\alpha') d\alpha' \epsilon^{-f(\alpha')} \text{quando} \epsilon \to 0 \tag{1.34}$$

Suponha que tenhamos $p_i(\epsilon)$, para vários tamanhos de caixa. Seja

$$X_i(\epsilon) = \ln p_i(\epsilon) \tag{1.35}$$

Constrói-se primeiro um histograma da distribuição de X_i para cada valor de ϵ considerado. Seja $[Xmin, Xmax]$ o intervalo de valores de X_i encontrado. Divide-se esse intervalo em Nd partes iguais de tamanho dX e define-se H_j como o número de X_i encontrados na j-ésima parte do intervalo $[Xmin, Xmax]$.

Calcula-se α para cada divisão do histograma, utilizando a eq. 1.33 diretamente, nos valores utilizados para dividir o intervalo $[Xmin, Xmax]$. Como as divisões feitas nesse intervalo são iguais, a variação de α de uma divisão para outra é constante.

Ao calcular-se $f(\alpha)$, verifica-se que existe uma correção logarítmica em ϵ, que pode ser determinada utilizando a eq. 1.19 quando $q = 1$.

$$1 = \lim_{\epsilon \to 0} \int \rho(\alpha') d\alpha' \epsilon^{\alpha' - f(\alpha')} \tag{1.36}$$

A equação acima é solucionada por ponto de sela, onde o termo dominante é aquele que possui o menor expoente. O próximo termo de ordem dominante é $|\ln \epsilon|^{-1/2}$ [37],

que neste caso, ainda é relevante. Tem-se então:

$$1 = C(\alpha)\epsilon^{\alpha - f(\alpha)}|\ln \epsilon|^{-1/2} \tag{1.37}$$

onde $C(\alpha)$ é uma função resultante da integração, e que depende basicamente de $\rho(\alpha)$. Das propriedades da curva $f(\alpha)$, sabemos que para $q = 1$, $\alpha = f(\alpha)$, e então podemos determinar a função $C(\alpha)$, em primeira ordem de aproximação:

$$C(\alpha) = |\ln \epsilon|^{1/2} \tag{1.38}$$

A análise dessa correção logarítmica é apresentada em outra abordagem no apêndice da referência [41].

Este cálculo sugere que a eq. 1.34 ao ser invertida, contenha essa correção logarítmica:

$$f(\alpha) = -\lim_{\epsilon \to 0} \frac{N/\sqrt{|\ln \epsilon|}}{\ln \epsilon} \tag{1.39}$$

onde $d\alpha$ já foi eliminado por ser constante nesse método.

Para substituir N, utilizamos a seguir uma igualdade entre as distribuições de número de caixas:

$$N d\alpha = -H_j dX \tag{1.40}$$

e elimina-se ϵ do numerador da expressão derivando-se a eq. 1.35 em α, depois de substituir p_i pela eq. 1.33:

$$dX = d\alpha \ln \epsilon \tag{1.41}$$

obtendo-se finalmente:

$$f(\alpha) = -\lim_{\epsilon \to 0} \frac{\ln(H_j/\sqrt{|dX|})}{\ln \epsilon}. \tag{1.42}$$

1.5.4 Método da integral de correlação

Neste método [33] o ponto de partida é a equação eq. 1.10, que define $\tau(q)$:

$$\tau(q) = \lim_{\epsilon \to 0} \frac{\ln \sum_i p_i^q}{\ln \epsilon} \tag{1.43}$$

Pode-se interpretar o termo $\sum_i p_i^q$ como uma média no espaço da grandeza p_i^{q-1}, pois p_i aproxima a medida invariante da dinâmica. Supondo que o sistema seja ergódico, a média espacial deve ser igual a média temporal, e a equação passa a ser:

$$\tau(q) = \lim_{\epsilon \to 0} \frac{\ln \left\langle P_j^{q-1} \right\rangle}{\ln \epsilon} \tag{1.44}$$

onde P_j é a fração de pontos do conjunto se situam a uma distância menor que ϵ centrada do j-ésimo ponto do conjunto. A média temporal é realizada centralizando a caixa de tamanho ϵ em todos os pontos da trajetória, ou seja, do conjunto.

Pode-se entender a mudança do expoente de q para $q-1$ observando-se que na definição original de $\tau(q)$ usa-se caixas em uma grade, ou seja, elas são elas estão distribuídas uniformemente no espaço, enquanto que na nova expressão as caixas são colocadas no espaço com a mesma distribuição que os pontos, que é aproximada por p_i.

As probabilidades P_j são calculadas contando-se quantas distâncias entre os pontos do conjunto e o j-ésimo ponto são menores que ϵ, de modo que, com esta definição de $\tau(q)$, este é um método de contagem de distâncias. A curva $f(\alpha)$ é obtida calculando-se a transformação de Legendre de $\tau(q)$. Este é, portanto, um método indireto.

1.5.5 Método da massa fixa

Neste método [4], ao invés de se considerar caixas de tamanho fixo e colecionar as diferentes probabilidades observadas, coleciona-se tamanhos de caixas centralizadas em pontos do conjunto que possuem um valor fixo de probabilidade.

Partindo-se da definição teórica de $\tau(q)$ dada pelas eqs. 1.13 e 1.14, e relaxando a otimização do recobrimento de caixas, temos:

$$\lim_{l\to 0}\sum_i \frac{p_i^q}{l_i^{\tau(q)}} = 1 \tag{1.45}$$

onde $l \geq l_i$.

Supondo-se que as probabilidades são fixas e possuem valor $p = p_i$, temos

$$\sum_i l_i^{\tau(q)} \approx p^{-q}, \quad \text{quando } p \to 0 \tag{1.46}$$

Por outro lado, vamos definir $\hat{P}(\alpha';p)$ como a probabilidade de se encontrar aleatoriamente uma esfera com probabilidade p e coeficiente de Lipschitz-Hölder α entre α' e $\alpha' + d\alpha'$. Aproximadamente, ela deve ser $\rho(\alpha')p$, pois as duas variáveis são independentes. Utilizando-se a definição de α dada na eq. 1.15, temos:

$$\sum_i p\, l_i^{-\tau(q)} \approx \int \hat{P}(\alpha;p)\, p^{-\tau(q)/\alpha} \tag{1.47}$$

de modo que, unindo as equações, temos:

$$\int \hat{P}(\alpha;p) p^{-\tau(q)/\alpha} \approx p^{1-q} \tag{1.48}$$

Da teoria multifractal, temos a relação 1.23:

$$\tau(q) = \min_{\alpha}\{q\alpha - f(\alpha)\} \tag{1.49}$$

Dessa maneira, se utilizamos o *ansatz*:

$$\hat{P}(\alpha; p) \approx |\ln p|^{1/2} p^{1-\frac{f(\alpha)}{\alpha}} \tag{1.50}$$

a igualdade da eq. 1.48 é obedecida, resolvendo a integral por aproximação de ponto de sela, quando $p \to 0$.

Suponha que o conjunto multifractal é dado por n pontos. Vamos definir como $P(\delta; p)$ a probabilidade de encontrarmos aleatoriamente uma esfera de tamanho δ com probabilidade p.

Podemos aproximar:

$$P(\delta; p)\, d\delta \sim \hat{P}(\alpha; p)\, d\alpha \tag{1.51}$$

Utilizando a eq. 1.15, podemos calcular a derivada:

$$\frac{d\alpha}{d\delta} = -\frac{\alpha^2}{p^{1/\alpha} \ln p} \tag{1.52}$$

e, desprezando contribuições logarítmicas, obter o resultado:

$$f(\alpha) \approx \alpha \left(\frac{\ln P(\delta; p)}{\ln p} - 1 \right) + 1 \tag{1.53}$$

Exatamente como é feito no método dos histogramas, colecionamos os valores $Y_i = \ln l_i$ e dividimos o intervalo que contém estes valores $[Ymin, Ymax]$ em Nd partes iguais. O número normalizado H_j de Y_i que pertencem a cada divisão é $P(\delta; p)$. Substituindo os valores encontrados em cada divisão do histograma nas eqs. 1.15, 1.53 obtemos a curva $f(\alpha)$.

Neste método não se optou por realizar regressões lineares em $\ln p$, de modo que as equações fornecem a curva $f(\alpha)$ diretamente. O limite de $p \to 0$ é resolvido da seguinte forma: admite-se que o valor das equações é dado a menos de uma constante. Lembrando que a curva $f(\alpha)$ sempre tangencia a bissetriz do primeiro quadrante em $\alpha = f(\alpha) = D_1$, calcula-se o valor de D_1 através de outro método e desloca-se a curva $f(\alpha)$ em ambos os eixos de modo que ela obedeça a essa condição.

Numericamente, supomos que $p = k/n$, e o tamanho l_i da esfera é considerado a k-ésima distância mais próxima do ponto central da mesma, de modo que este é um método de contagem de distâncias.

Capítulo 2

Conjuntos multifractais utilizados

Neste capítulo introduzimos os elementos utilizados para a comparação dos métodos numéricos para o cálculo de $f(\alpha)$ descritos no capítulo anterior: os conjuntos multifractais nos quais os métodos foram aplicados. Utilizamos neste estudo dois tipos de conjuntos unidimensionais, os Conjuntos de Cantor Generalizados e o atrator do mapa quadrático. Esses conjuntos foram escolhidos porque possuem solução teórica para o espectro de singularidades. Também utilizamos um conjunto bidimensional, o atrator do mapa de Hénon, por ser um sistema muito utilizado na literatura. A seguir, descrevemos detalhadamente esses conjuntos e os algoritmos usados para representá-los numericamente.

2.1 Conjuntos de Cantor generalizados

Define-se como conjunto de Cantor generalizado o conjunto de pontos formado pelo seguinte processo, repetido infinitas vezes (Veja fig. 2.1):

Toma-se o intervalo unitário (ou seja, com comprimento e medida unitários) e o dividimos em três segmentos:

1. O comprimento do primeiro segmento é l_1, com medida ou probabilidade p_1.

2. Subtrai-se o segundo segmento.

3. Associa-se ao terceiro, de comprimento l_2, uma medida $p_2 = 1 - p_1$.

Essa primeira construção é denominada *gerador* do conjunto. As gerações seguintes são construídas aplicando-se, a cada segmento, a regra acima e atribuindo-se a cada

novo segmento um comprimento e medida dados pelo produto do comprimento e medida, respectivamente, do segmento inicial e do segmento do gerador correspondente (fig.2.1).

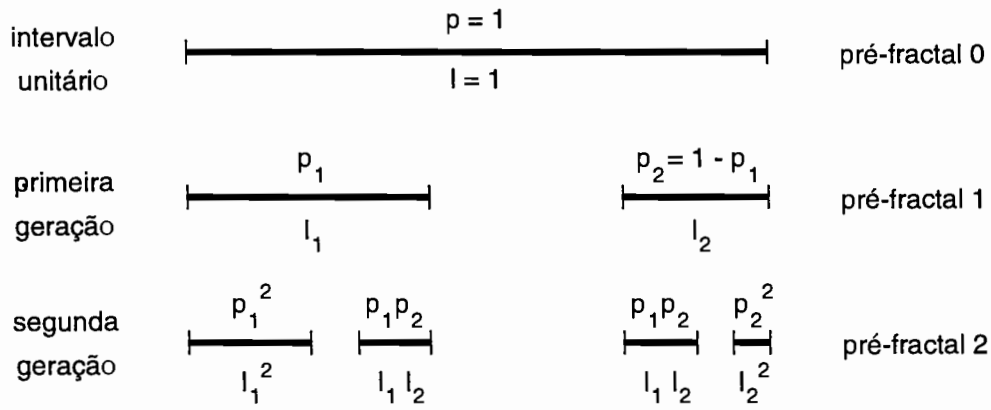


Figura 2.1: Primeiras gerações na construção do Conjunto de Cantor Generalizado.

Este é um dos poucos conjuntos para os quais existe solução teórica da curva $f(\alpha)$, dada pelas seguintes equações [19]:

$$q(\ln p_1 \ln l_2 - \ln p_2 \ln l_1) = \ln(n/m)\ln(l_1/l_2) - \ln(n/m-1)\ln l_1 \quad (2.1)$$

$$f(n/m) = \frac{(n/m-1)\ln(n/m-1) - n/m\ln(n/m)}{\ln l_1 + (n/m-1)\ln l_2} \quad (2.2)$$

$$\alpha(n/m) = \frac{\ln p_1 + (n/m-1)\ln p_2}{\ln l_1 + (n/m-1)\ln l_2} \quad (2.3)$$

onde a curva $f(\alpha)$ é dada parametricamente em n/m, que é um número real maior que 1.

Utilizamos duas maneiras de representar este conjunto numericamente. O primeiro é o sistema de funções iteradas, que o aproxima por um conjunto de pontos, que é a maneira como os conjuntos são usualmente representados, e o segundo é o método da escada do diabo, que calcula a massa do conjunto, e portanto fornece um valor mais preciso das probabilidades p_i. Descrevemos estas duas formas de representação em detalhe no apêndice A.

2.2 Atrator do mapa quadrático

Estudamos o atrator do mapa $x_{i+1} = \lambda(1 - 2x_i^2)$ com $\lambda = 0.837005134$, que é aproximadamente o valor onde termina a seqüência de bifurcações por duplicação de

período e inicia a janela caótica [19]. A trajetória com condição inicial $x = 0$ forma um Conjunto de Cantor como mostrado na fig. 2.2. A mesma quantidade de pontos cai em cada segmento, de modo que ao utilizar como recobrimento um dos pré-fractais desse Conjunto de Cantor, a probabilidade associada a cada segmento é igual.

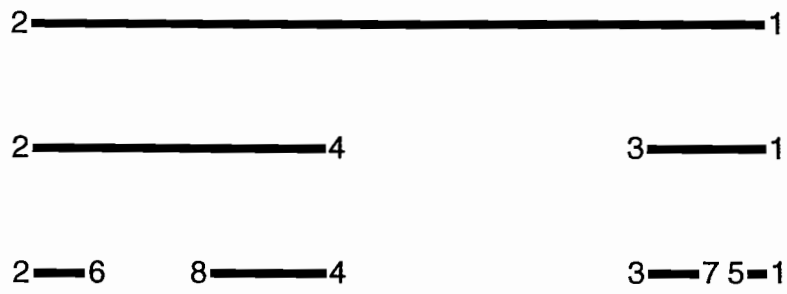


Figura 2.2: Pré-fractais do Conjunto de Cantor associado ao atrator do mapa quadrático no limiar da janela caótica. Os números indicam as iterações do mapa com condição inicial em $x = 0$.

A solução teórica [19] é então obtida gerando um grande conjunto de pontos (utilizamos 2^{11}), determinando os segmentos do pré-fractal do Conjunto de Cantor associado ao atrator do mapa quadrático e substituindo na eq. 1.14 seus comprimentos, determinando $\tau(q)$ numericamente, e a curva $f(\alpha)$ através da transformação de Legendre de $\tau(q)$.

2.3 Atrator do mapa de Hénon

O mapa de Hénon [20] é dado pelas expressões:

$$x_{i+1} = 1.0 - A\, x_i^2 + y_i \tag{2.4}$$

$$y_{i+1} = B\, y_i \tag{2.5}$$

onde $A = 1.4$ e $B = 0.3$ são os valores geralmente utilizados na literatura.

O espectro de singularidades do atrator do mapa de Hénon não possui uma solução teórica, mas é um sistema muito utilizado, e podemos comparar os resultados obtidos com outros cálculos numéricos existentes na literatura [4, 17, 2].

Capítulo 3

O problema das caixas espúrias

Neste capítulo vamos descrever o que, conforme concluímos no trabalho de mestrado, é a maior fonte de erro em cálculos das dimensões generalizadas e do espectro de singularidades através de métodos numéricos que utilizam o algoritmo de contagem de caixa, e que originou nosso trabalho de doutorado.

No trabalho de mestrado, mostramos que a utilização de uma grade para se obter o recobrimento de um conjunto nos métodos de contagem de caixa usuais permite o aparecimento de caixas que são quase totalmente preenchidas pelo vazio da estrutura multifractal, se situando no que poderíamos considerar como borda do conjunto em uma dada escala. Se o recobrimento fosse construído com o rigor descrito na teoria multifractal ao invés de se utilizar uma grade, essas caixas não existiriam, ou teriam menor tamanho, de modo que as denominamos de *caixas espúrias*. Acreditamos que essas caixas são responsáveis pelas distorções que ocorrem no ramo $q < 0$ das curvas D_q e $f(\alpha)$.

Nos cálculos das dimensões generalizadas e do espectro de singularidades realizados através de métodos de contagem de caixa observamos que, no momento do cálculo da regressão linear dos mesmos, nos valores do ramo $q < 0$ havia uma grande flutuação (fig. 3.1). Essa flutuação possuia as seguintes características:

- Deslocando-se a grade em relação ao conjunto, as flutuações continuam a existir, não se alterando, portanto, a situação qualitativamente. Porém, quantitativamente os valores apresentam uma grande sensibilidade quanto à posição da grade.

- Utilizando correções estatísticas, como o cálculo dos valores médios ou supremos

e ínfimos* com relação a diferentes grades, elimina-se a flutuação, mas o valor obtido não corresponde ao esperado teoricamente.

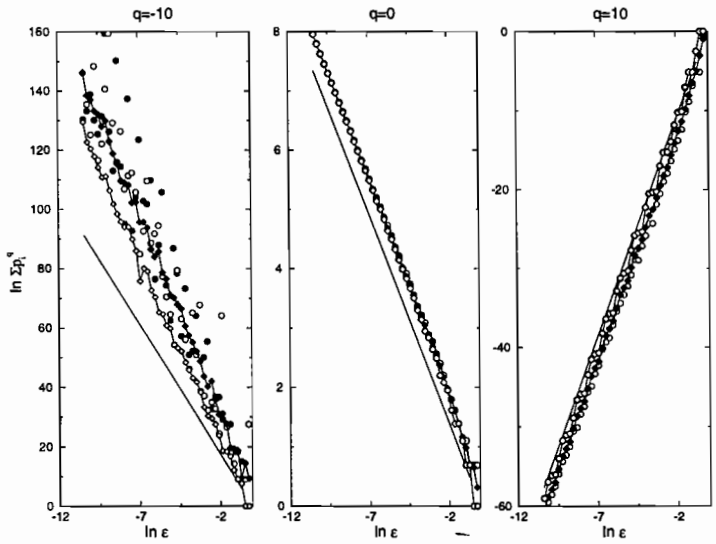


Figura 3.1: Flutuações para $q < 0$: No método microcanônico, $\tau(q)$ é dado pelos coeficientes angulares de $\sum p_i^q$ contra o $\ln \epsilon$. Para $q < 0$ observa-se uma grande flutuação nos valores, observada também nos outros métodos que utilizam o algoritmo de contagem de caixa. Foi utilizado o Conjunto de Cantor Generalizado $l_1 = 0.4$, $l_2 = 0.35$ e $p_1 = 0.6$, gerado pelo método da escada do diabo.
reta sólida: Inclinação calculada teoricamente.
círculos escuros: A grade inicia em $x = 0$.
círculos claros: Deslocamento de $3/5 \times 2^{-15}$.
diamantes escuros: Médias em 50 deslocamentos múltiplos de $2^{-15}/50$.
diamantes claros: Ínfimo/Supremo dos deslocamentos acima.

De onde concluímos que:

- Este é um problema independente da posição da grade em relação ao conjunto.
- O problema não possui uma causa simplesmente estatística, ele está associado a um erro sistemático do método.

Observando que, nas expressões utilizadas para o cálculo das dimensões generalizadas e do espectro de singularidades, sempre se calcula somatórias da q-ésima potência

*Esse critério tem sua origem na teoria multifractal. Na eqs. 1.12, escolhe-se entre os diferentes recobrimentos aquele que possui o menor valor de $\sum_i p_i^q$ para $q < 1$ e o maior para $q > 1$, o que eqüivale, no método microcanônico à escolha do ínfimo ou supremo dos valores de $\ln \sum_i p_i^q$.

das probabilidades, se existir alguma probabilidade muito pequena, quando $q < 0$, essa probabilidade contribuirá decisivamente na somatória. Assim, se de alguma forma, existirem no recobrimento caixas de pequena probabilidade que não representam uma área rarefeita real do conjunto, elas irão produzir uma grande distorção no resultado.

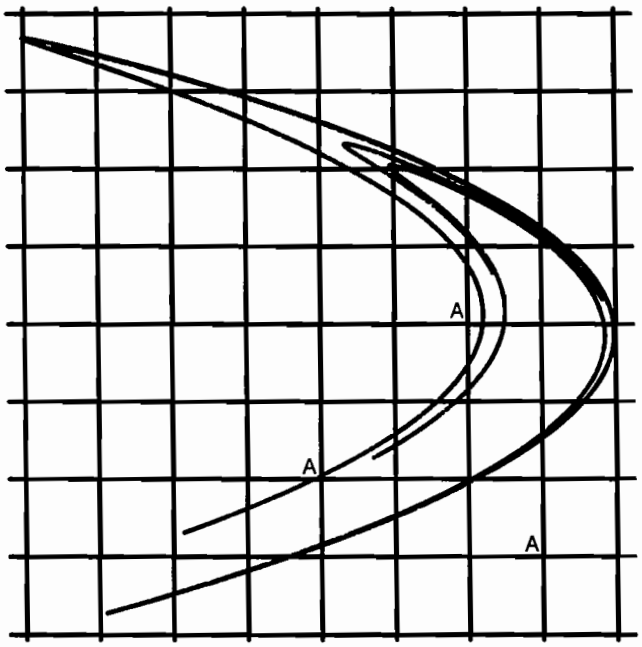


Figura 3.2: Recobrimento feito por uma grade: O recobrimento feito por uma grade faz com que existam caixas espúrias no mesmo. Essas caixas espúrias possuem um baixo valor de probabilidade não porque se situam sobre uma região de baixa probabilidade, mas por que estão quase preenchidas pelo espaço vazio da estrutura multifractal do conjunto. Algumas caixas espúrias são mostradas marcadas com a letra A.

Na fig. 3.2 mostramos um recobrimento feito através de uma grade em um conjunto multifractal típico. Observa-se que existem várias caixas da grade que estão quase completamente preenchidas pelo vazio da estrutura multifractal. Essas caixas possuem baixa probabilidade porque possuem uma pequena fração dos pontos em seu interior, mas não estão necessariamente em áreas rarefeitas do conjunto, mas nas "bordas" do mesmo. Se estivéssemos fazendo um recobrimento como descrito pela teoria, as caixas deveriam se ajustar ao conjunto, de maneira que não existiriam caixas como as observadas anteriormente. Por isso denominamos essas caixas de espúrias.

A existência dessas caixas explica as flutuações observadas nos valores do ramo

$q < 0$. Elas ocorrem em todas as escalas, porque como o objeto multifractal possui detalhes em todas as suas escalas de ampliação, existem vazios e "bordas" em todas as escalas. Dessa forma, ao contrário do que poderia se esperar a partir de uma análise ingênua, a proporção de caixas espúrias não se anula no limite de pequenos tamanhos de caixa.

A existência das caixas espúrias também não depende da posição da grade. Movimentando-a não se altera a situação do recobrimento qualitativamente: o deslocamento da grade apenas remove umas e gera outras caixas espúrias em diferentes locais. Mas os valores das probabilidades mudam, o que altera quantitativamente os valores da flutuação observada. Como um conjunto multifractal possui detalhes em todas as suas escalas de ampliação, mesmo pequenos deslocamentos são capazes de produzir alterações.

Ainda que o conjunto esteja sendo bem representado estatisticamente (ou seja, representado por um número suficientemente grande de pontos), existirão caixas espúrias, porque estas são geradas pela incompatibilidade entre a geometria do conjunto e da grade usada para seu recobrimento. Dessa forma, procedimentos estatísticos como o cálculo de médias ou supremos e ínfimos sobre posições da grade não são suficientes para se obter os resultados corretos. Porém, observamos que, em geral, o critério de ínfimos ou supremos fornece um valor mais próximo do esperado teoricamente.

Assim, depois que identificamos a causa do problema, adotamos duas estratégias para tentar resolvê-lo:

- Estudar métodos para o cálculo do espectro de singularidades que não utilizassem algoritmos de contagem de caixa, e se possível propor algum novo método.

- Tentar propor um novo algoritmo de contagem de caixa que construísse recobrimentos que evitassem ao máximo a existência das caixas espúrias.

Os resultados obtidos serão apresentados a seguir e constituem esta tese de doutoramento.

Capítulo 4

Métodos de Contagem de Caixa

Neste capítulo vamos apresentar e comparar entre si os resultados obtidos através dos métodos tradicionais que utilizam o algoritmo de caixa na sua forma original. Isto representa parte de um trabalho de sistematização desses métodos, que são encontrados dispersos na literatura.

Tendo em vista o problema das caixas espúrias, vamos apresentar nossos resultados para um conjunto onde os recobrimentos feitos pelo algoritmo de contagem de caixa usual possuam exatamente a mesma estrutura que ele. Nesses casos não deve-se esperar distorções devidas às caixas espúrias. Depois apresentaremos os resultados obtidos para outros conjuntos. Nesses conjuntos, verificamos que há deformações na curva $f(\alpha)$ provocadas pelas caixas espúrias.

4.1 Correções Estatísticas

Antes de apresentarmos a análise dos diferentes métodos para o cálculo de $f(\alpha)$, iremos introduzir alguns procedimentos adotados inicialmente em relação ao método de contagem de caixa e que correspondem a tentativas de eliminar, através de procedimentos meramente estatísticos, a dependência da posição relativa da grade ao conjunto, e representou inicialmente, uma forma ingênua de tentar reduzir o efeito das caixas espúrias. A isto denominamos genericamente de correções estatísticas para o algoritmo de contagem de caixa.

De acordo com a teoria, descrita na sec. 1.3.1, deve-se cobrir o conjunto estudado utilizando-se diferentes recobrimentos, adotando porém, os valores das probabilidades de um único recobrimento, escolhido de acordo com o critério de ínfimos e supremos do

valor da função $\Gamma(q, \tau', l)$. Calculando diferentes recobrimentos com posições relativas ao conjunto distintas, podemos então considerar como resultado final: (a) a média entre as diferentes grades (critério da média) ou (b) utilizar um critério de ínfimos e supremos similar ao definido teoricamente.

A média é efetuada sobre os valores que, submetidos à regressão linear, fornecem os valores de $\tau(q)$ ou $f(\alpha)$, que chamamos de $SomTau$, $SomAlfa$ e $SomFalfa$, nos métodos canônico e microcanônico. No método dos histogramas apenas o critério de médias é usado, através de uma média dos histogramas calculados em diferentes recobrimentos. O ínfimo ou supremo são sempre escolhidos entre os valores $SomTau$.

No decorrer deste trabalho, não pudemos determinar com exatidão qual dos critérios fornece o melhor resultado nos métodos canônico e microcanônico. Apesar de descrito pela teoria como o critério a ser usado na escolha de um recobrimento, o critério de ínfimos e supremos forma curvas a serem submetidas à regressão mais irregulares. O critério de médias por sua vez, forma curvas mais lisas e algumas vezes, mais precisas.

Nos cálculos realizados neste capítulo, onde o algoritmo de caixa usual é utilizado, na maior parte das vezes o critério de ínfimos e supremos fornece um melhor resultado final. No capítulo 7, onde os resultados do novo algoritmo de caixa são descritos, não se consegue determinar qual o melhor critério. Uma razão para isso é, no caso onde utiliza-se o algoritmo usual, a existência das caixas espúrias. Devido a elas, o critério de médias tende a gerar um resultado mais distorcido. A utilização do novo algoritmo gera melhores recobrimentos, de modo que a média dos valores destes pode se aproximar mais do resultado teórico, e assim seu resultado pode ser melhor do que o calculado com a utilização do critério de ínfimos e supremos.

4.2 Método Microcanônico

No método microcanônico, apresentado na sec. 1.5.1, o espectro de singularidades $f(\alpha)$ é calculado através da Transformação de Legendre do espectro de massa $\tau(q)$. Um exemplo da implementação deste método pode ser encontrada no Apêndice C.

A principal característica dos resultados obtidos através do método microcanônico é conseqüência da etapa do cálculo da derivada numérica de $\tau(q)$, que suaviza artificialmente a curva. Dessa forma, torna-se mais difícil verificar se o resultado está correto, caso não se possua uma maneira de compará-lo a outro mais confiável ou

exato, porque a curva adquire uma "aparência" de resultado correto.

De maneira geral, no método microcanônico, as barras de erro de $f(\alpha)$ são grandes. Isso ocorre devido à propagação de erro no cálculo da Transformação de Legendre. O termo $f(\alpha)$ possui uma parcela $q\alpha$, o que faz com que seu erro aumente com o módulo de q.

Um cálculo utilizando o Método Microcanônico é apresentado na fig. 4.1, onde o resultado é comparado ao obtido através de outros métodos que utilizam o algoritmo de contagem de caixa. O conjunto no qual o método foi aplicado é o Conjunto de Cantor Generalizado, com parâmetros $l_1 = l_2 = 0.5$ e $p_1 = 0.6$. O conjunto foi gerado utilizando o método da escada do diabo. Os parâmetros utilizados no método foram $q = 0, \pm 1, \pm 2, \ldots 10$ e $\epsilon = 2^{-i}(1 + 2^{-10}), i = 1, 2, \ldots 10$. Este conjunto é bem particular, não possuindo espaços vazios e tendo estrutura uniforme, sendo portanto compatível com a grade usada em seu recobrimento. Dessa forma, não aparecem caixas espúrias e não há distorções no resultado final.

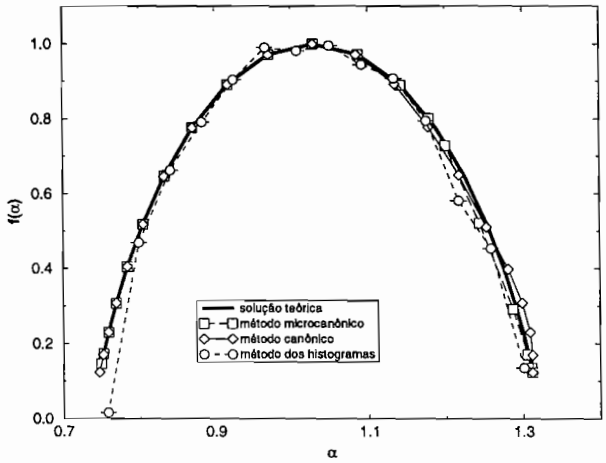


Figura 4.1: Curva $f(\alpha)$: Métodos para o cálculo de $f(\alpha)$ que utilizam algoritmos de contagem de caixa: As curvas foram calculadas para o Conjunto de Cantor Generalizado com parâmetros $l_1 = l_2 = 0.5$ e $p_1 = 0.6$. Os métodos calculam corretamente a curva $f(\alpha)$, mesmo utilizando o algoritmo usual de contagem de caixa, porque a estrutura desse conjunto é compatível com a estrutura de uma grade.
quadrados – Método Microcanônico
diamantes – Método Canônico
círculos – Método dos Histogramas

A seguir mostramos resultados para conjuntos com estruturas espaciais diferentes desta, onde espera-se que apareçam distorções devido às caixas espúrias. Na fig. 4.2 mostramos o resultado para o Conjunto de Cantor com parâmetros $l_i = 0.4$, $l_2 = 0.35$ e $p_1 = 0.6$. O conjunto foi representado por 200 000 e 500 000 pontos, gerados por IFS. Os parâmetros do método foram $q = 0, \pm 1, \pm 2 \ldots \pm 10$ e $\epsilon = 2^{-i}(1 + 2^{-10})$, com $i = 1, 2, \ldots 10$. Foi utilizado o critério de ínfimos e supremos em 51 deslocamentos de grade. Observa-se uma grande distorção no ramo $q < 0$ da curva $f(\alpha)$ calculada, exatamente aquele afetado pela presença de caixas espúrias. A utilização do critério

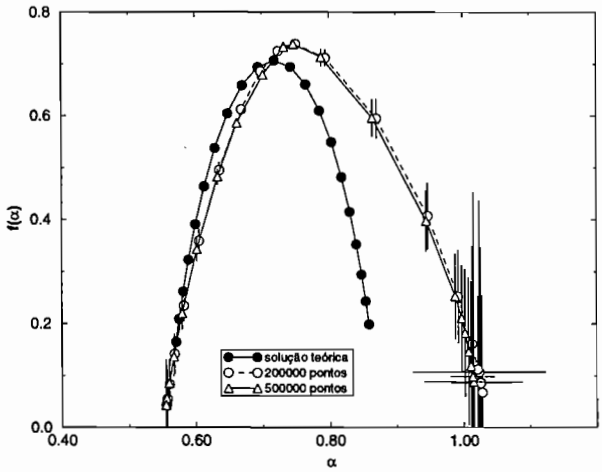


Figura 4.2: Método Microcanônico, utilizando o algoritmo usual de contagem de caixa: Curva $f(\alpha)$ do Conjunto de Cantor Generalizado com parâmetros $l_i = 0.4$, $l_2 = 0.35$ e $p_1 = 0.6$. A existência de caixas espúrias produz uma distorção no ramo $q < 0$ da curva $f(\alpha)$, mesmo com a utilização de correções estatísticas.

de ínfimos e supremos produz um resultado melhor do que a utilização do critério de médias, como podemos observar na fig. 4.3. Foram usados nesse cálculo os mesmos parâmetros do caso anterior. O conjunto foi representado por 500 000 pontos.

Na fig. 4.4 mostramos as curvas de $SomTau = \ln \sum p_i^q$ para valores de $q = -10, 0$ e 10, obtidos no cálculo anterior. O resultado obtido utilizando os critérios de médias e ínfimos ou supremos é apresentado. Para $q < 0$ o valor obtido através do critério de ínfimos e supremos se aproxima mais da inclinação esperada teoricamente. A diferença entre os valores obtidos calculando as médias e o calculado usando o critério de ínfimos e supremos se deve às flutuações existentes nos valores de $SomTau$ devido às caixas espúrias. Na região de $q < 0$ o valor de suas probabilidades dá uma

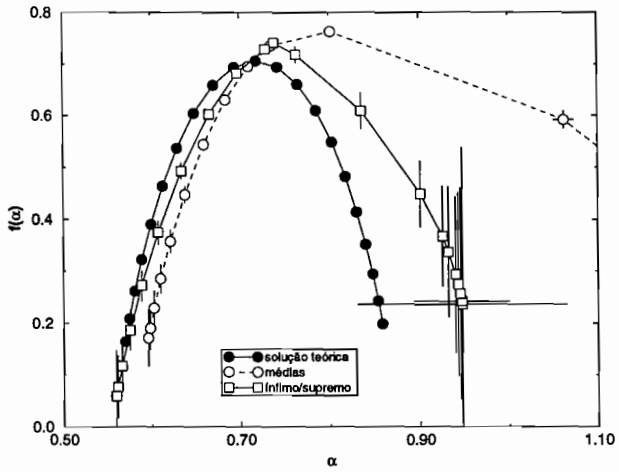


Figura 4.3: Método Microcanônico, utilizando o algoritmo usual de contagem de caixa: Curva $f(\alpha)$ do Conjunto de Cantor Generalizado com parâmetros $l_i = 0.4$, $l_2 = 0.35$ e $p_1 = 0.6$ gerado por IFS e representado por 500 000 pontos. O critério de ínfimos e supremos produz o melhor resultado, mas ambos critérios não corrigem a distorção das caixas espúrias no ramo $q < 0$.

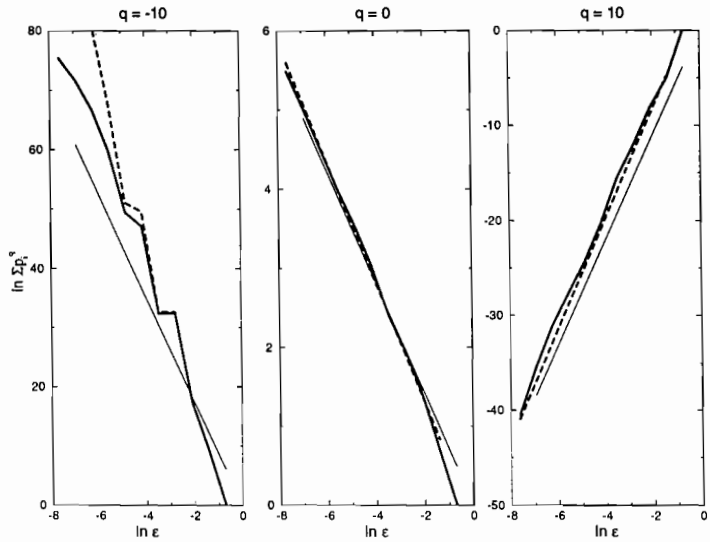


Figura 4.4: Método Microcanônico: Curvas $Somtau = \ln \sum p_i^q$ do Conjunto de Cantor Generalizado com parâmetros $l_i = 0.4$, $l_2 = 0.35$ e $p_1 = 0.6$. Foram utilizados 51 deslocamentos de grade para o cálculo de médias (pontilhado) e o critério de ínfimos e supremos (sólido). Existe uma diferença entre ambos e a inclinação esperada teoricamente (linha fina) quando $q < 0$, mas o critério de ínfimos e supremos fornece o melhor resultado.

grande contribuição à somatória. Como elas surgem irregularmente, dependendo da geometria do conjunto, os valores de $SomTau$ também se mostram irregulares.

Na fig. 4.5 vemos o resultado obtido para o atrator do mapa quadrático $x_{i+1} = \lambda(1 - 2x_i^2)$ com $\lambda = 0.837005134$. O conjunto foi representado por 300 000 e 700 000 pontos. Os parâmetros do método foram $q = 0, \pm 1, \pm 2 \ldots \pm 10$ e $\epsilon = 1.2 \times 2^{-i}$, com $i = 1, 2, \ldots 10$. Foi utilizado o critério de ínfimos e supremos em 51 deslocamentos de grade. Neste caso a simples seleção estatística do critério de ínfimos e supremos corrige em grande parte as distorções do ramo $q < 0$. A utilização desse critério produziu um

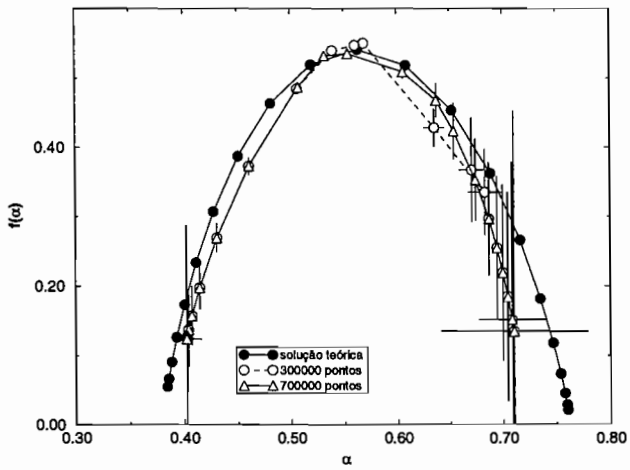


Figura 4.5: Método Microcanônico, utilizando o algoritmo usual de contagem de caixa: Curva $f(\alpha)$ do atrator do mapa quadrático $x_{i+1} = \lambda(1 - 2x_i^2)$ com $\lambda = 0.837005134$, Neste caso o critério parece eliminar a maior parte da interferência das caixas espúrias, pois a distorção do ramo $q < 0$ é pequena.

resultado melhor do que a utilização do critério de médias, como podemos observar na fig. 4.6. Foram usados neste cálculo os mesmos parâmetros do anterior, representando o conjunto por 700 000 pontos. Observando as curvas $SomTau = \ln \sum p_i^q$ dos cálculos anteriores, apresentadas na fig. 4.7 pode-se verificar que realmente o critério de ínfimos e supremos produziu uma curva com a inclinação mais próxima do cálculo teórico para valores $q < 0$.

Para o atrator do mapa de Hénon $x_{i+1} = 1.0 - A\,x_i^2 + y_i$ e $y_{i+1} = B\,y_i$, onde $A = 1.4$ e $B = 0.3$, temos o resultado mostrado na fig. 4.8. O conjunto foi representado por 200 000 e 1 200 000 pontos. Os parâmetros do método foram $q = 0, \pm 1, \pm 2 \ldots \pm 10$ e $\epsilon = 2.6 \times 2^{-i}$, com $i = 1, 2, \ldots 8$. Foi utilizado o critério de ínfimos e supremos

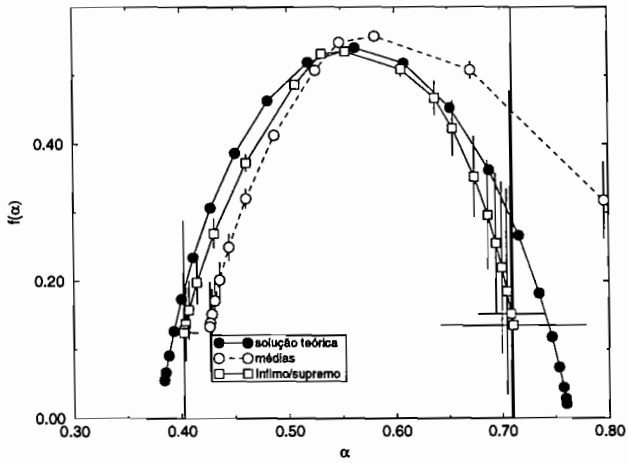


Figura 4.6: Método Microcanônico, utilizando o algoritmo usual de contagem de caixa: Curva $f(\alpha)$ do atrator do mapa quadrático $x_{i+1} = \lambda(1 - 2x_i^2)$ com $\lambda = 0.837005134$. O critério de ínfimos e supremos forneceu o melhor resultado. A distorção no ramo $q < 0$ é bem visível com a utilização do critério de médias.

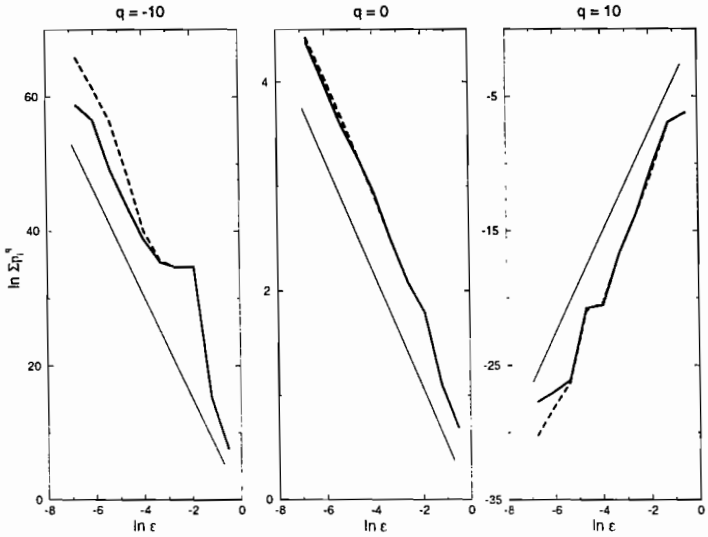


Figura 4.7: Método Microcanônico: Curvas $\ln \sum p_i^q$ do atrator do mapa quadrático $x_{i+1} = \lambda(1-2x_i^2)$ com $\lambda = 0.837005134$. Foram utilizados 51 deslocamentos de grade para o cálculo de médias (pontilhado) e o critério de ínfimos e supremos (sólido). O critério de ínfimos e supremos gerou uma curva com a inclinação mais próxima do valor teórico para $q < 0$.

em 97 deslocamentos de grade. Observamos que a distorção no ramo $q < 0$ da curva é grande, devido à existência das caixas espúrias. Comparando o resultado obtido

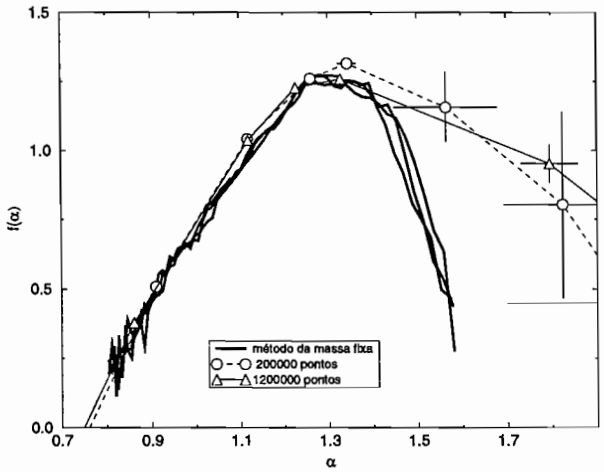


Figura 4.8: Método Microcanônico, utilizando o algoritmo usual de contagem de caixa: Curva $f(\alpha)$ do atrator do mapa de Hénon $x_{i+1} = 1.0 - A\,x_i^2 + y_i$ e $y_{i+1} = B\,y_i$, onde $A = 1.4$ e $B = 0.3$, Observa-se que a distorção no ramo $q < 0$ é grande, devido à existência das caixas espúrias.

utilizando os critérios de médias e ínfimo ou supremo, vemos que o ramo $q < 0$ continua muito distorcido em ambos os casos, como podemos observar na fig. 4.9. De fato, tínhamos uma expectativa de que, em um conjunto de dimensão maior que 1, os efeitos do aparecimento de caixas espúrias devido ao uso de uma grade fossem acentuados. Isso era esperado porque em duas dimensões é muito menos provável que o conjunto estudado possua uma geometria compatível com uma grade.

Observa-se que, em quase todos os resultados, permanece uma grande distorção no ramo de $q < 0$, mesmo após a utilização de correções estatísticas. Como explicamos anteriormente, concluímos que essas distorções são produzidas pelas caixas espúrias. Elas possuem uma baixa probabilidade não porque se situam em regiões pouco densas do conjunto, mas porque estão quase totalmente preenchidas pelos vazios inerentes à estrutura multifractal. Dessa forma, distorcem com maior intensidade o ramo de $q < 0$ da curva $f(\alpha)$.

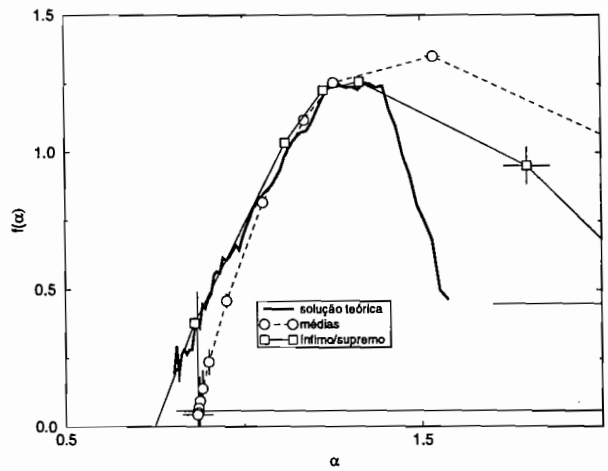


Figura 4.9: Método Microcanônico, utilizando o algoritmo usual de contagem de caixa: Curva $f(\alpha)$ do atrator do mapa de Hénon $x_{i+1} = 1.0 - A\,x_i^2 + y_i$ e $y_{i+1} = B\,y_i$, onde $A = 1.4$ e $B = 0.3$, representado por 1 200 000 pontos. Observa-se que a distorção no ramo $q < 0$ é grande, devido à existência das caixas espúrias.

4.3 Método Canônico

O método canônico, apresentado na sec. 1.5.2, calcula o espectro de singularidades $f(\alpha)$ parametricamente em q, usando a mesma definição teórica que o método microcanônico. Um exemplo da implementação deste método pode ser encontrada no Apêndice D.

Como o argumento teórico deste método é o mesmo que o do método microcanônico, é razoável esperar que os resultados sejam semelhantes. A diferença entre ambos é a eliminação do cálculo da derivada numérica. O argumento a favor desse método é que ele forneceria um melhor resultado para curvas $f(\alpha)$ que possuem algum tipo de descontinuidade. A desvantagem é que, neste método, é necessário calcular o dobro de regressões lineares em relação ao método microcanônico. Como em muitos casos se observam flutuações nas curvas submetidas à regressão, o maior número delas torna o método mais vulnerável. Um exemplo de um cálculo é mostrado na fig. 4.1, onde foi utilizado o Conjunto de Cantor Generalizado $l_1 = l_2 = 0.5$ e $p_1 = 0.6$ onde foram aplicados os mesmos parâmetros utilizados no método anterior. Esse conjunto é aquele que não possui uma estrutura de espaços vazios e cuja curva $f(\alpha)$ pode ser obtida sem problemas utilizando métodos de contagem de caixa usuais.

Repetimos os cálculos feitos através do Método Microcanônico utilizando agora o Método Canônico, mantendo sempre os mesmos parâmetros para os conjuntos utilizados:

Realizamos o cálculo para o Conjunto de Cantor com parâmetros $l_i = 0.4$, $l_2 = 0.35$ e $p_1 = 0.6$. Foi utilizado o critério de ínfimos e supremos em 51 deslocamentos de grade. Observa-se que, mesmo assim, o ramo $q < 0$ está bastante distorcido, devido à existência das caixas espúrias (fig. 4.10). A diferença na utilização do critério de

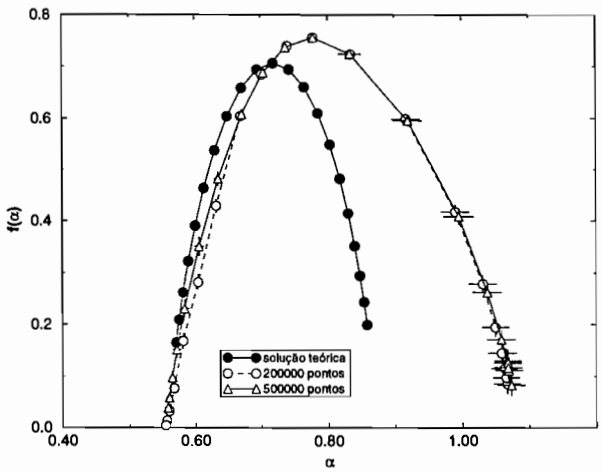


Figura 4.10: Método Canônico, utilizando o algoritmo usual de contagem de caixa: Curva $f(\alpha)$ do Conjunto de Cantor Generalizado com parâmetros $l_i = 0.4$, $l_2 = 0.35$ e $p_1 = 0.6$. A existência de caixas espúrias produz uma distorção no ramo $q < 0$ da curva $f(\alpha)$.

médias ou de ínfimos e supremos pode ser observada na fig. 4.11. Embora ambas as curvas estejam bem distorcidas no ramo $q < 0$, a utilização do critério de ínfimos e supremos conduz ao melhor resultado. Nesse cálculo foram usados 500 000 pontos do conjunto.

Na fig. 4.12 e 4.13 mostramos as curvas de $SomAlfa = \sum \mu_i \ln p_i$ e $SomFalfa = \sum \mu_i \ln \mu_i$, para valores de $q = -10, 0$ e 10, obtidos no cálculo anterior. O resultado obtido com as médias e o critério de ínfimos ou supremos em diferentes posições da grade é apresentado.

Note que, nesse caso, o critério de ínfimos e supremos não corresponde ao cálculo dos ínfimos e supremos de $SomAlfa$ e $SomFalfa$, mas sim de $SomTau$. Observe também que as curvas de $SomFalfa$ para módulo de q grande aparentam possuir uma grande flutuação, devido aos pequenos valores assumidos pelas mesmas. A

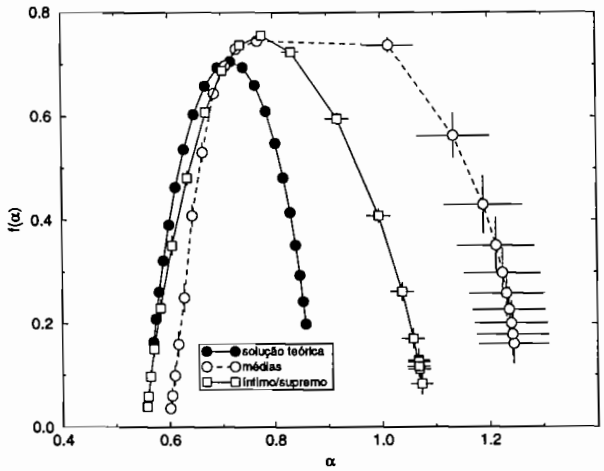


Figura 4.11: Método Canônico, utilizando o algoritmo usual de contagem de caixa: Curva $f(\alpha)$ do Conjunto de Cantor Generalizado com parâmetros $l_i = 0.4$, $l_2 = 0.35$ e $p_1 = 0.6$. A existência de caixas espúrias produz uma distorção no ramo $q < 0$ da curva $f(\alpha)$.

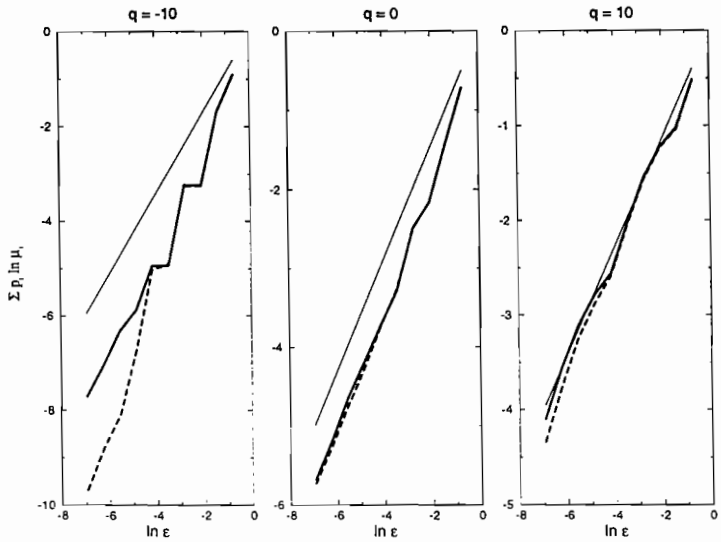


Figura 4.12: Método Canônico: Curvas $SomAlfa = \sum \mu_i \ln p_i$ do Conjunto de Cantor Generalizado com parâmetros $l_i = 0.4$, $l_2 = 0.35$ e $p_1 = 0.6$. Foram utilizados 51 deslocamentos de grade para o cálculo de médias (pontilhado) e o critério de ínfimos e supremos (sólido). O resultado obtido com o último se aproxima mais da inclinação esperada teoricamente (linha fina). A diferença entre as curvas para $q = -10$ deve-se à existência de flutuações devido à existência de caixas espúrias.

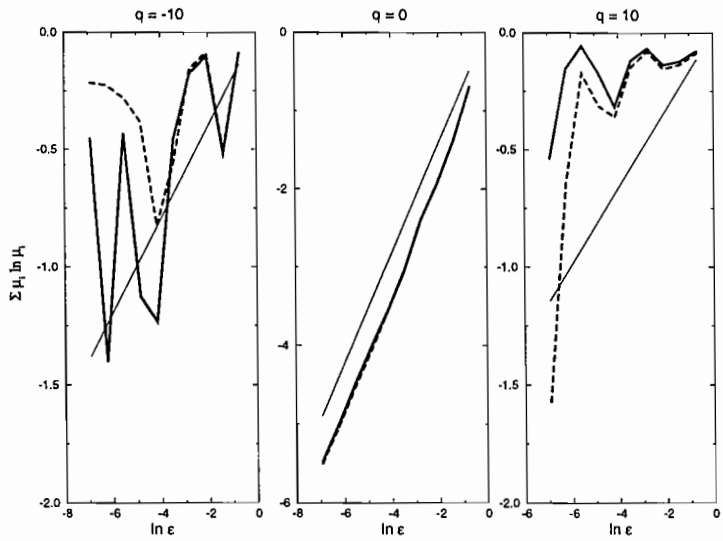


Figura 4.13: Método Canônico: Curvas $SomFalfa = \sum \mu_i \ln \mu_i$ do Conjunto de Cantor Generalizado com parâmetros $l_i = 0.4$, $l_2 = 0.35$ e $p_1 = 0.6$. Foram utilizados 51 deslocamentos de grade para o cálculo de médias (pontilhado) e o critério de ínfimos e supremos (sólido). A diferença entre as curvas para $q = -10$ deve-se à existência de flutuações devido à existência de caixas espúrias.

diferença entre os valores obtidos calculando as médias e usando o critério de ínfimos e supremos para $q = -10$ também reflete as flutuações existentes nos valores de $SomTau$ devido às caixas espúrias. Na região de $q < 0$ o valor de suas probabilidades dá uma grande contribuição às somatórias. Como elas surgem irregularmente, os valores de $SomAlfa$ e $SomFalfa$ também se mostram irregulares.

O resultado para o mapa quadrático é mostrado na fig. 4.14. Foi utilizado o critério de ínfimos e supremos em 51 deslocamentos de grade. O ramo $q < 0$ está um pouco mais distorcido, quando comparado ao resultado obtido através do método microcanônico. A utilização do critério de médias gerou um resultado com o ramo $q < 0$ mais distorcido (fig. 4.15). Novamente neste cálculo foram usados os mesmos parâmetros do cálculo anterior, sendo que o conjunto foi representado por 700 000 pontos. As respectivas curvas $SomAlfa$ e $SomFalfa$ são mostradas nas fig. 4.16, 4.17. Nota-se que o resultado obtido utilizando o critério de ínfimos e supremos é, novamente, mais próximo do teórico.

Na fig. 4.18 mostramos o resultado obtido para o atrator do mapa de Hénon. Foi utilizado o critério de ínfimos e supremos em 97 deslocamentos de grade. Novamente, a distorção no ramo $q < 0$ é grande. A solução obtida utilizando-se o critério de

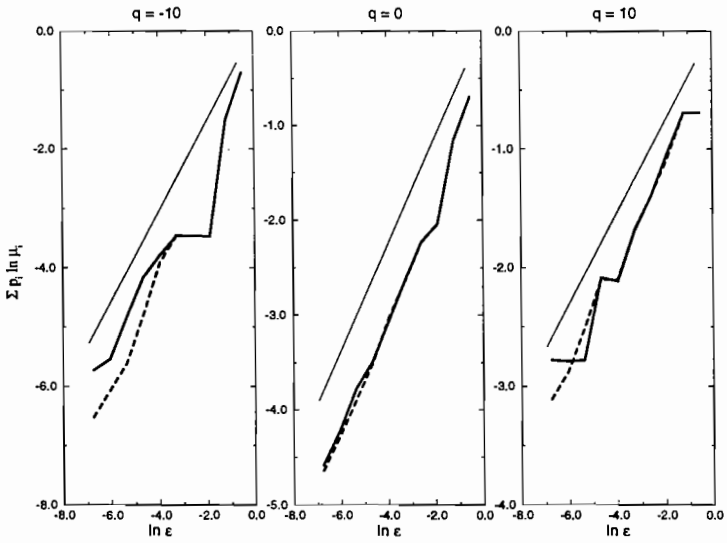


Figura 4.16: Método Canônico: Curvas $SomAlfa = \sum \mu_i \ln p_i$ do mapa quadrático $x_{i+1} = \lambda(1 - 2x_i^2)$ com $\lambda = 0.837005134$, representado por 700 000 pontos. Foram utilizados 51 deslocamentos de grade para o cálculo de médias (pontilhado) e o critério de ínfimos e supremos (sólido). O resultado obtido com o último se aproxima mais da inclinação esperada teoricamente (linha fina).

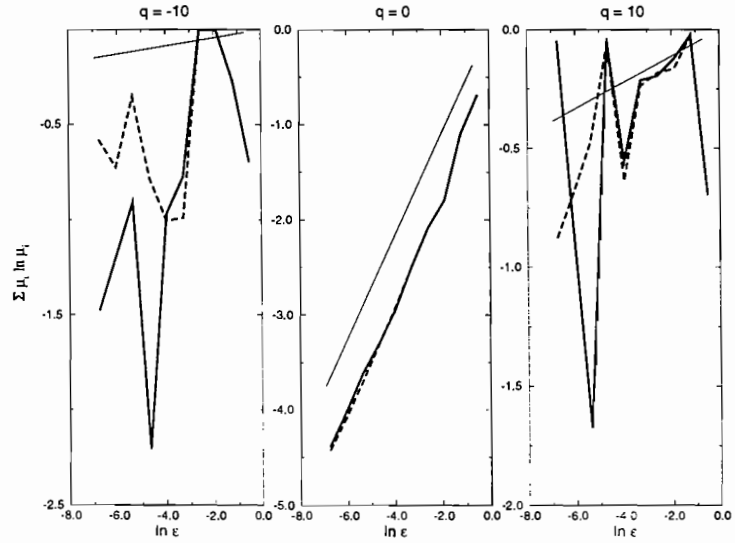


Figura 4.17: Método Canônico: Curvas $SomFalfa = \sum \mu_i \ln \mu_i$ do mapa quadrático $x_{i+1} = \lambda(1 - 2x_i^2)$ com $\lambda = 0.837005134$, representado por 700 000 pontos. Foram utilizados 51 deslocamentos de grade para o cálculo de médias (pontilhado) e o critério de ínfimos e supremos (sólido).

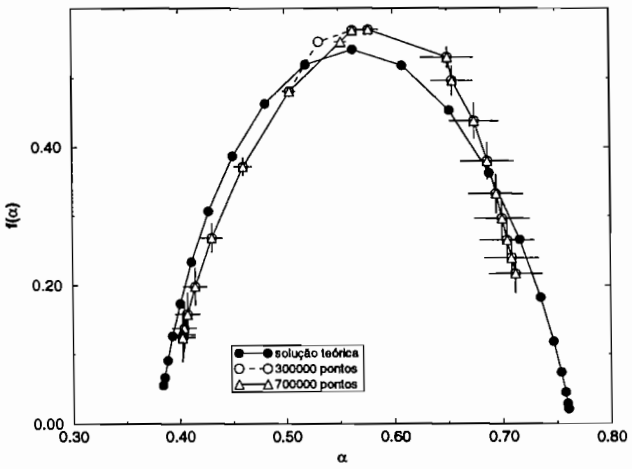


Figura 4.14: Método Canônico, utilizando o algoritmo usual de contagem de caixa: Curva $f(\alpha)$ do atrator do mapa quadrático $x_{i+1} = \lambda(1 - 2x_i^2)$ com $\lambda = 0.837005134$. Neste caso o critério de ínfimos e supremos parece ter sido eficiente em eliminar a interferência das caixas espúrias, pois a distorção do ramo $q < 0$ é menor.

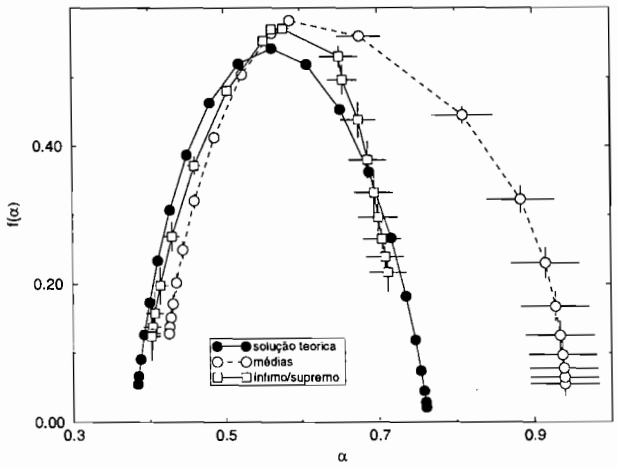


Figura 4.15: Método Canônico, utilizando o algoritmo usual de contagem de caixa: Curva $f(\alpha)$ do atrator do mapa quadrático $x_{i+1} = \lambda(1 - 2x_i^2)$ com $\lambda = 0.837005134$. O critério de ínfimos e supremos produziu um resultado mais próximo do previsto teoricamente.

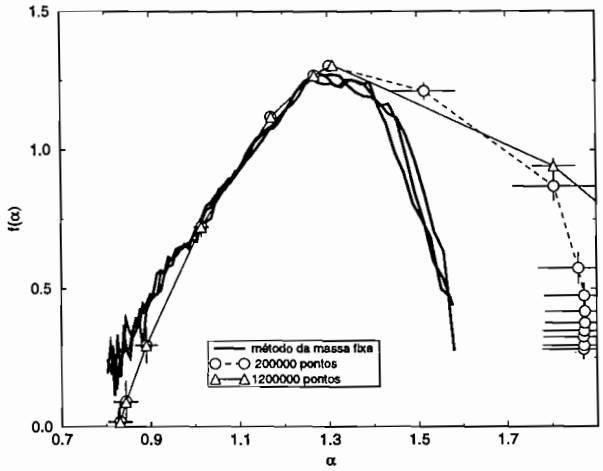


Figura 4.18: Método Canônico, utilizando o algoritmo usual de contagem de caixa: Curva $f(\alpha)$ do atrator do mapa de Hénon $x_{i+1} = 1.0 - A\,x_i^2 + y_i$ e $y_{i+1} = B\,y_i$, onde $A = 1.4$ e $B = 0.3$. Observa-se que a distorção no ramo $q < 0$ é grande.

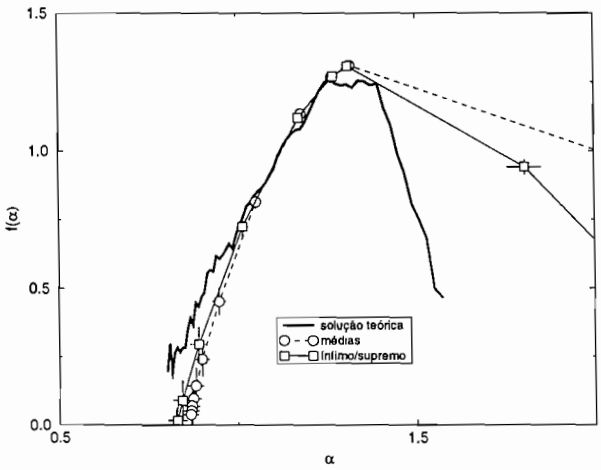


Figura 4.19: Método Canônico, utilizando o algoritmo usual de contagem de caixa: Curva $f(\alpha)$ do atrator do mapa de Hénon $x_{i+1} = 1.0 - A\,x_i^2 + y_i$ e $y_{i+1} = B\,y_i$, onde $A = 1.4$ e $B = 0.3$, representado por 1 200 000 pontos. Observa-se que a distorção no ramo $q < 0$ é grande, também devido à existência das caixas espúrias.

médias é muito semelhante à obtida utilizando-se o critério de ínfimos e supremos, como podemos observar na fig. 4.19. Nesse cálculo o conjunto foi representado por 1 200 000 pontos. Os demais parâmetros foram mantidos constantes.

De maneira geral, novamente o ramo de $q < 0$ das curvas $f(\alpha)$ se apresenta distorcido, devido à presença de caixas espúrias. Também vemos que os resultados são muito semelhantes aos obtidos utilizando o método microcanônico. Mas, nos casos apresentados, acreditamos que a diferença entre os resultados se deve mais às flutuações existentes nas curvas submetidas à regressão linear do que à etapa de derivação numérica.

4.4 Método dos Histogramas

Nesta seção vamos analisar outro método do cálculo de $f(\alpha)$, que também utiliza o algoritmo de contagem de caixa, o chamado método dos histogramas. Mostraremos como os parâmetros do mesmo se relacionam e exemplificaremos com o cálculo de um conjunto multifractal. A seguir, mostraremos os resultados obtidos da aplicação deste método.

4.4.1 Parâmetros do método

Os parâmetros do método dos histogramas, descrito na sec. 1.5.3, são o número de colunas nos histogramas Nd e o menor tamanho de caixa da grade utilizada na contagem de caixa ϵ. Mostraremos a relação entre esses parâmetros utilizando o Conjunto de Cantor Generalizado com valores $l_1 = l_2 = 0.5$ e $p_1 = 0.6$.

Suponha que utilizaremos um único tipo de recobrimento, com caixas de tamanho $\epsilon = 2^{-k}$, com k inteiro positivo, e de modo que o início da primeira caixa sempre seja coincidente ao início do conjunto. Dessa maneira, as probabilidades das caixas desse recobrimento sobre o Conjunto de Cantor escolhido terão valores conhecidos, cuja distribuição é binomial, ou seja, as $i+1$ probabilidades existentes serão $p_1^m(1-p_1)^{k-m}$, onde $m = 0, 1 \ldots k$ e cada uma destas probabilidades aparecerá em $\binom{k}{m}$ caixas.

Para construir o histograma na escala k, devemos agora encontrar a maior e a menor probabilidade, e dividir o intervalo formado pelo logaritmo de ambas por Nd. Verifica-se então a ocorrência de dois problemas: Primeiro, se $Nd > k + 1$, existirão colunas do histograma com valores H_j nulos, o que não é permitido no método, porque o valor de $f(\alpha)$ será calculado a partir do logaritmo de H_j. Segundo, ao considerarmos

diferentes escalas, como o valor de Nd se mantém constante, os valores de H_j mudam abruptamente, pois as probabilidades formam uma distribuição de picos. Dessa forma são produzidas alterações bruscas nas curvas cuja regressão linear resulta em $f(\alpha)$, o que pode dificultar, ou até mesmo impedir, o seu cálculo.

Dessa forma, percebe-se o porque da necessidade de efetuarmos médias em H_j sobre várias posições da grade sobre o conjunto. Esse procedimento alarga os picos de probabilidades encontrados na contagem de caixa, sendo uma tentativa de eliminar os dois problemas. Apesar da discussão acima ter sido baseada em um conjunto em particular, na prática verifica-se que esses problemas ocorrem no cálculo feito em outros conjuntos.

Dessa forma, devemos estabelecer relações entre os parâmetros do método dos histogramas: o menor tamanho de caixa deve ser tal que gere um número de picos maior que o número de colunas do histograma. E as escalas com tamanhos de caixa grandes devem ser ignoradas, porque certamente os histogramas correspondentes possuem colunas vazias.

4.4.2 Aplicando o método

No método dos histogramas (sec. 1.5.3), o algoritmo de contagem de caixa é utilizado para o levantamento das probabilidades p_i. São construídos histogramas do logaritmo das probabilidades encontradas, para cada tamanho de caixa considerado. A partir destes histogramas, a curva $f(\alpha)$ é numericamente calculada. Uma implementação deste método é mostrada no Apêndice E.

As dificuldades encontradas na utilização desse método são várias. Na teoria, é feita a suposição de que ΔX, a largura da coluna dos histogramas, é pequeno, ou seja, o histograma deve ser próximo da distribuição que representa. Assim, seria necessário um grande número de caixas cobrindo os pontos do conjunto. Dessa forma, uma vez que o número de caixas é associado ao tamanho das mesmas e à dimensão fractal, pequenos tamanhos de caixa deveriam ser considerados. Em geral, estudar tais sistemas com tamanhos de caixa menores exige um maior número de pontos representando o conjunto e maiores recursos e esforço computacional, o que em muitos casos não é possível conseguir.

Aplicamos o método dos histogramas ao Conjunto de Cantor Generalizado cujos parâmetros são $l_1 = l_2 = 0.5$ e $p_1 = 0.6$, cujas probabilidades p_i foram calculadas numericamente utilizando-se o método da escada do diabo. Foi calculada a média dos

histogramas sobre 51 deslocamentos uniformes da grade no intervalo unitário, onde o Conjunto de Cantor foi representado periodicamente em todo o espaço. Construímos histogramas com 14 colunas, para tamanhos de caixa variando como $\epsilon = 2^{-i}(1 + 2^{-10}), i = 1, 2, 3 \ldots 19$.

Os histogramas obtidos são mostrados na fig. 4.20. A curva $f(\alpha)$ correspondente, na fig. 4.1. Como este conjunto não possui espaços vazios em sua estrutura, não apresenta caixas espúrias nos recobrimentos feitos pelo algoritmo de contagem de caixa, e consequentemente os resultados encontrados estão muito próximos dos esperados teoricamente.

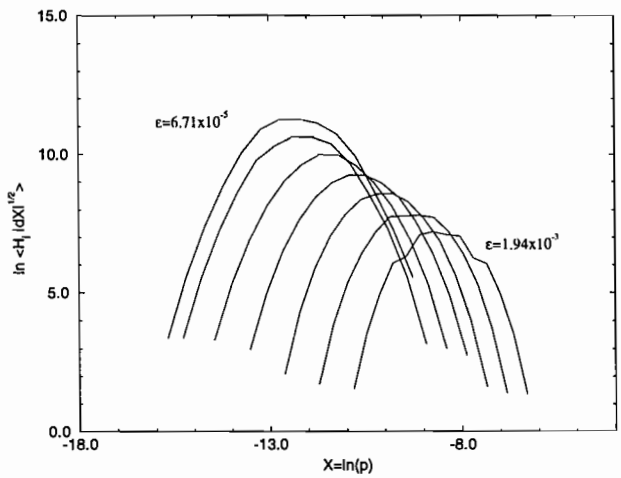


Figura 4.20: Método dos histogramas: Histogramas calculados para o Conjunto de Cantor Generalizado com parâmetros $l_1 = l_2 = 0.5$ e $p_1 = 0.6$. Este multifractal não possui uma estrutura de espaços vazios, de modo que não existem caixas espúrias. Os histogramas se apresentam aproximadamente simétricos.

Se repetirmos o cálculo em um conjunto com outra estrutura espacial, de modo geral pode-se esperar uma deformação da curva devido à existência de caixas espúrias. No método dos histogramas, quando há caixas espúrias, toda a curva é deformada, não ocorrendo apenas uma deformação maior no ramo $q < 0$ como nos métodos anteriores. Isso acontece porque os valores das probabilidades são usados para a construção de histogramas. Vamos mostrar a deformação nos histogramas e o seu reflexo na curva $f(\alpha)$ correspondente.

No Conjunto de Cantor Generalizado com parâmetros $l_i = 0.4$, $l_2 = 0.35$ e $p_1 = 0.6$ vemos a distorção provocada pela existência de caixas espúrias. Novamente as pro-

babilidades p_i foram calculadas numericamente utilizando o método da escada do diabo. Foi calculada a média dos histogramas sobre 99 deslocamentos uniformes da grade no intervalo unitário, onde o Conjunto de Cantor foi representado periodicamente em todo o espaço. Construímos histogramas com 14 colunas, para tamanhos de caixa $\epsilon = 2^{-i}(1+2^{-10}), i = 1,2,3\ldots 19$. Alguns desses histogramas são mostrados na fig. 4.21.

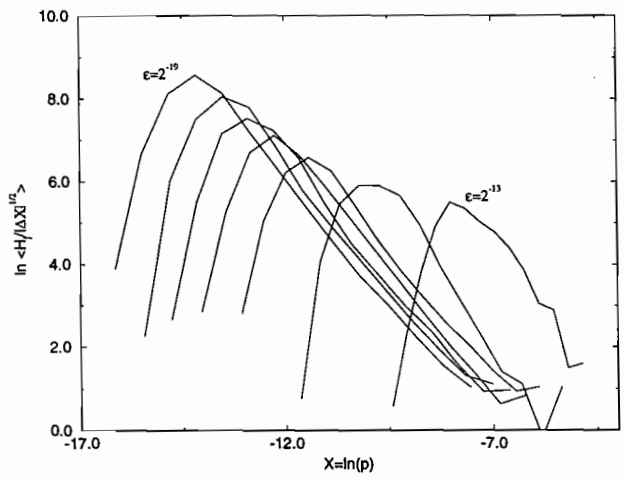


Figura 4.21: Método dos histogramas: Histogramas calculados para o Conjunto de Cantor Generalizado com parâmetros $l_i = 0.4$, $l_2 = 0.35$ e $p_1 = 0.6$. O valor do tamanho de caixa ϵ varia com fator 2 em cada histograma. Os histogramas estão distorcidos, mostram um grande número de caixas com baixa probabilidade. Esta distorção se deve à existência de caixas espúrias.

Esse conjunto possui uma estrutura de espaços vazios, de modo que aparecem muitas caixas espúrias, aumentando o número de caixas na região esquerda dos histogramas. A curva $f(\alpha)$ correspondente se apresenta bem distorcida (fig. 4.22). A diferença entre este método e os métodos canônico e microcanônico é que, nos últimos, os cálculos envolvem somatórias de potências do tipo p_i^q de modo que o ramo de $q > 0$ não é muito afetado pela presença das caixas espúrias, porque elas possuem baixa probabilidade, fornecendo uma contribuição menor à somatória. Em contrapartida, no método dos histogramas, o cálculo baseia-se na construção de um histograma das probabilidades dividindo em partes iguais o intervalo do logaritmo das probabilidades encontradas, de modo que a presença de caixas espúrias vai afetará também o resultado do ramo $q > 0$ da curva $f(\alpha)$, distorcendo-a globalmente.

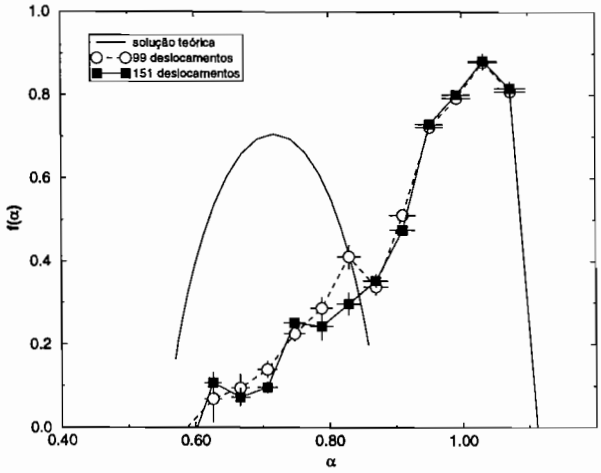


Figura 4.22: Método dos histogramas, utilizando o algoritmo usual de contagem de caixa: Curva $f(\alpha)$ calculada para o Conjunto de Cantor Generalizado com parâmetros $l_i = 0.4$, $l_2 = 0.35$ e $p_1 = 0.6$, relacionada aos histogramas da fig. 4.21. A curva se apresenta distorcida devido à existência das caixas espúrias.

Os histogramas obtidos aplicando este método ao mapa quadrático $x_{i+1} = \lambda(1 - 2x_i^2)$ com $\lambda = 0.837005134$ são mostrados na fig. 4.23. O conjunto foi representado por 10 000 000 pontos. Construimos histogramas com 14 colunas, para tamanhos de caixa $\epsilon = 1.2 \times 2^{-i}, i = 1, 2, 3 \ldots 19$. Foi feita a média entre 151 deslocamentos de caixa. Os histogramas também se apresentam distorcidos, devido à existência das caixas espúrias.

A curva $f(\alpha)$ correspondente é mostrada na fig. 4.24. Nota-se que as flutuações existentes nos histogramas foram transportadas à curva $f(\alpha)$. Tais flutuações podem indicar que neste caso seria necessário construir histogramas para tamanhos de caixa ainda menores, ou histogramas com um número menor de colunas, uma vez que a alteração do número de posições da grade pouco alterou o resultado. Um cálculo com um menor número de colunas nos histogramas é mostrado na fig. 4.25. As curvas $f(\alpha)$ possuem menos flutuações, mas ainda há uma grande distorção nas mesmas. Isto era esperado pois essas distorções se devem à existência de caixas espúrias, e não são eliminadas por esses procedimentos estatísticos.

Os histogramas obtidos da aplicação deste método no atrator do mapa de Hénon $x_{i+1} = 1.0 - A\,x_i^2 + y_i$ e $y_{i+1} = B\,y_i$, onde $A = 1.4$ e $B = 0.3$ são mostrados na fig. 4.26. Foram usados 1 000 000 pontos para representar o conjunto, e foram feitas

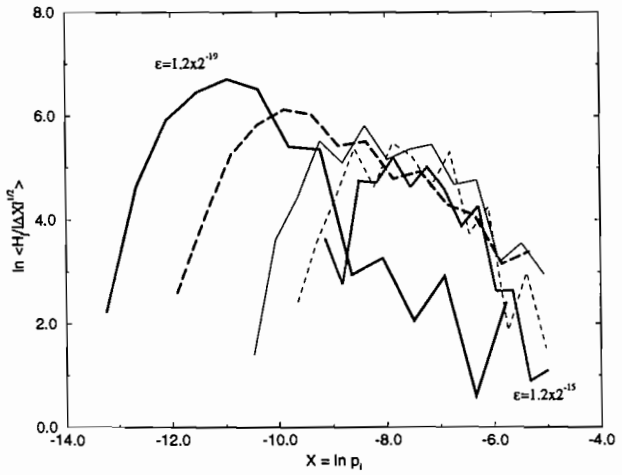


Figura 4.23: Método dos histogramas: Histogramas calculados para o mapa quadrático $x_{i+1} = \lambda(1 - 2x_i^2)$ com $\lambda = 0.837005134$. O valor do tamanho de caixa ϵ varia com fator 2 em cada histograma. Nota-se que os histogramas estão distorcidos, mostram um grande número de caixas com baixa probabilidade. Esta distorção se deve à existência de caixas espúrias.

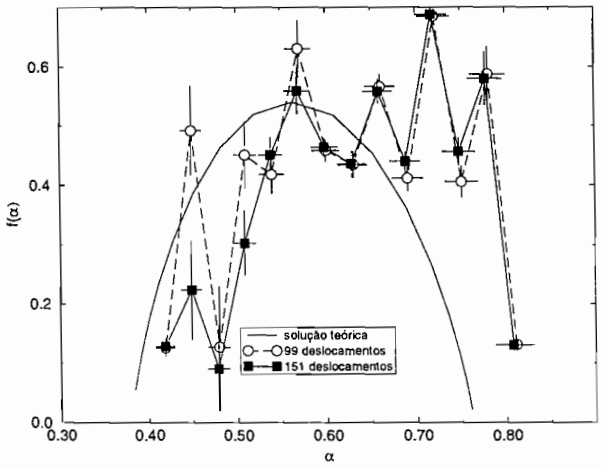


Figura 4.24: Método dos histogramas, utilizando o algoritmo usual de contagem de caixa: Curva $f(\alpha)$ calculada para o mapa quadrático $x_{i+1} = \lambda(1 - 2x_i^2)$ com $\lambda = 0.837005134$. Foram usados 97 e 151 deslocamentos de grade. As flutuações existentes nos histogramas são transmitidas à curva $f(\alpha)$ final.

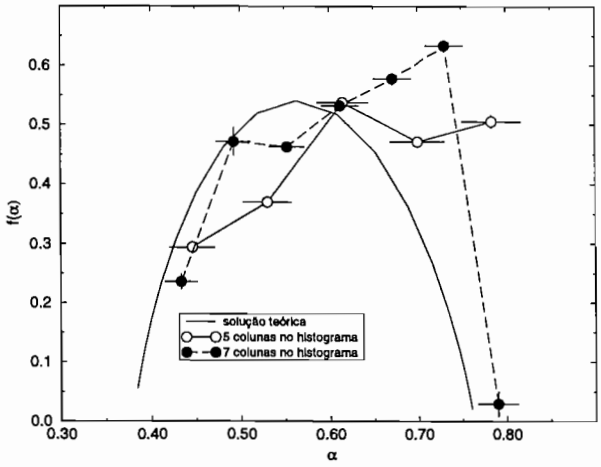


Figura 4.25: Método dos histogramas, utilizando o algoritmo usual de contagem de caixa: Curva $f(\alpha)$ calculada para o mapa quadrático $x_{i+1} = \lambda(1 - 2x_i^2)$ com $\lambda = 0.837005134$. Utilizando menos colunas nos histogramas, a curva fica mais suave, mas esse procedimento não elimina a distorção produzida pelas caixas espúrias.

médias em 151 deslocamentos de grade. Construímos histogramas com 14 colunas, para tamanhos de caixa $\epsilon = 2.6 \times 2^{-i}, i = 1, 2, 3 \ldots 8$. Pode-se ver a grande distorção gerada pelas caixas espúrias.

Essa distorção altera completamente a curva $f(\alpha)$ neste caso, mostrada na fig. 4.27.

Vemos que neste método, as curvas se apresentam distorcidas globalmente, porque a construção dos histogramas é prejudicada globalmente pela presença de caixas espúrias. Também pode-se notar flutuações nas curvas, devido ao fato dos histogramas possuírem colunas muito largas. Como pode-se ver pelos parâmetros usados neste método, ele é mais dispendioso do ponto de vista computacional do que os outros métodos que utilizam o algoritmo de contagem de caixa.

Em resumo, pudemos concluir que o uso do algoritmo usual de contagem de caixa afeta os três métodos estudados, e introduz um erro sistemático. Este erro não é eliminado através de correções estatísticas (critério de médias, critério de ínfimos e supremos, mudança do número de colunas do histograma). Esses critérios estatísticos melhoram o resultado, mas são incapazes de eliminar o efeito das caixas espúrias. No capítulo 7, retomaremos o problema das caixas espúrias, propondo um novo algoritmo de contagem de caixa.

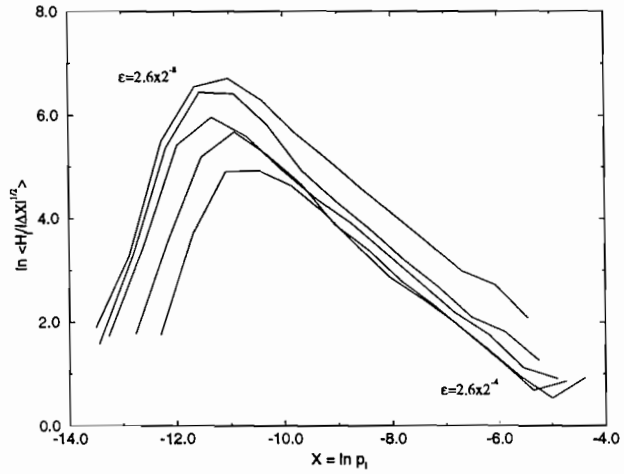


Figura 4.26: Método dos histogramas: Histogramas calculados para o atrator do mapa de Hénon $x_{i+1} = 1.0 - A\,x_i^2 + y_i$ e $y_{i+1} = B\,y_i$, onde $A = 1.4$ e $B = 0.3$. O valor do tamanho de caixa ϵ varia com fator 2 em cada histograma. Os histogramas estão distorcidos, mostram um grande número de caixas com baixa probabilidade. Esta distorção se deve à existência de caixas espúrias.

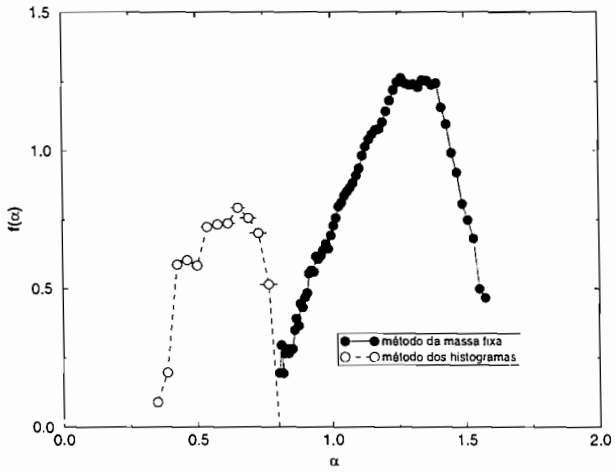


Figura 4.27: Método dos histogramas, utilizando o algoritmo usual de contagem de caixa: Curva $f(\alpha)$ calculada para o atrator do mapa de Hénon $x_{i+1} = 1.0 - A\,x_i^2 + y_i$ e $y_{i+1} = B\,y_i$, onde $A = 1.4$ e $B = 0.3$. Utilizando os parâmetros considerados, a curva $f(\alpha)$ se mostra bastante deformada.

Capítulo 5

Contagem de Distâncias

Como já expusemos anteriormente, acreditamos que os métodos de contagem de caixa produzem resultados distorcidos devido a problemas intrínsecos aos mesmos. Dessa forma, estudamos também métodos que calculam a função $f(\alpha)$ a partir de outra forma de analisar o conjunto. Os métodos cujos resultados vamos discutir a seguir, se baseiam na contagem e classificação de distâncias entre os pontos do conjunto. Com o auxílio desses métodos foram obtidos resultados muito bons para alguns dos conjuntos estudados, mas em outros casos encontramos dificuldades: Não conseguimos bons resultados nos Conjuntos de Cantor utilizando o Método da massa fixa, e no Atrator do mapa de Hénon no Método da Integral de Correlação.

5.1 Método da massa fixa

Nos algoritmos usuais de contagem de caixa, coleciona-se probabilidades de caixas iguais colocadas sobre o conjunto. O método da massa fixa parte de um princípio teórico diferente. Nele, armazena-se diversos tamanhos de caixa centradas em pontos do conjunto que contenham probabilidades iguais. Constrói-se um histograma da distribuição de tamanhos de caixa encontrados e a partir deste, a curva $f(\alpha)$.

São construídos histogramas para diferentes valores da probabilidade p para que se tenha noção do erro cometido no cálculo. Deve-se lembrar que, nesse método, o último passo é deslocar a curva de modo que ela tangencie a bissetriz do primeiro quadrante em $f(\alpha) = \alpha = D_1$, e que esse deslocamento também pode ser fonte de erros.

O principal argumento a favor deste método é que o mesmo fornece resultados

melhores no ramo negativo ($q < 0$) da curva $f(\alpha)$ do que no ramo positivo, e que tal diferença se acentua a medida que consideramos sistemas de maior dimensão. A razão dessa melhora pode ser entendida analisando o erro estatístico de se representar um conjunto multifractal como um conjunto finito de pontos. Esse erro é menos evidente nas regiões rarefeitas porque nelas, o raio da esfera de probabilidade p é maior, e uma variação no número de pontos produz, no raio, um erro percentual menor do que em regiões densas, onde os raios são menores.

Utilizamos o método da massa fixa em um Conjunto de Cantor Generalizado e no atrator do mapa de Hénon. Utilizando um algoritmo auxiliado por caixas, diminui-se consideravelmente o tempo de processamento numérico. Mostraremos porém, que o método não pode ser utilizado na sua forma original para caracterizar conjuntos do tipo poeira de Cantor.

Descrevemos em detalhe a implementação deste método no Apêndice I. Resumidamente, considere que o conjunto a ser estudado é representado por n pontos. Fixa-se uma probabilidade $p = k/n$ que as caixas devem conter. Escolhe-se um conjunto amostra de m pontos nos quais centraliza-se caixas que contém em seu interior k pontos e colecionamos os seus tamanhos. O tamanho da caixa é a distância do k-ésimo vizinho mais próximo do ponto de referência. Organizamos essas distâncias em um histograma, que por sua vez, será utilizado para calcular $f(\alpha)$.

O cálculo da curva $f(\alpha)$ para o atrator do mapa de Hénon foi mostrado na referência que apresenta o método [4]. O resultado que obtivemos neste caso (fig. 5.1) está de acordo com o obtido pelo autor. O conjunto foi representado por $n = 900\,000$ pontos. Foi considerada uma amostra de $m = 64\,000$ pontos para se construir o histograma. Foram feitos histogramas para caixas com $k = 100$, 75 e 50 pontos em seu interior, e eles contém 64 colunas.

Calculamos também a curva $f(\alpha)$ para o Conjunto de Cantor Generalizado com parâmetros $l_1 = l_2 = 0.5$ e $p_1 = 0.6$. Os parâmetros utilizados foram $n = 1\,000\,000$ pontos para representar o conjunto. Uma amostra de $m = 100\,000$ pontos foi utilizada para construir o histograma. Foram construídos histogramas associados à caixas com $k = 100$, 75 e 50 pontos em seu interior. Os histogramas possuem 30 colunas. O resultado obtido também é muito próximo do esperado (fig. 5.2).

Observamos porém, que o método não pode ser utilizado na sua proposta original em conjuntos do tipo poeira de Cantor. Estes conjuntos são compostos exclusivamente por pontos desconectados uns dos outros, possuindo vazios de vários tamanhos em sua

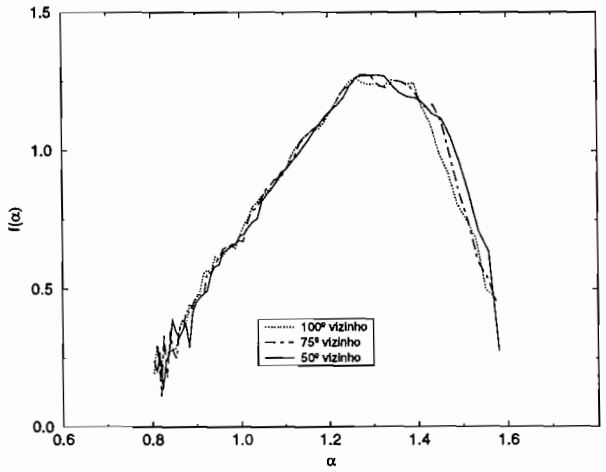


Figura 5.1: Método da massa fixa. Curva $f(\alpha)$: Atrator do mapa de Hénon. O resultado está de acordo com o apresentado no artigo que apresenta o método [4].

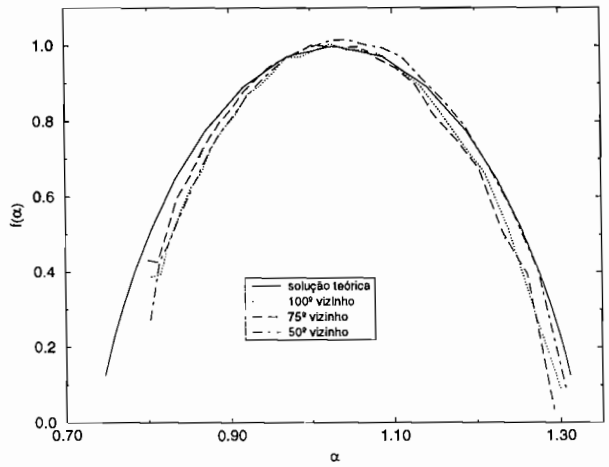


Figura 5.2: Método da massa fixa. Curva $f(\alpha)$: Conjunto de Cantor Generalizado com parâmetros $l_1 = l_2 = 0.5$ e $p_1 = 0.6$. O resultado obtido também está próximo ao resultado esperado.

estrutura. Dessa forma, nas distâncias relacionadas a esses vazios há uma diminuição do número de caixas encontradas no histograma.

Os dois conjuntos para os quais a curva $f(\alpha)$ foi calculada com este método anteriormente não são conjuntos do tipo poeira de Cantor. Eles são conjuntos conexos, ou seja, pode-se encontrar pelo menos um caminho que une quaisquer dois pontos do conjunto formado somente por pontos que pertencem ao conjunto. Dessa forma, o histograma de tamanhos de caixa não sofre nenhuma deformação. Por exemplo, no atrator do mapa de Hénon, utilizando os mesmos parâmetros descritos anteriormente, o histograma para $k = 50$ é mostrado na fig. 5.3.

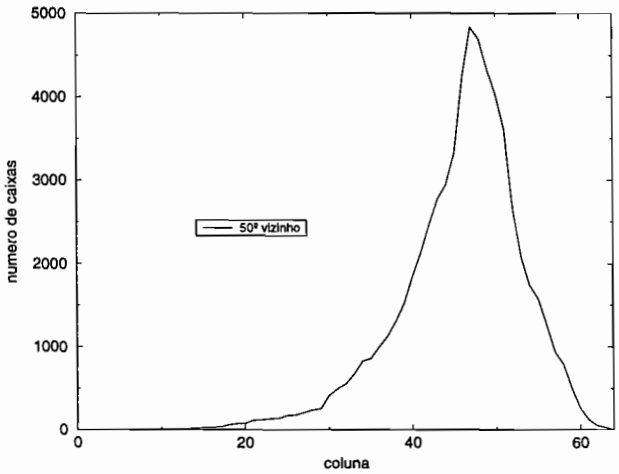


Figura 5.3: Método da massa fixa. Histograma do número de tamanhos de caixa contendo 50 pontos em seu interior, calculado para o atrator do Mapa de Hénon. O histograma não apresenta deformações porque o conjunto é conexo. O mesmo ocorre com o Conjunto de Cantor Generalizado cujo resultado foi mostrado nos gráficos anteriores.

Considere agora o Conjunto de Cantor Generalizado com parâmetros $l_1 = 0.4$, $l_2 = 0.35$ e $p_1 = 0.6$. Este conjunto é do tipo poeira de Cantor. O seu histograma é mostrado na fig. 5.4. Os parâmetros utilizados no cálculo foram $n = 500\,000$, $m = 64\,000$, $k = 50$. O histograma possui 64 colunas. As oscilações existentes no histograma se devem ao fato dos pontos do conjunto serem todos desconectados. Dessa forma, distâncias associadas aos espaços vazios que separam os pontos existirão em menor quantidade no histograma. Na fig. 5.5 mostramos um conjunto tipo poeira de Cantor típico. Ampliações revelam vazios cada vez menores na estrutura. Esses vazios em diferentes escalas geram os mínimos locais no histograma.

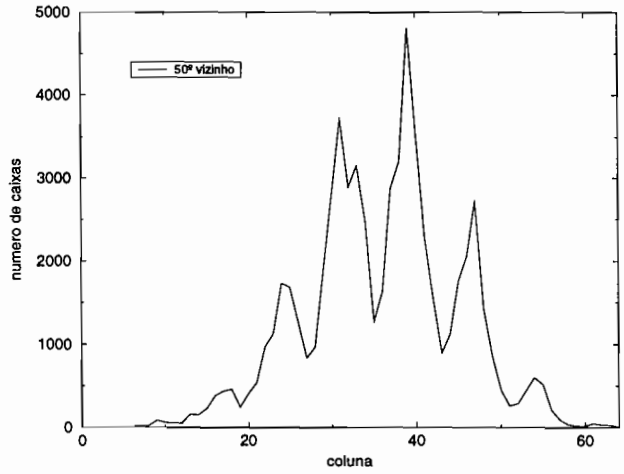


Figura 5.4: Método da massa fixa. Histograma do número de tamanhos de caixa contendo 50 pontos em seu interior, calculado para o Conjunto de Cantor Generalizado cujos parâmetros são $l_1 = 0.4$, $l_2 = 0.35$ e $p_1 = 0.6$. Este conjunto é do tipo poeira de Cantor, ou seja, todos os seus pontos são desconectados. Isso faz com que as distâncias associadas aos vazios que separam os pontos apareçam em menor número nos histogramas, produzindo oscilações que são transmitidas à curva $f(\alpha)$ final.

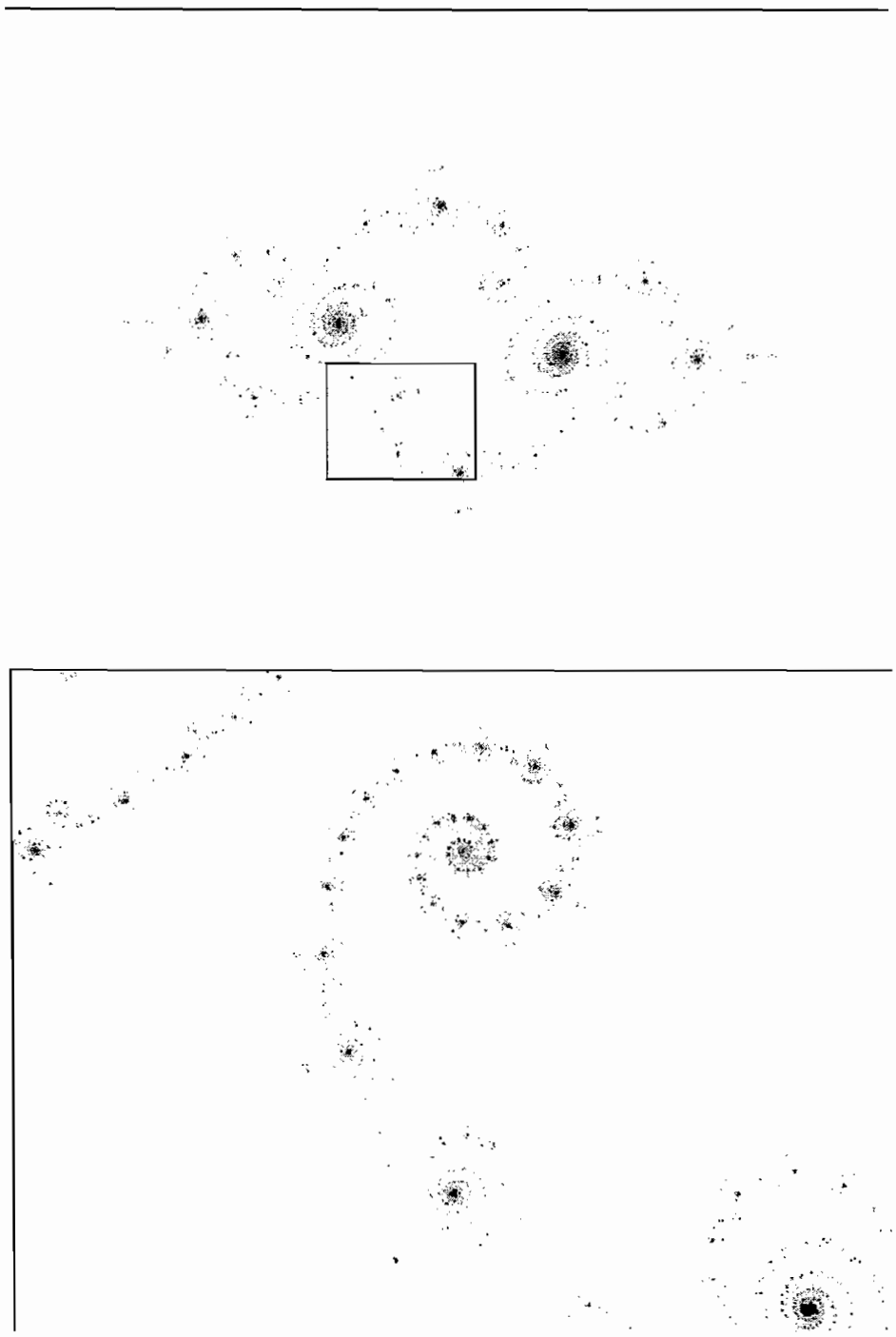


Figura 5.5: Conjunto de Julia associado ao mapa $z_{i+1} = z_i^2 + z_c$ com z complexo e $z_c = -0.760 + i0.135$. Esse conjunto é do tipo poeira de Cantor. Ampliações revelam vazios na estrutura aparentemente conexa. Esses vazios em diferentes escalas produzem as oscilações nos histogramas do método da massa fixa.

Essas oscilações não podem ser eliminadas alterando os parâmetros do método. Como os vazios que separam os pontos existem em todas as escalas, alterar o valor de k não elimina tais oscilações. Construir um histograma cujas colunas contenham a oscilação também não soluciona o problema, porque a curva $f(\alpha)$ resultante terá poucos pontos, de modo que no momento de deslocar a curva, haveria grande imprecisão. É necessário encontrar o ponto da curva com inclinação unitária e o cálculo da derivada não é preciso quando se possui poucos pontos.

Seria necessário utilizar alguma correção externa ao método para eliminar as oscilações. Entretanto, não obtivemos um bom resultado com algumas técnicas mais simples (convolução com função constante e gaussiana).

Há uma grande semelhança entre este método e o método dos histogramas. As duas diferenças mais importantes são: (a) O fato deste método colecionar tamanhos de caixa para probabilidades fixas, enquanto o segundo coleciona probabilidades para tamanhos de caixa fixos; (b) Neste método cada histograma gera uma curva, enquanto que no método dos histogramas, é necessário realizar regressões lineares dos valores contidos nas colunas dos histogramas.

O método da massa fixa poderia ser alterado para operar através de regressões lineares dos histogramas, e o método dos histogramas poderia obter a curva $f(\alpha)$através de um único histograma, de modo a eliminar a segunda diferença descrita entre ambos. No caso do método dos histogramas, a dificuldade estaria na construção do histograma, porque seria necessário um grande número de caixas. Isso implica em um tamanho de caixa muito pequeno, que não é facilmente alcançado na prática. No método da massa fixa, o histograma é obtido a partir de uma grande quantidade de pontos, o que para conjuntos teóricos é viável. A regressão linear possui problemas intrínsecos, como a escolha de uma região linear para o cálculo, de modo que a opção presente no método da massa fixa é mais conveniente. Ainda no caso da regressão linear dos histogramas, existe uma correção logarítmica a ser considerada, que é, portanto, evitada na proposta original do método.

Quanto aos resultados obtidos, o método apresentou bons resultados quando aplicado à conjuntos conexos. Dessa forma, ele poderia ser aplicado com êxito em fluxos. Uma outra vantagem é a rapidez no processamento numérico, desde que um algoritmo auxiliado por caixas seja utilizado. Concluímos também que este método não pode ser usado em conjuntos do tipo poeira de Cantor sem o auxílio de alguma técnica externa a ele.

5.2 Método da integral de correlação

O método da integral de correlação calcula a integral de correlação que lhe dá essa denominação através da contagem de distâncias entre pontos. Dados os pontos do conjunto, calculamos as distâncias existentes entre os mesmos e organizamo-las em um histograma, cujas colunas correspondem às diferentes escalas que desejamos considerar. A integral numérica de potências do histograma, partindo da distância nula até ϵ é a integral de correlação. Uma implementação deste algoritmo é descrita no Apêndice G.

Aplicamos este método aos Conjuntos de Cantor Generalizados, ao atrator do mapa quadrático no limiar da janela caótica e ao atrator do mapa de Hénon. Os resultados obtidos foram bons para o primeiro, e foram encontradas dificuldades nos outros casos. Além disso, esse é um método lento, mesmo quando se utiliza outros algoritmos e sugestões existentes na literatura para reduzir o tempo gasto no processamento. Essas sugestões se encontram descritas no Apêndice G. Uma delas, que consideramos nas análises abaixo é a escolha de um subconjunto de pontos para o cálculo das distâncias.

O resultado para o Conjunto de Cantor Generalizado com parâmetros $l_1 = 0.4$, $l_2 = 0.35$ e $p_1 = 0.6$ é mostrado na fig. 5.6. As escalas consideradas foram $\epsilon = 2^{i/2}$, com $i = 7, 8 \ldots 20$, para $q = 0, \pm 1, \pm 2 \ldots \pm 10$. Mostramos dois cálculos, que diferem apenas no número de pontos utilizado para representar o conjunto: 60 000 e 120 000.

As curvas $\ln < P_j^{q-1} >$ em função de $\ln \epsilon$ são apresentadas na fig. 5.7. São apresentados os valores para $q = 0, \pm 1, \pm 2 \ldots \pm 10$, sendo que o conjunto foi representado por 120 000 pontos. As curvas apresentam um comportamento aproximadamente linear, sem grandes flutuações. Dessa forma, os valores de $\tau(q)$ associados a essas curvas podem ser calculados com menor erro.

No caso, as amostras utilizadas são todos os pontos do conjunto. Tentamos também realizar o cálculo com subconjuntos menores, e obtivemos bons resultados para amostras maiores ou iguais a 5% para este Conjunto de Cantor Generalizado (fig. 5.8). Como o tempo de processamento é proporcional ao tamanho da amostra, é vantajoso realizar o cálculo com amostras pequenas. Com o aumento do tamanho da amostra, espera-se que o resultado obtido se aproxime ao calculado com uma amostra igual a todos os pontos. Para que esse recurso possa ser utilizado, deve-se verificar se o resultado obtido com determinada amostra já se encontra próximo a esse limite.

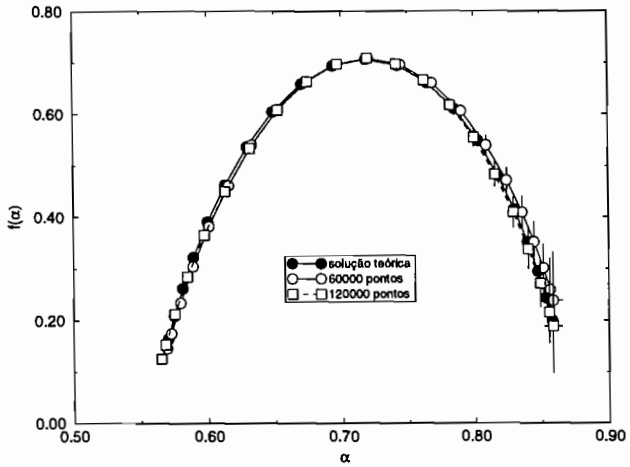


Figura 5.6: Método da Integral de correlação. Curva $f(\alpha)$. Calculada para o Conjunto de Cantor Generalizado cujos parâmetros são $l_1 = 0.4$, $l_2 = 0.35$ e $p_1 = 0.6$. O resultado obtido está muito próximo do teórico.

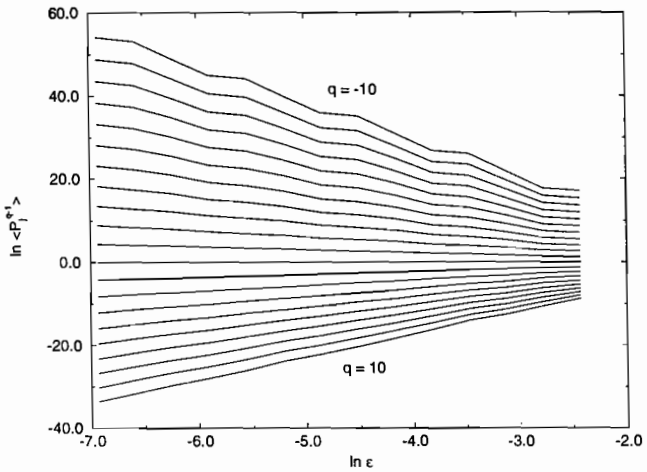


Figura 5.7: Método da Integral de correlação. Curvas $\ln < P_j^{q-1} >$. Calculada para o Conjunto de Cantor Generalizado cujos parâmetros são $l_1 = 0.4$, $l_2 = 0.35$ e $p_1 = 0.6$. De cima para baixo, estão as curvas correspondentes a $q = -10, -9, \ldots 10$. As curvas se apresentam quase lineares, com pouca flutuação.

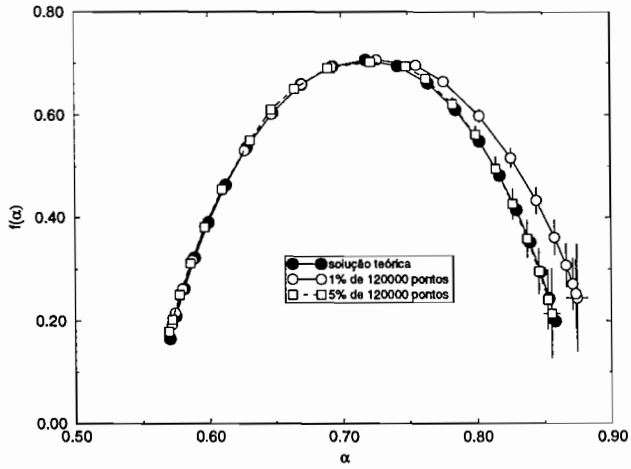


Figura 5.8: Método da integral de correlação. Curva $f(\alpha)$. Calculada para o Conjunto de Cantor Generalizado cujos parâmetros sao $l_1 = 0.4$, $l_2 = 0.35$ e $p_1 = 0.6$. O conjunto é representado por 120 000 pontos. Com uma amostra de 5% já se alcança um bom resultado. Utilizando amostras pequenas o tempo de processamento é menor, visto que ele é proporcional ao tamanho da amostra.

Espera-se que o tamanho mínimo da amostra na qual se obtém um bom resultado varie de conjunto para conjunto.

O resultado para o mapa quadrático no limiar da janela caótica $x_{i+1} = A(1 - x^2)$, com $A = 0.837005$ é mostrado na fig. 5.9. As escalas consideradas foram $\epsilon = 1.2 \times 2^{i/2}$, com $i = 7, 8 \ldots 20$, para $q = 0, \pm 1, \pm 2 \ldots \pm 10$. Mostramos dois cálculos, que diferem apenas no número de pontos utilizado para representar o conjunto: 50 000 e 100 000. Nota-se que o ramo $q < 0$ parece um pouco deslocado em relação à posição esperada teoricamente. não fomos capazes de determinar a origem desse deslocamento.

As curvas $\ln < P_j^{q-1} >$ em função de $\ln \epsilon$ são apresentadas na fig. 5.10. Foi usado o cálculo com 100 000 pontos. As curvas também são aproximadamente lineares, com poucas flutuações. Dessa maneira, parece improvável que o erro observado no ramo $q < 0$ da curva $f(\alpha)$ seja provocado por erros nas regressões lineares.

Também foi analisada a quantidade da amostra em relação ao número total de pontos do conjunto, cujo resultado é mostrado na fig. 5.11. Para amostras maiores que resultado que utilizando todos os pontos também para este conjunto.

Apresentamos também o resultado para o atrator do mapa de Hénon, na fig. 5.12. Foram feitos cálculos utilizando 90 000, 200 000 e 500 000 pontos. As escalas utilizadas

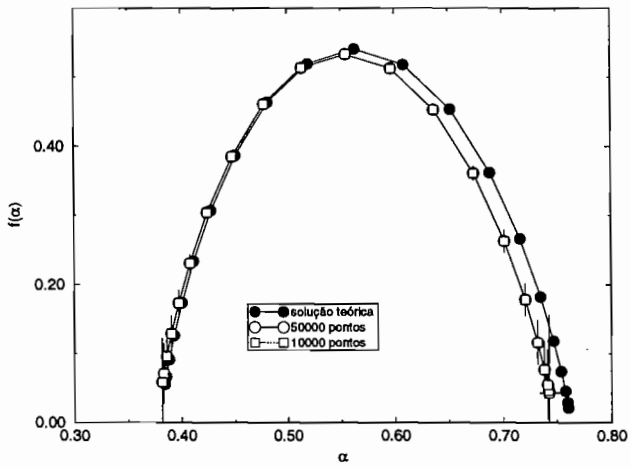


Figura 5.9: Método da integral de correlação. Curva $f(\alpha)$. Calculada para o atrator do mapa quadrático no limiar da janela caótica $x_{i+1} = A(1 - x^2)$, com $A = 0.837005$. Nota-se um deslocamento do ramo $q < 0$ neste caso.

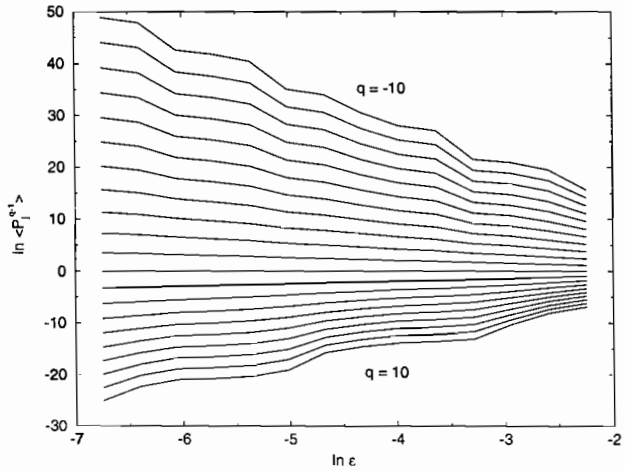


Figura 5.10: Método da Integral de correlação. Curvas ln $< P_j^{q-1} >$. Calculada para o atrator do mapa quadrático no limiar da janela caótica $x_{i+1} = A(1-x^2)$, com $A = 0.837005$. De cima para baixo, estão as curvas correspondentes a $q = -10, -9, \ldots 10$. As curvas se apresentam quase lineares, com pouca flutuação.

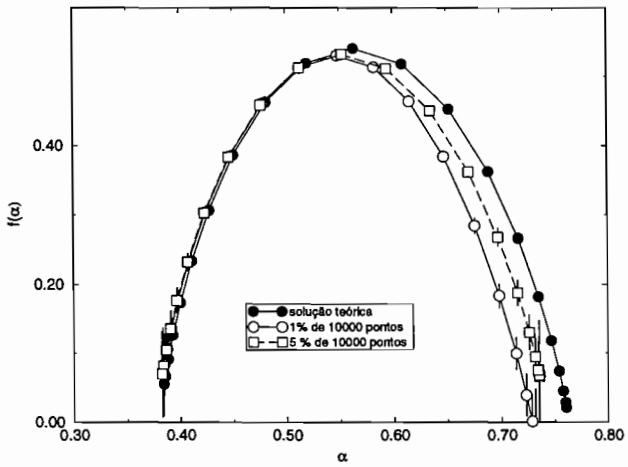


Figura 5.11: Método da Integral de correlação. Curva $f(\alpha)$. Calculada para o atrator do mapa quadrático no limiar da janela caótica $x_{i+1} = A(1 - x^2)$, com $A = 0.837005$. O conjunto é representado por 100 000 pontos. Com uma amostra de 5% já se alcança um bom resultado.

foram $\epsilon = 2.6 \times 2^{i/2}$, com $i = 7, 8 \ldots 20$, para $q = 0, \pm 1, \pm 2 \ldots \pm 10$.

Observamos uma grande distorção na curva $f(\alpha)$. Não pudemos determinar a origem dessa distorção, mas vemos que ela não parece diminuir conforme se aumenta o número de pontos usados para representar o conjunto.

As curvas $\ln < P_j^{q-1} >$ em função de $\ln \epsilon$ são apresentadas na fig. 5.13. Foi usado o cálculo com 200 000 pontos. As curvas também são aproximadamente lineares, com poucas flutuações. Dessa forma, a distorção observada não pode ser atribuída a um erro nas regressões lineares efetuadas.

Apresentamos também a curva $\tau(q)$ correspondente ao cálculo efetuado representando o conjunto por 1 200 000 pontos, na fig. 5.14. Ela apresenta um ponto de inflexão, que gera a deformação observada, após a aplicação da Transformação de Legendre na mesma.

Não foi possível determinar as causas da distorção observada. Podemos apenas apresentar algumas hipóteses: Talvez a distorção possa desaparecer em um cálculo onde se use escalas menores do que as que foram utilizadas. Outra possibilidade que pode explicá-la seria uma possível descontinuidade na derivada da curva $f(\alpha)$, e a distorção seria provocada por falta de precisão no método numérico no ponto de descontinuidade. Existe a conjectura de uma descontinuidade semelhante na curva

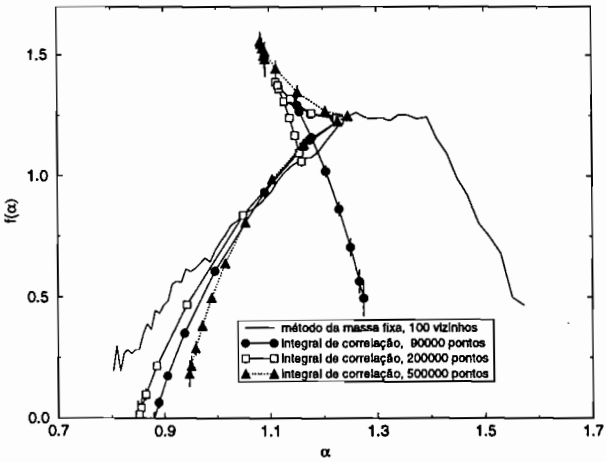


Figura 5.12: Método da integral de correlação. Curva $f(\alpha)$. Calculada para o atrator do mapa de Hénon. Observamos uma grande deformação na curva, de origem indeterminada.

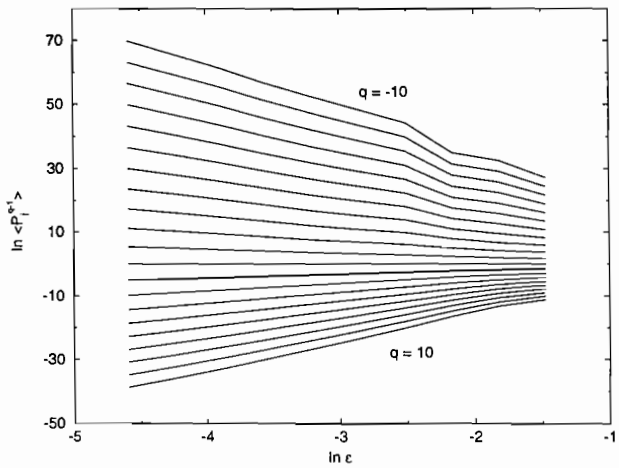


Figura 5.13: Método da Integral de correlação. Curvas $\ln < P_j^{q-1} >$. Calculado para o atrator do mapa de Hénon. De cima para baixo, estão as curvas correspondentes a $q = -10, -9, \ldots 10$. As curvas se apresentam quase lineares, com pouca flutuação.

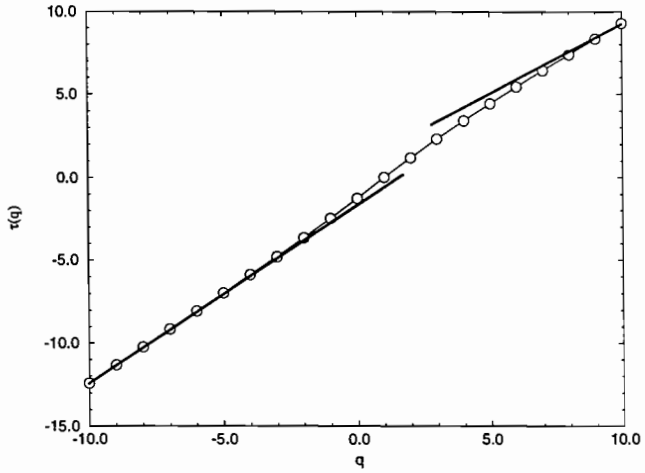


Figura 5.14: Método da Integral de correlação. Curva $\tau(q)$. Calculado para o atrator do mapa de Hénon. Os erros da curva são pequenos, da ordem do símbolo que representa os pontos. As retas mostradas evidenciam o ponto de inflexão na curva, que gera a distorção na curva $f(\alpha)$.

$\tau(q)$ do atrator do mapa de Hénon em $q \approx 2.3$ [17]. As curvas $f(\alpha)$ apresentadas nos capítulos posteriores parecem possuir características semelhantes.

Na literatura [8] encontram-se comparações de vários métodos para o cálculo de $f(\alpha)$, onde também se descreve a falta de precisão no cálculo do ramo $q < 0$ da curva utilizando métodos baseados na integral de correlação. Também não se encontram explicações para essa falta de precisão.

Em resumo, os resultados obtidos através do método da integral de correlação apresentam distorções no ramo $q < 0$ da curva $f(\alpha)$ em alguns dos conjuntos estudados. Tais distorções não são devidas à falta de pontos na representação do conjunto ou aos erros na regressão linear das curvas $\ln < P_j^{q-1} >$, uma vez que estas se apresentam com um comportamento próximo ao linear, mas devido a algum problema no próprio método, e talvez, na própria geometria do conjunto estudado.

Capítulo 6

Método canônico da integral de correlação

O Método canônico da integral de correlação [39] tem como partida a equação para $\tau(q)$ utilizada no método da integral de correlação, mas a transformação de Legendre é embutida nas equações como no método canônico. Assim, este é um método direto. Ele foi o primeiro método desenvolvido por nós, na tentativa de unir as vantagens observadas no método da integral de correlação (evitar as caixas espúrias) e no método canônico (ser um método direto).

Apresentamos a parte teórica do método, assim como um argumento para a criação do mesmo. Finalizamos o capítulo apresentando alguns resultados obtidos através da sua utilização.

6.1 Teoria

Observando com mais atenção os métodos encontrados na literatura, e comparando os métodos baseados em algoritmos de contagem de caixa e de classificação de distâncias, notamos que:

- O método microcanônico e o método da integral de correlação podem ser considerados análogos, pois ambos são métodos não-diretos.

- O método dos histogramas e o método da massa fixa possuem a construção de histogramas em comum.

Apenas o método canônico não encontra um método correspondente entre aqueles baseados na classificação de distâncias, e um método canônico baseado na integral de correlação uniria as principais vantagens de ambos os esquemas.

No método microcanônico [1] calculamos a curva $\tau(q)$, e através de uma transformação de Legendre obtemos a curva $f(\alpha)$. A única diferença entre ele e o método canônico [10], é que este fornece a curva $f(\alpha)$ diretamente, sem a necessidade da transformação de Legendre, que fica embutida no cálculo.

Usando o mesmo raciocínio, propusemos uma variação do método da integral de correlação [33] que fornece diretamente a curva $f(\alpha)$. Por essa razão o chamamos de método canônico da integral de correlação.

Partindo da expressão 1.44 para o cálculo da integral de correlação:

$$\tau(q) = \lim_{\epsilon \to 0} \frac{\ln \left\langle P_j^{q-1} \right\rangle}{\ln \epsilon} \tag{6.1}$$

A transformação de Legendre é dada pelas eq. 1.26 e 1.24:

$$\alpha(q) = \frac{d}{dq}\tau(q) \tag{6.2}$$

$$f(\alpha(q)) = \tau(q) - \alpha(q)q \tag{6.3}$$

Como apenas o numerador desta expressão depende de q, vamos denominá-lo *SomTau*. O denominador de $\alpha(q)$, que chamaremos de *SomAlfa*, é calculado segundo a eq. 1.26:

$$SomAlfa = \frac{d}{dq} SomTau = \frac{\left\langle P_j^{q-1} \ln P_j \right\rangle}{\left\langle P_j^{q-1} \right\rangle} \tag{6.4}$$

E o numerador da expressão para $f(\alpha(q))$, que chamaremos de *SomFalfa*, é calculado pela eq. 6.3.

$$SomFalfa = SomTau - q\, SomAlfa \tag{6.5}$$

6.2 Resultados Numéricos

O método canônico da integral de correlação se baseia nos mesmos argumentos teóricos do método da integral de correlação. A diferença é que este método obtém a curva $f(\alpha)$ sem realizar a transformação de Legendre da curva $\tau(q)$. É a mesma situação que ocorre ao compararmos o método microcanônico ao método canônico.

Espera-se que o método canônico forneça melhores resultados em cálculos onde há alguma descontinuidade na curva $f(\alpha)$ ou em sua primeira derivada, porque a transformação de Legendre envolve o cálculo de uma derivada numérica que elimina tais descontinuidades, distorcendo artificialmente o resultado.

Aplicamos este método aos Conjuntos de Cantor Generalizados, ao atrator do mapa quadrático no limiar da janela caótica e ao atrator do mapa de Hènon. Os resultados obtidos foram similares aos obtidos com o método da integral de correlação.

O resultado para o Conjunto de Cantor Generalizado com parâmetros $l_1 = 0.4$, $l_2 = 0.35$ e $p_1 = 0.6$ é mostrado na fig. 6.1. As escalas consideradas foram $\epsilon = 2^{i/2}$, com $i = 7, 8 \ldots 20$, para $q = 0, \pm 1, \pm 2 \ldots \pm 10$. Mostramos dois cálculos, que diferem apenas no número de pontos utilizado para representar o conjunto: 60 000 e 120 000.

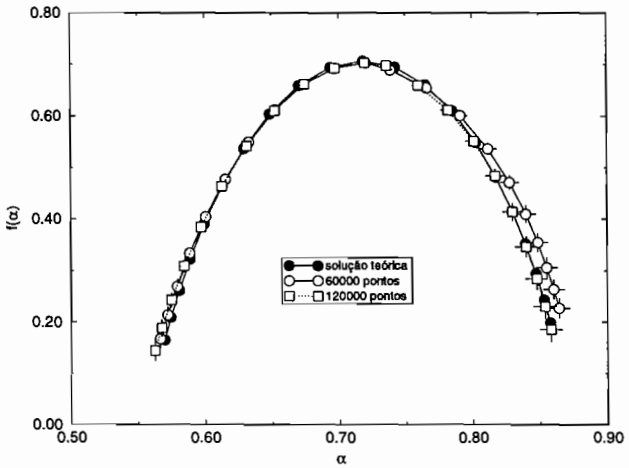


Figura 6.1: Método canônico da integral de correlação. Curva $f(\alpha)$. Calculada para o Conjunto de Cantor Generalizado cujos parâmetros são $l_1 = 0.4$, $l_2 = 0.35$ e $p_1 = 0.6$. Os resultado obtido representando o conjunto por 60 000 e 120 000 pontos está próximo do teórico.

Os gráficos das curvas $SomAlfa$ e $SomFalfa$ podem ser vistas nas figuras 6.2 e 6.3. Como pode-se observar, as curvas são compostas de flutuações ao redor de uma reta. Neste cálculo, foram utilizados 120 000 pontos para representar o conjunto.

No caso, as amostras utilizadas são todos os pontos do conjunto. Tentamos também realizar o cálculo com amostras menores, e obtivemos bons resultados para amostras maiores ou iguais a 5% neste Conjunto de Cantor Generalizado (fig. 6.4).

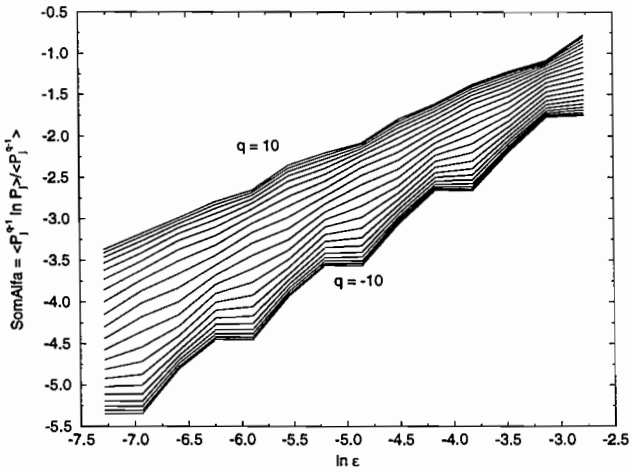


Figura 6.2: Método da Integral de correlação. Curvas $SomAlfa$. Calculada para o Conjunto de Cantor Generalizado cujos parâmetros são $l_1 = 0.4$, $l_2 = 0.35$ e $p_1 = 0.6$. De baixo para cima, estão as curvas correspondentes a $q = -10, -9, \ldots 10$. As curvas se apresentam quase lineares, com pouca flutuação.

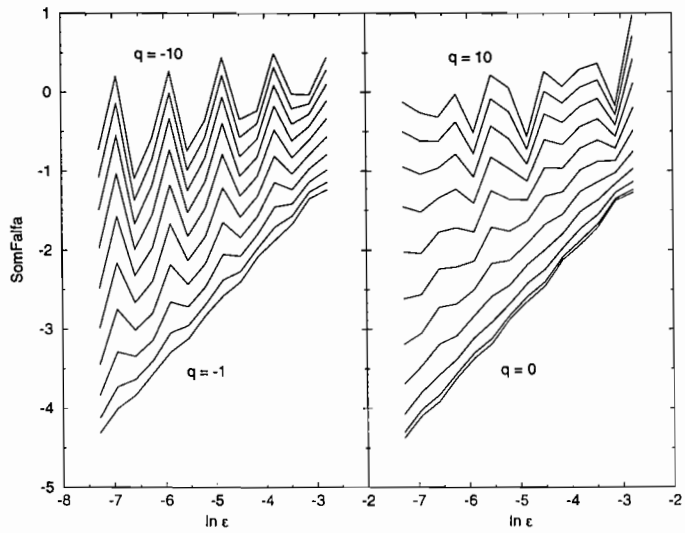


Figura 6.3: Método da Integral de correlação. Curvas $SomFalfa$. Calculada para o Conjunto de Cantor Generalizado cujos parâmetros são $l_1 = 0.4$, $l_2 = 0.35$ e $p_1 = 0.6$. À esquerda, se apresentam as curvas para $q < 0$. De cima para baixo, estão as curvas correspondentes a $q = -10, -9, \ldots -1$. À direita, se apresentam as curvas para $q \geq 0$. De baixo para cima, estão as curvas correspondentes a $q = 0, 1, \ldots 10$. As curvas se apresentam quase lineares, com pouca flutuação.

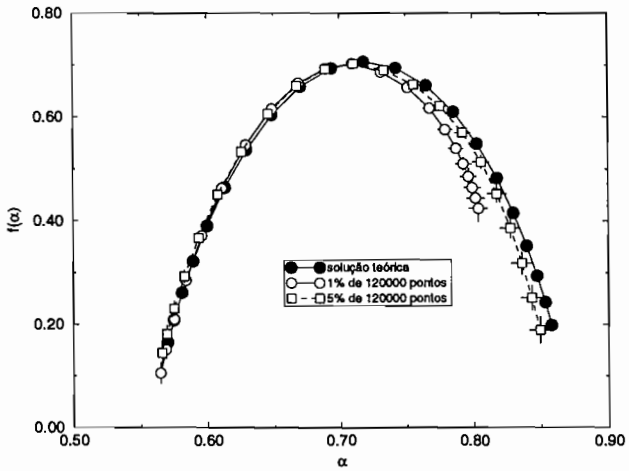


Figura 6.4: Método canônico da integral de correlação. Curva $f(\alpha)$. Calculada para o Conjunto de Cantor Generalizado cujos parâmetros são $l_1 = 0.4$, $l_2 = 0.35$ e $p_1 = 0.6$. O conjunto é representado por 120 000 pontos. Com uma amostra de 5% já se alcança um bom resultado. Utilizando amostras pequenas o tempo de processamento é menor, visto que ele é proporcional ao tamanho da amostra.

O resultado para o mapa quadrático no limiar da janela caótica $x_{i+1} = A(1 - x^2)$, com $A = 0.837005$ é mostrado na fig. 6.5. Há uma diferença no ramo $q < 0$ cuja origem não pudemos determinar. As escalas consideradas foram $\epsilon = 1.2 \times 2^{i/2}$, com $i = 7, 8 \ldots 20$, para $q = 0, \pm 1, \pm 2 \ldots \pm 10$. Mostramos dois cálculos, que diferem apenas no número de pontos utilizado para representar o conjunto: 50 000 e 100 000. Dessa forma, constatamos que a diferença observada na curva não deve ter origem na falta de pontos que representam o conjunto.

Os gráficos das curvas $SomAlfa$ e $SomFalfa$ podem ser vistas nas figuras 6.6 e 6.7. Como pode-se observar, as curvas são compostas de flutuações ao redor de uma reta. Neste cálculo, foram utilizados 100 000 pontos para representar o conjunto. As retas são razoavelmente bem definidas, de modo que não é possível atribuir à etapa de regressão linear o desvio encontrado na curva $f(\alpha)$ final.

Nesse cálculo, as amostras utilizadas são todos os pontos do conjunto. Tentamos também realizá-lo com amostras menores, e obtivemos bons resultados, mostrados na fig. 6.8, para amostras maiores ou iguais a 5%.

Apresentamos também o resultado para o atrator do mapa de Hènon, na fig. 6.9. Foram feitos cálculos utilizando 90 000, 200 000 e 500 000 pontos. As escalas utilizadas

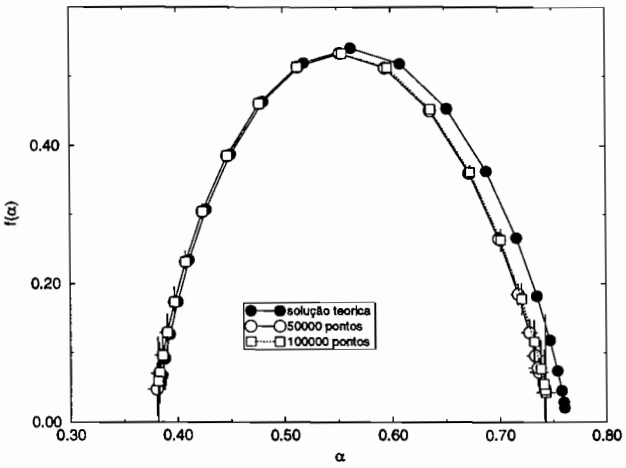


Figura 6.5: Método canônico da integral de correlação. Curva $f(\alpha)$. Calculada para o atrator do mapa quadrático no limiar da janela caótica $x_{i+1} = A(1-x^2)$, com $A = 0.837005$. O cálculo foi realizado com 50 000 e 100 000 pontos.

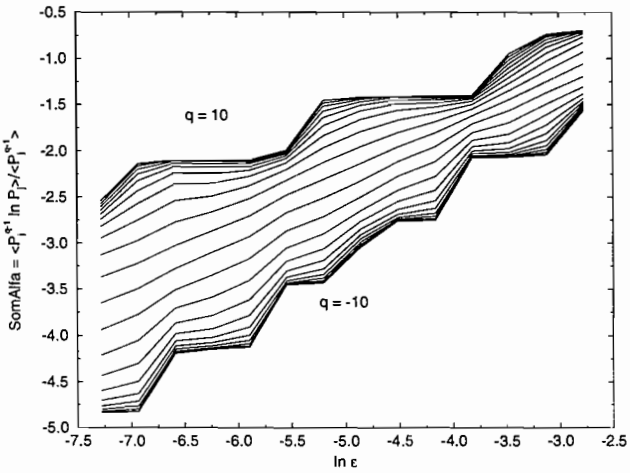


Figura 6.6: Método da Integral de correlação. Curvas $SomAlfa$. Calculada para o atrator do mapa quadrático no limiar da janela caótica $x_{i+1} = A(1 - x^2)$, com $A = 0.837005$. De baixo para cima, estão as curvas correspondentes a $q = -10, -9, \ldots 10$. As curvas se apresentam quase lineares, com pouca flutuação.

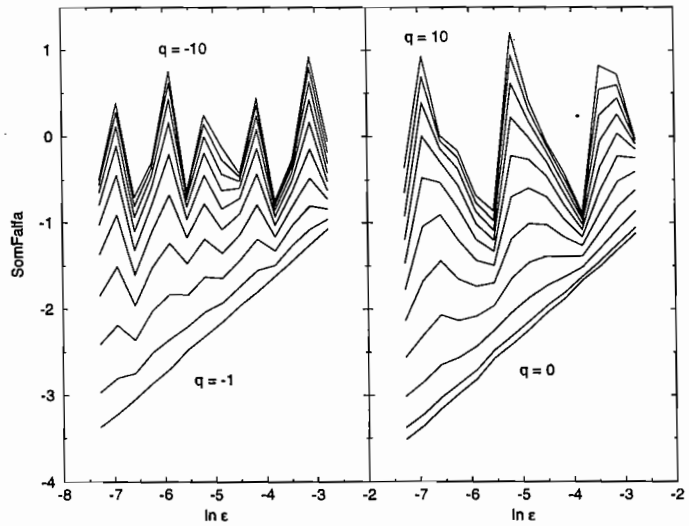


Figura 6.7: Método da Integral de correlação. Curvas $SomFalfa$. Calculada para o atrator do mapa quadrático no limiar da janela caótica $x_{i+1} = A(1 - x^2)$, com $A = 0.837005$. À esquerda, se apresentam as curvas para $q < 0$. De cima para baixo, estão as curvas correspondentes a $q = -10, -9, \ldots -1$. À direita, se apresentam as curvas para $q \geq 0$. De baixo para cima, estão as curvas correspondentes a $q = 0, 1, \ldots 10$. As curvas se apresentam quase lineares, com pouca flutuação.

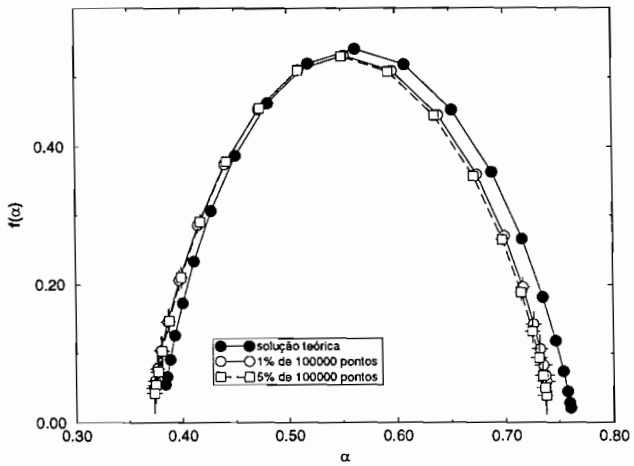


Figura 6.8: Método canônico da integral de correlação. Curva $f(\alpha)$. Calculada para o atrator do mapa quadrático no limiar da janela caótica $x_{i+1} = A(1-x^2)$, com $A = 0.837005$. O conjunto é representado por 100 000 pontos. Com uma amostra de 5% já se alcança um bom resultado. Utilizando amostras pequenas o tempo de processamento é menor, visto que ele é proporcional ao tamanho da amostra.

foram $\epsilon = 2.6 \times 2^{i/2}$, com $i = 7, 8 \ldots 20$, para $q = 0, \pm 1, \pm 2 \ldots \pm 10$.

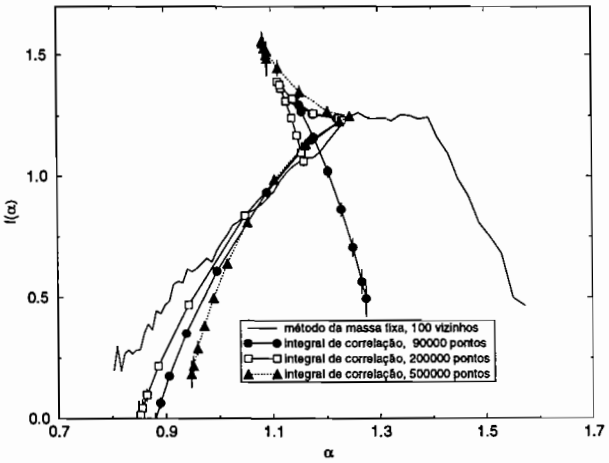


Figura 6.9: Método canônico da integral de correlação. Curva $f(\alpha)$. Calculada para o atrator do mapa de Hènon. Há uma grande deformação no ramo $q < 0$ da curva, cuja origem não pudemos determinar.

Uma grande deformação na curva é observada. Não pudemos determinar a origem dessa deformação, mas podemos observar que ela não parece diminuir com o aumento do número de pontos que representa o conjunto.

Os gráficos das curvas $SomAlfa$ e $SomFalfa$ podem ser vistas nas figuras 6.10 e 6.11. Neste cálculo, foram utilizados 500 000 pontos para representar o conjunto. Como pode-se observar, as curvas são compostas de flutuações ao redor de uma reta, de modo que a deformação da curva $f(\alpha)$ final não pode ser explicada por erros durante a etapa de regressão linear.

Note que todos os resultados são equivalentes aos obtidos com o método da integral de correlação, o que é esperado, uma vez que eles possuem o mesmo princípio teórico.

Comparando aos dois métodos baseados em algoritmos de contagem de caixa, o método canônico e o método microcanônico, esses possuem resultados que diferem mais entre si do que os obtidos através de métodos baseados na integral de correlação. Essa diferença pode ser causada por flutuações que ocorrem nas curvas submetidas à regressão linear que levam a pequenos desvios que também podem ser ampliados pela tranformação de Legendre. Em cálculos onde a classificação de distâncias é usada, essas flutuações são menores, o que levaria a resultados mais semelhantes. Essa semelhança confirma a hipótese que de os cálculos feitos através de classificações de

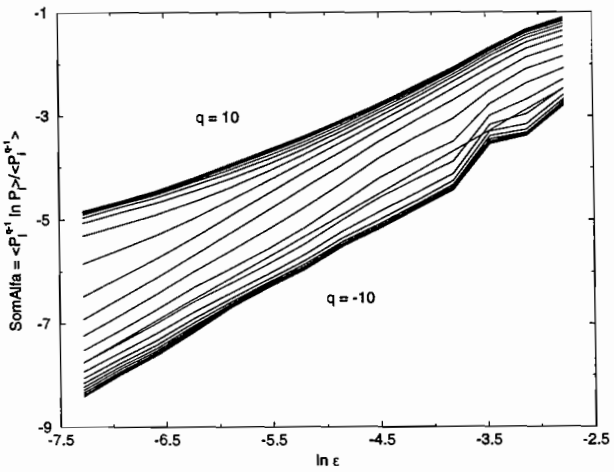


Figura 6.10: Método da Integral de correlação. Curvas $SomAlfa$. Calculada para o atrator do mapa de Hènon. À esquerda, se apresentam as curvas para $q < 0$. De baixo para cima, estão as curvas correspondentes a $q = -10, -9, \ldots 10$. As curvas se apresentam quase lineares, com pouca flutuação.

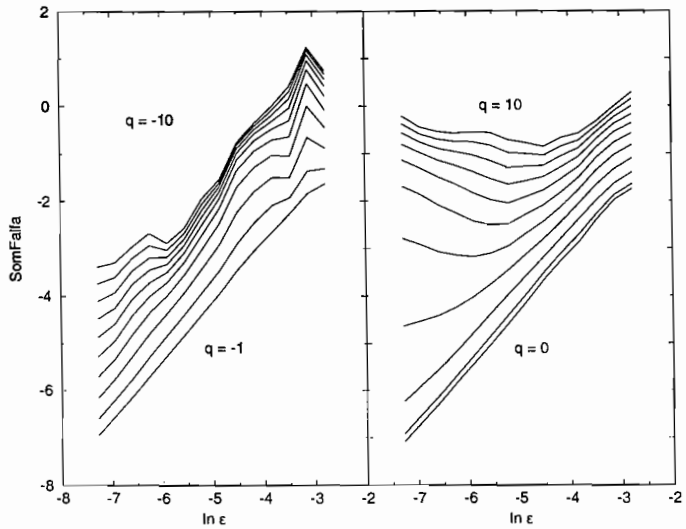


Figura 6.11: Método da Integral de correlação. Curvas $SomFalfa$. Calculada para o atrator do mapa de Hènon. De cima para baixo, estão as curvas correspondentes a $q = -10, -9, \ldots -1$. À direita, se apresentam as curvas para $q \geq 0$. De baixo para cima, estão as curvas correspondentes a $q = 0, 1, \ldots 10$. As curvas se apresentam quase lineares, com pouca flutuação.

distâncias são mais robustos do que os feitos através de algoritmos de contagem de caixa.

Capítulo 7

Um novo algoritmo de contagem de caixa

Nos algoritmos de contagem de caixa usuais o recobrimento é realizado utilizando uma grade. Como discutimos anteriormente, este recobrimento produz caixas espúrias. Neste capítulo propomos um novo algoritmo de contagem de caixa, que reduz sensivelmente o número de caixas espúrias, melhorando os resultados obtidos nos métodos de cálculo de $f(\alpha)$ baseados em algoritmos de contagem de caixa.

7.1 Introdução

O algoritmo que vamos apresentar é um aprimoramento do algoritmo de Oiwa e Ferrara [30] para se construir um recobrimento de um conjunto dado. Nele, os recobrimentos ainda são constituidos de caixas de mesmo tamanho e orientação no espaço, mas não necessariamente se organizam em uma grade. Dessa maneira, elimina-se um grande número de caixas espúrias.

Mas o algoritmo original divide o conjunto em partes isoladas, tratando cada uma como um novo conjunto, perdendo a noção da geometria do conjunto inicial como um todo. Essa característica faz com que ainda existam caixas espúrias no recobrimento. Propomos uma modificação que localiza as caixas espúrias que restam e que tenta retirá-las movimentando um grupo de caixas vizinhas a elas. Dessa maneira, se recupera uma visão mais global do conjunto e se elimina um número bem maior de caixas espúrias. Exemplos de recobrimentos obtidos utilizando esse algoritmo são mostrados na fig. 7.1.

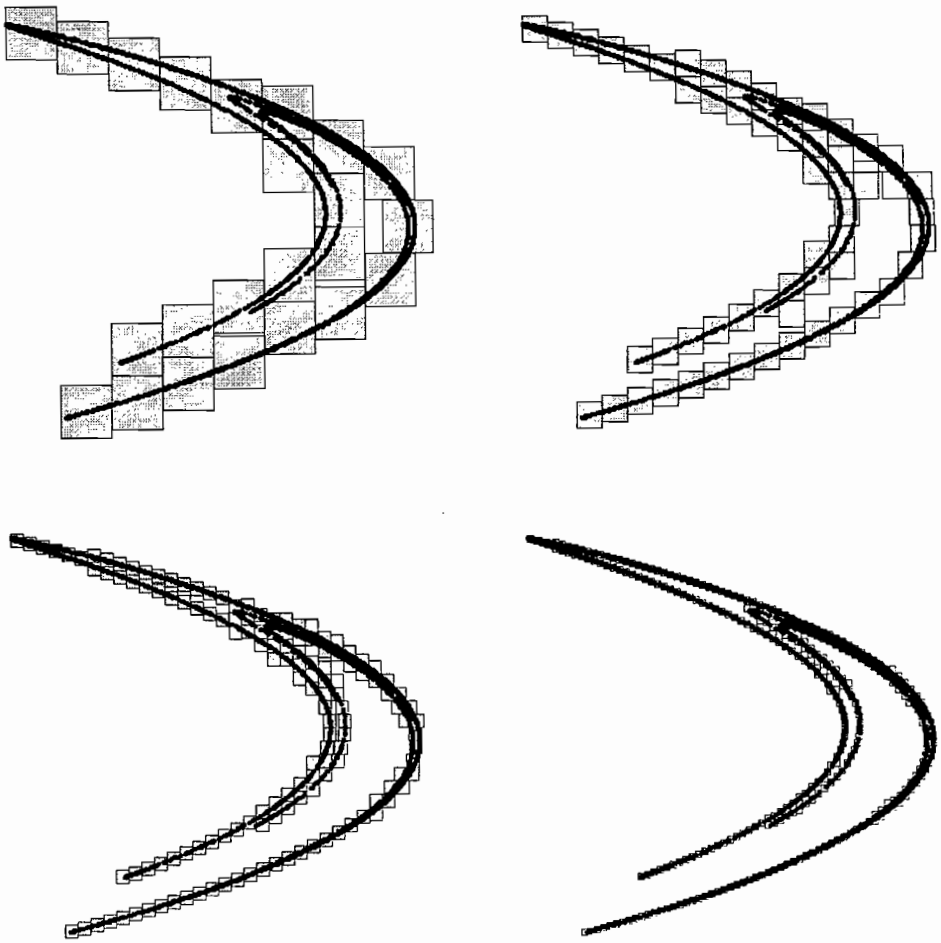


Figura 7.1: Recobrimentos do atrator de Hénon obtidos utilizando o algoritmo de contagem de caixa que propomos, para diferentes tamanhos de caixa.

Dessa maneira, o algoritmo é divido em duas partes: Na primeira os recobrimentos do conjunto são construídos. Esta parte consiste no algoritmo de Oiwa em si. Na segunda parte, caixas espúrias são eliminadas do recobrimento a partir do movimento de caixas vizinhas a elas.

O algoritmo de Oiwa se inicia considerando um recobrimento de apenas uma única caixa que contém todo o conjunto. Um novo recobrimento é gerado a partir deste, e assim, sucessivamente, novos recobrimentos são formados a partir dos antigos. O novo recobrimento é formado por caixas com a metade do tamanho do recobrimento anterior. As caixas em cada recobrimento possuem o mesmo tamanho e orientação no espaço, mas não pertencem necessariamente a uma grade.

Na segunda parte do algoritmo, tentamos melhorar o recobrimento eliminando caixas espúrias do mesmo. As caixas espúrias são aquelas que possuem seu interior quase totalmente preenchido pelos vazios da estrutura do multifractal. Dessa forma, muitas vezes é possível removê-las com deslocamentos de caixas vizinhas a ela, de tal forma que as condições da existência de um recobrimento não sejam violadas.

Esta parte do algoritmo pode ser dividida em duas etapas: Na primeira, localiza-se uma caixa no recobrimento que contenha uma parte do conjunto de tal forma que ocupe menos da metade de seu tamanho. Essa provavelmente é uma caixa espúria. Tenta-se construir um grupo caixas vizinhas que ao ser deslocado elimine a caixa espúria. Na segunda etapa dessa parte do algoritmo, tenta-se efetivamente mover esse grupo de caixas. Esse movimento não pode violar as condições para a existência do recobrimento, que são:

- As caixas do recobrimento não podem se sobrepor.
- As caixas do recobrimento devem conter todos os pontos do conjunto estudado.

O algoritmo de Oiwa é simples de ser aplicado em sistemas de qualquer dimensão. A segunda parte do algoritmo, porém, depende da dimensão do espaço considerada quando se define o grupo a ser deslocado. Nós consideramos sistemas de até duas dimensões apenas.

O grupo de caixas que consideramos nesse algoritmo é bem simples, mas muito comum, de modo que um grande número de caixas espúrias é removido. Cada caixa do grupo é vizinha de duas caixas do grupo, exceto as duas nas extremidades. Além disso, todas são vizinhas na mesma componente do espaço. As duas caixas nas extremidades do grupo devem possuir espaço vazio adjacente ao lado que não é vizinho a caixas

do grupo, que somado, deve ser maior que o tamanho da caixa. Dessa forma pode ser possível eliminar uma caixa do grupo através do seu deslocamento. Um exemplo desse grupo é mostrado na fig. 7.2.

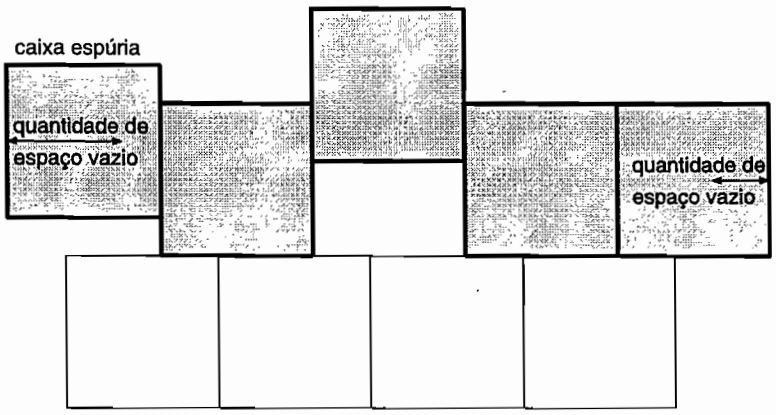


Figura 7.2: Exemplo de grupo construido pelo algoritmo para eliminar caixas do recobrimento. O grupo de caixas está desenhado em cinza. Através de um deslocamento do grupo pode ser possível eliminar a caixa indicada como espúria.

No espaço bidimensional existem outros tipos de grupos que não consideramos. Dessa forma, ainda restam algumas caixas espúrias que poderiam ser eliminadas utilizando-se de um deslocamento de caixas vizinhas. Um exemplo de grupo que não é considerado pelo algoritmo é mostrado na fig. 7.3. A caixa espúria poderia ser removida com o deslocamento indicado.

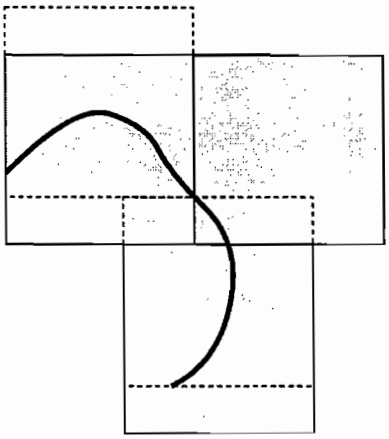


Figura 7.3: Exemplo de grupo (em cinza) não considerado pelo algoritmo para efetuar a eliminação de caixas do recobrimento. A caixa poderia ser eliminada através do deslocamento indicado em linhas pontilhadas.

Esse algoritmo está descrito em detalhe no ap. L.

7.2 Resultados

Comparamos o resultado obtido com nosso algoritmo no cálculo de $f(\alpha)$ através de métodos que usam contagem de caixa com os resultados anteriores apresentados no Capítulo 4. Sabemos que o recobrimento gerado pelo nosso algoritmo não é totalmente livre de caixas espúrias, por isso, consideramos diferentes orientações das caixas para construir o recobrimento. Utilizamos os critérios de ínfimos ou supremos e de médias sobre os diferentes recobrimentos para o cálculo da curva $f(\alpha)$. Vale ressaltar que, mesmo utilizando essas correções, podem restar algumas caixas espúrias no recobrimento.

Na utilização do algoritmo usual de contagem de caixa, a aplicação do critério de ínfimos e supremos fornece o melhor resultado. Com a utilização do novo algoritmo, não podemos prever qual o critério a fornecer o melhor resultado.

7.2.1 Método Microcanônico

Aplicamos o algoritmo de contagem de caixa usual e o nosso, para efeito de comparação, no método microcanônico, aos conjuntos que estudamos. De maneira geral, os resultados obtidos utilizando o novo algoritmo são melhores do que aqueles calculados através do algoritmo de contagem de caixa usual.

Inicialmente estudamos o Conjunto de Cantor Generalizado, definido na sec.2.1. O resultado obtido é mostrado na fig. 7.4. O conjunto foi representado por 500 000 pontos, e os parâmetros do método são $q = 0, \pm 1, \pm 2, \ldots 10$. $\epsilon = 2^{-i}, i = 1, 2, \ldots 10$. 51 deslocamentos de grade são utilizados no algoritmo usual, e 51 rotações no algoritmo proposto dentro de um ângulo de 45°. Foi utilizado o critério de médias nos deslocamentos.

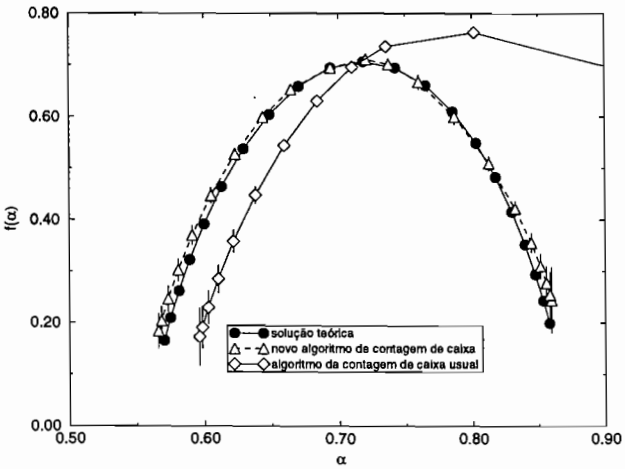


Figura 7.4: Método microcanônico, utilizando o novo algoritmo de contagem de caixa. Curva $f(\alpha)$: Conjunto de Cantor Generalizado com parâmetros $l_1 = 0.4$, $l_2 = 0.35$, e $p_1 = 0.6$. Foi utilizado o critério de médias. O resultado obtido com a utilização do algoritmo proposto é melhor do que o encontrado com o algoritmo usual e é muito próxima do resultado esperado teoricamente. A curva realizada com o algoritmo usual é a mesma da fig. 4.3.

O resultado obtido utilizando o critério de médias é melhor do que o obtido através do critério de ínfimos e supremos, que mostramos na fig. 7.5. Observando as curvas cujas inclinações geram os pontos da curva, mostradas na fig. 7.6, nota-se que utilizando o critério de médias ou o critério de ínfimos e supremos as inclinações das

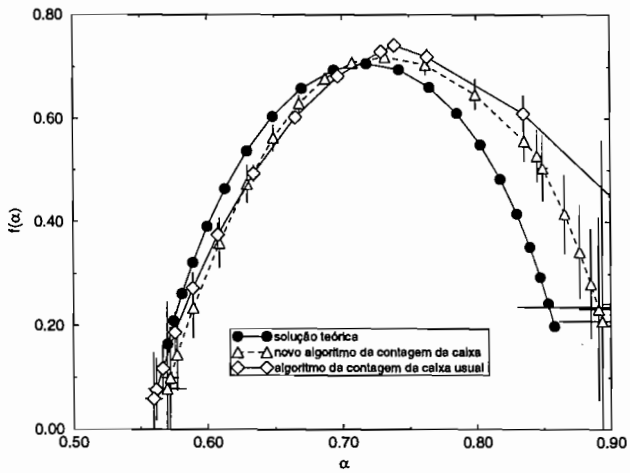


Figura 7.5: Método microcanônico, utilizando o novo algoritmo de contagem de caixa. Curva $f(\alpha)$: Conjunto de Cantor Generalizado com parâmetros $l_1 = 0.4$, $l_2 = 0.35$, e $p_1 = 0.6$. Foi utilizado o critério de ínfimos e supremos. O resultado obtido com a utilização do algoritmo proposto é melhor do que o encontrado com o algoritmo usual, mas não está muito próximo do esperado teoricamente. O resultado obtido utilizando o critério de médias é melhor. A curva realizada com o algoritmo usual é a mesma da fig. 4.3.

curvas se aproximam muito do valor esperado teoricamente. Essa é uma forte indicação de que o desvio produzido pelas caixas espúrias foi eliminado. Essa diferença

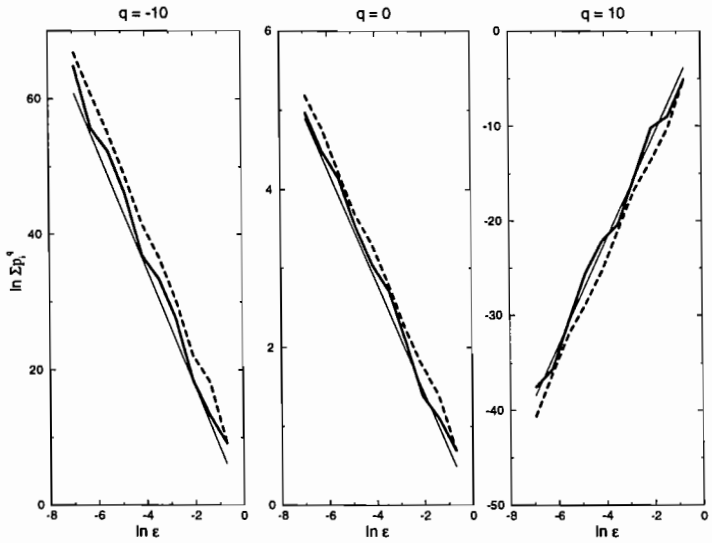


Figura 7.6: Método microcanônico. Curvas de $SomTau = \ln \sum p_i^q$: Conjunto de Cantor Generalizado com parâmetros $l_1 = 0.4$, $l_2 = 0.35$, e $p_1 = 0.6$. As curvas obtidas com ambos critérios possuem uma inclinação muito próxima da esperada teoricamente, indicando que a distorção produzida pelas caixas espúrias foi eliminada. Este resultado pode ser comparado ao da fig. 4.4, onde o algoritmo usual de contagem de caixa é utilizado.
linha sólida - critério de ínfimos e supremos
linha pontilhada - critério de médias

pode indicar que restam outros fatores a serem considerados nos recobrimentos que geram flutuações ao redor da média. Nesse caso, a utilização do critério de ínfimos e supremos pode não ser apropriada.

O resultado para o atrator do mapa quadrático, descrito na sec. 2.2, está apresentado na fig. 7.7. O conjunto foi representado por 700 000 pontos, e os parâmetros do método são $q = 0, \pm 1, \pm 2, \ldots 10$. $\epsilon = 1.2 \times 2^{-i}, i = 1, 2, \ldots 10$. 51 deslocamentos de grade são utilizados no algoritmo usual, e 51 rotações no algoritmo proposto dentro de um ângulo de 45°. Foi utilizado o critério de ínfimo ou supremo nos deslocamentos. Observa-se que o resultado utilizando este critério é melhor que o resultado obtido usando o critério de médias, mostrado na fig. 7.8.

A utilização do critério de ínfimos e supremos neste conjunto produziu um resultado próximo do teórico mesmo utilizando o algoritmo usual. Podemos ver que a utilização do algoritmo proposto gerou uma curva $f(\alpha)$ ainda mais próxima do

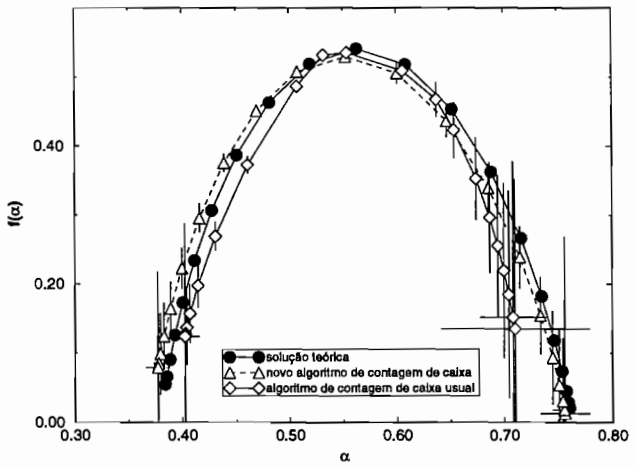


Figura 7.7: Método microcanônico, utilizando o novo algoritmo de caixa. Curva $f(\alpha)$: Atrator do mapa quadrático $x_{i+1} = \lambda(1-2x_i^2)$ com $\lambda = 0.837005134$. Foi utilizado o critério de ínfimos e supremos. Comparamos o algoritmo de caixa usual com o novo algoritmo que propomos. O resultado com o novo algoritmo praticamente coincide com o resultado teórico. A curva realizada com o algoritmo usual é a mesma da fig. 4.6.

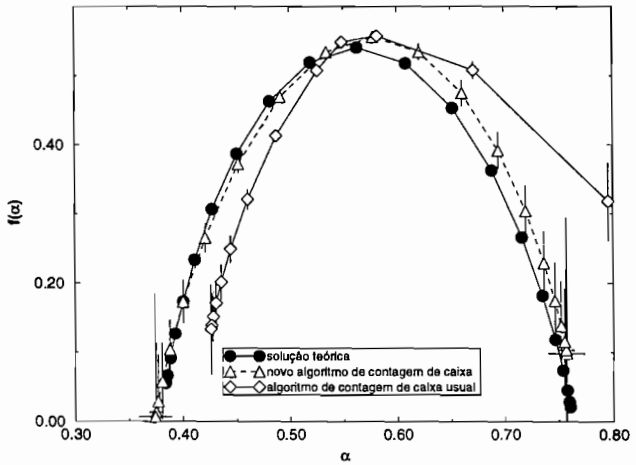


Figura 7.8: Método microcanônico, utilizando o novo algoritmo de contagem de caixa. Curva $f(\alpha)$: Atrator do mapa quadrático $x_{i+1} = \lambda(1 - 2x_i^2)$ com $\lambda = 0.837005134$. Foi utilizado o critério de médias. Comparamos o algoritmo de caixa usual com o novo algoritmo que propomos. O resultado com o novo algoritmo se aproxima do resultado esperado teoricamente, embora o resultado obtido utilizando o critério de ínfimos e supremos seja melhor. A curva realizada com o algoritmo usual é a mesma da fig. 4.6.

calculado teoricamente.

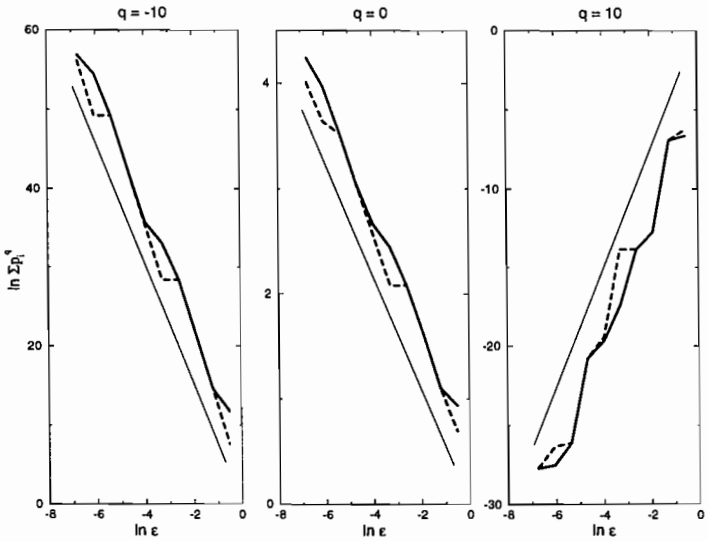


Figura 7.9: Método microcanônico. Curvas de $SomTau = \ln \sum p_i^q$: Atrator do mapa quadrático $x_{i+1} = \lambda(1 - 2x_i^2)$ com $\lambda = 0.837005134$. Com a utilização do novo algoritmo, observa-se que utilizando tanto o critério de médias quanto o critério de ínfimos e supremos obtém-se curvas com inclinação muito próxima da teórica. Essas curvas podem ser comparadas ao da fig. 4.7, onde o algoritmo usual foi utilizado.
linha sólida - critério de ínfimos e supremos
linha pontilhada - critério de médias

As curvas $SomTau = \ln \sum p_i^q$ obtidas através da utilização do critério de ínfimos e supremos para $q < 0$ possuem uma inclinações muito próximas à esperada teoricamente, como pode ser observado na fig. 7.9. Esta é uma indicação de que a distorção das caixas espúrias foi eliminada. No caso, ambos critérios geraram curvas com flutuações. Essas flutuações são provocadas pela escolha das escalas consideradas nos recobrimentos. A periodicidade do comportamento das curvas sugere que elas seriam praticamente lineares se as escalas variassem de um fator de 16 ao invés de um fator de 2.

A seguir testamos o algoritmo no atrator do mapa de Hénon, descrito na sec. 2.3 O conjunto foi representado por uma série de 1 200 000 pontos. Os parâmetros usados no método são $q = 0, \pm 1, \pm 2, \ldots 10$. $\epsilon = 2.6 \times 2^{-i}, i = 1, 2, \ldots 8$. Novamente, 97 deslocamentos de grade são realizados no algoritmo usual e 97 rotações em um ângulo de 90°. O resultado, apresentado na fig. 7.10, obtido com o novo algoritmo, se aproxima do encontrado na literatura. Foi utilizado o critério de ínfimos e supremos

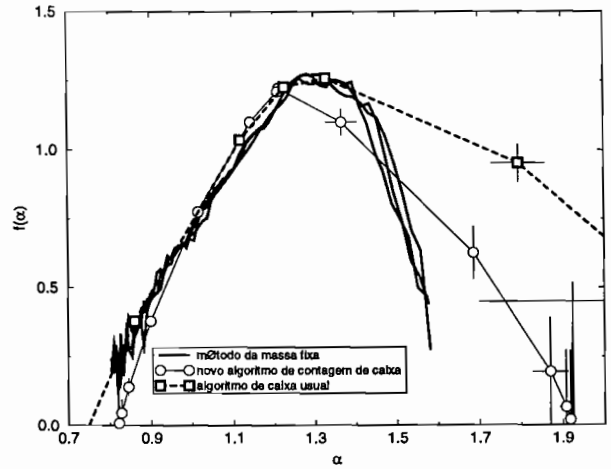


Figura 7.10: Método microcanônico, utilizando o novo algoritmo de caixa. Curva $f(\alpha)$: Atrator do mapa de Hénon. O resultado obtido utilizando o novo algoritmo com o critério de ínfimos e supremos está um pouco mais próximo do calculado através do método da massa fixa. A curva obtida utilizando o algoritmo de contagem de caixa usual é a mesma apresentada na fig. 4.9.

nos deslocamentos. A curva foi comparada à obtida através do método da massa fixa, uma vez que não há previsão teórica para a mesma neste caso. Podemos ver que o resultado, mostrado na fig. 7.11, possui um ramo $q < 0$ um pouco mais próximo do resultado obtido no método da massa fixa, mas ainda está muito distorcido. As curvas $SomTau = \ln \sum p_i^q$ obtidas podem ser observadas na fig. 7.12. Vemos que para $q = -10$, o melhor valor (ínfimo) foi alcançado pelo uso do novo algoritmo e do critério de ínfimos e supremos.

É interessante notar que existe uma semelhança entre essas curvas e aquelas obtidas através dos métodos de integral de correlação. No mesmo ponto onde foi observada uma deformação nos cálculos da integral de correlação, parece existir uma descontinuidade na derivada da curva $f(\alpha)$ nas curvas obtidas com o novo algoritmo de contagem de caixa.

Essa semelhança entre os cálculos pode ser devida a alguma propriedade específica do atrator do mapa de Hénon, mas poderia ter outras causas, como caixas espúrias ainda remanescentes. Descartamos a falta de pontos na representação do atrator, pois observamos que o resultado melhora com a diminuição do número de pontos utilizando do critério de ínfimos e supremos, o que é o oposto do usualmente espe-

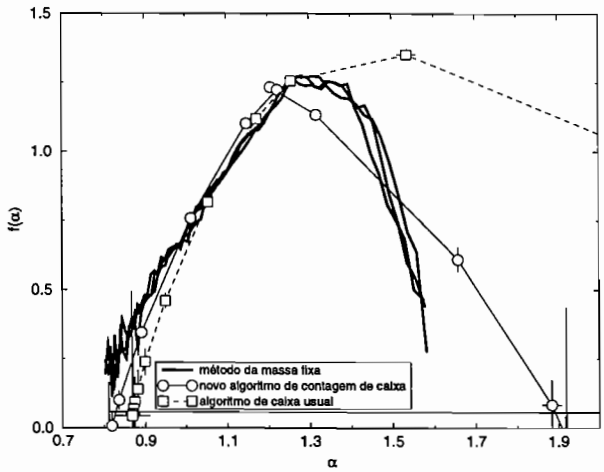


Figura 7.11: Método microcanônico, utilizando o novo algoritmo de caixa. Curva $f(\alpha)$: Atrator do mapa de Hénon. Foi utilizado o critério de médias. A utilização do critério de ínfimos e supremos gerou uma curva semelhante à calculada com esse critério. A curva obtida utilizando o algoritmo de contagem de caixa usual é a mesma apresentada na fig. 4.9.

rado. Mostramos um cálculo onde o conjunto foi representado por 200 000 pontos na fig. 7.13.

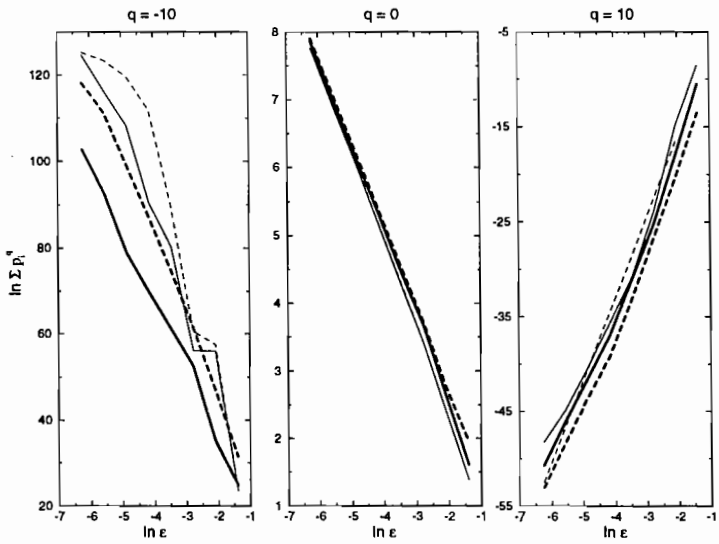


Figura 7.12: Método microcanônico. Curvas de $SomTau = \ln \sum p_i^q$: Atrator do mapa de Hénon. Os valores para $q < 0$ são os mais sensíveis. O melhor resultado foi obtido através do critério de ínfimos e supremos e utilizando o novo algoritmo (valores mais baixos).
linha sólida fina - algoritmo de contagem de caixa usual, critério de ínfimos e supremos
linha pontilhada fina - algoritmo de contagem de caixa usual, critério de médias
linha sólida grossa - algoritmo de contagem de caixa proposto, critério de ínfimos e supremos
linha pontilhada grossa - algoritmo de contagem de caixa proposto, critério de médias

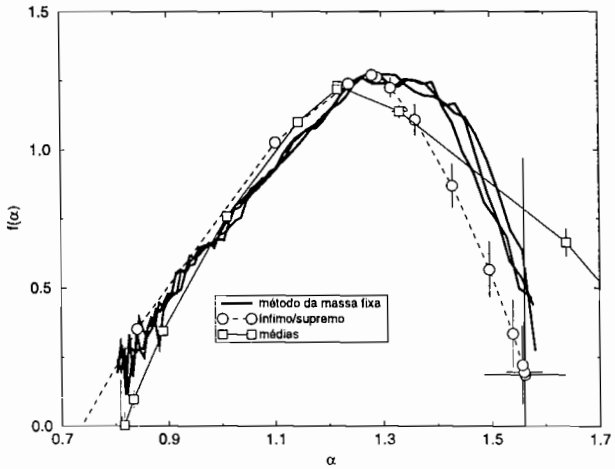


Figura 7.13: Método microcanônico, utilizando o novo algoritmo de caixa. Curva $f(\alpha)$: Atrator do mapa de Hénon. Diminuindo o número de pontos, há uma melhora no ramo $q < 0$ da curva quando se utiliza o critério de ínfimos e supremos, o que não é o comportamento esperado com essa variação.

7.2.2 Método Canônico

Aplicamos o nosso algoritmo de contagem de caixa ao método canônico, descrito na sec. 1.5.2, e o comparamos ao algoritmo de caixa usual.

Aplicamos o método ao Conjunto de Cantor Generalizado com parâmetros $l_1 = 0.4$, $l_2 = 0.35$, e $p_1 = 0.6$. O conjunto foi representado por uma série de 200 000 pontos. Os parâmetros usados no método são $q = 0, \pm 1, \pm 2, \ldots 10$. $\epsilon = 2^{-i}, i = 1, 2, \ldots 10$. Novamente, 51 deslocamentos de grade são realizados no algoritmo usual e 51 rotações em um ângulo de 90°. Na fig. 7.14 foi utilizado o critério de médias entre deslocamentos, onde foi obtido um resultado melhor do que com a aplicação do critério de infimos e supremos, cujo resultado é apresentado na fig. 7.15.

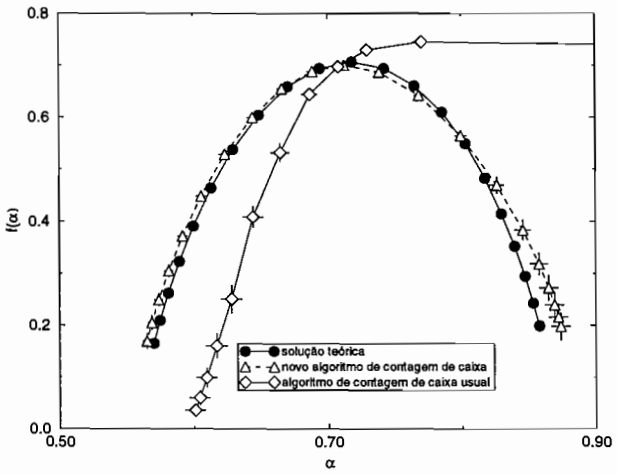


Figura 7.14: Método canônico, utilizando o novo algoritmo de caixa. Curva $f(\alpha)$: Conjunto de Cantor Generalizado com parâmetros $l_1 = 0.4$, $l_2 = 0.35$, e $p_1 = 0.6$. Foi utilizado o critério de médias. Comparamos o algoritmo de caixa usual com o novo algoritmo que propomos. O resultado obtido com o novo algoritmo é bem mais próximo dos valores esperados teoricamente. A curva onde foi utilizado o algoritmo usual é a mesma apresentada na fig. 4.11.

As respectivas curvas $SomAlfa = \sum \mu_i \ln p_i$ e $SomFalfa = \sum \mu_i \ln \mu_i$ são mostradas nas figs. 7.16, 7.17. Principalmente nas primeiras nota-se que em ambos os critérios as inclinações são muito próximas das esperadas teoricamente. De maneira semelhante ao método microcanônico, as curvas geradas utilizando o critério de médias possuem flutuações menores do que as obtidas utilizando o critério de

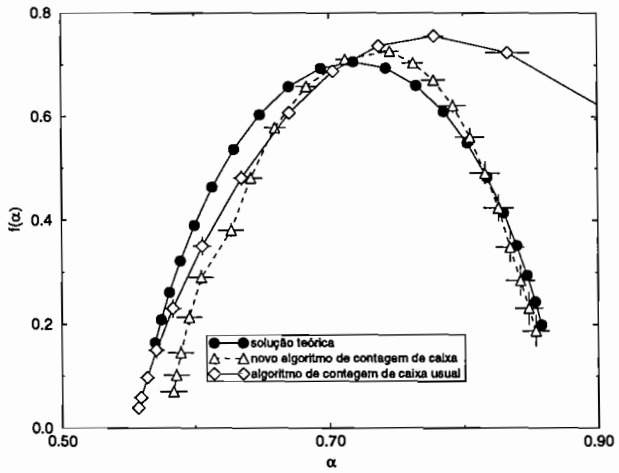


Figura 7.15: Método canônico, utilizando o novo algoritmo de caixa. Curva $f(\alpha)$: Conjunto de Cantor Generalizado com parâmetros $l_1 = 0.4$, $l_2 = 0.35$, e $p_1 = 0.6$. Foi utilizado o critério de ínfimos e supremos. Comparamos o algoritmo de caixa usual com o novo algoritmo que propomos. O resultado obtido com o novo algoritmo é mais próximo dos valores esperados teoricamente, mas o melhor resultado foi obtido através da utilização do critério de médias. A curva onde foi utilizado o algoritmo usual é a mesma apresentada na fig. 4.11.

ínfimos e supremos. Para as curvas $SomFalfa$, deve-se lembrar que para módulo de q grande, seus valores e suas inclinações são muito pequenos, de modo que as flutuações se destacam.

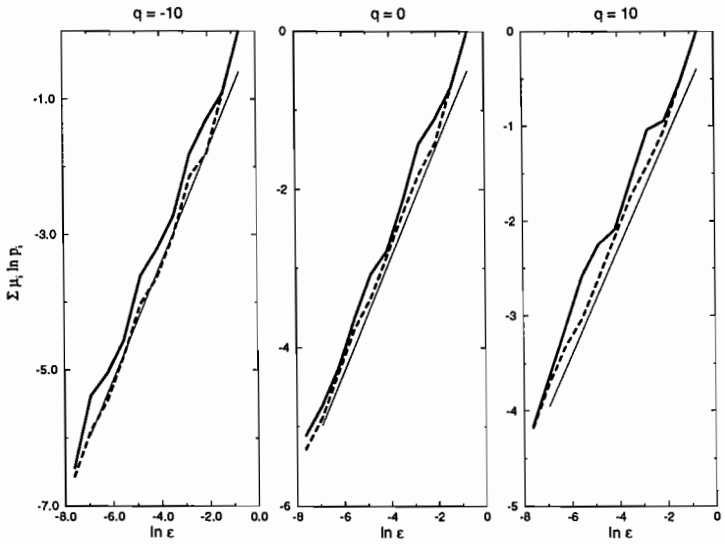


Figura 7.16: Método canônico. Curvas de $SomAlfa = \sum \mu_i \ln p_i$: Conjunto de Cantor Generalizado com parâmetros $l_1 = 0.4$, $l_2 = 0.35$, e $p_1 = 0.6$. Com a utilização do novo algoritmo, observa-se que utilizando tanto o critério de médias quanto o critério de ínfimos e supremos obtém-se curvas com inclinação muito próxima da teórica. Essas curvas podem ser comparadas àquelas apresentadas na fig. 4.12, onde o algoritmo usual foi utilizado.
linha sólida - critério de ínfimos e supremos
linha pontilhada - critério de médias

Resumindo os resultados obtidos aplicando o novo algoritmo de contagem de caixa ao Conjunto de Cantor Generalizado, como o método canônico e o método microcanônico possuem o mesmo princípio teórico, diferindo apenas no fato do método canônico ser um método direto e o método microcanônico em método indireto, em geral, há uma expectativa de que o resultado gerado pelo método canônico seja mais preciso exatamente por não estar sujeito a etapas como o cálculo da derivada numérica, presente no outro método. Apesar disso, para o Conjunto de Cantor Generalizado apresentado, o melhor resultado foi obtido utilizando o método microcanônico.

O resultado para o atrator do mapa quadrático, descrito na sec. 2.2, está apresentado na fig. 7.18. O conjunto foi representado por 700 000, e os parâmetros do método são $q = 0, \pm 1, \pm 2, \ldots 10$. $\epsilon = 1.2 \times 2^{-i}, i = 1, 2, \ldots 10$. 51 deslocamentos de grade são utilizados no algoritmo usual, e 51 rotações no algoritmo proposto dentro

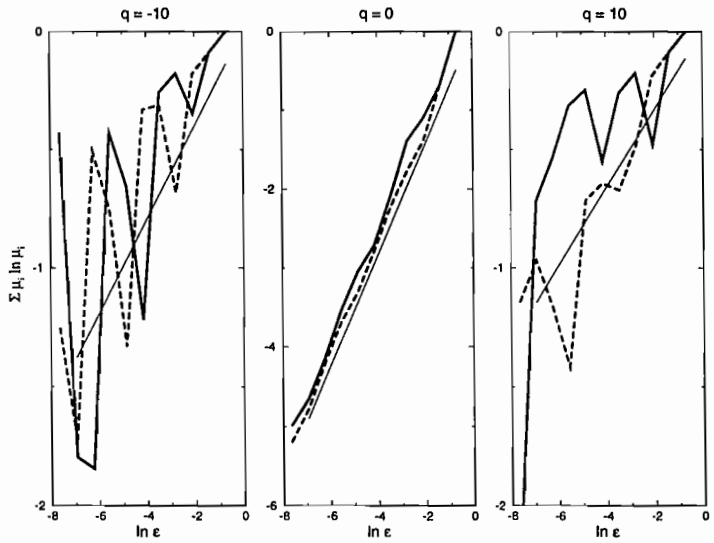


Figura 7.17: Método canônico. Curvas de $SomFalfa = \sum \mu_i \ln \mu_i$: Conjunto de Cantor Generalizado com parâmetros $l_1 = 0.4$, $l_2 = 0.35$, e $p_1 = 0.6$. Com a utilização do novo algoritmo, observa-se que utilizando tanto o critério de médias quanto o critério de ínfimos e supremos obtém-se curvas com inclinação próxima da teórica, apesar da flutuação existente. Essas curvas podem ser comparadas àquelas apresentadas na fig. 4.13, onde o algoritmo usual foi utilizado.

linha sólida - critério de ínfimos e supremos

linha pontilhada - critério de médias

de um ângulo de 45°. Foi utilizado o critério de médias em diferentes recobrimentos, cuja aplicação gerou um resultado melhor do que a aplicação do critério de ínfimos e supremos, mostrado na fig. 7.19.

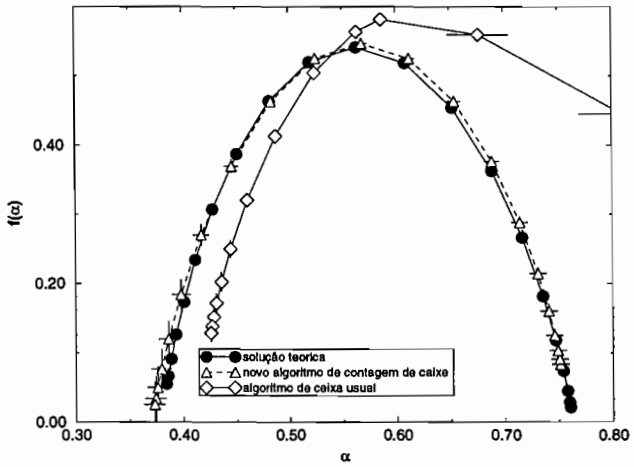


Figura 7.18: Método canônico, utilizando o novo algoritmo de caixa. Curva $f(\alpha)$: Atrator do mapa quadrático $x_{i+1} = \lambda(1 - 2x_i^2)$ com $\lambda = 0.837005134$. Foi utilizado o critério de médias. Comparamos o algoritmo de caixa usual com o novo algoritmo que propomos. O resultado com o novo algoritmo praticamente coincide com o resultado teórico. A curva onde o algoritmo usual foi utilizado é a mesma apresentada na fig. 4.15.

Notamos que mesmo usando o algoritmo usual de contagem de caixa, o resultado obtido através do critério de ínfimos e supremos é bom. No entanto, usando o algoritmo proposto, o resultado obtido é ainda melhor.

As curvas $SomAlfa = \sum \mu_i \ln p_i$ e $SomFalfa = \sum \mu_i \ln \mu_i$ são mostradas nas figs. 7.20, 7.21. Novamente, observamos que com ambos os critérios (médias ou ínfimos e supremos) as inclinações se aproximam dos valores esperados teoricamente, exceto pelas flutuações encontradas nas curvas de $SomFalfa$.

No caso do atrator do mapa logístico, o melhor resultado foi obtido utilizando o critério de ínfimos e supremos. Os métodos microcanônico e canônico geraram resultados equivalentes.

Utilizando o atrator do mapa de Hénon, descrito na sec. 2.3, obtivemos o resultado apresentado na fig. 7.22. O conjunto foi representado por uma série de 1 200 000 pontos. Os parâmetros usados no método são $q = 0, \pm 1, \pm 2, \ldots 10$. $\epsilon = 2.6 \times 2^{-i}, i = 1, 2, \ldots 8$. Novamente, 97 deslocamentos de grade são realizados no algoritmo usual

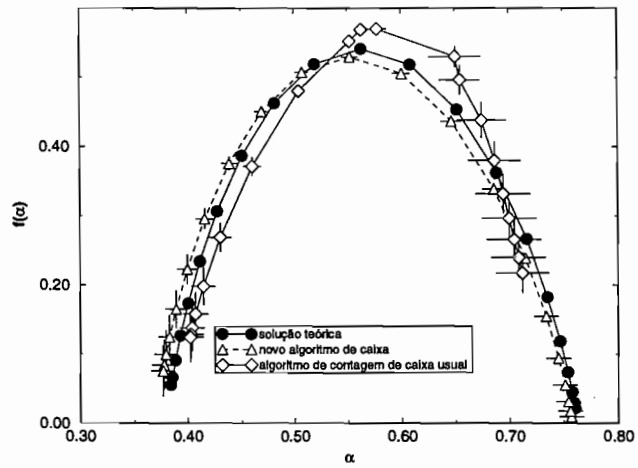


Figura 7.19: Método canônico, utilizando o novo algoritmo de caixa. Curva $f(\alpha)$: Atrator do mapa quadrático $x_{i+1} = \lambda(1 - 2x_i^2)$ com $\lambda = 0.837005134$. Foi utilizado o critério de ínfimos e supremos. Comparamos o algoritmo de caixa usual com o novo algoritmo que propomos. O resultado com o novo algoritmo é muito próximo do resultado teórico, mas não tanto quanto o obtido através da utilização do critério de médias. A curva onde o algoritmo usual foi utilizado é a mesma apresentada na fig. 4.15.

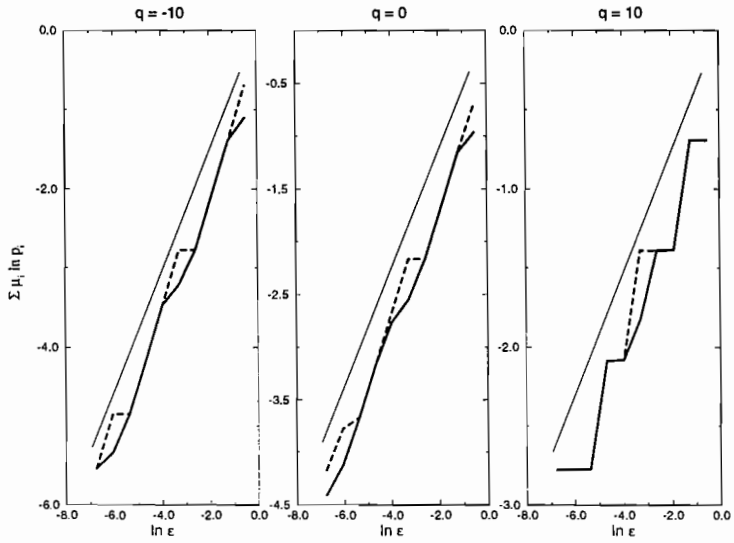


Figura 7.20: Método canônico. Curvas de $SomAlfa = \sum \mu_i \ln p_i$: Atrator do mapa quadrático $x_{i+1} = \lambda(1 - 2x_i^2)$ com $\lambda = 0.837005134$. Com a utilização do novo algoritmo, observa-se que tanto utilizando o critério de médias (pontilhado) ou de ínfimos e supremos (sólido) obtém-se curvas com inclinação muito próxima da teórica. Essas curvas podem ser comparadas às apresentadas na fig. 4.16, onde o algoritmo usual foi utilizado.

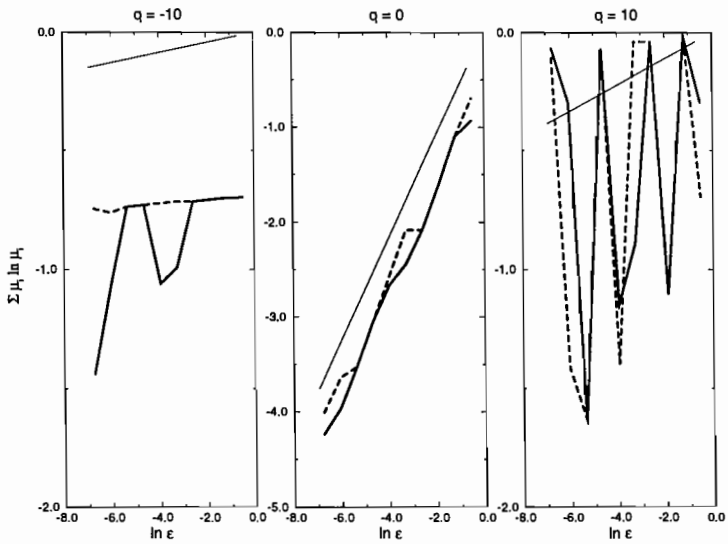


Figura 7.21: Método canônico. Curvas de $SomFalfa = \sum \mu_i \ln \mu_i$: Atrator do mapa quadrático $x_{i+1} = \lambda(1 - 2x_i^2)$ com $\lambda = 0.837005134$. Com a utilização do novo algoritmo, observa-se que tanto utilizando o critério de médias (pontilhado) ou de ínfimos e supremos (sólido) obtém-se curvas com inclinação muito próxima da teórica. Essas curvas podem ser comparadas às apresentadas na fig. 4.17, onde o algoritmo usual foi utilizado.

e 97 rotações em um ângulo de 90°. Foi utilizado o critério de ínfimos ou supremos entre deslocamentos. O resultado obtido com o algoritmo usual possui o ramo $q < 0$ mais próximo do obtido através do método da massa fixa, mas ainda se apresenta distorcido, como vemos na fig. 7.22, onde o critério de ínfimos e supremos foi utilizado, e na fig 7.23, onde o critério de médias foi utilizado.

Observe que este resultado com a distorção é semelhante ao obtido através do Método da Integral de Correlação, apresentado na fig. 5.12, e através do Método Canônico da Integral de Correlação, apresentado na fig. 6.9.

As curvas $SomAlfa = \sum \mu_i \ln p_i$ e $SomFalfa = \sum \mu_i \ln \mu_i$ são mostradas nas figs. 7.24, 7.25. Pode-se notar a diferença entre as curvas para $q < 0$ ao se utilizar os diferentes critérios e algoritmos. Podemos ver grande diferença entre as curvas para $q = -10$, indicando a influência do novo algoritmo no ramo $q < 0$ da curva $f(\alpha)$.

Analogamente ao ocorrido com o método microcanônico, ao reduzir o número de pontos que representam o conjunto o resultado tem uma melhora na curva $f(\alpha)$ quando se utiliza o critério de ínfimos e supremos. Mostramos um cálculo onde o conjunto foi representado por 200 000 pontos na fig. 7.26. Dessa forma, concluimos que a distorção observada nas curvas não tem origem na falta de pontos que representam

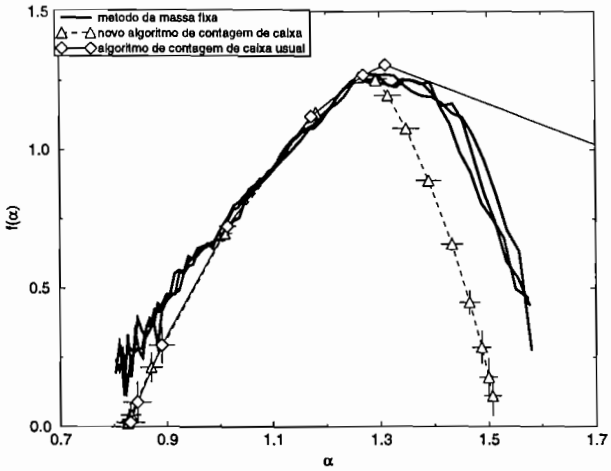


Figura 7.22: Método canônico, utilizando o novo algoritmo de caixa. Curva $f(\alpha)$: Atrator do mapa de Hénon. Foi utilizado o critério de ínfimos e supremos. O resultado obtido com o novo algoritmo é melhor do que o obtido com o algoritmo de caixa usual, mas ainda se apresenta um pouco deslocado. A curva utilizando o algoritmo de caixa usual é a mesma apresentada na fig. 4.19.

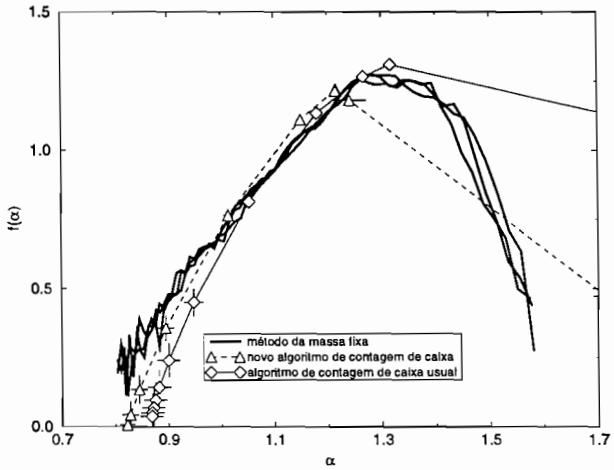


Figura 7.23: Método canônico, utilizando o novo algoritmo de caixa. Curva $f(\alpha)$: Atrator do mapa de Hénon. Foi utilizado o critério de médias. O resultado obtido com o novo algoritmo não foi muito próximo do calculado através do método da massa fixa. A curva utilizando o algoritmo de caixa usual é a mesma apresentada na fig. 4.19.

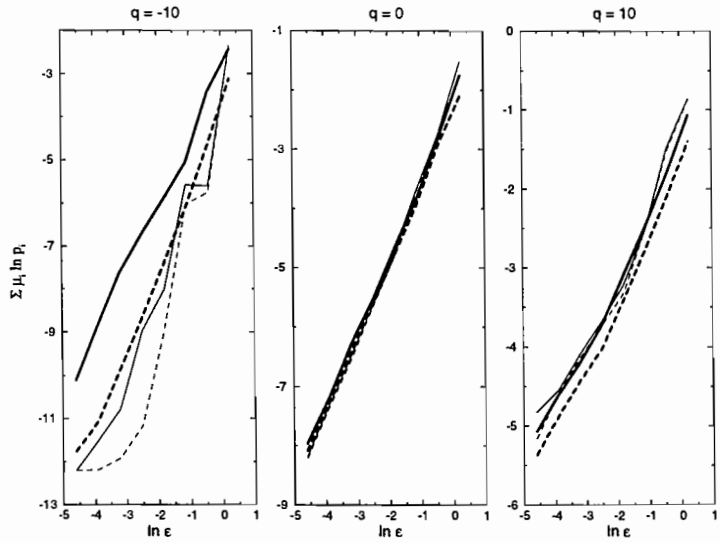


Figura 7.24: Método canônico. Curvas de $SomAlfa = \sum \mu_i \ln p_i$: Atrator do mapa de Hénon. A utilização dos diferentes critérios e algoritmos altera as curvas no ramo $q < 0$.
linha sólida fina - algoritmo de contagem de caixa usual, critério de ínfimos e supremos
linha pontilhada fina - algoritmo de contagem de caixa usual, critério de médias
linha sólida grossa - algoritmo de contagem de caixa proposto, critério de ínfimos e supremos
linha pontilhada grossa - algoritmo de contagem de caixa proposto, critério de médias

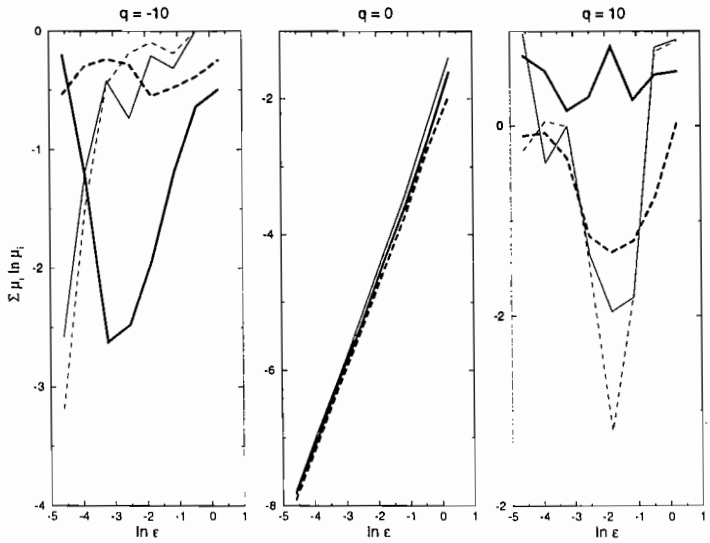


Figura 7.25: Método canônico. Curvas de $SomFalfa = \sum \mu_i \ln \mu_i$: Atrator do mapa de Hénon. A utilização dos diferentes critérios e algoritmos altera as curvas no ramo $q < 0$. Foi utilizada a mesma legenda do gráfico anterior.

o conjunto.

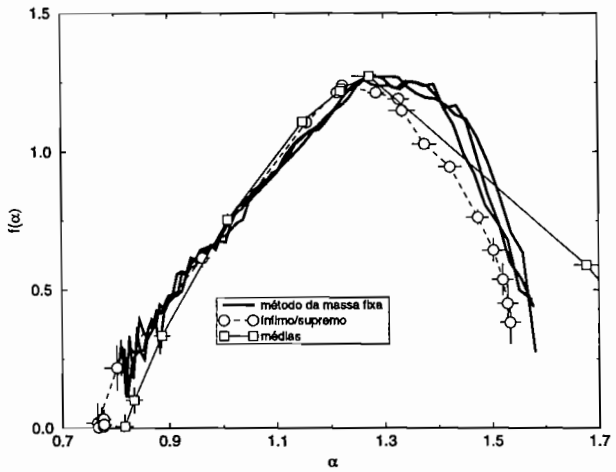


Figura 7.26: Método Canônico, utilizando o novo algoritmo de caixa. Curva $f(\alpha)$: Atrator do mapa de Hénon. Diminuindo o número de pontos, há uma melhora no ramo $q < 0$ da curva quando se utiliza o critério de ínfimos e supremos, o que não é o comportamento esperado com essa variação.

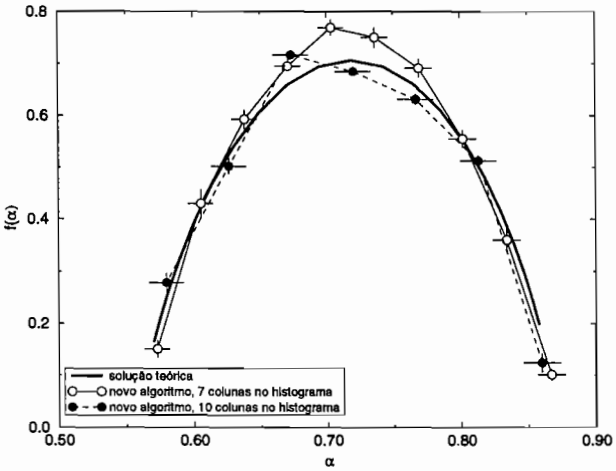
7.2.3 Método dos Histogramas

O algoritmo de contagem de caixa que propomos também foi usado no Método dos Histogramas. Obtivemos resultados um pouco melhores utilizando menos colunas nos histogramas. Nossa interpretação é de que como o nosso algoritmo produz recobrimentos que se ajustam ao conjunto, as distribuições de probabilidades obtidas também são próximas das esperadas teoricamente, que são muito irregulares. O método dos histogramas funciona bem quando as distribuições são suaves e uma maneira de tornar as distribuições mais suaves é reduzir o número de colunas.

Porém, o desenvolvimento teórico do método envolve uma aproximação para valores pequenos de δX, a largura da coluna do histograma, no cálculo de $f(\alpha)$. Ao reduzir o número de colunas, aumentamos esse valor, de modo que se espera erros maiores em $f(\alpha)$. De fato, encontramos uma diferença significativa em alguns dos cálculos que fizemos para valores de $f(\alpha)$ próximos à dimensão fractal, justamente os maiores valores de $f(\alpha)$.

Para o Conjunto de Cantor Generalizado, com parâmetros $l_1 = 0.4$, $l_2 = 0.35$,

e $p_1 = 0.6$. O conjunto foi representado por 1 000 000 pontos, as escalas utilizadas foram $\epsilon = 2^{-i}, i = 0, 1, \ldots 15$. No novo algoritmo de contagem de caixa, foram feitas 51 rotações no conjunto, que foi imerso em um espaço bidimensional. No algoritmo de caixa usual utilizamos 51 deslocamentos de grade. A curva é mostrada na fig. 7.27. Neste caso, podemos observar que o resultado é próximo do esperado, e melhor do que o obtido com o algoritmo usual, mostrado na fig. 4.22. Como já dissemos, os valores próximos à dimensão fractal são aqueles que mais diferem dos valores teóricos.



UNIVERSIDADE DE SÃO PAULO
ex. 1
SERVIÇO DE BIBLIOTECA E INFORMAÇÃO
2662
INSTITUTO DE FÍSICA

Figura 7.27: Método dos histogramas. Curva $f(\alpha)$: Conjunto de Cantor Generalizado com parâmetros $l_1 = 0.4$, $l_2 = 0.35$, e $p_1 = 0.6$. Foram construidos histogramas com 7 e 10 colunas. Observa-se que o resultado se aproxima do esperado teoricamente. Esse resultado pode ser comparado ao obtido utilizando o algoritmo usual, mostrado na fig. 4.22.

Podemos ver que os histogramas obtidos usando 10 colunas em cada um, mostrados na fig. 7.28, são regulares e não possuem as distorções observadas quando se utiliza o algoritmo usual (fig. 4.21).

O resultado para o atrator do mapa quadrático no limiar da janela caótica $x_{i+1} = A(1 - 2x_i^2)$, com $A = 0.837005$ é mostrado na fig. 7.29. Foram utilizados 1 000 000 pontos para representar o conjunto, e foram feitas médias em 51 deslocamentos de grade. Os parâmetros do método foram $\epsilon = 1.2 \times 2^{-i}, i = 1, 2, \ldots 15$ e usou-se 5 e 7 colunas nos histogramas. A solução obtida é muito irregular. Nesse caso, há várias oscilações nos histogramas, como observa-se na fig. 7.30. Não foi possível reduzir as flutuações dos histogramas até um ponto onde estes estivessem mais suaves e o cálculo pudesse ser realizado sem problemas.

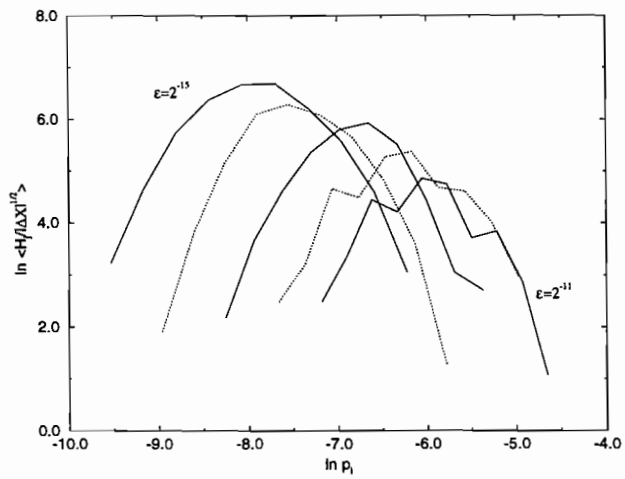


Figura 7.28: Método dos histogramas. Histogramas: Conjunto de Cantor Generalizado com parâmetros $l_1 = 0.4$, $l_2 = 0.35$, e $p_1 = 0.6$. Os histogramas são regulares e não apresentam as distorções observadas nos resultados obtidos através do algoritmo de caixa usual. Tais distorções teriam como causa a existência das caixas espúrias, que foram portanto, eliminadas. Esses histogramas podem ser comparados aos obtidos através do algoritmo usual de contagem de caixa, mostrado na fig. 4.21.

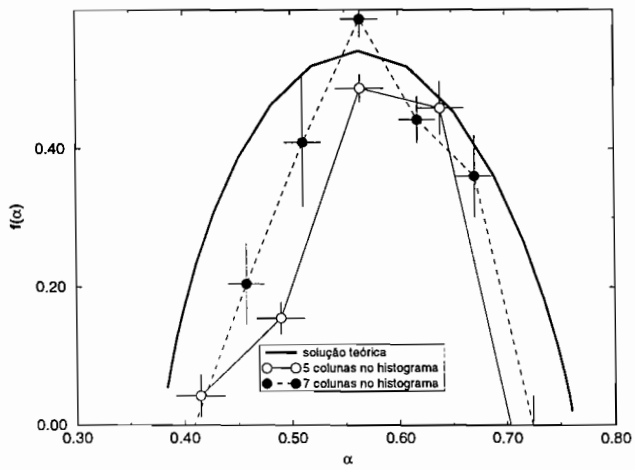


Figura 7.29: Método dos histogramas. Curva $f(\alpha)$: Atrator do mapa quadrático $x_{i+1} = A(1 - 2x_i^2)$, com $A = 0.837005$. Observa-se que os resultados obtidos se assemelham à solução teórica, mas são muito irregulares. O resultado pode ser comparado ao obtido utilizando o algoritmo de caixa usual, mostrado na figs. 4.24 e 4.25.

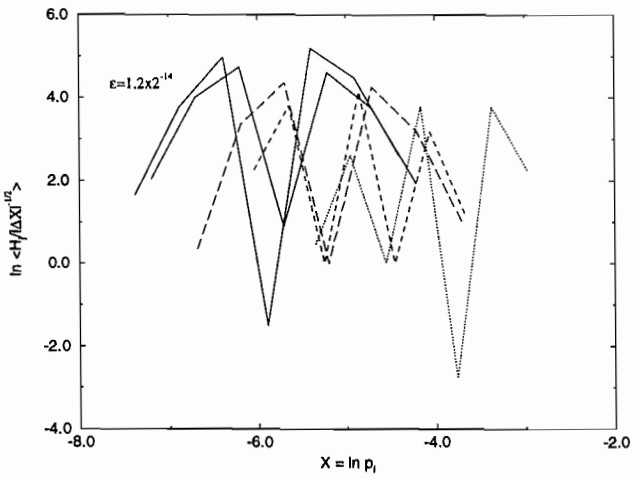


Figura 7.30: Método dos histogramas. Histogramas: Atrator do mapa quadrático $x_{i+1} = A(1-2x_i^2)$, com $A = 0.837005$. Foram usadas 7 colunas nos histogramas. Observa-se grandes oscilações nos histogramas, que explica os resultados obtidos anteriormente. Os histogramas podem ser comparados aos obtidos utilizando o algoritmo de caixa usual, mostrados na fig. 4.23.

Também aplicamos o método ao atrator do mapa de Hénon, dado por $x_{i+1} = 1 - Ax_i^2 + y_i$ e $y_{i+1} = Bx_i$, com $A = 1.4$ e $B = 0.3$. O conjunto foi representado por 1 000 000 pontos. Os parâmetros do método foram 15 colunas nos histogramas e $\epsilon = 2^{-i}, i = 1, 2, \ldots 8$. Foram utilizadas 91 rotações do conjunto no novo algoritmo de contagem de caixa (fig. 7.31). Para esse conjunto havia uma expectativa de que os histogramas fossem mais suaves porque o conjunto precisa ser estudado no espaço bidimensional, de modo que o arranjo de caixas é mais complexo, fazendo com que os valores das probabilidades fossem mais variados. Entretanto, ainda assim não foi possível construir histogramas muito suaves com um grande número de colunas.

Os histogramas relacionados a esse cálculo podem ser vistos na fig. 7.32. Os histogramas são um pouco mais regulares do que os obtidos através do algoritmo usual. As distorções devido às caixas espúrias foram eliminadas.

De maneira em geral, o método dos histogramas é de difícil utilização. Quando utilizado com o algoritmo de caixa usual, a existência das caixas espúrias distorce os histogramas. Quando se utiliza o algoritmo proposto de contagem de caixa, os histogramas se aproximam muito da distribuição real do conjunto, que é muito irregular, gerando curvas $f(\alpha)$ igualmente irregulares.

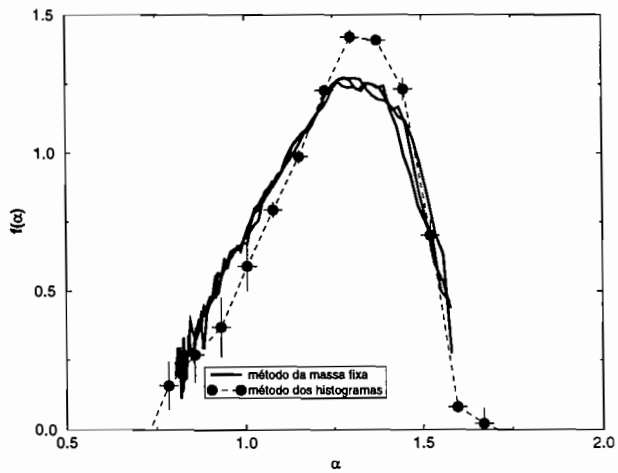


Figura 7.31: Método dos histogramas. Curva $f(\alpha)$: Atrator do mapa de Hénon $x_{i+1} = 1 - Ax_i^2 + y_i$ e $y_{i+1} = Bx_i$, com $A = 1.4$ e $B = 0.3$. O resultado é comparável ao obtido através do método da massa fixa. Porém, o máximo da curva ultrapassou o valor da dimensão fractal do conjunto. O resultado pode ser comparado ao obtido utilizando o algoritmo de caixa usual, mostrado na fig. 4.27.

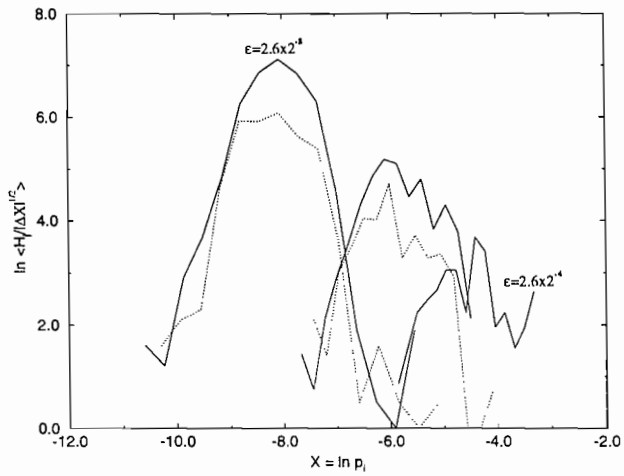


Figura 7.32: Método dos histogramas. Histogramas: Atrator do mapa de Hénon $x_{i+1} = 1 - Ax_i^2 + y_i$ e $y_{i+1} = Bx_i$, com $A = 1.4$ e $B = 0.3$. Podemos observar que os histogramas obtidos são regulares, não possuindo distorções devido à caixas espúrias. Esses histogramas podem ser comparados aos obtidos utilizando o algoritmo de caixa usual, mostrados na fig. 4.26.

Argumenta-se que as flutuações dos histogramas diminuem conforme se constrói recobrimentos com tamanhos de caixa bem pequenos. Não realizamos cálculos com tamanhos de caixa muito pequenos porque como o algoritmo proposto não elimina todas as caixas espúrias do recobrimento, elas voltam a produzir deformações. Em escalas menores, as caixas espúrias também existem em maior número, sendo difícil de evitar os recobrimentos que as contenham, mesmo usando diferentes orientações de caixa.

Como conclusão, vemos que nosso algoritmo melhora significantemente o resultado em todas as situações apresentadas.

Comentários e Conclusões

Como já dissemos, em nosso trabalho de mestrado havíamos concluido que os métodos do cálculo da função $f(\alpha)$ que se utilizam de algoritmos usuais de contagem de caixa para o levantamento das probabilidades p_i conduzem a resultados distorcidos no ramo de $q < 0$ da curva obtida, devido à existência de caixas espúrias. Essas caixas são aquelas que em um dado recobrimento, se situam próximo à "borda" do conjunto, sendo preenchidas em grande parte pelo espaço vazio existente na estrutura multifractal. Essa é uma das maiores fontes de erro nesses métodos, e atinge todo um conjunto dos mesmos, de modo que continuamos a estudar esse problema no doutoramento.

Foram consideradas duas estratégias para o estudo das caixas espúrias: Uma possibilidade foi contornar o algoritmo de contagem de caixa estudando métodos que utilizam outra abordagem para o cálculo das probabilidades p_i. Assim, estudamos métodos que aproximam as probabilidades p_i através de classificação de distâncias, entre esses, um método de nossa autoria. A outra possibilidade foi desenvolver um novo algoritmo de contagem de caixa que não utiliza grades em seu recobrimento.

Ao analisarmos os métodos que classificam distâncias entre pontos do multifractal estudado, chegamos às seguintes conclusões:

- O método da massa fixa não pode ser utilizado em conjuntos que sejam do tipo poeira de Cantor, devido à estrutura desses conjuntos, que faz com que distribuições obtidas sejam oscilantes. Essas oscilações são transmitidas à curva $f(\alpha)$ final, distorcendo o resultado. Entretanto, os resultados obtidos da aplicação desse método em conjuntos conexos é muito próximo ao previsto teoricamente, e a sua rapidez em termos de processamento numérico são pontos positivos desse método.

- O método da integral de correlação não produz um bom resultado no ramo de $q < 0$ da curva $f(\alpha)$ em todos os cálculos realizados, como já mencionado na

literatura. As razões desse problema ainda são desconhecidas e necessitam de mais estudos.

- O método canônico da integral de correlação, desenvolvido por nós, produz resultados muito próximos aos obtidos através do método da integral de correlação. Isso indica uma robustez maior das probabilidades p_i calculadas através de contagem de distâncias, quando comparada àquelas obtidas através dos métodos que utilizam o algoritmo de contagem de caixa, que são o método canônico e o método microcanônico.

Os resultados obtidos através do novo algoritmo de contagem de caixa são melhores do que os obtidos utilizando o algoritmo de contagem de caixa usual, nos três métodos estudados: o método microcanônico, o método canônico e o método dos histogramas. Essa é mais uma indicação de que as caixas espúrias são realmente a causa principal da distorção observada na curva $f(\alpha)$ no ramo $q < 0$, pois o que diferencia nosso algoritmo dos algoritmos usuais são etapas onde tenta-se eliminar o máximo possível as caixas espúrias. Esse resultado também indica que os algoritmos de cálculo da função $f(\alpha)$ através de algoritmos de contagem de caixa só vão produzir melhores resultados se os recobrimentos construídos sobre os conjuntos forem melhores.

Ainda com relação aos resultados obtidos através dos métodos que utilizam algoritmos de contagem de caixa, acrescentamos que:

- O método dos histogramas necessita de muito mais recursos numéricos e estatísticos para fornecer um bom resultado, sendo difícil aplicá-lo. Mas os extremos da curva $f(\alpha)$ são calculados com boa precisão, de modo que este é um método mais indicado para se obter esses valores.

- Apesar do critério de ínfimos e supremos ser definido na teoria multifractal como aquele que realmente escolhe o melhor recobrimento feito sobre o conjunto estudado em uma certa escala, o critério de médias forneceu algumas vezes um resultado final mais próximo do teórico. Isso pode ocorrer devido a problemas no momento da regressão linear das curvas cujos valores resultam em $f(\alpha)$. Um dos problemas seria a escolha correta dos valores de tamanhos de caixa para se realizar o cálculo, pois nem todos geram um bom comportamento linear, produzindo muitas oscilações.

Como perspectiva de estudos futuros, existem os seguintes pontos a serem ainda examinados:

- Investigar a razão das distorções no ramo $q < 0$ da curva $f(\alpha)$ que ocorrem nos resultados dos métodos de integral de correlação.

- Estudar a razão da causa da distorção do ramo $q < 0$ da curva $f(\alpha)$ do atrator do mapa de Hènon, que aparece bastante distorcida nos métodos de integral de correlação e aparentemente com um ponto de descontinuidade utilizando o novo algoritmo de contagem de caixa. Os métodos em questão são bem distintos, indicando a possibilidade de existência de alguma propriedade do atrator que explicaria esses resultados.

- Ainda é possível melhorar os recobrimentos feitos sobre o multifractal estudado no algoritmo de contagem de caixa proposto. Esse algoritmo não consegue eliminar alguns tipos de caixas espúrias, o que pode ser necessário em cálculos em sistemas de maior dimensão ou com menores tamanhos de caixa do que os estudados.

- Existe outro método para o cálculo de $f(\alpha)$ que se assemelha aos analisados neste trabalho: o spanning tree method [42], cujos resultados poderiam ser comparados aos dos métodos aqui estudados, com a finalidade de aumentar o conhecimento relacionado ao cálculo numérico da função $f(\alpha)$.

Apêndice A

Conjuntos de Cantor Generalizados

Utilizamos duas maneiras de representar os Conjuntos de Cantor Generalizados numericamente. A primeira, através dos Sistemas de Funções Iteradas, que o fornece como um conjunto de pontos. Dessa forma, podemos dar ao conjunto o mesmo tratamento dado aos mapas. Na segunda, o método da escada do diabo, o conjunto é representado pela sua massa, de modo que é uma maneira mais precisa de calcular suas probabilidades p_i.

A.1 Sistema de funções iteradas

O Sistema de Funções Iteradas (Iterated Function System) é um método geral de construção randômica de multifractais. Vamos apresentá-lo a seguir, de maneira bastante simplificada [6].

Uma transformação afim $w : R^2 \rightarrow R^2$ é definida por:

$$w \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a_{11}x + a_{12}y + b_1 \\ a_{21}x + a_{22}y + b_2 \end{bmatrix} \tag{A.1}$$

onde a_{ij} e b_i são números reais e (x, y) é o par ordenado no plano.

Dada uma transformação afim, pode-se encontrar um número s não-negativo tal que:

$$d(w(\vec{r_1}) - w(\vec{r_2})) \leq s \; d(\vec{r_1} - \vec{r_2}) \tag{A.2}$$

para todo vetor $\vec{r_1}$ e $\vec{r_2}$ no plano, onde $d(\vec{x}, \vec{y})$ é o módulo da soma dos dois vetores (distância euclidiana entre os dois pontos). O menor s que satisfaz esta relação para w chamamos de constante de Lipschitz.

Define-se um sistema de funções iteradas como um conjunto de transformações afins $\{w_1, w_2, \ldots w_N\}$ e um conjunto de probabilidades $\{p_1, p_2, \ldots p_N\}$, com p_i real, positivo e tal que $\sum_i p_i = 1$, onde cada p_i é associado a w_i. Ele deve satisfazer à condição de contração média

$$s_1^{p_1} s_2^{p_2} s_3^{p_3} \ldots s_N^{p_N} < 1 \tag{A.3}$$

A.1.1 Algortimo

O algoritmo para a construção de um multifractal é, então, similar a um mapa, como descrito na fig. A.1.

Sistema de funções iteradas — Algoritmo Geral

Escolhe-se um ponto no plano, como estado inicial do mapa.
Normalmente ignoram-se os primeiros pontos da trajetória, gerados pelo loop descrito a seguir, que pertencem a um transiente da mesma.
Sorteia-se um número aleatório de 0 a 1, verificando-se em que região do conjunto de probabilidades dado ele pertence, isto é, se o número aleatório z é tal que $$\sum_{i=1}^{m} p_i \leq z < \sum_{i=1}^{m+1} p_i,$$ z pertence à região de p_m.
Aplica-se a transformação afim correspondente à escolha do número aleatório no estado atual do sistema, obtendo o próximo ponto do mapa, ou seja, aplicamos w_m.
Repete-se um número suficientemente grande de vezes.

Figura A.1: Diagrama de Nassi-Schneiderman do algoritmo que constrói uma representacçao de um multifractal utilizando IFS.

Seja $K \subset R^2$ o multifractal resultante. Ele possui a propriedade

$$K = w_1(K) \cup w_2(K) \cup w_3(K) \ldots w_N(K) \tag{A.4}$$

e cada parte do conjunto $w_i(K)$ possui massa p_i, em relação ao todo.

Na verdade, a definição de sistemas de funções iteradas é bem mais ampla do que esta. No entanto, a definição apresentada atende às nossas necessidades (para uma discussão mais completa, veja [5]).

No Conjunto de Cantor Generalizado, toda a estrutura no intervalo unitário é reproduzida no intervalo $[0, l_1]$ com probabilidade p_1, e no intervalo $[1 - l_2, 1]$ com probabilidade $1 - p_1$, no espaço unidimensional. Assim, o sistema de funções iteradas correspondente é:

$$\begin{cases} x' = x * l_1, & \text{probabilidade} \quad p_1 \\ x' = 1 - (1 - x) * l_2, & \text{probabilidade} \quad p_2 = 1 - p_1 \end{cases} \tag{A.5}$$

O algoritmo para a gerar os pontos que representam um Conjunto de Cantor Generalizado é mostrado na fig. A.2.

Sistema de Funções Iteradas — Conjunto de Cantor Generalizado

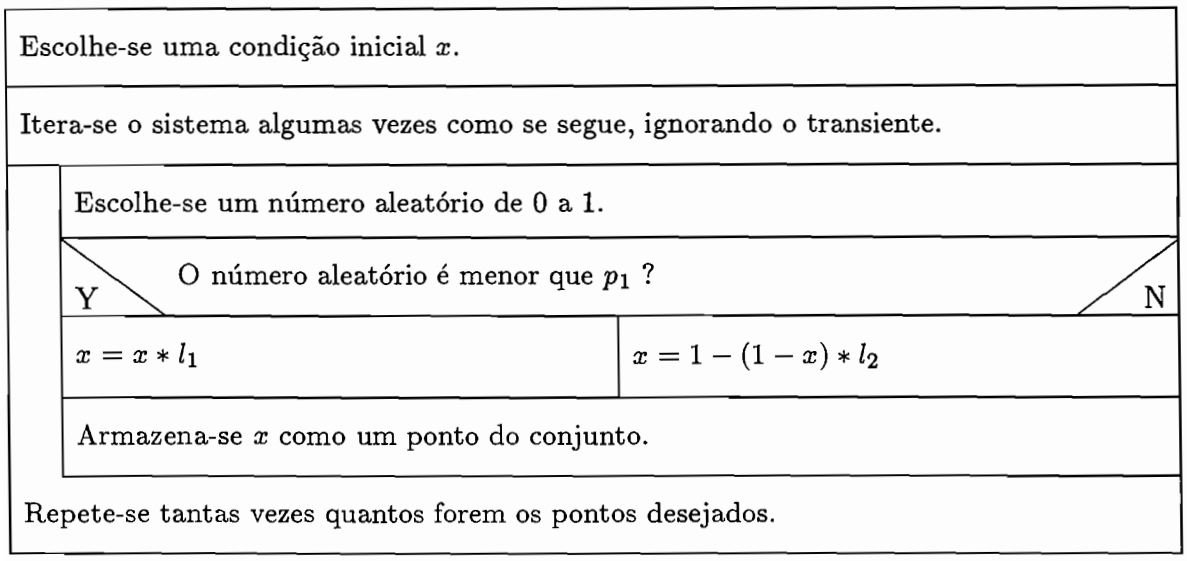


Figura A.2: Diagrama de Nassi-Scheiderman da construção de um Conjunto de Cantor gerado por IFS.

A.2 Método da escada do diabo

O método da escada do diabo (devil's staircase) é uma tentativa de reproduzir numericamente a massa de um Conjunto de Cantor Generalizado em um intervalo

que inicie na origem. A função gerada possui a estrutura de uma escada do diabo.

Partindo de um Conjunto de Cantor Generalizado, como descrito na sec. 2.1, definiremos uma função $M(x)$, com domínio no intervalo unitário, representando a massa do conjunto no intervalo $[0, x]$.

Dessa forma, a função pode ser construída paralelamente aos pré-fractais que geram o Conjunto de Cantor Generalizado, considerando os espaços vazios que vão surgindo (fig. A.3). Por exemplo, considere o espaço representado pelo intervalo $[l_1, 1 - l_2]$. Como ele é vazio, a função $M(x)$ é constante em seu interior e igual a p_1, que é o valor da massa do segmento $[0, l_1]$.

Assim, a forma final da função $M(x)$ é um conjunto infinito de patamares. Ela é uma função constante, exceto em um conjunto enumerável de pontos. A essa estrutura se dá o nome de escada do diabo.

Este método de representação do conjunto só pode ser utilizado em métodos de contagem de caixa. A probabilidade p em uma caixa representada pelo intervalo $[x_1, x_2]$ é $M(x_2) - M(x_1)$.

A.2.1 Algoritmo

O algoritmo que apresentamos calcula o valor da função $M(x)$, dados x e as características do Conjunto de Cantor Generalizado l_1, l_2, p_1. Utiliza-se um algoritmo do tipo pesquisa binária, que se apóia na propriedade da estrutura do conjunto se repetir nas escalas. Em resumo, como o valor da função fica definido assim que se encontra qual é a primeira geração do conjunto para a qual x pertence a um intervalo vazio, na verdade o algoritmo 'amplia' sucessivamente o segmento do gerador que contém x, até que ele seja visto em um vazio.

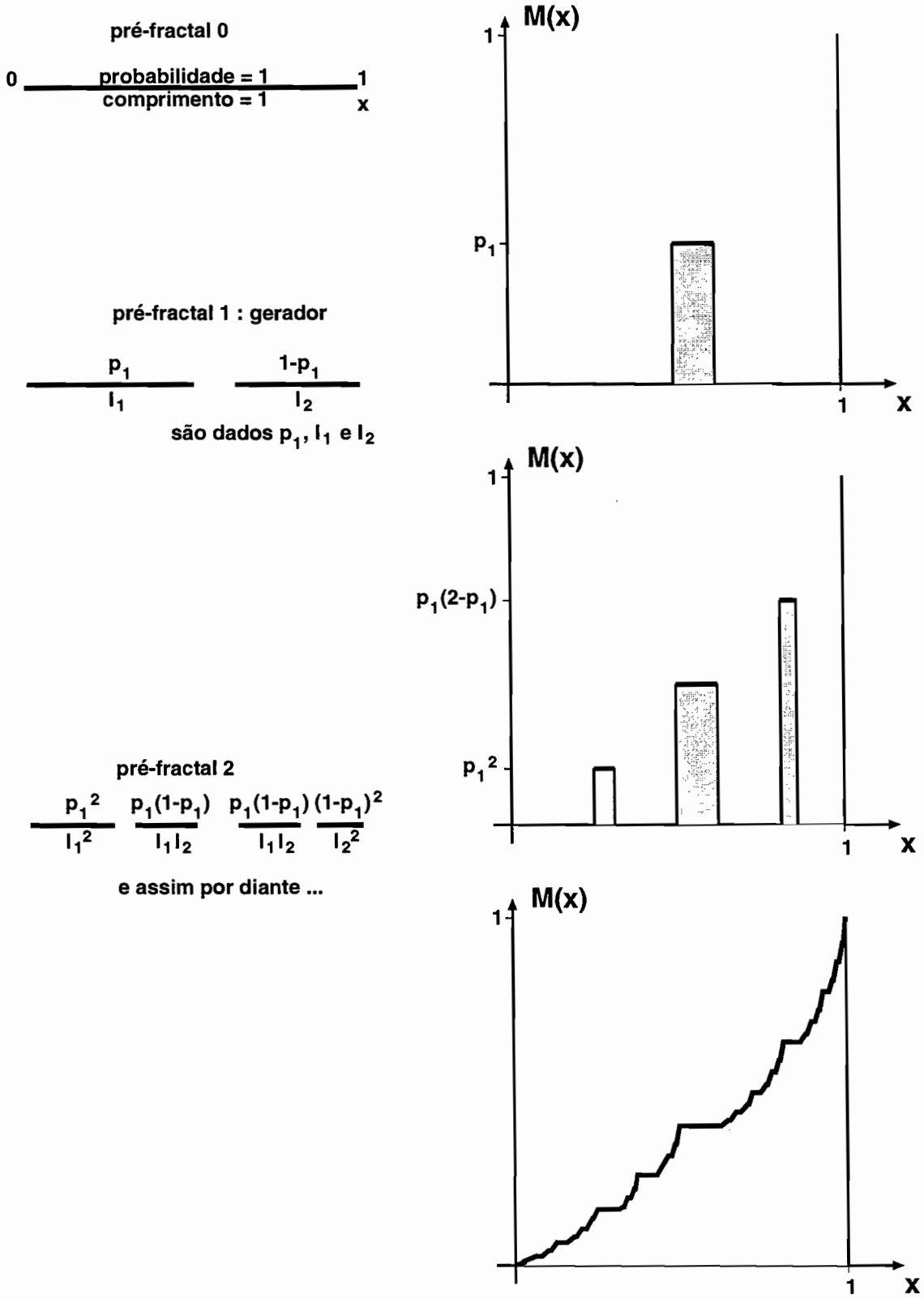


Figura A.3: construção da função $M(x)$, a massa de um Conjunto de Cantor Generalizado.

Método da Escada do Diabo — Conjunto de Cantor Generalizado

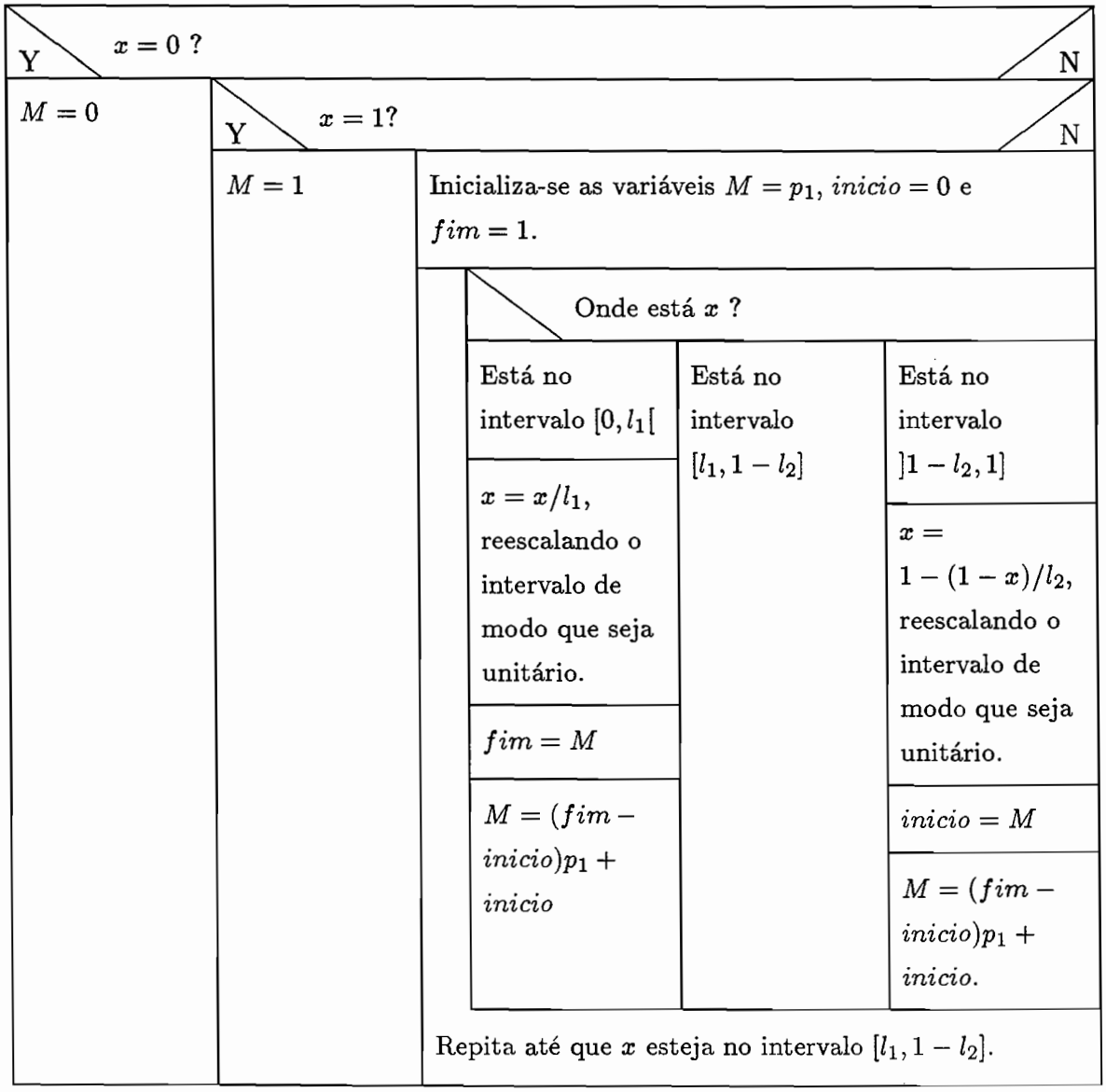


Figura A.4: Diagrama de Nassi-Schneiderman do algoritmo para o cálculo de $M(x)$. x é o valor de entrada, e M o valor de saída.

Apêndice B

Algoritmo de contagem de caixa

O método que descrevemos a seguir foi apresentado por Block [7]. A rapidez na execução deste algoritmo se deve a uma particular utilização da representação binária dos computadores, de modo que a contagem de caixa propriamente dita é realizada apenas para o menor ϵ. A única restrição é feita sobre a variação do tamanho das caixas, que é dividido por 2 a cada escala.

B.1 Método numérico

Considere a grade para o menor tamanho de caixa $\epsilon_{min} = \epsilon_0/2^{IPot}$. Seja o número de caixas em cada dimensão do espaço $NCel = 2^{IPot}$. A única 'contagem de caixa' é realizada nesta escala.

Cada caixa da grade recebe um rótulo. O mais simples vamos chamar de rótulo cartesiano. Na grade bidimensional das figuras B.1, B.2, ele seria a fila e a coluna onde se localiza a caixa. A informação da relação entre as caixas de diferentes ϵ no algoritmo de Block, está justamente no rótulo utilizado. Ele é gerado intercalando os bits da notação binária do rótulo cartesiano. Por exemplo, na figura B.1, a caixa de rótulo cartesiano $5 = (0101)_2$ para a fila e $4 = (0100)_2$ para a coluna tem como rótulo neste método $50 = (00110010)_2$.

Se repetirmos o processo, rotulando as caixas para a grade com $\epsilon = 2\epsilon_{min}$, (veja figura B.2) percebe-se com clareza a relação entre as caixas das duas escalas:

Uma caixa na grade $2\epsilon_{min}$ contém as quatro caixas da grade ϵ_{min} cuja parte inteira da divisão dos rótulos destas por quatro seja igual ao seu rótulo. Generalizando, no espaço d-dimensional, substitui-se quatro por 2^d.

ϵ_{min}

42	43	46	47	58	59	62	63	7
40	41	44	45	56	57	60	61	6
34	35	38	39	50	51	54	55	5
32	33	36	37	48	49	52	53	4
10	11	14	15	26	27	30	31	3
8	9	12	13	24	25	28	29	2
2	3	6	7	18	19	22	23	1
0	1	4	5	16	17	20	21	0
0	1	2	3	4	5	6	7	

$$NCel = 8 = 2^3$$

Figura B.1: grade para ϵ_{min}.

$2\epsilon_{min}$

10	11	14	15	3
8	9	12	13	2
2	3	6	7	1
0	1	4	5	0
0	1	2	3	

$$NCel = 4 = 2^2$$

Figura B.2: grade para $2\epsilon_{min}$.

Dessa forma, armazenando os rótulos das caixas que possuem pontos na grade ϵ, pode-se obter os rótulos das caixas que possuem pontos na grade 2ϵ, ou seja, apenas com a contagem para ϵ_{min}, podemos determinar, sem realizá-la de fato, a contagem para o restante do grupo de grades.

O diagrama de Nassi-Scheiderman desse algoritmo é mostrado na fig.B.3.

Algoritmo de Block — Contagem de caixa

Recobre-se o conjunto multifractal utilizando uma grade cujas caixas possuam tamanho ϵ_{min}. Este recobrimento é representado por dois vetores, um deles contendo uma lista ordenada dos rótulos da caixas, dados pela regra descrita anteriormente, e o outro contendo as probabilidades de cada caixa.

O recobrimento para um tamanho de caixa duas vezes maior é obtido dividindo cada rótulo por 2^d e mantendo apenas a parte inteira. Como a lista de rótulos de caixas é ordenada, uma nova lista de caixas pode ser facilmente construída verificando se o próximo rótulo na lista é igual ao considerado, após essa divisão. Se são encontrados rótulos iguais, as respectivas probabilidades devem ser somadas, e depois tais repetições devem ser eliminadas, pois se referem à mesma caixa.

Calcula-se grandezas relacionadas à $f(\alpha)$.

As probabilidades podem ser utilizadas nesta etapa, conforme o método para o cálculo de $f(\alpha)$ ou D_q que está sendo aplicado ao conjunto.

Repete-se o procedimento até que o tamanho de caixa seja da ordem do tamanho do conjunto.

Figura B.3: Diagrama de Nassi-Scheiderman do algoritmo de Block.

Apêndice C

Algoritmo para o Método Microcanônico

O método microcanônico é um método indireto para o cálculo de $f(\alpha)$. Calcula-se $\tau(q)$ utilizando-se o algoritmo de contagem de caixa, como o apresentado no ap. B. Através de uma transformação de Legendre, obtém-se $f(\alpha)$, como descrito na sec. 1.5.1. Uma das etapas da transformação de Legendre é a derivada de $\alpha(q)$, realizada numericamente, utilizando o procedimento descrito no ap. K.

C.1 Método numérico

A expressão de $\tau(q)$ é:

$$\tau(q) = \lim_{\epsilon \to 0} \frac{\ln \sum p_i^q}{\ln \epsilon}$$

Em todo algoritmo de contagem de caixa, os valores de probabilidades p_i encontrados são discretos, uma vez que são números de pontos em uma caixa. Devem existir muitos valores repetidos, de modo que é conveniente construir a distribuição destas probabilidades antes de realizar o cálculo. Essa distribuição será representada por dois vetores $Prob$ e $Freq$, que contém respectivamente o valor da probabilidade e o número de vezes que cada uma é encontrada, e o número $NProb$ de diferentes valores de probabilidades. É conveniente que as probabilidades em $Prob$ sejam ordenadas.

Comumente os valores utilizados de q possuem a forma $q = \pm k\Delta q$, para $k = 0, 1, 2, \ldots N_q$, e dessa forma podemos calcular a somatória das potências p_i^q mais eficientemente se ao invés de calcularmos potências, utilizarmos produtos por $p_i^{\Delta q}$ para obtê-las.

Armazenaremos em um vetor $SomTau[q, \epsilon]$ o numerador $\ln \sum p_i^q$.

C.2 Algoritmo

Iniciamos o cálculo obtendo as probabilidades p_i através do algoritmo de contagem de caixa. É vantajoso organizá-las na forma de distribuição de probabilidades, que vamos indicar por dois vetores $Prob$ e $Freq$, e o número $Nprob$ de diferentes probabilidades encontradas. $Prob$ Contém os valores ordenados das diferentes probabilidades, e $Freq$ quantas vezes cada valor foi encontrado. Para o cálculo de $\tau(q)$ é necessário somar potências da forma p_i^q, para diversos valores de q. Supondo-se que valores de q são dados por $q_k = \pm k\Delta q$, para $k = 0, 1, 2, \ldots Nq$, pode-se evitar o cálculo repetido de potências, substituindo-o por produtos. Para isso são utilizados dois vetores de tamanho $Nprob$. Vamos chamar esses dois vetores de $Pdelq$ e Pq. Vamos armazenar em $Pdelq$ os valores de $p_i^{\Delta q}$, em Pq os valores de p_i^q, mas de tal forma que ele sempre seja calculado recorrentemente, ou seja, se nele estiver contido o valor de $p_i^{q_k}$, calculamos o valor para q_{k+1} multiplicando cada valor do vetor por $Pdelq$. O valor para os valores de q negativo são obtidos invertendo os respectivos valores de Pq.

Suponha que o número total de caixas ocupadas pelo menos por um ponto do conjunto é N. As etapas da parte central do algoritmo são apresentadas na fig. C.1:

Depois de realizar todas as passagens acima (desde a contagem de caixa) para todos os valores de ϵ desejados, deve-se realizar uma regressão linear de $SomTau[q, \epsilon]$ contra o logaritmo de ϵ, como descrito no ap.J. Os valores dos coeficientes angulares obtidos serão uma estimativa para a curva $\tau(q)$.

Calcula-se $\alpha(q)$ através dos procedimentos mostrados no ap. K. $f(\alpha(q))$ é calculado através da expressão $q\alpha(q) - \tau(q)$.

Método Microcanônico — Núcleo

Inicializa-se $Pdelq$ com os valores de $Prob^{\Delta q}$ e Pq com valores 1.

$SomTau[q, \epsilon]$ deve ser inicializados com 0.

Faz-se $SomTau[0, \epsilon] = \ln N$.

Varia-se q, através de j, que inicia em 1 e termina em Nq.

Multiplica-se cada valor de Pq pelo seu correspondente valor em $PDelq$, para todos os $NProb$ valores de diferentes probabilidades.

Calcula-se as somatórias:
$SomTau[q, \epsilon] = SomTau[q, \epsilon] + Freq * Pq$
$SomTau[-q, \epsilon] = SomTau[-q, \epsilon] + Freq/Pq$.

para todos as $NProb$ probabilidades.

Figura C.1: Diagrama de Nassi-Scheiderman do núcleo do Método Microanônico.

Apêndice D

Algoritmo para o Método Canônico

O método canônico é um método direto para o cálculo de $f(\alpha)$, descrito na sec. 1.5.2. Calcula-se probabilidades associadas ao conjunto recoberto por grades utilizando-se um algoritmo de contagem de caixa. A partir delas, são calculadas grandezas cujo comportamento linear com o logaritmo do tamanho de caixa ϵ da grade fornece a curva $f(\alpha)$ parametricamente.

D.1 Método numérico

As expressões utilizadas por este método são:

$$\mu_i(\epsilon, q) = \frac{p_i^q(\epsilon)}{\sum_j p_j^q(\epsilon)} \tag{D.1}$$

$$\alpha(q) = \lim_{\epsilon \to 0} \frac{\sum_i \mu_i \ln p_i}{\ln \epsilon} \tag{D.2}$$

$$f(\alpha(q)) = \lim_{\epsilon \to 0} \frac{\sum_i \mu_i \ln \mu_i}{\ln \epsilon} \tag{D.3}$$

Para cada valor de q e ϵ, é conveniente calcular as seguintes variáveis, na sequência apresentada:

- $ESomTau = \sum_i p_i^q$
- $SomAlfa[q, \epsilon] = 1/ESomTau \sum_i p_i^q \ln p_i$
- $SomFalfa[q, \epsilon] = qSomAlfa[q, \epsilon] - \ln ESomTau$

Utilizamos o mesmo tipo de estratégia de cálculo que no método microcanônico (veja ap. C), onde as probabilidades p_i encontradas através do algoritmo de contagem de caixa são organizadas em vetores $Prob$ e $Freq$, que contém respectivamente

uma lista ordenada das probabilidades encontradas e o número de vezes que cada uma foi encontrada. O número de valores de probabilidades diferentes encontradas é armazenado na variável $NProb$.

Também vamos supor que os valores de q utilizados são da forma $q = \pm k\Delta\, q$, para $k = 0, 1, 2, \ldots$, calculando as potências p_i^q através de produtos sucessivos de $p_i^{\Delta\, q}$.

D.2 Algoritmo

Inicia-se o cálculo obtendo as probabilidades p_i através do algoritmo de contagem de caixa. É vantajoso organizá-las na forma de distribuição de probabilidades, que vamos indicar por dois vetores $Prob$ e $Freq$, e o número $Nprob$. Para o cálculo de $\tau(q)$ é necessário somar potências da forma p_i^q, para diversos valores de q. Supondo que valores de q são dados por $q = \pm k\Delta q$, para $k = 0, 1, 2, \ldots Nq$, pode-se evitar o cálculo repetido de potências, substituindo-o por produtos. Para isso são utilizados dois vetores de tamanho $Nprob$. Vamos chamar esses dois vetores de $Pdelq$ e Pq. Vamos armazenar em $Pdelq$ os valores de $p_i^{\Delta q}$, em Pq os valores de p_i^q, mas de tal forma que ele sempre seja calculado recorrentemente, ou seja, se nele estiver contido o valor de $p_i^{q(k)}$, calculamos o valor para q_{k+1} multiplicando cada valor do vetor por $Pdelq$. O valor para os valores de q negativo são obtidos invertendo os respectivos valores de Pq.

Vamos supor que o número total de caixas ocupadas pelo menos por um ponto do conjunto é N. As somatórias $SomAlfa[q, \epsilon]$ e $SomFalfa[q, \epsilon]$ são a saída da parte central do algoritmo, cujas etapas são dadas na fig. D.1 :

Depois de realizar todas essas passagens (desde a contagem de caixa) para todos os valores de ϵ desejados, deve-se calcular uma regressão linear de $SomAlfa[q, \epsilon]$ e $SomFalfa[q, \epsilon]$ contra $\ln \epsilon$, como descrito no ap.J.

No caso da utilização do critério de ínfimos e supremos, o denominador das probabilidades μ_i ($ESomTauP$ e $ESomTauN$) é exatamente a grandeza a ser maximizada ou minimizada. Ela deverá ser armazenada e utilizada para escolher os recobrimentos que correspondem ao critério.

Método Canônico — Núcleo

Inicializa-se $Pdelq$ com os valores de $Prob^{\Delta q}$ e Pq com valores 1.

$SomAlfa[q, \epsilon]$ e $SomFalfa[q, \epsilon]$ devem ser inicializados com 0.

Faz-se $SomFalfa[0, \epsilon] = -\ln N$.

Calcula-se a somatória $SomAlfa[0, \epsilon] = SomAlfa[0, \epsilon] + Freq \ln Prob$, para todas as $NProb$ probabilidades e a dividimos por N.

Varia-se q, através de j, que inicia em 1 e termina em Nq.

- Multiplica-se cada valor de Pq pelo seu correspondente valor em $PDelq$, para todos os $NProb$ valores de diferentes probabilidades.
- $ESomTauP$ e $ESomTauN$ são inicializados com 0.
- Calcula-se as somatórias:
  $ESomTauP = ESomTauP + Freq * Pq$
  $ESomTauN = ESomTauN + Freq/Pq$.

  para todas as $NProb$ probabilidades.
- Calcula-se as somatórias:
  $SomAlfa[q, \epsilon] = SomAlfa[q, \epsilon] + Freq * Pq/ESomTauP \ln Prob$
  $SomAlfa[-q, \epsilon] = SomAlfa[-q, \epsilon] + Freq/Pq/ESomTauN \ln Prob$.

  para todas as $NProb$ probabilidades.
- calcula-se as somatórias:
  $SomFalfa[q, \epsilon] = j * \Delta q \; SomAlfa[q, \epsilon] - ESomTauP$
  $SomFalfa[-q, \epsilon] = -j * \Delta q \; SomAlfa[-q, \epsilon] - ESomTauN$.

Figura D.1: Diagrama de Nassi-Scheiderman do núcleo do Método Canônico.

Apêndice E

Algoritmo para o Método dos histogramas

O método dos histogramas é um método direto para o cálculo de $f(\alpha)$, descrito na sec. 1.5.3. Calcula-se probabilidades associadas ao conjunto em grades utilizando-se um algoritmo de contagem de caixa, como por exemplo, o apresentado no ap. B. A partir delas, são construídos histogramas, cujo comportamento linear das variáveis envolvidas com o logaritmo do tamanho de caixa ϵ da grade fornece a curva $f(\alpha)$ parametricamente.

E.1 Método numérico

Seja Nd o número de colunas dos histogramas, cujas variáveis são:

- $X = lnp_i$ na ordenada.
- O número de pontos H_j que pertencem à j-ésima parte do intervalo $[Xmin, Xmax]$, que é dividido em Nd partes iguais. $Xmin$ é o menor valor de X encontrado na contagem de caixa considerada. $Xmax$ é o maior valor.

Os valores extremos de α, α_{max} e α_{min} são calculados a partir das regressões lineares de $Xmin$ e $Xmax$ em função de $\ln \epsilon$, respectivamente. Valores intermediários correspondentes a uma das Nd colunas do histograma são obtidos calculando proporcionalmente seu valor, através de uma expressão do tipo:

$$Alfa[i] = (i - 0.5)/Nd\,(\alpha_{max} - \alpha_{min}) + \alpha_{min} \tag{E.1}$$

Os valores de $f(\alpha)$ correspondentes são calculados a partir das regressões lineares de $\ln(H_j/\sqrt{|dX|})$ contra $\ln \epsilon$.

Foi utilizado o mesmo tipo de estratégia de cálculo que no método microcanônico (veja ap. C), onde as probabilidades p_i encontradas através do algoritmo de contagem de caixa são organizadas em vetores $Prob$ e $Freq$, que contém respectivamente uma lista ordenada das probabilidades encontradas e o número de vezes em que cada uma foi encontrada. O número de valores de probabilidades diferentes encontradas é armazenado em $NProb$.

E.2 Algoritmo

Iniciamos o cálculo obtendo as probabilidades p_i através do algoritmo de contagem de caixa. É vantajoso organizá-las na forma de distribuição de probabilidades, representada por dois vetores $Prob$ e $Freq$ e o número $Nprob$. Como $Prob$ é ordenado, $Xmin = Prob[1]$ e $Xmax = Prob[Nprob]$. A largura da coluna do histograma é $Xint = (Xmax - Xmin)/Nd$.

Como $Prob$ está ordenado, não há nenhuma dificuldade em calcular H_j. Basta somar todos os valores de $Freq$ para probabilidades $Prob$ entre $\exp(Xmin + i\ Xint)$ e $\exp(Xmin + (i+1)Xint)$, para i de 0 a $N_d - 1$. Esta parte central do algoritmo se encontra na fig. E.1.

Os valores de $H_j/\sqrt{Xint}$, $Xmin$ e $Xmax$ devem ser calculados para vários valores de ϵ e armazenados. Todo o cálculo deve ser repetido para várias posições da grade em relação ao conjunto, e deve ser armazenado o valor médio das variáveis nessas várias posições, pois este valor é aquele a ser utilizado nas regressões lineares correspondentes às eqs. 1.42, 1.33.

Método dos histogramas — Núcleo

O vetor de probabilidades $prob$ é ordenado. Então $Xmin = Prob[1]$ e $Xmax = Prob[Nprob]$.

A largura da coluna do histograma é $Xint = (Xmax - Xmin)/Nd$

Inicialize $H_j = 0$ para todos os j. Seja $k = 1$

Para $j = 1$ até $j = Nd - 1$, acrescendo de 1 em 1

- $lim = \exp(Xmin + j\ Xint)$.
- Enquanto $Prob[k] < lim$
  - $H_j = H_j + Freq[k]$
  - $k = k + 1$

Enquanto $k < Nprob$

- $H_{Nd} = H_{Nd} + Freq[k]$
- $k = k + 1$

Divida todos os H_j por $\sqrt{Xint}$.

Figura E.1: Diagrama de Nassi-Scheiderman do núcleo do Método dos Histogramas.

Apêndice F

Algoritmo auxiliado por caixas

Este algoritmo pode ser utilizado em todos os métodos para o cálculo do espectro de singularidades que necessitam de uma contagem de distâncias entre pontos do conjunto estudado. A finalidade do mesmo é diminuir o número de distâncias calculadas, ignorando aquelas que são maiores do que um determinado limite. Como tipicamente as expressões da caracterização multifractal devem ser utilizadas no limite de pequenas distâncias, não é necessário calcular as grandes distâncias. Para isso é necessária uma organização prévia dos pontos, o que consumirá um certo tempo e memória. Entretanto, o tempo que se utilizaria para calcular distâncias desnecessárias compensa essa organização.

F.1 Método numérico

Recobre-se o conjunto multifractal a ser estudado com uma grade, cujas caixas possuem tamanho ϵ, exatamente como é realizado em algoritmos de contagem de caixa. Calculando-se as distâncias entre os pontos que pertencem à mesma caixa ou a caixas vizinhas tem-se a certeza de que todas as distâncias menores que ϵ foram consideradas. Dessa forma, evita-se o cálculo das distâncias grandes, que tipicamente não são necessárias, ganhando-se tempo no cálculo.

Porém, nos métodos de integral de correlação com a aplicação deste algoritmo, os valores que são obtidos nos passam a ser válidos somente para escalas menores do que ϵ. No método da massa fixa, onde não são feitas regressões lineares, não há perda de informação ao se utilizar este algoritmo.

Para que o algoritmo seja mais eficiente, é necessário realizar uma boa organização

dos pontos, para que se possa localizar com facilidade os pontos que existem em uma dada caixa da grade. O primeiro algoritmo auxiliado por caixas é de autoria de Theiler [36], mas o algoritmo que escolhemos para implementação, e que será descrito adiante, foi criado por Grassberger [14].

F.2 Algoritmo

Este algoritmo auxiliado por caixas é de autoria de Grassberger [14]. Para se poder encontrar com facilidade quais pontos pertencem a uma dada caixa da grade, são necessários dois vetores:

- *grade*, que representa a grade que recobre o conjunto. Ele deve possuir a mesma dimensão do espaço no qual o conjunto está imerso, e conter tantos inteiros quantos forem as caixas de tamanho ϵ existentes no recobrimento.
- *fila*, que deve conter tantos inteiros quantos forem os pontos do conjunto. Ele servirá para apontar pontos, associando um ponto a outro.

Vamos apresentar um exemplo, que preenche os valores dos vetores, para melhor compreensão do algoritmo. Considere um conjunto imerso no espaço bidimensional, representado por dez pontos, e cuja grade, depois de determinado o valor de ϵ, possui as 25 caixas mostradas na fig. F.1. Inicialmente, todos os valores de *grade* e *fila* contém zeros.

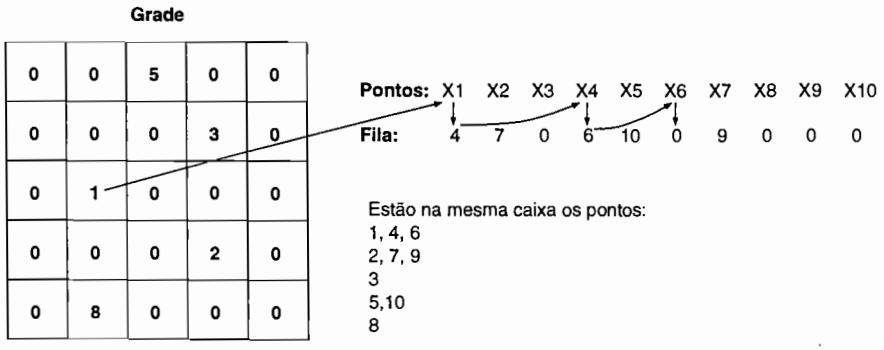


Figura F.1: Exemplo do algoritmo auxiliado por caixas. Dada uma caixa, pode-se encontrar rapidamente os pontos que nela estão contidos verificando as posições armazenadas nos vetores *grade* e *fila*.

Inicia-se o processo de organização dos pontos considerando-se um ponto do conjunto e calculando a qual caixa ele pertence. Verifica-se qual o valor de *grade* nesta

caixa. Se for zero, o substituímos pelo ordinal, ou posição no vetor de pontos do conjunto, do ponto considerado. No exemplo, isso ocorre com os pontos 1, 2 e 3.

Suponha que o ponto 4 está na mesma caixa que o ponto 1. Nesse caso, o valor de *grade* para esta caixa já contém o valor 1. Então, colocamos na posição 1 de *fila* o valor 4 que é a posição do ponto tratado.

E assim, sucessivamente. Vamos completando os valores de *grade* e *lista* com as posições dos pontos. Suponha que o ponto 6 também pertence à mesma caixa que o ponto 1. Colocamos o valor 6 na posição 4 de *fila*.

Para encontrarmos os pontos de uma dada caixa, procuramos pelo valor da respectiva caixa em *grade*, e seguimos a seqüência das posições armazenadas em *lista*, até encontrarmos o valor zero, que indica que não há mais pontos nessa caixa.

O algoritmo completo, e mais rápido, para organizar os pontos do conjunto consiste nas etapas descritas na fig. F.2 :

Algoritmo auxiliado por caixas — Organizando pontos do conjunto

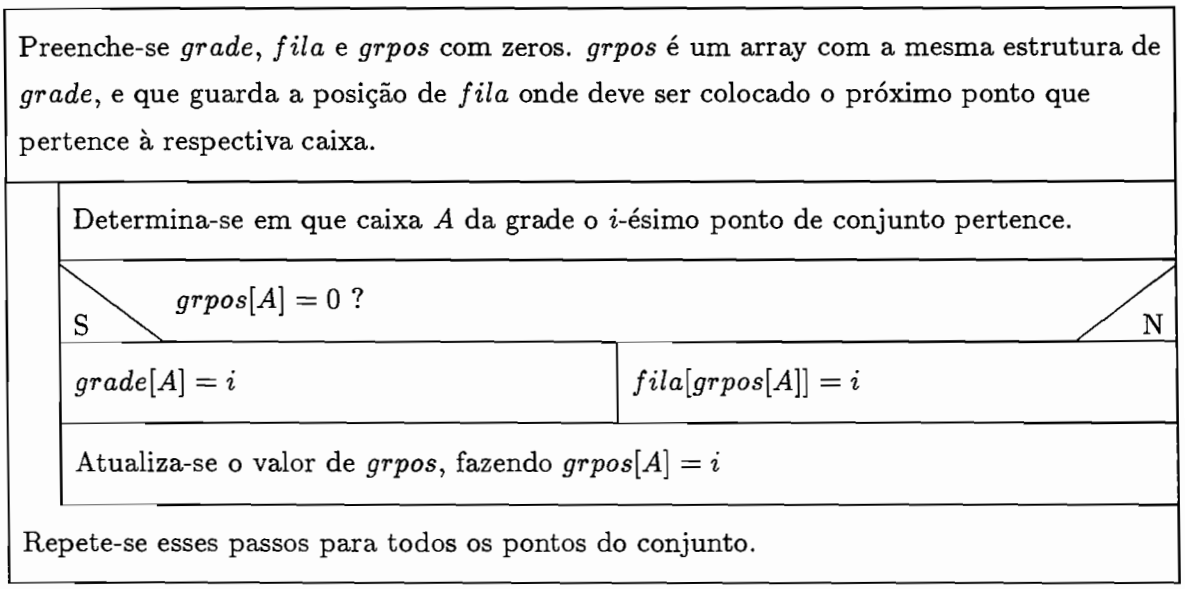


Figura F.2: Diagrama de Nassi-Scheiderman da etapa que organiza os pontos do conjunto no algoritmo auxiliado por caixas.

A saída desse algoritmo é *grade* e *fila*. Nota-se que o consumo de memória exigido por este algoritmo é o menor possível, pois todas as posições de *grade* e *fila* contém alguma informação a respeito da organização dos pontos do conjunto. A recuperação dessa informação também é muito rápida, e deve ser realizada utilizando as etapas

mostradas na fig. F.3 :

Algoritmo auxiliado por caixas — Utilizando os pontos de uma caixa

Determina-se a caixa A que contém os pontos que se deseja utilizar.
Cria-se um ponteiro $p = grade[A]$. Ele indica o ordinal, ou a posição do ponto que pertence à caixa.
Utiliza-se o ponto para o cálculo de interesse:
O p-ésimo ponto contido no vetor de pontos pertence à caixa considerada. A maneira como ele será utilizado depende do método que está sendo aplicado.
faz-se $p = fila[p]$. Ou seja, considera-se o próximo ponto da caixa A.
Repete-se até que $p = 0$, o que indica que não existem mais pontos na caixa A.

Figura F.3: Diagrama de Nassi-Scheiderman do algoritmo de utilização dos pontos organizados pelo algoritmo auxiliado por caixas.

Apêndice G

Algoritmo para o Método da integral de correlação

O método da integral de correlação é um método direto para o cálculo de $f(\alpha)$, descrito na sec. 1.5.4. Calcula-se distâncias entre os pontos do conjunto, organizando-os em um histograma, utilizado para calcular a integral de correlação. A partir da regressão linear da integral de correlação contra $\ln \epsilon$, se obtém a curva $\tau(q)$. A curva $f(\alpha)$ é obtida através de uma transformação de Legendre desta.

G.1 Método numérico

A integral de correlação é dada por:

$$C(\epsilon) = \ln \left\langle P_i^{q-1} \right\rangle \tag{G.1}$$

onde P_i é a probabilidade de se encontrar um ponto do conjunto em uma circunferência de tamanho ϵ centrada no i-ésimo ponto do conjunto.

Devemos então, calcular todas as distâncias do ponto i aos outros pontos do conjunto, e classificá-las em um histograma H_j. É apropriado que as colunas do histograma sejam construídas em intervalos do tipo b^{-n} com n inteiro positivo, e b um número real positivo maior que um a ser utilizado como base. Dessa forma, Na regressão linear de $C(\epsilon)$ contra $\ln \epsilon$ teremos pontos espaçados por $\ln b$.

A probabilidade P_i para um circunferência de tamanho b^{-n} será a soma de todos os valores H_j do histograma cujo ϵ é menor que b^{-n}.

Se a curva vai ser calculada para variáveis do tipo $q = \pm k\Delta q$, para $k = 0, 1, 2, \ldots N_q$, pode-se evitar o cálculo das potências em q, substituindo-as por produ-

tos, como descrevemos no Método Microcanônico. Neste caso isto é um pouco mais complicado porque as potências são em $q-1$.

Depois, basta repetir o cálculo das distâncias para vários pontos i do conjunto, calculando o valor médio de P_i^{q-1}.

G.2 Algoritmo

A princípio, todas as distâncias entre pares de pontos devem ser calculadas. Como isso torna o algoritmo muito lento, adotam-se as seguintes estratégias:

- Utiliza-se um algoritmo auxiliado por caixas. Veja Ap. F. Nesse caso, se o tamanho das caixas da grade usada no algoritmo é ϵ', será perdida a informação a respeito das distâncias maiores que ϵ'.

- Toma-se um subconjunto de pontos i escolhido aleatoriamente como referência, ou seja, apenas as distâncias dos pontos deste subconjunto aos pontos de todo o conjunto são consideradas. Deve-se realizar testes para verificar se este conjunto representa bem o conjunto estatisticamente. O tempo de processamento é proporcional ao tamanho desse subconjunto, de modo que pode-se reduzir em muito o tempo de processamento consumido por este método. Mas cada caso é um caso e sempre necessitamos dos testes iniciais.

Ao calcular uma distância, deve-se classificá-la em um histograma. Em geral, como a regressão linear é feita contra $\ln \epsilon$, as distâncias são classificadas em intervalos regulares no logaritmo de seu valor. Dessa forma o logaritmo de cada distância deve ser calculado, o que na prática torna o algoritmo lento. Para otimizar o tempo utilizado pelo algoritmo, sugere-se a utilização das seguintes estratégias:

- Todos os números reais são representados no computador na forma mantissa e expoente na base 2. Dessa forma, pode-se retirar diretamente o expoente dessa representação, de modo que se temos o valor da distância ao quadrado, obter esse expoente é equivalente a tirar o logaritmo na base $\sqrt{2}$. Os computadores que obedecem o padrão ANSI IEEE 754-1985 armazenam o expoente da seguinte forma:

precisão	total de bits	localização	diferença
simples	32	23-30	127
dupla	64	52-62	1024

na localização, os bits são contados da direita para a esquerda, iniciando com 0. O expoente é dado pelo número em binário encontrado na posição menos a diferença.

Dessa forma, podemos construir um histograma cujas colunas, no intervalo do logaritmo das distâncias, possuem largura $\log\sqrt{2}$. O expoente obtido da distância ao quadrado é usado como índice de um vetor. Basta acrescentar a esse valor, e obtemos o histograma.

- O procedimento descrito no item anterior é o mais rápido. Porém, os valores do logaritmo ficam associados à base $\sqrt{2}$. Se for necessário utilizar outro valor, é mais interessante construir uma tabela com as potências da base escolhida e comparar os valores do quadrado da distância com eles, evitando o cálculo do logaritmo.

Uma média $<x>_N = 1/N \sum_i^N x_i$ sempre deve ser realizada numericamente iterando uma expressão do tipo:

$$<x>_{i+1} = [i/(i+1)] <x>_i + x_i/(i+1) \tag{G.2}$$

evitando estouro na representação da variável e reduzindo erros.

Nas etapas do cálculo da integral de correlação $IntCorr[q, \epsilon]$ levamos em consideração o fato de que o intervalo em q é do tipo $[-Nq\ \Delta q, Nq\ \Delta q]$, mas o expoente na expressão da integral de correlação é $q-1$. Escolhemos Δq de modo que seu inverso é um número inteiro, e $q = 1$ sempre um valor a ser considerado. Veja as etapas do algoritmo na fig. G.1.

Método da Integral de Correlação — Núcleo

- Inicializa-se $IntCorr[q, \epsilon] = 0$ para todo q e ϵ.
- $IntCorr[1, \epsilon] = 1$, para todos os ϵ considerados.
- Para cada ponto de referência (i) repita:
  - Calcule todas as distâncias do ponto de referência aos pontos do conjunto e classifique-as em um histograma.
  - Calcule a integral do histograma, $Cq[\epsilon]$.
  - Inicializa-se $Pdelq = Cq^{\Delta q}$ e $Pq = 1$ para todos os ϵ.
    - Variamos q, usando j de 1 até $N_q - 1/\Delta q$
      - Multiplicamos cada valor de Pq pelo seu correspondente valor em $PDelq$.
      - Realizamos a média $IntCorr[(1+j)\Delta q, \epsilon] = [(i-1)/i]IntCorr[(1+j)\Delta q, \epsilon] + Pq/i$
      - Realizamos a média $IntCorr[(1-j)\Delta q, \epsilon] = [(i-1)/i]IntCorr[(1-j)\Delta q, \epsilon] + Pq/i$
    - Variamos q, usando j de $N_q - 1/\Delta q + 1$ até $N_q + 1/\Delta q$
      - Multiplicamos cada valor de Pq pelo seu correspondente valor em $PDelq$, para todos os valores de ϵ.
      - Realizamos a média $IntCorr[(1-j)\Delta q, \epsilon] = [(i-1)/i]IntCorr[(1-j)\Delta q, \epsilon] + Freq/Pq/i$
  - Repita para todos os ϵ

Figura G.1: Diagrama de Nassi-Scheiderman do núcleo do Método da Integral de Correlação.

Apêndice H

Algoritmo para o Método canônico da integral de correlação

O método canônico da integral de correlação é um método direto para o cálculo de $f(\alpha)$, descrito na sec. 6.1. Calcula-se distâncias entre os pontos do conjunto, organizando-as em um histograma, utilizado para calcular a integral de correlação. A curva $f(\alpha)$ é obtida parametricamente através das regressões lineares de grandezas associadas à integral de correlação.

H.1 Método numérico

A integral de correlação é dada por:

$$C(\epsilon) = \ln \left\langle P_i^{q-1} \right\rangle \tag{H.1}$$

onde P_i é a probabilidade de se encontrar um ponto do conjunto em uma circunferência de tamanho ϵ centrada no i-ésimo ponto do conjunto.

A partir dela são calculadas as grandezas:

$$SomAlfa = \frac{\left\langle P_i^{q-1} \ln P_i \right\rangle}{C(\epsilon)} SomFalfa = C(\epsilon) - q\, SomAlfa \tag{H.2}$$

Valem, portanto, todas as observações feitas com relação ao Método da integral de correlação no Ap. G.

H.2 Algoritmo

O algoritmo é o mesmo que o utilizado no Método da integral de correlação, exceto que é necessário armazenar a grandeza $InCorLog[q, \epsilon] = \left\langle P_i^{q-1} \ln P_i \right\rangle$ paralelamente à integral de correlação $IntCorr[q, \epsilon]$. A partir dessas duas, as variáveis relacionadas à curva $f(\alpha)$ podem ser determinadas.

Apêndice I

Algoritmo para o Método da massa fixa

O método da massa fixa é um método direto para o cálculo de $f(\alpha)$, descrito na sec. 1.5.5. Calcula-se distâncias entre os pontos de um subconjunto de referência, escolhido aleatoriamente, e os pontos do conjunto, construindo um histograma das distâncias dos k-ésimos vizinhos mais próximos. A curva $f(\alpha)$ é obtida através deste histograma. Neste método não é necessário realizar regressões lineares, mas deve-se deslocar a curva final no plano para que ela atinja a posição correta.

I.1 Método numérico

Deseja-se calcular uma aproximação para a distribuição $P(\delta, p)$, ou seja, aquela que é associada à probabilidade p de se encontrar um ponto do conjunto em uma circunferência de raio δ com centro em outro ponto.

Em um conjunto representado por N pontos, aproximamos $p = k/N$, onde k é o número de pontos no interior da circunferência. Assim, o raio da circunferência δ é a distância do ponto do subconjunto de referência ao seu k-ésimo vizinho mais próximo.

O algoritmo auxiliado por caixas (Ap. F) é muito útil neste algoritmo, reduzindo consideravelmente o tempo de processamento utilizado, porque o número de distâncias realmente necessárias para se calcular o histograma para cada ponto de referência é k, mas um algoritmo mais ingênuo precisaria calcular N. N é muito maior do que k, pela própria condição de probabilidade p pequena exigida pelo método.

Exatamente como no método dos histogramas (ap. E), o histograma deve ser construido dividindo uniformemente o intervalo do logaritmo das distâncias encontradas.

I.2 Algoritmo

O algoritmo para o método da massa fixa é colocado na fig. I.1. Ele deve ser repetido para diferentes valores de k. A utilização do algoritmo auxiliado por caixas altera o laço que calcula as distâncias entre o ponto de referência e os pontos do conjunto para um cálculo das distâncias entre o ponto de referência e os pontos de caixas próximas a ele.

Método da Massa Fixa — Núcleo

Escolhe-se aleatoriamente um conjunto de referência com $M \leq N$ pontos.

Repita para cada ponto do conjunto de referência

- Inicialize um vetor $Viz[\,]$, com k elementos, usando algum valor alto. Esse vetor vai armazenar as menores distâncias do ponto de referência aos do conjunto.
- Repita para cada ponto do conjunto
  - Calcule a distância entre o ponto de referência e o ponto do conjunto, e compará-la a $Viz[k]$. Para conjuntos em espaços de dimensão maior que um, compara-se o valor das componentes da distância.
  - A componente da distância é menor que $Viz[k]$?
    - Y:
    - N:
  - A distância é menor que $Viz[k]$?
    - Y: A distância deve ser inserida no vetor $Viz[\,]$ de modo que os valores se mantenham ordenados. O valor de $Viz[k]$ será descartado.
    - N:
- armazena-se o valor de $Viz[k]$ a construção do histograma.

O intervalo $[Ymin, Ymax]$ do logaritmo das distâncias encontradas deve ser divido em Nd partes iguais Y_j. O histograma H_j é o número normalizado de distâncias encontradas na j-ésima parte do intervalo.

Para todos os j, calcule:
$alpha[j] = \ln(k/N)/Y_j$
$falpha[j] = alpha[j](\ln H_j / \ln(k/N) - 1) + 1$

Calcule D_1 através de outro método e desloque a curva $f(\alpha)$ de modo que ela tangencie a bissetriz $\alpha = f(\alpha)$.

Figura I.1: Diagrama de Nassi-Scheiderman do núcleo do Método da Massa Fixa.

Apêndice J

Regresão Linear

Dado um conjunto de N pontos (x_i, y_i), deseja-se determinar a reta, de equação $y = ax + b$, que melhor se ajusta ao conjunto. Esta reta é obtida utilizando o *método dos mínimos quadrados*, que estabelece que a reta aproximada deve resultar no mínimo da soma quadrática da diferença entre a coordenada y dos pontos e sua correspondente na reta [21]:

$$\sum_i (y_i - ax_i - b)^2 \text{ deve ser mínimo}$$

A solução desta relação conduz ao sistema de equações:

$$\begin{bmatrix} N & \sum_i x_i \\ \sum_i x_i & \sum_i x_i^2 \end{bmatrix} \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} \sum_i y_i \\ \sum_i x_i y_i \end{bmatrix}$$

onde denominamos a matriz quadrada de M.

As soluções para a reta são:

$$a = \frac{N \sum_i x_i y_i - \sum_i x_i \sum_i y_i}{N \sum_i x_i^2 - (\sum_i x_i)^2} \tag{J.1}$$

$$b = \frac{\sum_i y_i - a \sum_i x_i}{N} \tag{J.2}$$

Os erros associados a a e b são dados pelos coeficientes da diagonal da inversa da matriz M, e um fator vindo da distribuição χ^2 [23].

$$\sigma_a^2 = \frac{\sum_i (y_i - ax_i - b)^2}{N \sum_i x_i^2 - (\sum_i x_i)^2} \frac{N}{N-2} \tag{J.3}$$

$$\sigma_b^2 = \frac{\sum_i (y_i - ax_i - b)^2}{N \sum_i x_i^2 - (\sum_i x_i)^2} \frac{\sum_i x_i^2}{N-2} \tag{J.4}$$

Apêndice K

Derivada numérica

O método, que apresentamos a seguir, para o cálculo de derivadas numéricas foi apresentado em [24]. Desejamos calcular numericamente a derivada, de uma função $f(x)$ dada em um conjunto discreto de pontos $x_i = x_0 + i\delta x$, com i inteiro.

Aproxima-se a derivada $f'(x_0)$ a partir da série de Taylor para a função $f(x)$ em potências de δx. A série deve ser truncada em alguma ordem de δx, e admite-se que o erro associado à derivada seja o primeiro termo da série de Taylor que foi desprezado.

A expansão em série de Taylor da função $f(x)$, é dada por :

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + f''(x_0)\frac{(x - x_0)^2}{2!} + \ldots f^{n'}(x_0)\frac{(x - x_0)^n}{n!} \tag{K.1}$$

onde $f^{n'}(x)$ é a n-ésima derivada da função no ponto x.

Para $x = x_1$, temos:

$$f(x_1) = f(x_0) + f'(x_0)\delta x + f''(x_0)\frac{\delta^2 x}{2} + (0\delta^3 x)$$

invertendo, temos nossa primeira estimativa para $f'(x_0)$:

$$f'(x_0) = \frac{f(x_1) - f(x_0)}{\delta x} \tag{K.2}$$

onde o erro é da ordem de:

$$\left| f''(x_0)\frac{\delta x}{2} \right| \tag{K.3}$$

Subtraindo a equação K.1 para $x = x_1$ e $x = x_{-1}$,

$$\begin{aligned} f(x_1) &= f(x_0) + f'(x_0)\delta x + f''(x_0)\frac{\delta^2 x}{2} + f'''(x_0)\frac{\delta^3 x}{6} + (0\delta^4 x) \\ f(x_{-1}) &= f(x_0) - f'(x_0)\delta x + f''(x_0)\frac{\delta^2 x}{2} - f'''(x_0)\frac{\delta^3 x}{6} + (0\delta^4 x) \end{aligned}$$

temos como resultado,

$$f(x_1) - f(x_{-1}) = 2f'(x_0)\delta x + f'''(x_0)\frac{\delta^3 x}{3} + (0\delta^4 x)$$

invertendo, temos uma estimativa mais precisa de $f'(x_0)$:

$$f'(x_0) = \frac{f(x_1) - f(x_{-1})}{2\delta x} \tag{K.4}$$

com erro da ordem de

$$\left| f'''(x_0)\frac{\delta^2 x}{6} \right| \tag{K.5}$$

De maneira análoga, utilizando as expansões de Taylor para $x = x_2$, $x = x_1$, $x = x_{-1}$ e $x = x_{-2}$,

$$\begin{aligned}
f(x_2) &= f(x_0) + 2f'(x_0)\delta x + 2f''(x_0)\delta^2 x + \\
&+ \frac{4}{3}f'''(x_0)\delta^3 x + \frac{2}{3}f''''(x_0)\delta^4 x + \frac{4}{15}f'''''(x_0)\delta^5 x + (0\delta^6 x) \\
8f(x_1) &= 8f(x_0) + 8f'(x_0)\delta x + 4f''(x_0)\delta^2 x + \\
&+ \frac{4}{3}f'''(x_0)\delta^3 x + \frac{1}{3}f''''(x_0)\delta^4 x + \frac{1}{15}f'''''(x_0)\delta^5 x + (0\delta^6 x) \\
8f(x_{-1}) &= 8f(x_0) - 8f'(x_0)\delta x + 4f''(x_0)\delta^2 x + \\
&- \frac{4}{3}f'''(x_0)\delta^3 x + \frac{1}{3}f''''(x_0)\delta^4 x - \frac{1}{15}f'''''(x_0)\delta^5 x + (0\delta^6 x) \\
f(x_2) &= f(x_0) - 2f'(x_0)\delta x + 2f''(x_0)\delta^2 x + \\
&- \frac{4}{3}f'''(x_0)\delta^3 x + \frac{2}{3}f''''(x_0)\delta^4 x - \frac{4}{15}f'''''(x_0)\delta^5 x + (0\delta^6 x)
\end{aligned}$$

temos a expressão para a derivada:

$$f'(x_0) = \frac{-f(x_2) + 8f(x_1) - 8f(x_{-1}) + f(x_2)}{12\delta x} \tag{K.6}$$

com erro da ordem de

$$\left| \frac{2}{5}f'''''(x_0)\delta^4 x \right| \tag{K.7}$$

No cálculo dos erros dados pelas eqs. K.3, K.5 e K.7, as derivadas superiores são calculadas em primeira ordem de aproximação, ou seja

$$f^{(n+1)'}(x_i) = \frac{f^{n'}(x_{i+1}) - f^{n'}(x_i)}{\delta x} \tag{K.8}$$

Por indução, pode-se mostrar que o resultado geral é:

$$f^{(n+1)'}(x_0) = \frac{\sum_{i=0}^{n+1}(-1)^{n+i}\begin{pmatrix} n+1 \\ i \end{pmatrix} f(x_i)}{\delta^{n+1}x} \tag{K.9}$$

Normalmente apenas os índices dessa equação são importantes, pois a derivada também pode ser escrita na forma

$$f^{(n+1)'}(x_i) = \frac{f^{n'}(x_i) - f^{n'}(x_{i-1})}{\delta x} \tag{K.10}$$

de modo que existe uma certa liberdade em se atribuir a derivada n-ésima ao ponto x_i. Por exemplo,

$$\begin{aligned} f''(x_0) &= \frac{f(x_2) - 2f(x_1) + f(x_0)}{\delta^2 x} \\ f''(x_0) &= \frac{f(x_1) - 2f(x_0) + f(x_{-1})}{\delta^2 x} \end{aligned}$$

são formas corretas, que podem ser deduzidas combinando as eqs. K.8, K.10.

Em particular, para os erros no cálculo numérico de $f'(x_0)$, as derivadas de ordem superior utilizadas são:

$$f''(x_0) = \frac{f(x_2) - 2f(x_1) + f(x_0)}{\delta^2 x} \tag{K.11}$$

$$f'''(x_0) = \frac{f(x_3) - 3f(x_2) + 3f(x_1) - f(x_0)}{\delta^3 x} \tag{K.12}$$

$$f''''(x_0) = \frac{f(x_5) - 5f(x_4) + 10f(x_3) - 10f(x_2) + 5f(x_1) - f(x_0)}{\delta^5 x} \tag{K.13}$$

Apêndice L

Novo algoritmo de contagem de caixa

Neste apêndice, vamos descrever em detalhe o novo algoritmo de contagem de caixa que utilizamos. Sua principal característica é o fato dos recobrimentos não serem realizados por grades, e serem construidos de tal modo que muitas caixas espúrias são eliminadas.

L.1 Introdução

O algoritmo que utilizamos pode ser dividido em duas partes: A primeira é uma modificação do proposto por Oiwa[30], e a segunda tenta eliminar as caixas espúrias do recobrimento construido na primeira parte. Vamos descrever inicialmente as variáveis consideradas no algoritmo e depois detalharemos as duas partes do mesmo.

L.1.1 Variáveis

É necessário fornecer a dimensão do espaço, armazenada na variável dim, e os N pontos do conjunto, armazenados em $points[N]$. Cada caixa do recobrimento será representada pelo seguinte conjunto de variáveis:

- $boxpos[dim]$ contém a posição do vértice cuja coordenada em todas as componenetes do espaço é menor.
- $size$ contém o tamanho da maior aresta da caixa (ela pode ser retangular).

- $table[dim]$ contém o valor 1 se a respectiva componente possui a aresta de tamanho $size$, e 0 se a aresta possui a metade desse tamanho.

- $pntmin[dim]$, $pntmax[dim]$ contém o menor intervalo que possui todos os pontos no interior da caixa, em todas as componentes.

- $spclft[dim]$, $spcrgt[dim]$ contém a quantidade de espaço vazio (que não possui outras caixas) ao redor da caixa considerada.

- $mypnt$ contém a posição em $points[]$ de um ponto que está contido na caixa, substituindo $grade$ no algoritmo auxiliado por caixas, descrito no ap. F, que é utilizado para armazenar os pontos contidos em cada caixa. Dessa forma, definimos também um vetor $list[N]$ para todo o recobrimento, no lugar do vetor $fila$ exigido por esse algoritmo.

Cada caixa é armazenada como uma estrutura, que contém ainda, dois ponteiros para outras duas caixas. Dessa forma, um recobrimento é armazenado como uma cadeia de caixas, onde cada estrutura é ligada através dos ponteiros à caixa anterior e a posterior na cadeia.

Devido a erros de arredondamento que podem ocorrer durante a construção dos recobrimentos, todas as variáveis que descrevem as caixas são definidas como números 'inteiros' através de uma discretização do espaço contínuo. Nós utilizamos uma unidade de discretização igual a um milésimo do menor tamanho de caixa usado para se construir os recobrimentos.

L.2 Primeira parte do algoritmo

Dado um recobrimento, constrói-se um novo conjunto de caixas cujo tamanho é a metade do tamanho das caixas do recobrimento anterior. Isso é feito através da divisão ou encolhimento de cada caixa. Cada componente do espaço é analisada separadamente. Se o intervalo de pontos é maior que a metade do tamanho da caixa ela é dividida ao meio, caso contrário a caixa é encolhida até a metade de seu tamanho. Note que durante o processo serão geradas caixas retangulares. Em um espaço de n dimensões, será necessário considerar todas as componentes do espaço para obter um recobrimento com caixas iguais.

L.2.1 Algoritmo

A primeira parte do algoritmo consiste nos passos descritos nas figs. L.1, L.2. Eles devem ser repetidos *dim* vezes para se obter um novo recobrimento.

Um recobrimento real, onde todas as caixas são quadradas, é gerado repetindo-se o processo anterior *dim* vezes. Um exemplo de um recobrimento realizado utilizando esta parte do algoritmo é mostrada na fig. L.3. Os números indicam a ordem na qual os particionamentos forma realizados nas caixas existentes. Um recobrimento de caixas quadradas é obtido depois de duas séries de particionamentos.

L.3 Segunda parte do algoritmo

A segunda parte do algoritmo tenta eliminar caixas espúrias do recobrimento através do deslocamento de um grupo de caixas vizinhas, como descrevemos anteriormente. Dividimos esta parte do algoritmo em duas etapas: na primeira tenta-se construir o grupo de caixas, que quando deslocado, permite eliminar uma delas, e na segunda parte tenta-se realizar de fato tal deslocamento.

Esta parte do algoritmo depende da dimensão *dim* do sistema. Descreveremos a seguir considerando um conjunto imerso no espaço bidimensional.

Vamos representar o grupo de caixas por um array $box[\,]$. No algoritmo, tenta-se verificar se um grupo com as propriedades desejadas pode ser construido. Caso contrário, desiste-se de construir esse grupo, tentando considerar outro. Nós representamos o cluster por um array de estruturas $box[\,]$. Os passos para a construção desse são apresentadas na fig. L.4. Alguns detalhes dos passos são descritos abaixo:

- a direção do deslocamento do grupo de caixas é determinado verificando a posição de espaço vazio na caixa $box[0]$. Se $pntmin[comp] - boxpos[comp] > size/2$, a direção do deslocamento que será aplicado ao grupo será para a esquerda. Se $pntmax[comp] - boxpos[comp] < size/2$, a direção do deslocamento será para a direita. Nos dois casos a caixa é meio-vazia.

- A componente *trans* é estudada considerando a distância entre os vértices das caixas que formam o grupo, $dist = box[i-1] \rightarrow boxpos[trans] - box[i] \rightarrow boxpos[trans]$. Se *dist* é positivo, o espaço para a caixa que irá ficar posicionada no lugar de $box[i]$ e $box[i-1]$ é $dist = size - dist + box[i] \rightarrow spcrgt[trans] + box[i-1] \rightarrow spclft[trans]$. Caso contrário, $dist = size - dist + box[i-1] \rightarrow$

Algoritmo de Contagem de Caixa — Primeira parte

Repita *dim* vezes

Repita para todas as caixas *boxold* do recobrimento

Escolha uma componente *comp* do espaço na caixa *boxold*. O critério para essa escolha que utilizamos é o maior $interv = pntmax[comp] - pntmin[comp]$, e se $table[comp] = 0$.

O intervalo *interv* é maior que $size/2$?

Y	N
A caixa será particionada na direção *comp*. Esta parte é descrita na figura L.2.	A caixa será encolhida na direção *comp*. Uma cópia *boxnew* dessa caixa *boxold* é adicionada no recobrimento.
	Aparecerá um espaço vazio. $boxnew \rightarrow scprgt[comp]$ aumentará de $size/2$.

As caixas são centralizadas em seu intervalo de pontos. O algoritmo é similar ao utilizado para particionar a caixa. As variáveis $boxpos[comp]$, $spclft[comp]$ e $spcrgt[comp]$ podem ser alteradas.

$table[comp]$ é igualado a 1, indicando que este lado da caixa é menor. É necessário checar se a caixa é quadrada. Ela é quadrada se todos os valores de $table[\,]$ são iguais a 1. Nesse caso, *size* é dividido por 2, e todos os valores de $table[\,]$ são igualados a 0. Ou então, não são alterados.

O recobrimento antigo é apagado, e o recém calculado passa a ser o antigo para efeito do cálculo do próximo recobrimento calculado quando se repete o procedimento acima. Deve-se repeti-lo *dim* vezes para que todas as caixas sejam quadradas.

Figura L.1: Diagrama de Nassi-Scheiderman da primeira parte do algoritmo.

Algoritmo de Contagem de Caixa — Particionando a caixa

Duas novas estruturas $boxlft$ e $boxrgt$, representando as novas caixas, são criadas. Seus valores são todos iguais aos de $oldbox$, exceto $boxlft \rightarrow spcrgt[comp] = boxrgt \rightarrow spclft[comp] = 0$, porque as caixas são vizinhas.
A caixa é divida o mais próximo possível do ponto central de $interv$. A variável $mute$ é criada. $mute = pntmin[comp] + interv/2 - (boxpos[comp] + size/2)$. Se $mute > spcrgt[comp]$, faça $mute = spcrgt[comp]$. Se $-mute > spclft[comp]$, faça $mute = -spclft[comp]$.
Ambas as caixas são deslocadas de $mute$, ou seja, $boxlft \rightarrow boxpos[comp] = boxold \rightarrow boxpos[comp] + mute$ e $boxrgt \rightarrow boxpos[comp] = boxlft \rightarrow boxpos[comp] + size/2$. Os espaços ao redor das caixas muda também. $boxrgt \rightarrow spcrgt$ diminui de $mute$, $boxlft \rightarrow spclft$ aumenta de $mute$.
Os pontos devem ser separados para as respectivas caixas. Pontos cuja coordenada na componente $comp$ são menores que $boxrgt \rightarrow boxpos[comp]$ pertencem a $boxlft$, caso contrário pertencem a $boxrgt$. Dessa forma os valores de $mypnt$ ficam determinados nas duas estruturas. Isso também altera $row[]$.
O intervalo que contém todos os pontos no interior da caixa devem ser calculados. $pntmin[]$ and $pntmax[]$ são definidos em ambas estruturas.

Figura L.2: Diagrama de Nassi-Scheiderman do procedimento para particionar caixas, usado na primeira parte do algoritmo.

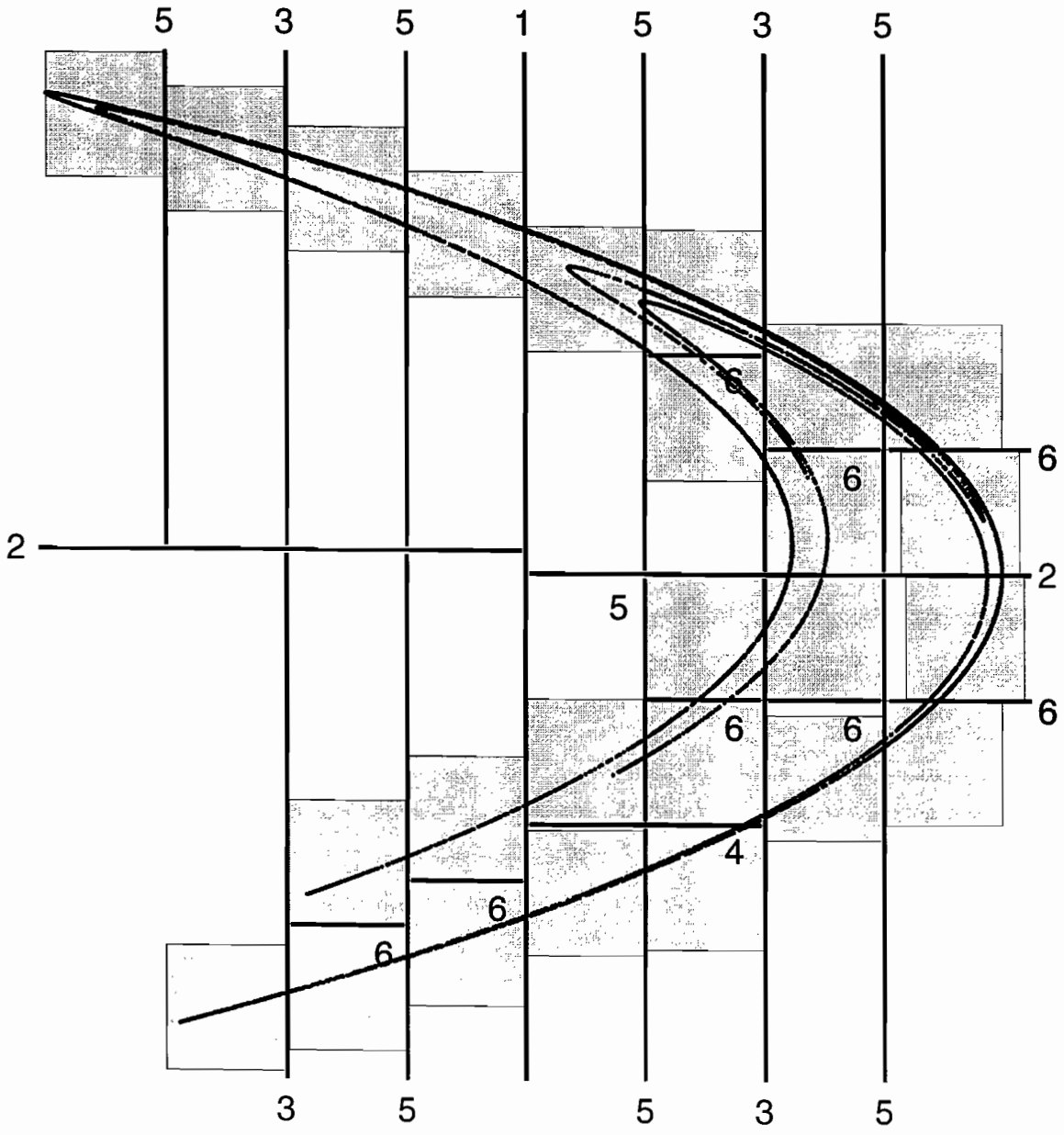


Figura L.3: Atrator do mapa de Hènon. Este é um recobrimento obtido utilizando o algoritmo de Oiwa.

Algoritmo de Contagem de Caixa — Construindo um Cluster

repita para todas as caixas do recobrimento e componentes do espaço.

Uma caixa meio vazia $box[0]$ na componente *comp* deve ser encontrada. a direção na qual a caixa será movimentada depende do espaço vazio no interior da caixa. Seja $i = 1$.

Na componente *comp* e direção definidas acima, procura-se caixas vizinhas.

Uma caixa vizinha de $box[i-1]$ foi encontrada ?

Y

N

Se apenas uma caixa vizinha $box[i]$ é encontrada, o número de caixas vizinhas desta caixa no sentido oposto ao do deslocamento do cluster deve ser determinado.

A caixa $box[i]$ é a única vizinha de $box[i-1]$?

Y

N

Se $box[i]$ possui apenas $box[i-1]$ como vizinha, os espaços na componente transversal *trans* precisam ser checados. (Veja fig. L.5a).

Existe espaço na componente *trans* o suficiente para uma caixa ?

Y

N

A caixa $box[i]$ é aceita no cluster. O cluster pode ser encerrado se existe espaço em $box[i]$ e $box[0]$ o suficiente para se retirar uma caixa. Faça $i = i + 1$.

O cluster pode ser terminado ?

Y

N

O cluster é deslocado.

Veja o algoritmo na fig. L.7.

Até a formação do cluster ou a violação de alguma de suas condições.

Figura L.4: Diagrama de Nassi-Scheiderman do procedimento de criação de um cluster, usado na segunda parte do algoritmo.

$spcrgt[trans] + box[i] \rightarrow spclft[trans]$. Se $dist$ é maior que $size$, há espaço suficiente para a caixa.

- Para encerrar o grupo, as seguintes condições devem ser verdadeiras: Se a direção na qual se moverá o grupo é para a direita, $box[0] \rightarrow (pntmin[comp] - boxpos[comp]) - box[i] \rightarrow (pntmax[comp] - boxpos[comp]) > 0$. Caso contrário, é $box[i] \rightarrow (pntmin[comp] - boxpos[comp]) - box[0] \rightarrow (pntmax[comp] - boxpos[comp]) > 0$.

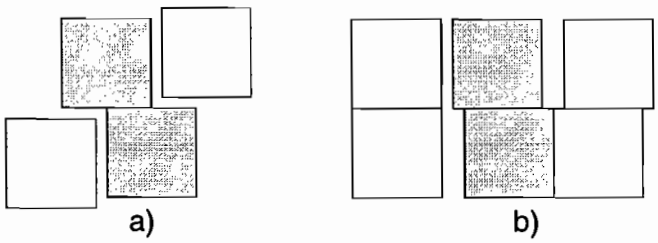


Figura L.5: a) Etapa do algoritmo para construção de grupos para eliminar caixas espúrias. O grupo (cinza) não pode ser trocado por uma caixa devido à posição das caixas próximas. b) Correção do espaço na direção transversal ao movimento das caixas deslocadas. Do lado esquerdo das caixas deslocadas, há caixas alinhadas, então nehuma alteração é necessária. Do lado direito, A menor distância entre as caixas próximas deve ser considerada.

No espaço bidimensional movimentar um grupo de caixas pode violar as condições que definem um recobrimento, ou seja, podem ocorrer sobreposições de caixas ou surgir pontos que não estão contidos em caixa alguma. Não conseguimos, com as informações existentes saber se tais violações irão ocorrer, por isso, é feita uma cópia do vetor $list[\,]$, e do grupo de caixas $box[\,]$, antes do deslocamento. Se esse deslocamento for inválido, restaura-se o estado das variáveis através das cópias.

Para mover o grupo de N caixas, seguimos o procedimento mostrado na fig. L.7.

Agora as probabilidades associadas às caixas podem ser determinadas. Por exemplo, as somas $\sum_i p_i \ln \mu_i$ e $\sum_i \mu_i \ln \mu_i$ podem ser calculados para o tamanho de caixa relacionado a esse recobrimento se o método canônico está sendo aplicado ao conjunto multifractal. Um exemplo de um recobrimento após o deslocamento de grupos de caixas se encontra na fig. L.8.

Como mencionamos anteriormente, não consideramos todos os grupos contendo caixas espúrias. Na fig. 7.3, mostramos um grupo básico que não é deslocado. Os grupos que consideramos são parecidos com estruturas unidimensionais no espaço

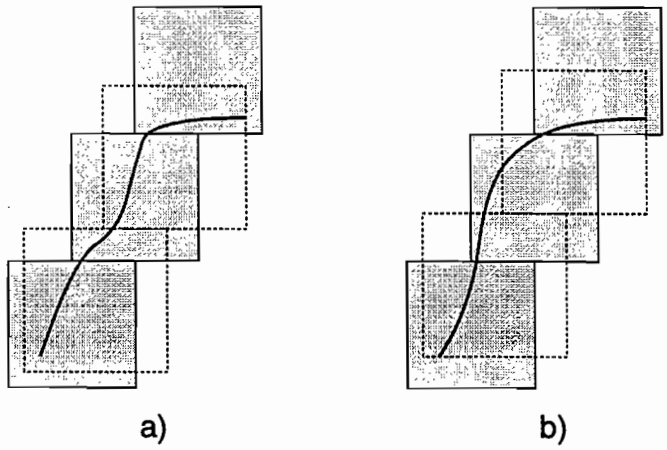


Figura L.6: Espaço bidimensional: Depois de deslocar o grupo de caixas, é necessário verifica se há sobreposição de caixas ou pontos do conjunto fora das caixas. Em a) a distribuição dos pontos na caixa central do grupo (cinza) torna possível substituir o cluster pelas duas caixas (pontilhado). Em b) isso não é possível.

bidimensional. Há outros grupos que poderiam ser considerados pelo algoritmo, mas esses que estudamos são mais comuns, de modo que os resultados não chegam a ficar comprometidos por essa restrição.

Algoritmo de Contagem de Caixa — Deslocando o grupo

Se o grupo de caixas deve ser deslocado à esquerda, inverta todas as posições de $box[\,]$. Agora, todos os grupos se movem para a direita.

O deslocamento $disp$ que será aplicado ao grupo deve centralizá-lo, no intervalo de pontos nele contido.
$disp = box0 \rightarrow (pntmin[comp] - boxpos[comp]) - box[N] \rightarrow (pntmax[comp] - bospos[comp])/2$.
Seja $i = 1$.

As caixas devem ser consideradas em duplas $box[i]$ e $box[i+1]$, porque pontos em uma caixa podem ter que passar para a outra. Adicione $disp$ a $box[i] \rightarrow boxpos[comp]$;
$box[i] \rightarrow boxpos[comp] + size$ será usado como limite para separar os pontos. Se um ponto de $box[i+1]$ tem seu valor na componente $comp$ maior que esse limite, ele permanecerá na caixa. Caso contrário, será movido para a caixa $box[i]$. A redistirbuição dos pontos pode alterar row e $box[i+1] \rightarrow mypnt$.

Os valores dos intervalos que contém todos os pontos na caixa $pntmin[\,]$ e $pntmax[\,]$ de $box[i]$ são recalculados. $i = i + 1$.

Repita para todos os pares de caixas vizinhas que pertencem ao grupo, ou até $i = N - 1$

A última caixa, $box[N]$, deve conter nenhum ponto agora. Esta será a caixa a ser eliminada. Os ponteiros do recobrimento que apontavam para esta caixa devem ser alterados.

Devido à eliminação dessa caixa, um espaço vazio vai aparecer. $box[0] \rightarrow spclft[comp]$ será aumentado de $disp$, e $box[N-1] \rightarrow spcrgt[comp]$ será aumentado de $size - disp$.

Pode ser necessário alterar os espaços na componente transversal das novas caixas, veja fig. L.5b. As caixas do lado esquerdo do cluster estão alinhadas, então nenhuma alteração é mecessária. Do lado direito, porém, a menor distância das caixas mais próximas deve ser colocada na variável $spcrgt[trans]$.

As caixas devem ser centralizadas no intervalo de pontos que existe dentro delas. Isto pode alterar $boxpos[\,]$, $spclft[\,]$ e $spcrgt$.

Figura L.7: Diagrama de Nassi-Scheiderman do procedimento de deslocamento do grupo, usado na segunda parte do algoritmo.

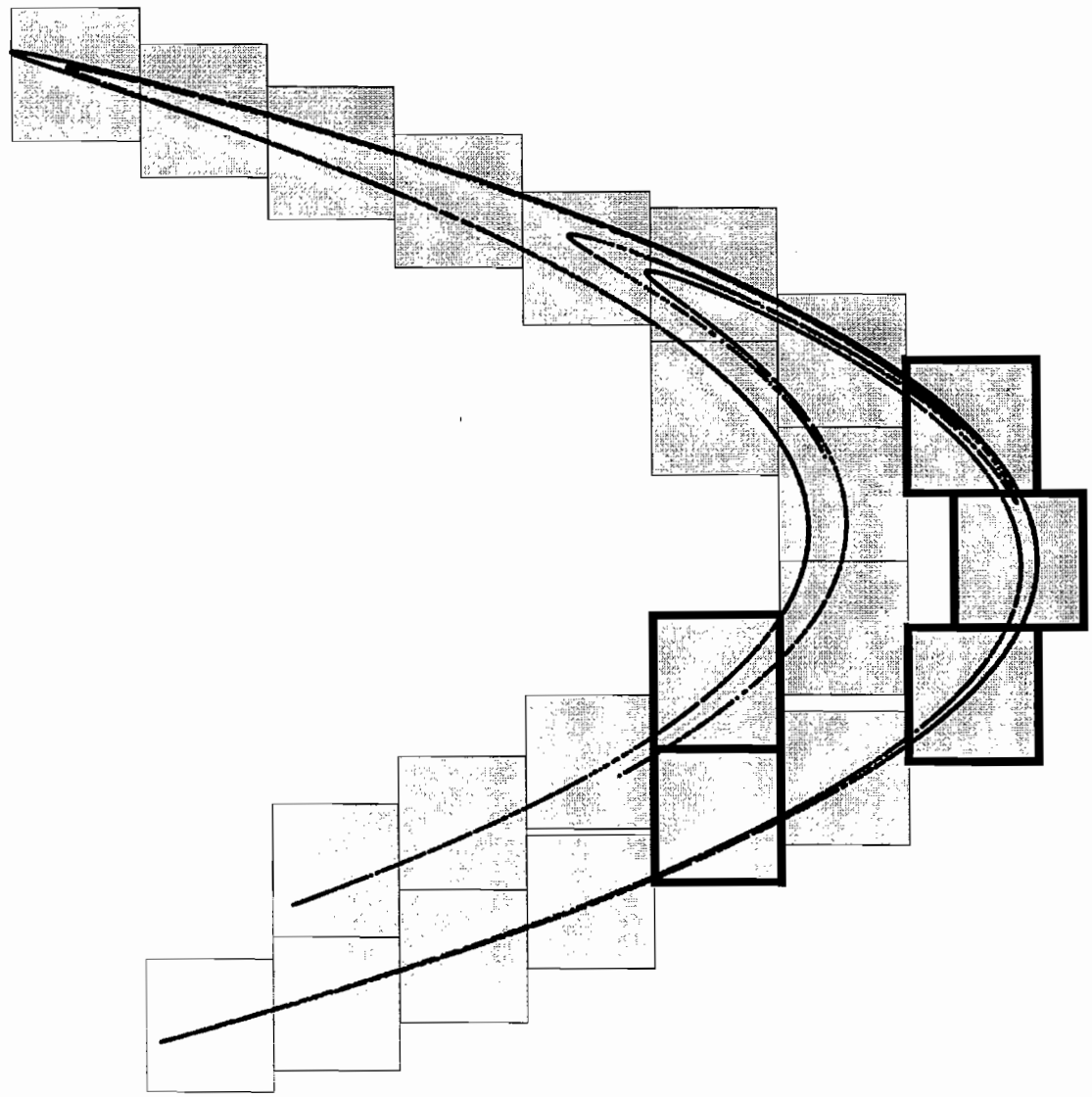


Figura L.8: Atrator do mapa de Hènon. O mesmo recobrimento da figura anterior após a movimentação de grupos de caixas, assinalados com traços mais fortes.

Apêndice M

Tempos de Processamento

Mostramos abaixo alguns tempos de processamento consumidos em alguns dos cálculos apresentados neste trabalho, de maneira ilustrativa. O formato do tempo utilizado no processamento é (dias horas :minutos :segundos .décimos).

Os seguintes cálculos foram feitos em uma estação SUN SPARC2:

Algoritmo de contagem de caixa usual Método Microcanônico	
figura	tempo
4.3, ínfimo/supremo	49:44.91
4.3, médias	48:59.89
4.6, ínfimo/supremo	1:06:13.75
4.6, médias	1:04:12.07
4.9, ínfimo/supremo	2:48:56.82
4.9, médias	2:43:32.25

Algoritmo de contagem de caixa usual Método Canônico	
figura	tempo
4.11, ínfimo/supremo	1:02:32.63
4.11, médias	1:00:30.38
4.15, ínfimo/supremo	1:14:52.72
4.15, médias	1:12:37.13
4.19, ínfimo/supremo	3:11:07.87
4.19, médias	3:03:33.32

Algoritmo de contagem de caixa usual Método dos Histogramas	
figura	tempo
4.22, 99 deslocamentos de grade	1:07:27.46
4.24, 97 deslocamentos de grade	1:29:36.17
4.27	2:04:28.56

Algoritmo proposto de contagem de caixa Método Microcanônico	
figura	tempo
7.5	2:30:43.02
7.4	2:31:52.76
7.7	2:15:57.68
7.8	2:19:47.97
7.10	21:12:32.87
7.11	21:04:35.29

Algoritmo proposto de contagem de caixa Método Canônico	
figura	tempo
7.15	2:40:35.50
7.14	2:39:00.51
7.19	2:14:23.15
7.18	2:23:17.02
7.22	19:36:51.86
7.23	18:43:31.84

Algoritmo proposto de contagem de caixa Método dos Histogramas	
figura	tempo
7.27, 10 colunas no histograma	2:57:42.72
7.29, 7 colunas no histograma	3:16:46.92
7.31	14:15:57.93

Método da Massa Fixa	
figura	tempo
5.2	13:59.26
5.1	10:05.35

Os seguintes cálculos foram feitos em uma estação HP série 7500, cerca de 5 vezes mais veloz que a anterior.

Método da Integral de Correlação	
figura	tempo
5.6, com 120 000 pontos	1:47:03.62
5.9, com 100 000 pontos	1:37:58.95
5.12, com 500 000 pontos	6 08:44:36.63

Método Canônico da Integral de Correlação	
figura	tempo
6.1, com 120 000 pontos	2:02:10.80
6.5, com 100 000 pontos	1:48:17.02
6.9, com 500 000 pontos	7 00:38:00.33

Bibliografia

[1] F. Argoul, A. Arneodo, G. Grasseau *Z. Angew. Math. Mech.* **68**, 519 (1988)

[2] A. Arneodo, G. Grasseau, E.J. Kostelich *Phys. Lett. A* **124**, 426 (1987)

[3] H. Bai-Lin *Chaos, An introduction and reprints volume*, World Scientific, (1984)

[4] R.Baddi, G. Broggi *Phys. Lett. A* **131**, 339 (1988)

[5] M.F. Barnsley, S. Demko *Proc. R. Soc. London A* **399**, 245 (1985)

[6] M.F. Barnsley, et al. *The Science of Fractal Images*, Springer-Verlag, NewYork (1988)

[7] A. Block, W. von Bloh, H.J. Schellnhuber *Phys Rev. A* **42**, 1869 (1990)

[8] S. Borgani, G. Murante, A. Provenzale, R. Valdarnini *Phys. Rev. E* **47**, 3879 (1993)

[9] J.L. Mc Cauley *Phys. Rep.* **189**, 225 (1990)

[10] A.B. Chhabra, C. Meneveau, K.R. Srenivasan *Phys. Rev. A* **40**, 5284 (1989)

[11] J.D.Farmer, Ott, Yorke *Physica* **7D**, 153 (1983)

[12] J. Feder *Fractals*, Plenum, New York (1988)

[13] G. Gunaratne, I. Proccacia *Phys. Rev. Lett.* **59**, 1377 (1987)

[14] P. Grassberger, I. Procaccia *Phys.* **9D**, 189 (1983)

[15] P. Grassberger *Phys. Lett. A* **148** 63 (1990)

[16] P. Grassberger *Int. J. of Mod. Phys. C* 4 515 (1993)

[17] P. Grassberger, R. Badii, A. Politi *J. of Stat. Phys.* **51**, 135 (1988)

[18] J. Guckenheimer, Holmes *Nonlinear Oscilations, Dynamical Systems, and Bifurcations of Vector Fields*, Spring Verlag, (1983)

[19] T.C. Halsey, M.H. Jersen, L.P. Kadanoff, I. Procaccia, B.I. Shraiman *Phys. Rev. A 33*, 1141 (1986)

[20] M. Hénon *Commun. Math. Phys.* **50**, 69 (1976)

[21] O.A.M. Helene, V.R. Vanin *Tratamento Estatístico de Dados em Física Experimental*,

[22] H.G.E. Hentschel, I. Procaccia *Phys.* **8D**, 435 (1983)

[23] R. Hoffmann, S. Vieira *Análise de Regressão*, HUCITEC-USP, São Paulo, (1977)

[24] S.E. Koonin *Computational Physics*, Addison-Wesley,(1986) Cap. 1

[25] F.H. Ling, G. Schmidt *J. Comp. Phys.* **99**, 196 (1992)

[26] B.B. Mandelbrot *Fractals, Form, Chance and Dimension*, Freeman, New York, (1977)

[27] B.B. Mandelbrot *The fractal geometry of nature*, Freeman, New York, (1983)

[28] C. Meneveau, K.R. Sreivasan *Phys. Lett. A* **137**, 103 (1981)

[29] T. C. A. Molteno *Phys. Rev. E* **48**, R3263 (1993)

[30] N. Oiwa, N.F. Ferrara *Comunicação interna*

[31] A.M. Ozorio de Almeida *Hamiltonian Systems: Chaos and Quantization*, Cambridge University Press, (1988)

[32] G. Paladin, A. Vulpiani *Phys. Rep.* **156**, 148 (1987)

[33] K. Pawelzik, H.G. Schuster *Phys. Rev. A* **35**, 481 (1987)

[34] H. G. Schuster *Deterministic Chaos, an introduction*, VCH, Alemanha, 2nd edition, 1988.

[35] T. Tél, A. Fülöp, T. Vicsek *Phys. A* **159**, 155 (1989)

[36] J. Theiler *Phys. Rev. A* **36**, 4456 (1987)

[37] C.J. Thompson *Classical Equilibrium Statistical Mechanics*, Clarendon Press, Oxford (1988) Cap. 3, Ex. 2

[38] M. Yamaguti *Influência da grade no algoritmo de contagem de caixa*, Tese de mestrado defendida no Instituto de Física da USP (1992)

[39] M. Yamaguti, C.P.C. Prado Phys. Lett. A **206**, 318 (1995)

[40] M. Yamaguti, C.P.C. Prado Phys. Rev. E **55**, 7726 (1997)

[41] W. Water, P. Schram *Phys. Rev. A* **37**, n8, 3118 (1988)

[42] R. Weygaert, B.J.T. Jones, V.J. Martinez *Phys. Lett. A* **169**, 145 (1992)