

Computação Gráfica e Sistemas Dinâmicos mais um pouquinho de Teoria Ergódica e Processos Estocásticos

Gonzalo Contreras

Departamento de Matemática
PUC/RJ
Rua Marquês de São Vicente, 225
22.453 Rio de Janeiro - RJ

1. Introdução.

Recentemente uma área abstrata da Matemática, de viva pesquisa no momento, tem dado uma aplicação importante à computação gráfica. É um método de gerar imagens complexas para o qual já se está desenhando hardware especializado para micro-computadores. Trata-se dos Sistemas Iterados de Funções (SIF), um objeto de estudo em Sistemas Dinâmicos e Teoria Ergódica.

Em [7] B. Mandelbrot chamou a atenção para a semelhança entre conjuntos de dimensão fracionária (que chamou de "fractais") e objetos existentes na Natureza. Essa semelhança mostra que esses conjuntos são mais adequados para modelar fenômenos naturais do que as curvas (dimensão 1) e superfícies (dimensão 2) utilizadas na área de modelagem geométrica.

Neste trabalho pretendemos descrever uma técnica de gerar "fractais" que tem sido objeto de estudo em Sistemas Dinâmicos e Teoria Ergódica. Essa técnica se baseia nos *sistemas iterados de funções* (SIF). Vários matemáticos têm utilizado essa técnica para gerar imagens realísticas da Natureza, de grande complexidade.

Inversamente, dada uma imagem, pode-se reproduzi-la com o grau de exatidão desejado usando SIF's, codificá-la por um conjunto de números (que definem os SIF's) e guardá-la num espaço de memória muito menor que o dos métodos de codificação usuais.

Uma vez armazenados os parâmetros de vários SIF's, pode-se dispôr de uma biblioteca que fica imediatamente ordenada de acôrdo com a proximidade dos (parâmetros dos) SIF's. O fato é que, como veremos nos teoremas da Colagem e da estabilidade estrutural, SIF's próximos dão imagens semelhantes. Esse fato indica que essas técnicas também podem ser úteis na área de reconhecimento de padrões.

Outra conseqüência dos teoremas da colagem e da estabilidade estrutural é que podem-se simular certos movimentos gerando uma seqüência de SIF's que ocile dentro de uma curva do espaço de parâmetros e assim, literalmente, ver a samambaia da Fig. 1 vibrar pela ação do vento.

Veremos também que os erros computacionais das máquinas não são de maior importância neste contexto. A estrutura do artigo é a seguinte: na seção 2 daremos a definição de um SIF. Na seção 3 faremos a descrição do algoritmo, em forma simples. É importante que o leitor entenda bem o exemplo 1, dado nessa seção, e consulte-o durante toda a leitura do artigo; isto lhe permitirá gerar suas próprias imagens. Uma técnica para reproduzir imagens e fazer novos "fractais" se apresenta na seção 6 no "teorema da colagem". Na seção 4 apresentamos esboços de provas que permitem ao leitor reconstruí-las, e dão uma boa idéia do tipo de argumentos e ferramentas que se usam na teoria dos SIF. A seção 5, "SIF e Teoria Ergódica", é menos elementar. Decidi incluir estes esboços porque a literatura dos SIF está dispersa em revistas de distintos gêneros e especialidades, com referências como "Veja [1], [4], [9], [7]" em lugar das provas. De fato foi devido a este problema que conheci os SIF's. O Teorema 4 é particularmente interessante para os que estudam os Sistemas Dinâmicos porque mostra que este método realmente gera conjuntos de Julia, e sua prova diz que o conjunto dos pontos transitivos de um shift de Bernoulli tem medida total. Como este tema envolve várias áreas da Matemática, vi-me obrigado a introduzir muitos conceitos. Só as definições em cursiva serão usadas adiante.



Figura 1.

Espero que, após a leitura deste trabalho, os artigos das referências possam ser lidos por um estudante do último ano da graduação em Matemática.

Quero agradecer a Javier Perales da PUC do Perú por introduzir-me neste tema. Agradeço também aos autores e editores dos livros [9] e [10] de onde foram tiradas as figuras deste artigo.

2. Sistemas Iterados de Funções.

Um SIF é um espaço métrico compacto (que pode ser $K = [0, 1] \times [0, 1]$, a tela do computador) com uma coleção finita de funções injetivas $F = \{f_1, \dots, f_n\}$, $f_i: K \rightarrow K$, tais que exista um $0 < \lambda < 1$ para o qual $d(f_i(x), f_i(y)) \leq \lambda d(x, y)$ para todo $x, y \in K$ e todo $i = 1, \dots, n$. Eventualmente um SIF pode conter um vetor de probabilidades $\bar{p} = (p_1, \dots, p_n)$ associado às f_i 's, i.e. $\sum_{i=1}^n p_i = 1$ e $p_i \geq 0 \forall i = 1, \dots, n$.

Nas aplicações à computação gráfica é mais eficiente exigir que o sistema F seja composto de transformações afins $F = \{T_1, \dots, T_n\}$, da forma

$$T_i \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a_i & b_i \\ c_i & d_i \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} u_i \\ v_i \end{bmatrix}, \quad \forall \begin{bmatrix} x \\ y \end{bmatrix} \in K = [0, 1]^2,$$

para alguns $\begin{bmatrix} u_i \\ v_i \end{bmatrix} \in \mathbb{R}^2$, $a_i, b_i, c_i, d_i \in \mathbb{R}$.

Deste modo, cada T_i fica caracterizado por 6 números e o sistema por $7n$ números, contando as probabilidades. Podemos associar a F uma aplicação de conjuntos $F: \mathcal{P}(K) \rightarrow \mathcal{P}(K)$ onde $\mathcal{P}(K) = \{A/A \subset K\}$, definida por $F(A) = \{f_i(x)/x \in A, f_i \in F\}$. Representamos por $F^2(A)$ o conjunto $F(F(A))$ e assim $F^n(A) = F(F(\dots(F(A))\dots))$, n -vezes, $F^0(A) := A$. O fato é que a "órbita" de cada ponto $x: \{F^n(x)/n \geq 0\}$ se acumula num "atrator" Λ que é a imagem que queremos obter. Antes de analisar a dinâmica de F vejamos o algoritmo que permite gerar essas imagens no computador.

3. O Algoritmo.

Os nossos dados são $F = \{T_1, \dots, T_n\}$ ($6n$ números) $\bar{p} = (p_1, \dots, p_n)$ (n -números). A tela é $K = [0, 1] \times [0, 1]$.

(1) Inicialize $x = \frac{1}{2}$, $y = \frac{1}{2}$.

(2) De $m = 1$ a 2500 repita os passos (3) a (6)

(3) Escolha k em $\{1, \dots, n\}$ com probabilidades (p_1, \dots, p_n)

(4) Calcule $(\tilde{x}, \tilde{y}) = T_k(x, y)$

(5) Faça $(x, y) := (\tilde{x}, \tilde{y})$

(6) Se $m > 20$, desenhe (x, y) .

O passo (3) pode ser implementado dividindo $[0, 1]$ em n subintervalos de comprimentos p_1, \dots, p_n e escolhendo um número

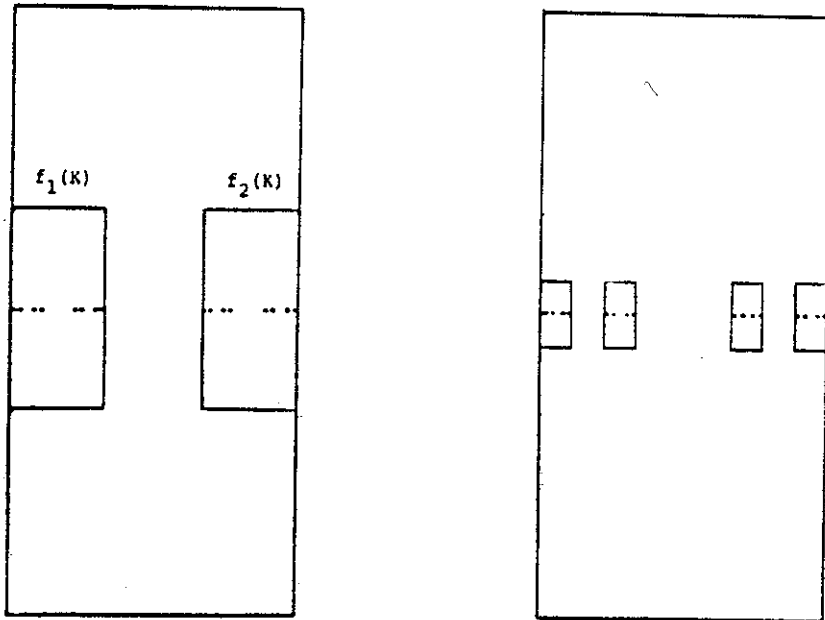


Figura 2.

RND aleatoriamente em $[0, 1]$: se escolhe k de acordo com o subintervalo onde está RND.

Os números 2500 e 20 podem variar de acordo com o sistema particular e as especificações da máquina; o seu significado ficará claro adiante. É recomendável uma execução prévia do programa para ter uma idéia dos valores máximo e mínimo que (x, y) pode tomar e assim determinar a janela da imagem a aparecer na tela.

Uma implementação do algoritmo acima usando a linguagem basic pode ser vista em [4].

Exemplo 1. Tomemos $K = [0, 1] \times [-1, 1]$, $F = \{f_1, f_2\}$,
 $f_1(x, y) = (\frac{1}{3}x, \frac{1}{3}y)$, $f_2(x, y) = (\frac{2}{3} + \frac{1}{3}x, \frac{1}{3}y)$, $\bar{p} = (\frac{1}{2}, \frac{1}{2})$.
 $F(K) = f(K) \cup g(K)$ são dois retângulos de diâmetros $\frac{1}{3}$ diam K ,
 $F^2(K)$ são 4 retângulos de diâmetro $\frac{1}{3^2}$ diam K . $F^n(K)$ são 2^n
retângulos disjuntos de diâmetros $\frac{1}{3^n}$ diam K . De fato $F^n(K)$ é
uma união de retângulos que cobrem o conjunto de Cantor usual
extraíndo intervalos centrais de proporção $\frac{1}{3}$. A "órbita" de um
ponto : $\{F^n(x)/n \geq 0\}$ $x \in K$, é tal que $F^n(x)$ contém um ponto
em cada retângulo de $F^n(K)$. De fato $\langle F^n(x) \rangle_{n \geq 0}$ se acumula

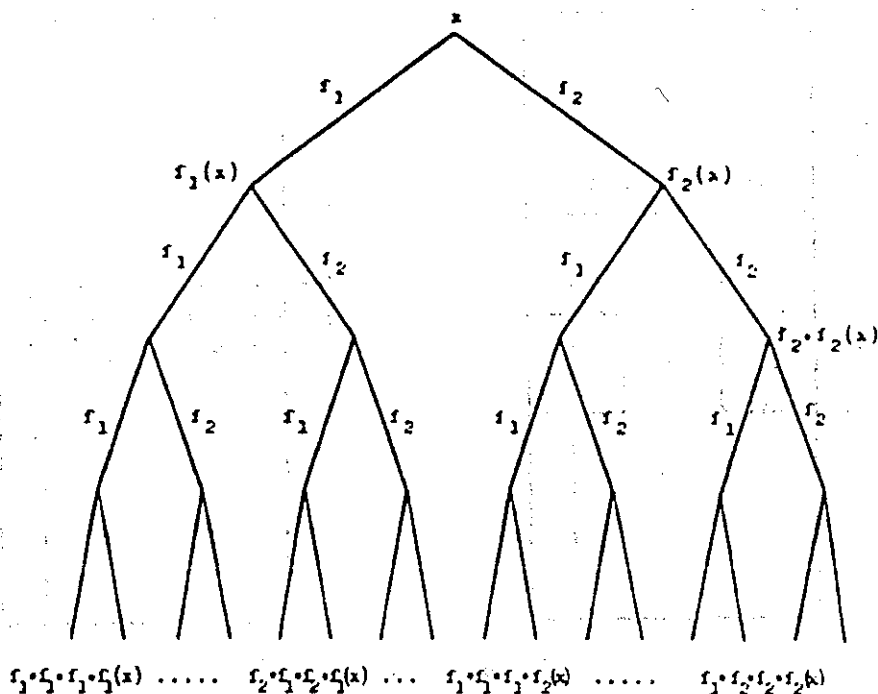


Figura 3.

no conjunto de Cantor $\Lambda = \bigcap_{n \geq 0} F^n(K)$. O algoritmo desenha a seqüência de pontos $f_{k(m)} \circ f_{k(m-1)} \circ \dots \circ f_{k(2)} \circ f_{k(1)}(x) \in F^n(K)$, onde os índices $k(j)$ são escolhidos de forma aleatória.

Veremos que com esta escolha dos $k(j)$ o conjunto de pontos de acumulação desta seqüência é o conjunto de Cantor Λ . O algoritmo executa 20 iterações antes de começar a desenhar para que os retângulos fiquem suficientemente pequenos e a seqüência esteja perto de Λ desde o começo.

Teorema 2 (Ver [2]). *Seja (K, F) um SIF; então existe um compacto $\phi \neq \Lambda \subseteq K$ tal que*

(a) Λ é completamente F -invariante, i.e. $F(\Lambda) = \Lambda = F^{-1}(\Lambda)$.

(b) Λ é o único conjunto fechado tal que $F^{-1}(\Lambda) = \Lambda$.

(c) $\Lambda = \bigcap_{n \geq 0} F^n(K)$.

(d) Para qualquer $x \in K$,

$\Lambda = \{ \lim_{m \rightarrow \infty} \pi(m, \bar{k}, x) / \bar{k} = (k_1, k_2, \dots, k_m, \dots) \}$;

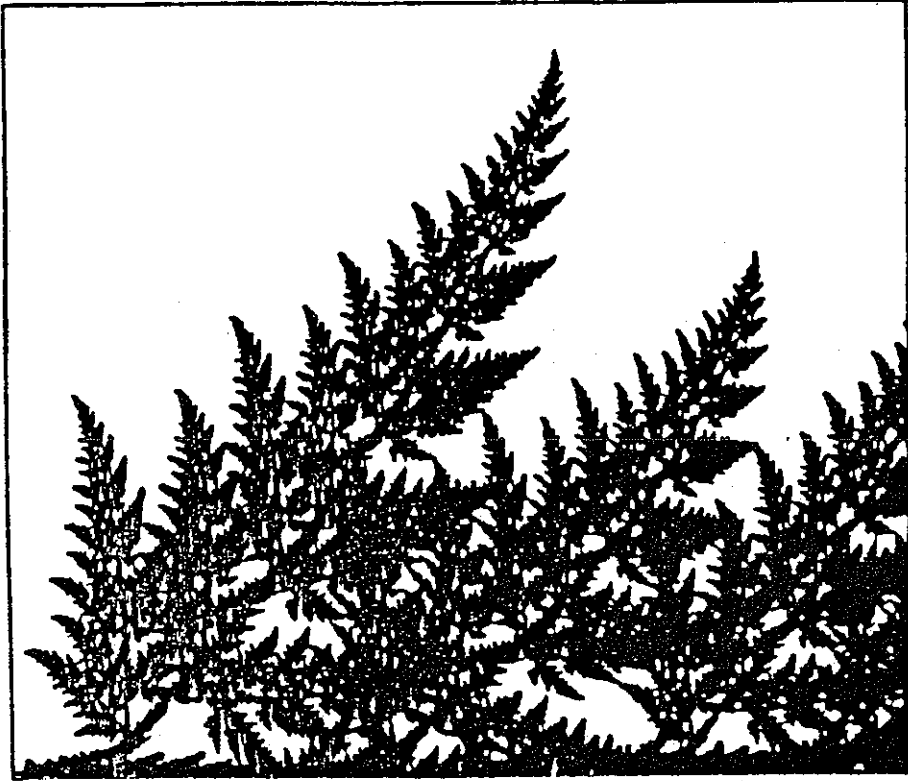


Figura 4.

$$k_i \in \{1, 2, \dots, n\}$$

$$\text{onde } \pi(m, \bar{k}, x) = f_{k(1)} \circ f_{k(2)} \circ \dots \circ f_{k(m)}(x).$$

A idéia da demonstração deste teorema é que, como no exemplo, sabemos que $F^n(K)$ cobre Λ com pequenos compactos de diâmetros $\leq \frac{1}{\lambda^n} \text{diam}(K)$.

4. SIF com Probabilidades.

Vemos então que o nosso algoritmo nos dá uma "caricatura" de Λ ; nosso objetivo agora é mostrar que a imagem da tela do computador nos dá uma boa aproximação de *todo* o Λ . Isto se

deve à nossa escolha aleatória das transformações T_k . Se tivéssemos escolhido uma seqüência fixa de T_k 's, só veríamos pequenas manchas ao redor de uma quantidade finita de pontos. Explicaremos isto no teorema a seguir. Como não temos uma referência a dar, o provaremos.

Antes precisamos de certas observações e alguma nomenclatura: Seja $\Sigma := \{(x_k)_{k \in \mathbb{Z}} / x_k \in \{1, \dots, n\}, \forall k\} = \prod_{-\infty}^{+\infty} \{1, \dots, n\}$, o espaço das seqüências bi-infinitas de números em $\{1, \dots, n\}$. A cada seqüência de Σ pode-se fazer corresponder uma escolha das transformações f_k do SIF. Σ se chama *O Espaço de Shift*.

Um *cilindro* de Σ é um conjunto da forma $[x_m, x_{m+1}, \dots, x_{m+k}]_m^{m+k} := \{\bar{y} = (y_j)_{j \in \mathbb{Z}} \in \Sigma / y_m = x_m, y_{m+1} = x_{m+1}, \dots, y_{m+k} = x_{m+k}\}$. Atribuindo a $\{1, \dots, n\}$ a topologia discreta e a Σ a topologia produto, sendo produto de compactos, Σ é compacto (Teorema de Tychonoff, [6]). A coleção dos cilindros é uma base para esta topologia que é metrizável com a distância

$$d(\bar{x}, \bar{y}) = \frac{1}{2^k}, \quad k = \max\{r \geq 0 / (x_{-r}, \dots, x_r) = (y_{-r}, \dots, y_r)\}.$$

O *shift positivo* é $\Sigma^+ = \{(x_k)_{k \geq 0} / x_k \in \{1, \dots, n\} \forall k\} = \prod_0^{\infty} \{1, \dots, n\}$ e o *shift negativo* é $\Sigma^- = \prod_{-1}^{-\infty} \{1, \dots, n\}$. Eles, com a métrica e a topologia induzidas de Σ , também são compactos.

Proposição 3. *Existe uma aplicação $\pi = \Sigma^+ \rightarrow \Lambda$ tal que*

(a) *π é dada por $\pi(\bar{x}) = \lim_{m \rightarrow \infty} f_{x(0)} \circ f_{x(1)} \circ \dots \circ f_{x(m)}(z)$ onde $z \in K$ é fixo e arbitrário.*

(b) *π é contínua e sobrejetora.*

(c) *Quando os $f_1(K), f_2(K), \dots, f_n(K)$ são disjuntos, π é injetiva.*

Prova: A existência do limite em (a) decorre do fato de que os conjuntos $f_{x(0)}(\dots f_{x(m)}(K))$ formam uma seqüência encaixante de compactos e $\text{diam } f_{x(0)}(\dots (f_{x(m)}(K))) \leq \lambda^m \text{diam}(K)$; daqui também decorre que π é contínua pois leva m -cilindros em conjuntos de diâmetro $\leq \lambda^m \text{diam}(K)$. Da existência do limite e da parte (c) do Teorema 1 decorre que $\pi(\Sigma) = \Lambda$, π é sobrejetora porque como $F^{-1}(\Lambda) = \Lambda$, para cada ω , existe $\omega_0 \in \Lambda$ e $f_{k_0} \in F$

com $f_{k_0}(\omega_0) = \omega$, e assim seqüências $\omega_0, \omega_1, \dots, \omega_m, \dots \in \Lambda$ e $f_{k_0}, f_{k_1}, \dots \in F$, com $f_{k_m}(\omega_m) = \omega_{m-1}$; por argumentos similares aos anteriores, $\omega = \lim_{m \rightarrow \infty} f_{k_0} \circ f_{k_1} \circ \dots \circ f_{k_m}(\omega_m) = \lim_{m \rightarrow \infty} f_{k_0} \circ f_{k_1} \circ \dots \circ f_{k_m}(z) = \pi(k_0, k_1, \dots)$. ■

Nos interessa saber que o algoritmo realmente dá uma figura aproximada de Λ . Se, por exemplo, a escolha das transformações for $f_1 f_2 f_1 f_2 \dots$, correspondente à seqüência periódica 1212..., então obteríamos na tela só uma seqüência de pontos acumulando nos pontos fixos das contrações $f_1 \circ f_2$ e $f_2 \circ f_1$.

O algoritmo corresponde ao seguinte processo (estocástico)

$$z \mapsto \text{Algum } \{f_1(z), f_2(z), \dots, f_n(z)\} \quad [P]$$

escolhidos com probabilidades p_1, p_2, \dots, p_n

i.e. obtemos uma seqüência $\langle z_k \rangle_{k \geq 0}$ de variáveis aleatórias com $\text{Prob}[z_{k+1} = f_i(z) / \text{dado que } z_k = z] = p_i$. Indicamos por $P(A|x)$ a probabilidade de que $z_{k+1} \in A$ dado que $z_k = x$, e chamamos a $P(\cdot|x)$ "probabilidade de transição". Quando, como neste caso, $P(z_k \in A / z_{k-1} = x_{k-1}) = P(z_k \in A / z_{k-1}, z_{k-2}, \dots, z_0 = x_{k-1}, x_{k-2}, \dots, x_0)$, a seqüência de v.a. $\langle z_k \rangle_{k \geq 0}$ se chama "Processo Estocástico de Markov".

Teorema 4. Para o processo de Markov $[P]$ $\langle z_k \rangle_{k \geq 0}$, com z_0 arbitrário e probabilidades de transição $P(z_{k+1} = f_i(\omega) / z_k = \omega) = p_i$, $i = 1, \dots, n$; temos que a probabilidade do evento [pontos de acumulação de $\langle z_k \rangle_{k \geq 0} = \Lambda$] é igual a 1.

Quer dizer, a probabilidade de que nosso algoritmo não funcione é 0. Falaremos mais disto nos comentários finais.

Agora vem as observações e alguns pre-requisitos. Note a diferença entre as seqüências consideradas no teorema 2d e na proposição 3a:

$$[+] \quad f_{x_0} \circ f_{x_1} \circ \dots \circ f_{x_m}(z),$$

e as seqüências que aparecem no algoritmo:

$$[-] \quad f_{x_m} \circ f_{x_{m-1}} \circ \dots \circ f_{x_1}(z).$$

(Brinque com o exemplo 1). Enquanto no primeiro caso $\langle f_{x_0} \circ f_{x_1} \circ \dots \circ f_{x_m}(K) \rangle_{m \geq 1}$ é uma coleção de compactos encaixados cada vez menores, todos dentro de f_{x_i} , por exemplo, $f_{x_m} \circ f_{x_{m-1}} \circ \dots \circ f_{x_1}(z)$

é um ponto que, quando m cresce, se move por distintos lugares em K . A idéia é modelar composições da forma $\{+\}$ por Σ^+ e composições da forma $\{-\}$ por Σ^- .

Precisamos da aplicação "shift" $\sigma: \Sigma \rightarrow \Sigma, (\sigma(\bar{x}))_j = x_{j-1}$, para $\bar{x} = (x_j)_{j=1}^\infty$, que empurra a seqüência um lugar para a direita. Podemos falar de $\sigma: \Sigma^- \rightarrow \Sigma^-, (\sigma(\bar{x}))_j = x_{j-1}$, que além de empurrar a seqüência, trunca o que ficaria no lugar $j = 0$. Notemos que $\sigma: \Sigma^- \leftarrow$ não é injetiva; de fato, para cada $\bar{x} \in \Sigma^-, \sigma^{-1}\{\bar{x}\}$ consiste em n seqüências da forma $(y_k)_{k=-1}^\infty$ com y_{-1} arbitrário e $y_k = x_{k+1}, k = -2, -3, \dots$

1. Modelo do Processo.

Uma escolha de uma seqüência de contrações no processo será modelada por um cilindro $[\bar{x}]_{-\infty}^{-1} = [\dots, x_{-2}, x_{-1}]_{-\infty}^{-1} \subset \Sigma$. A idéia é que a parte de índices negativos modela o futuro de z_0 : as contrações que serão aplicadas para obter os z_k 's, e a parte positiva modela a posição de z_0 por meio de $\pi: \Sigma^+ \rightarrow \Lambda$; como o teorema diz que z_0 pode ser arbitrário, consideramos cilindros que deixam a parte positiva livre. Modelamos o processo pelo shift do seguinte modo: $\sigma^m[\bar{x}] = [\dots, \underline{x_{-m}}, x_{-m+1}, \dots, x_{-1}] \subset \Sigma$, onde o lugar $(\sigma^m \bar{x})_{j=0}$ está sublinhado. A parte de índices positivos $\sigma^m[\bar{x}] \cap \Sigma^+ = [x_{-m}, \dots, x_{-1}] \stackrel{\text{def.}}{=} (\sigma^m[x])^+$ corresponderá a $z_m = f_{x_{-m}} \circ f_{x_{-m+1}} \circ \dots \circ f_{x_{-1}}(z_0)$ no processo. Para $z_0 = \pi(\bar{x}) \in \Lambda, \bar{x} \in \Sigma^+$ fixado, nos será útil a aplicação $\tau_{\bar{x}}: \Sigma^- \rightarrow \{\text{Resultados do Processo}\}, \tau_{\bar{x}}(\bar{y}) = \langle z_m \rangle_{m \geq 0}$ onde $z_0 = \pi(\bar{x}), z_m = \pi(\sigma^m(\bar{y}\bar{x})^+) = \pi(y_{-m}y_{-m+1} \dots y_{-1}x_0x_1 \dots) = f_{y_{-m}} \circ f_{y_{-m+1}} \circ \dots \circ f_{y_{-1}}(z_0)$. Definimos $\tau: \Sigma \rightarrow \{\text{Resultados do Processo}\}$, por $\tau(\bar{y}\bar{x}) = \tau_{\bar{x}}(\bar{y})$ para $\bar{y} \in \Sigma^-, \bar{x} \in \Sigma^+$.

Agora definimos uma probabilidade nos borelianos de Σ^- dando a medida dos cilindros de Σ^- : $\mu[x_m, x_{m+1}, \dots, x_{m+k}]_{m-1}^{m+k} = p_{x_m} p_{x_{m+1}} \dots p_{x_{m+k}}$. Uma propriedade importante desta medida é a de ser invariante por σ , isto é, $\mu(\sigma^{-1}(A)) = \mu(A)$ para todo boreliano A . Com efeito, basta mostrá-lo para cilindros e nesse caso temos $\sigma^{-1}[x_m, x_{m+1}, \dots, x_{m+k}]_{m-1}^{m+k} = [x_m, x_{m+1}, \dots, x_{m+k}]_{m-1}^{m+k-1}$, que tem a mesma medida. O terno (Σ, σ, μ) se chama "shift de Bernoulli".

O ponto fundamental do teorema é agora uma simples observação; no nosso processo [P], as probabilidades das escolhas de contrações correspondem à probabilidade μ . Quer dizer, a probabili-

dade de escolher f_{k_1}, f_{k_2}, \dots nessa ordem, com $k_m = x_m, k_{m+1} = x_{m+1}, \dots, k_{m+r} = x_{m+r}$, é a medida $\mu([x_m, x_{m+1}, \dots, x_{m+r}]_m^{m+r}) = p_{x_m} p_{x_{m+1}} \dots p_{x_{m+r}}$.

Vamos precisar de um lema da teoria de probabilidades. Uma seqüência de eventos $\langle A_m \rangle_{m \in \mathbb{N}}$ é independente se para toda subsequência finita $A_{m_1}, A_{m_2}, \dots, A_{m_k}$, temos que a probabilidade $P(A_{m_1} \cap A_{m_2} \cap \dots \cap A_{m_k}) = P(A_{m_1})P(A_{m_2}) \dots P(A_{m_k})$. Vamos chamar de *posto de um cilindro* $[x_m, x_{m+1}, \dots, x_{m+k}]_m^{m+k}$ ao conjunto $\{m, m+1, \dots, m+k\}$. É fácil ver que uma coleção de cilindros cujos postos são disjuntos é independente.

Lema de Borel-Cantelli ([5] 4.4).

(a) Se $\sum_n P(A_n)$ converge, então $P\left[\bigcap_{m=1}^{\infty} \bigcap_{k=m}^{\infty} A_k\right] = 0$.

(b) Se $\{A_n\}_{n \geq 1}$ é uma seqüência independente de eventos e

$\sum_n P(A_n)$ diverge, então $P\left[\bigcap_{m=1}^{\infty} \bigcup_{k=m}^{\infty} A_n\right] = 1$.

Indiquemos por $A_c(B)$ o conjunto de pontos de acumulação de B .

Demonstração do Teorema 4: Para uma seqüência $\langle z_k \rangle_{k \geq 0}$ do processo, seja $\bar{x} = (x_j)_{j \geq 1} \in \Sigma$ dado por $f_{x_j}(z_{j-1}) = z_j$. Como $\text{diam } f_{x_m} \circ f_{x_{m-1}} \circ \dots \circ f_1(K) \leq \lambda^m \text{diam } K$, temos que os pontos de acumulação

$$\begin{aligned} & \text{Ac}\langle f_{x_m} \circ f_{x_{m-1}} \circ \dots \circ f_1(K) \rangle_{m \geq 1} \\ &= \text{Ac}\langle f_{x_m} \circ f_{x_{m-1}} \circ \dots \circ f_1(\omega) \rangle_{m \geq 1} \end{aligned}$$

para quaisquer $z, \omega \in K$. Logo podemos escolher $z_0 \in \Lambda$ e bastará provar que $\text{Prob}(E) = 1$, com $E := [\langle z_k \rangle_{k \geq 0} \in \text{Processo}[P] / \text{Ac}\langle z_k \rangle_{k \geq 0} = \Lambda]$. Queremos caracterizar E . Como π é sobrejetiva, podemos supor $z_0 = \pi(\bar{x})$, $\bar{x} \in \Sigma^+$. Seguindo o nosso modelo, a cada seqüência de valores corresponde uma seqüência $\bar{y} \in \Sigma^-$, tal que $\tau(\bar{y}\bar{x}) = \langle z_k \rangle_{k \geq 0}$, onde $z_m = \sigma^m(\bar{y}\bar{x})^+$. Como $\pi: \Sigma^+ \rightarrow \Lambda$ é sobrejetiva e contínua, ela leva conjuntos $D \subset \Sigma^+$, com $\text{Ac}(D) = \Sigma$, em $\pi(D) \subset \Lambda$ com $\text{Ac}(\pi(D)) = \Lambda$. Logo $\tau(A) \subset E$, onde:

$$\begin{aligned} A &= \{\bar{y} \in \Sigma^- / \text{Ac}\langle \sigma^m(\bar{y}\bar{x})^+ \rangle_{m \geq 0} = \Sigma^+\} \\ &= \{\bar{y} \in \Sigma^- / \forall U \text{ cilindro de } \Sigma^+, \forall M > 0 \exists \\ & \quad m > M : \sigma^m(\bar{y}\bar{x})^+ \in U\}, \end{aligned}$$

pois $\mathcal{C} = \{U_r\}_{r \in \mathbf{N}}$, a coleção de cilindros (de posto finito) em Σ^+ , é uma base de Σ^+ . Como $\sigma^m(\bar{y}\bar{x})^+ = (\sigma^m\bar{y})^+(\sigma^m\bar{x}) \in U$, quando m é grande ($> \text{máx}(\text{posto de } U)$), equivale a $(\sigma^m\bar{y})^+ \in U$, ou $\bar{y} \in \sigma^{-m}U$, temos

$$\begin{aligned} A &= \bigcap_{r \in \mathbf{N}} \bigcap_{M \in \mathbf{N}} \bigcup_{m \geq M} \sigma^{-m}(U_r) \\ &\supseteq \bigcap_{r \in \mathbf{N}} \bigcap_{M \in \mathbf{N}} \bigcup_{Rm > M} \sigma^{-Rm}(U_r) \end{aligned}$$

onde $R := (\text{comprimento do posto de } U) + 2$. Mas agora os eventos $\sigma^{-Rm}(U)$, $m \in \mathbf{Z}$ são independentes para a nossa probabilidade μ (são cilindros com postos disjuntos), e como μ é invariante por σ : $\mu(\sigma^{-Rm}(U)) = \mu(U)$. Assim $\sum_{m > M/R} \mu(\sigma^{-Rm}(U)) = \infty$, e o

lema de Borel-Cantelli se aplica para obter $\mu(A) = 1$. ■

5. SIF's e Teoria Ergódica.

Consideremos um processo estocástico de Markov [M]: $\langle z_n \rangle_{n \geq 0}$ em um espaço métrico compacto K com probabilidade de transição $P(\cdot|x)$, $x \in K$.

Uma *medida invariante* para o processo (*distribuição estacionária*, na linguagem de probabilidade) é uma medida de probabilidade μ , nos borelianos de K , que satisfaz

$$\mu(A) = \int_K P(A|x) d\mu(x), \quad \forall A \text{ boreliano de } K.$$

Consideremos o espaço

$$\Sigma_M = \{(x_n)_{n \geq 0} / x_n \in K, x_{n+1} \in \text{suporte de } P(\cdot|x_n)\} \subset K^{\mathbf{N}}$$

e outra aplicação shift $s: \Sigma_M \leftarrow$, $(s(\bar{x}))_n = x_{n+1}$, $\bar{x} = (x_n)_{n \geq 0}$, que empurra as seqüências para a esquerda. Para $A_n, \dots, A_{n+r} \subset K$, um cilindro $[A_n, \dots, A_{n+r}]_n^{n+r} \stackrel{\text{def}}{=} \{(x_m)_{m \geq 0} \in \Sigma_M / x_j \in A_j, j = n, \dots, n+r\}$. Temos que os cilindros $[x, A]_{k-1}^k$, $x \in K$, formam uma partição "mensurável" do cilindro $[A]_k^k$.

Se temos uma medida de probabilidades ν nos borelianos de Σ_M , ν restrita a $[A]_k^k$: $\nu|_{[A]_k^k}$ se decompõe em uma série de *medidas*

condicionais ([5] ou [11]) ν_x , $x \in K$ ($\nu - qtpx$) que são únicas ($\nu - qtp$) e definidas por

(a) ν_x probabilidade em $[x, A]_{k-1}^k$, $\nu - qtpx$

(b) $\forall B \subset A$ $\nu([B]_k^k) = \int_K \nu_x(B) d(\nu|_{[A]_k})(x)$

A unicidade significa que se temos outro sistema de probabilidades condicionais η_x (definidas $\nu - qtpx$) que satisfaz (a) e (b), então $\nu_x = \eta_x$, $\nu - qtpx$.

Se temos, por acaso, que $\nu_x = P(\cdot|x)$ e ν é invariante pelo shift i.e. $\nu(s^{-1}A) = \nu(A)$, $\forall A$ boreliano de Σ_M , então a invariância de ν diz que

$$\nu(s^{-1}[A]_n^n) = \nu([A]_{n+1}^{n+1}) = \int_K \nu_x(A) d(\nu|_{[K]_n^n}(x))$$

i.e., $\mu(A) = \nu([A]_n^n)$, para qualquer n , nos dá uma medida invariante pelo processo. Inversamente, podemos obter ν a partir de μ : fazendo $\mu(A) = \nu([A]_n^n)$, e pela unicidade das medidas condicionais obtemos $\nu_x(A) = P(A|x)$. Agora que temos uma medida invariante por uma transformação mensurável podemos usar toda a maquinaria da Teoria Ergódica. Esta observação foi-me dita por Maria Eulalia Vares (IMPA).

Se ν é ergódica, o teorema de Birkhoff diz que $\frac{1}{n} \sum_{k=1}^{n-1} \delta_{\sigma^k(\bar{x})} \xrightarrow{w^*} \nu$, $\nu - qtp\bar{x} \in \Sigma_M$. A "moral" do Teorema 6 é que a distribuição estacionária descreve as probabilidades do processo, como o fazia a medida da seção 4.

Um subconjunto $A \subset K$ é *invariante* se para todo $x \in A$, $P(A|x) = 1$. Tal conjunto é modelado por $\hat{A} = A^{\mathbb{N}} \cap \Sigma_M$ que é invariante pelo shift: $s(\hat{A}) = \hat{A}$ e satisfaz $q_k(\hat{A}) = A \pmod{0}$ para todo k , onde $q_k: \Sigma_M \subset \prod_0^{\infty} K \rightarrow K$ é a projeção sobre o k -ésimo fator. A *distribuição estacionária* μ é *ergódica* se todo subconjunto mensurável invariante de K , tem medida 0 ou 1. Isto equivale a dizer que a probabilidade associada ν no shift é ergódica (a medida de um conjunto invariante em Σ_M é 0 ou 1).

As probabilidades de transição do processo [P] podem ser escritas como

$$P(\cdot|x) = \sum_{i=1}^n p_i \delta_{f_i(x)}$$

onde $\delta_y(A) = 1$ se $y \in A$, $\delta_y(A) = 0$ se $y \notin A$ (δ_y é a medida de Dirac concentrada em y). Para uma função contínua $\varphi: K \rightarrow \mathbb{C}$ seja $T\varphi: K \rightarrow \mathbb{C}$, definida por

$$(T\varphi)(x) = \sum_{i=1}^n p_i \varphi(f_i(x))$$

Definição. Um SIF com probabilidade é contínuo se T tem a propriedade $T: \mathcal{C}(K) \rightarrow \mathcal{C}(K)$ onde $\mathcal{C}(K) = \{\varphi: K \rightarrow \mathbb{C} / \varphi \text{ e cont\u00ednua}\}$.

Por exemplo, o SIF do exemplo 1 é cont\u00ednuo para qualquer vetor de probabilidades $\bar{p} = (p_1, p_2)$. O SIF (K, F, p) onde $K = \{z \in \mathbb{C} / \frac{1}{4} \leq |z| \leq 4\}$, $F = \{f_1, f_2\}$ onde f_1, f_2 s\u00e3o as solu\u00e7\u00f5es de $[f_i(z)]^2 = z$, s\u00f3 \u00e9 cont\u00ednuo para $p = (\frac{1}{2}, \frac{1}{2})$.

Teorema 5 [2]. Seja (K, \bar{F}, \bar{P}) um SIF cont\u00ednuo, ent\u00e3o para o processo $[P]$:

- (a) Existe uma \u00fanica distribui\u00e7\u00e3o estacion\u00e1ria μ ;
- (b) μ \u00e9 atratora i.e. $(T^n)^* \nu \xrightarrow{w^*} \mu$ quando $\nu \rightarrow \infty$, na topologia (fraca)*, para toda medida de probabilidade ν em K ;
- (c) Para todo $\varphi \in \mathcal{C}(K)$: $\|T^n \varphi - \int_K \varphi d\mu\| \rightarrow 0$ quando $n \rightarrow \infty$ onde $\|\varphi\| = \sup_{x \in K} |\varphi(x)|$;
- (d) Suporte $(\mu) \subset \Lambda$
- (e) μ \u00e9 "mixing" no seguinte sentido: $\lim_n \int_K (T^n \varphi) \psi d\mu =$

$$\left(\int_K \varphi d\mu \right) \left(\int_K \psi d\mu \right).$$

Em particular, μ \u00e9 erg\u00f3dica.

A id\u00e9ia da demonstra\u00e7\u00e3o \u00e9 observar primeiro que a medida μ_Σ do shift de Bernoulli (Σ^+, \bar{p}) definida por $\mu_\Sigma([x_n, \dots, x_{n+k}]_n^{n+k}) = p_n \dots p_{n+k}$ \u00e9 a tal μ para o SIF (Σ^+, U, \bar{p}) com $U = \{u_1, \dots, u_n\}$, $u_i(\bar{x}) = (i, \bar{x}) \in \Sigma^+$, $\bar{x} \in \Sigma^+$. Em seguida usar a proje\u00e7\u00e3o $\pi: \Sigma^+ \rightarrow \Lambda$, $\pi(\bar{x}) = \lim_m f_{x_1} \circ f_{x_2} \circ \dots \circ f_{x_m}(z)$, do teorema 2. A diferen\u00e7a entre Σ^+ e Λ \u00e9 que π n\u00e3o \u00e9 sobrejetiva quando os $f_i(K)$ n\u00e3o s\u00e3o disjuntos. Mas isto \u00e9 salvo pela condi\u00e7\u00e3o de continuidade do teorema que, usando as propriedades de π , permite ver que $\pi^* \mu_\Sigma$ \u00e9 a μ deste teorema em $(\Lambda, \bar{F}, \bar{p})$.

Não sei o que acontece quando (Λ, F, \bar{p}) não é contínuo. Isto é interessante porque quando se usa o teorema da colagem é moroso fazer com que o sistema seja contínuo. O algoritmo desenharia Λ (com probabilidade 1), mas não nos daria necessariamente uma idéia das distribuições estacionárias (se existirem).

O teorema 5 diz que para $z \in K$ arbitrário e δ_z a medida de Dirac em z , $T^{n*}\delta_z \rightarrow \mu$. Notemos que $T^*\delta_z = \sum_{i=1}^n p_i \delta_{f_i(z)}$, $T^{2*}\delta_z = \sum_{i,j=1}^n p_i p_j \delta_{f_i \circ f_j(z)}$, $T^{m*}\delta_z$ é a medida suportada em $F^m(z)$ (que tem $\leq n^m$ pontos) dando os pesos correspondentes a cada um dos seus pontos. Aproximar μ por este método em um computador requeriria gerar a árvore $\cup\{F^m(z)/0 \leq m \leq M\}$ (M grande) da órbita de z . Nosso algoritmo só toma um ponto de cada $F^m(z)$; em princípio ao formar-se Λ na tela, veríamos apenas uma caricatura grosseira de μ . O seguinte teorema diz que não é este o caso.

Teorema 6. *Seja (Λ, F, \bar{p}) um SIF contínuo e consideremos o processo $[P] \langle z_n \rangle_{n \geq 0}$. Então, com probabilidade 1. Temos*

$$\frac{1}{m} \sum_{k=0}^{m-1} \delta_{z_k} \xrightarrow{w^*} \mu.$$

Demonstração: Indiquemos por μ_Λ a distribuição estacionária do processo $\langle z_k \rangle_{k \geq 0}$. Observemos que pelos argumentos do teorema 3, basta fixar um $z_0 = \pi(\bar{y}) \in \Lambda$ com $\bar{y} \in \Sigma^+$, $\pi: \Sigma^+ \rightarrow \Lambda$. Consideremos a medida μ do shift de Bernoulli (Σ^+, \bar{p}) . Pelo teorema 5 ela é a distribuição estacionária do processo (Σ^+, U, \bar{p}) . Consideremos a medida ν do shift de Bernoulli (Σ^-, \bar{p}) . Pelas considerações da seção 4 ela realiza as probabilidades das escolhas no processo $[P]$. Consideremos a medida η no shift completo, definida por

$$\eta([x_n, \dots, x_{n+k}]_n^{n+k}) = p_{x_n} \dots p_{x_{n+k}}.$$

η é ergódica e invariante pelo shift $(\sigma(\bar{x}))_n = x_{n-1}$. O Teorema de Birkhoff [8] diz que

$$\frac{1}{n} \sum_{k=0}^{n-1} \varphi(\sigma^k \bar{x} \bar{y}) \xrightarrow{n \rightarrow \infty} \int \varphi d\eta,$$

$\eta - qtp \bar{x}, \bar{y} \in \Sigma^- \times \Sigma^+ = \Sigma$, para toda $\varphi: \Sigma \rightarrow \mathbb{C}$ contínua. Isto equivale a

$$[*] \quad \frac{1}{n} \sum_{k=0}^{n-1} \delta_{\sigma^k \bar{x} \bar{y}} \xrightarrow{w^*} \eta, \quad \eta - qtp \bar{x}, \bar{y} \in \Sigma.$$

Notemos que $\nu(A) = \eta(A \times \Sigma^+)$ para todo boreliano $A \subset \Sigma^-$. Logo, fixado y , $\eta - qtp \bar{x} \bar{y}$ é $\nu - qtp \bar{x}$ (i.e. ν é a medida condicional de η com respeito à partição $\Sigma^- \times \{y\}$, $\bar{y} \in \Sigma^+$). Agora consideremos $\mathcal{P}: \text{Prob}(\Sigma) \rightarrow \text{Prob}(\Sigma^+)$, $(\mathcal{P}\lambda)(A) = \lambda(\Sigma^- \times A)$. Para $\text{Prob}(\Sigma)$, $\text{Prob}(\Sigma^+)$ com as topologias (fracas)*, \mathcal{P} é contínua. Aplicando \mathcal{P} à equação [*] obtemos

$$\frac{1}{n} \sum_{k=0}^{n-1} \delta_{(\sigma^k \bar{x} \bar{y})^+} \xrightarrow{w^*} \mathcal{P}(\eta) = \mu, \quad \nu - qtp \bar{x} \in \Sigma^-.$$

Agora aplicando $\pi^*: \text{Prob}(\Sigma^+) \rightarrow \text{Prob}(\Lambda)$, $\pi^* \lambda(A) = \lambda(\pi^{-1} A)$, temos

$$\frac{1}{n} \sum_{k=1}^{n-1} \delta_{z_k} \xrightarrow{w^*} \mu_\Lambda \quad \text{para quase todo } \langle z_k \rangle_{k \geq 0}$$

$$z_k = \pi(\sigma^k \bar{x} \bar{y})^+ = \lim_{m \rightarrow \infty} f_{x_k} \circ \dots \circ f_{x_1} \circ f_{y_0} \circ \dots \circ f_{y_n}(z_0)$$

É importante repetir que o algoritmo desenha uma aproximação de Λ independentemente das probabilidades p_i . As probabilidades fazem com que o algoritmo também “desenhe” uma aproximação de μ .

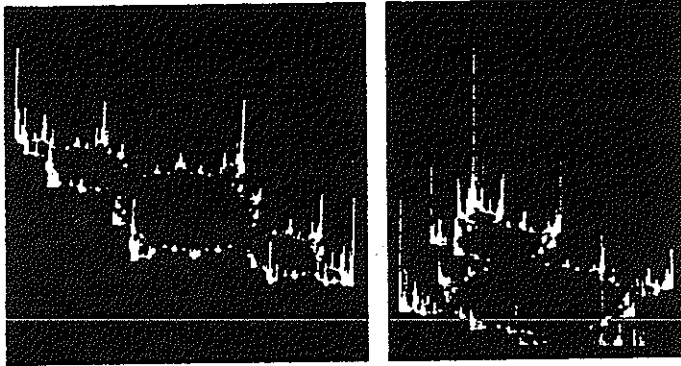


Figura 5. Distribuição estacionária μ (medida de máxima entropia) do conjunto de Julia de $f(z) = z^2 - 1$.

6. O Teorema da Colagem.

Queremos enfrentar o seguinte problema: dada uma imagem \mathcal{F} , obter um SIF que a reproduza ao grau de fidelidade desejado. O próximo teorema nos dará a resposta. O método é como segue: escolhe-se contrações (que podem ser afins) T_j tais que $\cup_j T_j(\mathcal{F})$ aproxima \mathcal{F} , onde \mathcal{F} é a figura original a desenhar-se. O atrator do SIF será tanto mais parecido à figura original quanto $\cup_j T_j(\mathcal{F})$ esteja perto de \mathcal{F} . Compare-se com a caracterização do atrator no teorema 2b. Usando mais contrações T_k , talvez mais contrativas, pode-se fazer $\cup_j T_j(\mathcal{F})$ tão próximo a \mathcal{F} quanto se queira.

A formalização disto precisa de mais um conceito: dados dois compactos A e B de K , a *distância de Hausdorff* entre eles é

$$h(A, B) = \max_{x \in A} \min_{y \in B} d(x, y) + \max_{y \in B} \min_{x \in A} d(x, y).$$

Se todas as transformações T_k satisfazem

$$\forall k \quad d(T_k(x), T_k(y)) \leq \lambda d(x, y), \quad 0 < \lambda < 1$$

temos o seguinte

Teorema 7 (da Colagem) [3]. *Dado um compacto $\mathcal{F} \subset K$, se se escolhe as contrações $\{T_1, \dots, T_n\}$ de modo que*

$$h(\mathcal{F}, \bigcup_{i=1}^n T_i(\mathcal{F})) < \epsilon,$$

para algum $\epsilon > 0$, então

$$h(\mathcal{F}, \Lambda) < \frac{\epsilon}{1 - \lambda}$$

onde Λ é o atrator do SIF $(K, \{T_1, \dots, T_n\})$.

Demonstração: Sejam $H(K) := \{A \subset K/A \text{ fechado}\}$ com a distância de Hausdorff e $F: H(K) \rightarrow H(K)$ dada por $F(A) = \bigcup_{i=1}^n T_i(A)$. Vejamos que F é uma contração, i.e., $h(F(A), F(B)) \leq \lambda h(A, B) \forall A, B \in K$:

$$\begin{aligned} h(F(A), F(B)) &= \sup_{x \in A, i} \inf_{y \in B, j} d(T_i(x), T_j(y)) \\ &\quad + \sup_{y \in B, i} \inf_{x \in A, j} d(T_i(y), T_j(x)) \\ &\leq \sup_{x \in A, i} \inf_{y \in B} d(T_i(x), T_i(y)) \\ &\quad + \sup_{y \in B, i} \inf_{x \in A} d(T_i(x), T_i(y)) \\ h(F^m(\mathcal{F}), \mathcal{F}) &\leq \sum_{k=1}^m h(F^k(\mathcal{F}), F^{k-1}(\mathcal{F})) \\ &\leq \sum_{k=1}^m \lambda^{k-1} h(F(\mathcal{F}), \mathcal{F}) \\ &\leq \frac{1 - \lambda^m}{1 - \lambda} h(F(\mathcal{F}), \mathcal{F}). \end{aligned}$$

Fazendo $m \rightarrow \infty$: $h(\Lambda, \mathcal{F}) \leq \frac{1}{1 - \lambda} h(F(\mathcal{F}), \mathcal{F})$. ■

Outra versão do Teorema da Colagem pode ser a seguinte: Consideremos $\mathcal{H}(K) = \{f: K \rightarrow K | f \text{ é homeomorfismo sobre sua imagem}\}$ com a métrica $d(f, g) = \sup_{x \in K} d(f(x), g(x))$. Consideremos

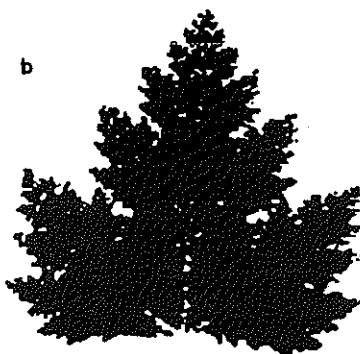
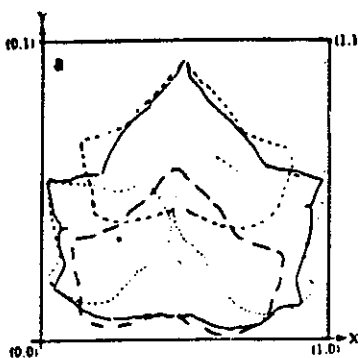


Figura 6. Exemplo do Teorema da Colagem

a métrica em $\prod_{i=1}^n \mathcal{H}(K)$, dada por $d((f_1, \dots, f_n), (g_1, \dots, g_n)) = \max_{i=1, \dots, n} d(f_i, g_i)$.

Teorema 8 (da Estabilidade Estrutural). *Seja (K, F) , $K = [0, 1]^d$, um SIF tal que $f_1(K), \dots, f_n(K)$ são disjuntos e seja $0 < \lambda < 1$ tal que $d(f_i(x), f_i(y)) < \lambda d(x, y)$, para todo $i = 1, \dots, n$ e todos $x, y \in K$. Então existe uma vizinhança $\mathcal{U} \subset \prod_{i=1}^n \mathcal{H}(K)$ de $F = (f_1, \dots, f_n)$ tal que para cada $G = (g_1, \dots, g_n) \in \mathcal{U}$, existe*

um homeomorfismo $h_G: K \rightarrow K$ tal que -

$$[**] \quad h_G \quad \begin{array}{ccc} K & \xrightarrow{f_i} & K \\ \downarrow & & \downarrow \\ K & \xrightarrow{g_i} & K \end{array} \quad h_G \text{ comuta para todo } i = 1, \dots, n.$$

Além disso $\|h_G - id_K\| \leq \frac{2}{1-\lambda} \|G - F\|$ e $h_G(\Lambda_F) = \Lambda_G$, o atrator do SIF (K, G) . Em particular $h(\Lambda_G, \Lambda_F) \leq \frac{4}{1-\lambda} \|G - F\|$.

Prova: Tomemos uma vizinhança \mathcal{U} de F tal que para todo $G = (g_1, \dots, g_n) \in \mathcal{U}$, os $g_i(K)$'s são disjuntos e $d(g_i(x), g_i(y)) < \lambda d(x, y)$ para todo $i = 1, \dots, n$ e todos os $x, y \in K$. Como os f_i, g_i são homeomorfismos sobre suas imagens, e os $f_i(K)$'s (e os $g_i(K)$'s) são disjuntos: $f_i(K) \stackrel{\text{homeo}}{\approx} g_i(K) \stackrel{\text{homeo}}{\approx} K$, e existe um homeomorfismo $h_G|_R$ entre $R \stackrel{\text{def}}{=} K - \bigcup_{i=1}^n f_i(K)$ e $S = K - \bigcup_{i=1}^n g_i(K)$

que satisfaz **[**]** em ∂K , e $d(h_G(x), x) \leq 2\|F - G\| \stackrel{\text{def}}{=} a$. Agora note-se que $K = \cup f_{i_1} \circ f_{i_2} \circ \dots \circ f_{i_k}(R)$, onde a união é sobre todas as $(i_1, i_2, \dots, i_k) \in \{1, \dots, n\}^k$ para todo $k \geq 1$, e é uma união disjunta. Pode-se estender h_G a K por $h_G(f_{i_1} \circ \dots \circ f_{i_k}(x)) = g_{i_1} \circ \dots \circ g_{i_k}(h_G(x))$, $x \in R$. Deste modo h é uma bijeção contínua entre espaços métricos compactos e, portanto, um homeomorfismo. **[**]** é satisfeita imediatamente; também $h_G(\Lambda_F) = \Lambda_G$ pelo teorema 3.

Temos $d(f_i(x), h_G(f_i(x))) = d(f_i(x), g_i(h_G(x))) \leq d(f_i(x), f_i(h_G(x))) + d(f_i h_G(x), g_i h_G(x)) \leq \lambda a + a$ para $x \in R$. Indutivamente $d(f_{i_1} \circ \dots \circ f_{i_k}(x), h_G(f_{i_1} \circ \dots \circ f_{i_k}(x))) \leq (1 + \lambda + \dots + \lambda^k)a$. A caracterização anterior de K nos dá $\min_{y \in \Lambda_G} d(x, y) \leq d(x, h_G(x)) \leq \frac{2}{1-\lambda} \|F - G\|$ para obter $h(\Lambda_F, \Lambda_G) \leq \frac{4}{1-\lambda} \|F - G\|$. ■

Exemplos: As seguintes tábuas correspondem a SIF's (K, F, \bar{P}) com $\bar{P} = (p_1, \dots, p_n)$, $F = \{T_1, \dots, T_n\}$ da forma

$$T_i \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a_i & b_i \\ c_i & d_i \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} e_i \\ f_i \end{bmatrix}$$

1. Samambaia (Fig. 1)

Ti	a	b	c	d	e	f	p
1	0	0	0	0.16	0	0	0.01
2	0.2	-0.26	0.23	0.22	0	1.6	0.07
3	-0.15	0.28	0.26	0.21	0	0.11	0.07
4	0.85	0.01	-0.01	0.85	0	1.6	0.85

$$(K = [0, 5] \times [0, 10])$$

2. Folha de outono

Ti	a	b	c	d	e	f	p
1	0.6	0	0	0.6	0.18	0.36	0.25
2	0.6	0	0	0.6	0.18	0.12	0.25
3	0.4	0.3	-0.3	0.4	0.27	0.36	0.25
4	0.4	-0.3	+0.3	0.4	0.27	0.09	0.25

3. Triângulo de Sierpinski

Ti	a	b	c	d	e	f	p
1	0.5	0	0	0.5	0	0	0.33
2	0.5	0	0	0.5	1	0	0.33
3	0.5	0	0	0.5	0.5	0.5	0.34

4. Árvore Fractal

Ti	a	b	c	d	e	f	p
1	0	0	0	0.5	0	0	0.05
2	0.1	0	0	0.1	0	0.2	0.15
3	0.42	-0.42	0.42	0	0	0.2	0.4
4	0.42	0.42	-0.42	0.42	0	0.2	0.4

7. Conjuntos de Julia da família $z^2 + \lambda$.

Podem-se desenhar como o atrator de um SIF (K, F, \bar{p}) com $\bar{p} = (1/2, 1/2)$, $K \subset \mathbb{C}$, $F = (f_1, f_2)$ onde f_1, f_2 são as soluções de $z = f_i(z)^2 + \lambda$, $\lambda \in \mathbb{C}$ fixo. Neste caso o algoritmo aproxima a distribuição estacionária μ que é a "medida de máxima entropia" do conjunto de Julia. Algumas sugestões são:

- (a) $\lambda = -1$
- (b) $\lambda = 0.11031 - 0.67037i$, Poeira de Fatou
Janela: $K : -1.5 \leq \text{Re } z \leq 1.5, -1.5 \leq \text{Im } z \leq 1.5$
- (c) $\lambda = i$, "Dentrito"
- (d) $\lambda = 0.32 + 0.043i$, "Dragão".
Janela: $K : -2 \leq \text{Re } z \leq 2, -1.5 \leq \text{Im } z \leq 1.5$.
- (e) $\lambda = -0.194 + 0.6557i$
Janela: $K : -1,8 \leq \text{Re } z \leq 1,8, -1.5 \leq \text{Im } z \leq 1.5$

8. Mais Conjuntos de Julia.

Para uma aplicação racional (quociente de dois polinômios) $R(z) = \frac{p(z)}{q(z)}$, $R : \mathbb{C} \cup \{\infty\} \leftarrow$ sempre existe um "conjunto de Julia", J_R , onde a complexidade da dinâmica de R se concentra.

Toda aplicação racional induz um SIF (K, F, \bar{p}) (mas talvez $\lambda \geq 1$ em alguns lugares) para um $K \subset \mathbb{C} \cup \{\infty\}$ apropriado onde $F = \{f_1, \dots, f_n\}$, $d = \max\{\text{grau}(p), \text{grau}(q)\}$ e os f_i 's são as soluções de $R(f_i(z)) = z$ e $\bar{p} = (\frac{1}{d}, \dots, \frac{1}{d})$. Com este \bar{p} o sistema é contínuo. O atrator de (K, F, \bar{p}) é o conjunto de Julia J_R .

Se consideramos o processo $[P]$ para este sistema, então para todo $z_0 \in \mathbb{C} \cup \{\infty\} - \{\text{no máximo 3 pontos excepcionais}\}$, $\{z_k\}_{k \geq 0}$ se acumula em J_R . A distribuição estacionária de $[P]$ se chama, neste caso, medida de máxima entropia de J_R .

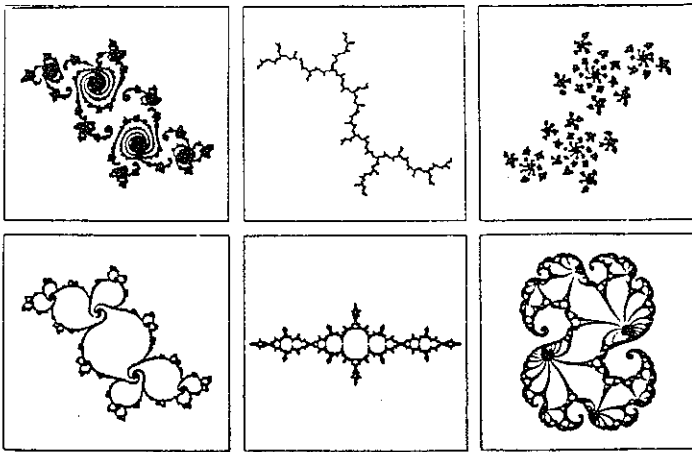


Figura 7. Alguns conjuntos de Julia da família $f_c(z) = z^2 + c$.

Normalmente em um SIF temos que esperar um tempo sem pintar para que os iterados do ponto inicial z_0 se aproximem do atrator. No caso dos conjunto de Julia podemos começar com um ponto que esteja no atrator (no Julia), sem ter que esperar para pintar. Para isto basta pegar como z_0 um ponto periódico repulsor i.e., $f^n(z_0) = z_0$, $|(f^n)'(z_0)| > 1$, para algum $n \geq 0$.

Excelentes referências para isto são [12], para a parte Teórica, e [9], para algoritmos de desenhos em Sistemas Dinâmicos.

9. Conjunto de Julia do método de Newton.

Para achar raízes de uma função real $f(x)$ geralmente se usa o método de Newton $x_{n+1} = x_n - f(x_n)(f'(x_n))^{-1}$. No caso em que $f(x)$ é um polinômio $f(x) = p(x)$, pode-se generalizar a mesma fórmula com $z \in \mathbb{C}$ para achar raízes complexas de $p(z)$. A primeira pergunta é se para qualquer escolha de z_0 , $\langle z_n \rangle$ converge a uma raiz. A resposta é não: Já no caso real não é difícil achar exemplos onde $\langle z_n \rangle$ possa aproximar-se a um ciclo finito. Visto em termos da aplicação racional $N(z) = z - \frac{p(z)}{p'(z)}$, este ciclo corresponde a uma órbita periódica atratora. Existe um teorema que diz que para todo polinômio de grau ≥ 3 sempre existem tais órbitas. Mas se o período é pequeno, elas são reconhecíveis em uma execução do método de Newton. As raízes de p sempre são pontos fixos "superatratores". Um ponto fixo atrator que não seja raiz é também reconhecível. Mas se $z_0 \in J_N$, $\langle z_n \rangle$ se mantém dentro de J_N e não converge a uma raiz. Por sorte em muitos casos J_N tem medida de Lebesgue nula i.e. tem-se probabilidade zero de começar com $z_0 \in J_N$.

Para aplicações racionais não se sabe com que generalidade J_R tem medida nula. A conjectura da estabilidade diz que $\text{Leb}(J_R) = 0$ é genérico.

Um conjunto de Julia interessante (Fig. 8) corresponde ao método de Newton para o polinômio $p(x) = x^3 - 1$. Ele soluciona o seguinte problema: um planeta de habitantes equitativos é dividido em tres países com áreas iguais de tal forma que cada ponto da fronteira de um deles é um ponto da fronteira das três. Uma tal fronteira é o conjunto de Julia. Ele contém os pontos $\infty, 0, x_1, x_2, x_3$ da esfera $\mathbb{C} \cup \{\infty\}$, onde $x_k = -\sqrt[3]{2} \exp(2\pi i k/3)$ são as soluções de $N(x_k) = 0$. A prova da solução é que $N(z) = D \circ N \circ D^{-1}(z)$ onde D é a rotação de 120° .

10. Comentários Diversos.

Vimos no teorema 4 que em princípio o algoritmo sempre desenha uma boa aproximação de Λ . Se assumirmos que o erro de computação ao calcular os f_i 's está sempre limitado por $\epsilon > 0$, i.e. $d(f_i(x), f_{i,\text{computador}}(X_{\text{computador}})) < \epsilon$, então por argumentos similares ao teorema 7 da colagem ou o teorema 8, podemos ver

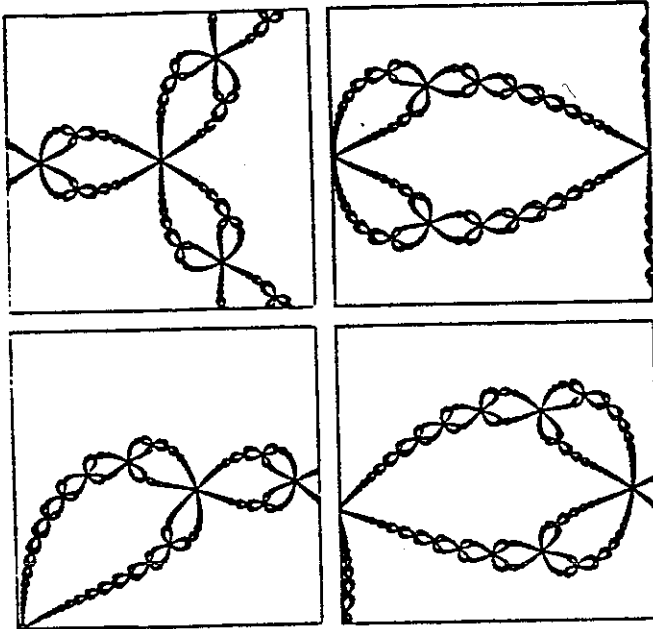


Figura 8. Algumas ampliações do conjunto de Júlia do método de Newton para $p(z) = z^3 - 1$.

que a distância

$$h(\Lambda_{\text{real}}, \Lambda_{\text{computador}}) < \frac{\epsilon}{1 - \lambda_{\text{comput.}}}$$

O teorema da colagem nos permite reconstruir uma figura dada. É interessante ver em [4] a reprodução de um retrato de uma menina peruana usando o teorema da colagem, com 120 contrações em 2000 bytes. O teorema da Estabilidade Estrutural ou o teorema da colagem dizem que dois SIF's com listas de contrações próximas entre si dão figuras (atratores) semelhantes: homeomorfas e próximas. Isto permite fazer uma biblioteca como vimos na introdução. Conjectura-se que objetos naturais de "classes" distintas (um pinheiro e um carvalho, por exemplo) correspondem a conjuntos de parâmetros $\left(\begin{bmatrix} a_i & b_i \\ c_i & d_i \end{bmatrix} \right)$ das contrações distantes; e que seria natural tomar como "iguais" figuras que estejam numa mesma vizinhança do espaço de parâmetros. É por isto que se pensa

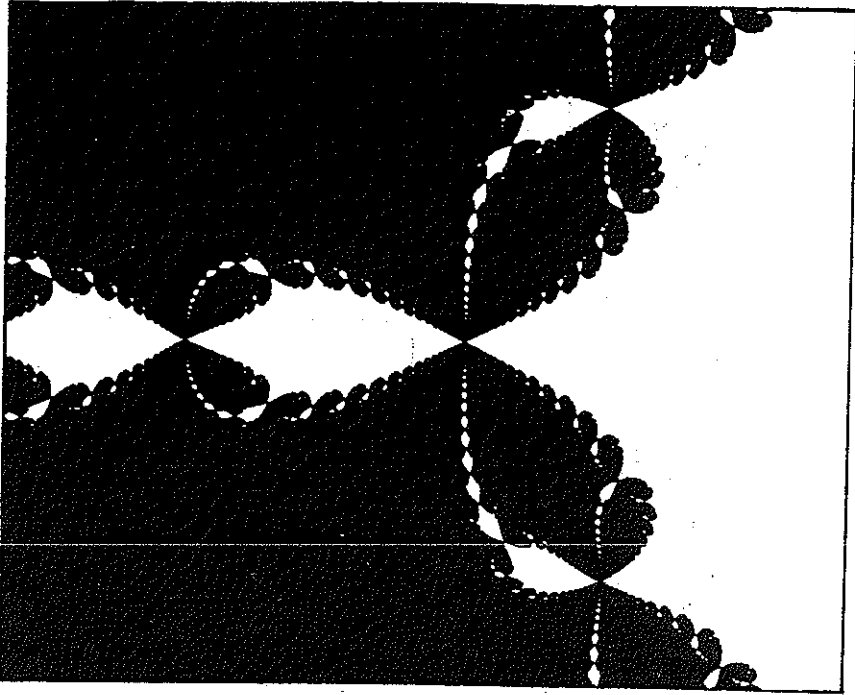


Figura 9. A bacia de atração da raiz $z_0 = 1$ aparece em branco.

que os SIF's podem ser úteis para o reconhecimento automático de figuras.

Vimos no teorema 6 que as probabilidades \bar{p} de um SIF não influem no fato de que o atrator Λ finalmente seja desenhado. Elas fazem com que os pontos que o computador obtém tenham certa distribuição (de probabilidade) em torno do Λ . Mas elas de fato influem no processo de pintura do Λ . Por exemplo, em um conjunto de Julia as "pontas exteriores" estarão mais carregadas destes pontos que as "pontas interiores" que de fato são difíceis de desenhar. São necessários muitos iterados para que um conjunto de Julia conexo apareça conexo no computador. Isto se deve ao fato de que a distribuição de probabilidade no conjunto de Julia corresponde à distribuição de carga elétrica de uma superfície cuja seção é esse conjunto. Geralmente se dá as probabilidades p_i proporcionais às áreas $\text{Área}(f_i(K))$. Porém isto pode ser modificado

para que partes importantes de baixa probabilidade sejam pintadas. Por exemplo, pode-se viciar o SIF de tal forma que se escolha a raiz $z_{n+1} = f_i(z_n)$ de $z_n = z_{n+1}^2 + c$ que dê o maior módulo da derivada $w_{n+1} = \frac{dz_{n+1}}{dz_n}$. Esta derivada se pode calcular iterativamente usando a fórmula $2z_{n+1}w_{n+1} = w_n$.

Já existem projetos de hardware especializado (cf. [4]) para agilizar os cálculos de decodificar figuras por este método. Espera-se poder decodificar imagens em um computador, em tempo real, usando SIF's.

Uma nota mais sobre os erros no computador. As equações diferenciais governam nosso mundo. Sua discretização é o estudo das órbitas $\{f^n(x)/n \geq 0\}$ de uma aplicação $f: M \leftarrow$. Resulta que muitos deles são hipersensíveis às condições iniciais, por isso tem sido chamados "sistemas caóticos". (O problema dos tres corpos é um deles). Alguns deles tem sido muito bem estudados. Mas na prática, para modelar algum sistema, tem-se que fazer cálculos (aproximados) em um computador. Por exemplo, para $f: [-1, 1] \leftarrow f(x) = 1 - 2x^2$, $f^{1000} \left(\frac{\sqrt{5}-1}{2} \right)$ pode dar resultados diferentes em computadores diferentes. Existem pesquisas recentes sobre um modelo de como faria o computador ao calcular os iterados. Trata-se de considerar um processo estocástico de Markov $x_n \mapsto$ algum x_{n+1} em $B_\epsilon(f(x_n))$ com probabilidade de transição $P(\cdot|x)$ suportada em $B_\epsilon(f(x))$, ao invés de considerar o "processo determinístico" $x_{n+1} = f(x_n)$. Nosso problema é: será que isto realmente modela o comportamento de um computador? Ou seja, será que os erros $e_n = d(f(x_n), x_{n+1})$ são realmente independentes? Se tem indícios de que isto não é verdade. Gostaria de ver alguns experimentos. Pode-se fazer $x_{n+1}^d = f^{(d)}(x_n^d)$, $y_{n+1}^s = f^{(s)}(x_n^d)$ onde "s" e "d" significam "precisão simples" e "precisão dupla". Logo $e_n = d(x_n^d, y_n^s)$ dão certa idéia sobre a distribuição dos erros no computador. Logo pode-se por exemplo fazer o gráfico de $e_{2n} \times e_{2n+1}$ e ver se os erros são independentes. Ou se poderia fazer algum trabalho estatístico para determinar o grau de dependência de e_m sobre e_n , $n \leq m$ ou $n = m - 1$. Alguns exemplos onde isto pode ser testado são: (a) $f: [0, 1] \leftarrow$, $f(x) = 2x \pmod{1}$, (b) Anosov no Toro: $T: [0, 1]^2 \leftarrow$, $T(x, y) = (2x + y \pmod{1}, x + y \pmod{1})$; (c) Aplicação de Hénon: $f(x, y) = (1 + y - ax^2, bx)$, $1 \leq a \leq 2$, $b > 0$.

Estas observações me foram dadas por Rafael de la Llave (Princeton University).

BIBLIOGRAFIA

1. Barnsley, M.F., *Making Chaotic Dynamical Systems to Order em "Chaotic Dynamics and Fractals"*, Notes and Reports in Mathematics in Science and Engineering, Academic Press. Barnsley & Demko (editores) 2 (1986).
2. Barnsley, M.F. & Demko, S., *Iterated Function Systems and the Global Construction of Fractals*, Proceedings of the Royal Society of London A 399 (1985), 243-275.
3. Barnsley, M.F.; Ervin, V.; Hardin, D. & Lancaster, J., *Solution of an Inverse Problem for Fractals and Other Sets*, Proceedings of the National Academy of Science 83 (April 1985).
4. Barnsley, M.F. & Sloan, A.D., *A Better way to Compress Images*, BYTE (January 1988), 215-223.
5. Billingsley, P., "Probability and Measures," Wiley series in Probability and Mathematical Statistics, John Wiley & Sons Inc, 1979.
6. Lima, E.L., "Elementos de Topologia Geral," Ao Livro Técnico, Rio de Janeiro, 1970.
7. Mandelbrot, B., "The Fractal Geometry of Nature," W.H. Freeman and Co., San Francisco CA, 1982.
8. Mañé, R., "Teoria Ergódica," Projeto Euclides, IMPA, 1983.
9. Peitgen, H.O. & Richter, P.H., "The Beauty of Fractals. Images of complex dynamical systems," Springer-Verlag, Berlin, Heidelberg, 1986.
10. Peitgen, H.O. & Saupe, D. (editores), "The Science of Fractal Images," Artigos de M.F. Barnsley, H.O. Peitgen e Y. Fisher, Springer-Verlag, 1988.
11. Rohlin, V.A., *On the fundamental ideas in measure theory*, A.M.S. translations 1, 10 (1962), 1-54.
12. Sad, P., "Introdução à dinâmica das funções racionais na esfera de Riemann," Apostilas - 14º Colóquio Brasileiro de Matemática, IMPA, Rio de Janeiro, 1983.