A DYNAMICAL MEANING OF FRACTAL DIMENSION

BY

STEVE PELIKAN

IMA Preprint Series # 73

May 1984

INSTITUTE FOR MATHEMATICS AND ITS APPLICATIONS
UNIVERSITY OF MINNESOTA
514 Vincent Hall
206 Church Street S.E.
Minneapolis, Minnesota 55455

A DYNAMICAL MEANING OF FRACTAL DIMENSION

Steve Pelikan

Institute for Mathematics and its Applications
University of Minnesota
Minneapolis, Minnesota 55455

A DYNAMICAL MEANING OF FRACTAL DIMENSION

Abstract

When two attractors of a dynamical system have a common basin boundary B, small changes in initial conditions which lie near B can result in radically different long-term behavior of the trajectory. A quantitative description of this phenomenon is obtained in terms of the fractal dimension of the basin boundary B.

Introduction

This paper contains theorems concerning the fractal dimension of certain invariant sets of dynamical systems. Rather than computing the dimension of invariant sets, we describe a phenomenon observed in a variety of systems which admits a quantitative description in terms of the fractal dimension of an invariant set.

The sets of interest form a boundary between the basins of attraction of two attractors. The map $f: \mathbb{R} \to \mathbb{R}$ shown in figure 1 provides a good example. The map f has two attracting fixed points, A and B. Almost every initial condition $x_0 \in \mathbb{R}$ tends to one of these two points under the action of f. The exceptions to this rule are the points of a Cantor set Λ contained in $[0, 1]$.

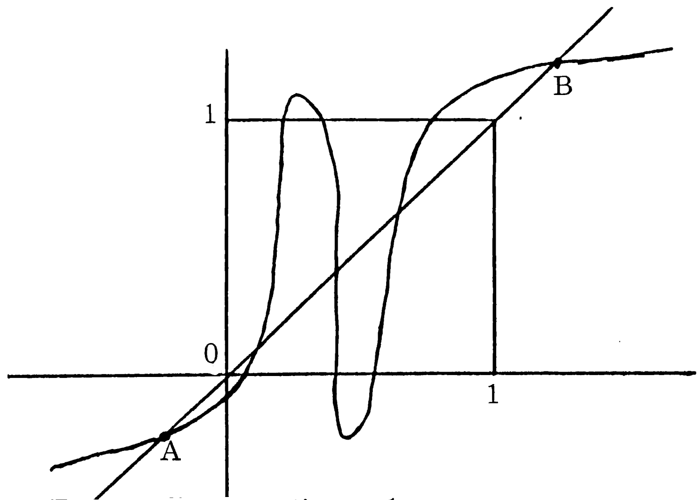

Figure 1

In Section 2 we consider the following question. Suppose x and y are chosen at random from $[0, 1]$ subject only to the condition $|x-y| < \epsilon$. What is the probability p_ϵ that x and y tend to different attractors? If f describes a physical system this question can be phrased as: Suppose there is a uniform ϵ-error in determining the initial conditions of the system. What is the probability that our prediction of the long-term behavior of the system will be incorrect? We show (Theorem 2) that p_ϵ tends to zero like ϵ^{1-d} as ϵ goes to zero, where d is the Hausdorff dimension of Λ. Thus, when $d > 0$ (and especially when d is near 1) it may be very difficult to predict the long-term behavior of the system correctly, even though the long-term behavior of almost every trajectory is very simple.

This phenomenon, called "final state sensitivity," was studied numerically by the authors of [1]. They actually discuss two measures of the degree of sensitivity. The numerical experiments they describe are concerned with measuring the quantity p_ϵ. In addition they consider q_ϵ, which is the measure of the set of x for which there exists a y within ϵ of x so that x and y tend to different attractors. In [1] it is conjectured that $q_\epsilon \sim \epsilon^{1-d}$, where d is the dimension of Λ.

The motivation for studying p_ϵ (rather than q_ϵ) is primarily that it represents a quantity with a more natural interpretation; p_ϵ is the probability of making an incorrect prediction while q_ϵ is the probability of being able to make an incorrect prediction. In addition, as the numerical techniques of [1] show, p_ϵ can be computed quite easily numerically, while q_ϵ can not.

Section 3 describes two examples of dynamical systems in two dimensions for which theorems analogous to Theorem 2 are true. These are the "linear" horseshoe and certain rational maps of the Riemann sphere.

In Section 1 we show that for invariant sets such as Λ, the two most frequently used definitions of fractal dimension (i.e., the Hausdorff and capacity dimensions) are equal. The utility of this result comes from the fact that the capacity dimension is much easier to compute numerically than the Hausdorff dimension.

I. Fractal dimensions

Here $\Lambda \subset [0,1]$ is a compact set. The Hausdorff dimension of Λ, denoted $HD(\Lambda)$, is defined as follows:

$$\text{For } \alpha \in \mathbb{R}, \text{ let } h(\alpha) = \lim_{\epsilon \to 0^+} \inf_{\|\mathcal{a}\| < \epsilon} \sum_{\mathcal{a}} |A_i|^\alpha ,$$

where $|A|$ = diameter of A, $\mathcal{a}$ denotes a finite open cover of Λ, and $\|\mathcal{a}\| = \max_{A_i \in \mathcal{a}} |A_i|$.

One shows (see [2]) that there is a number α^* so that $h(\alpha) = 0$ if $\alpha > \alpha^*$ and $h(\alpha) = +\infty$ if $\alpha < \alpha^*$. The Hausdorff dimension of Λ is $HD(\Lambda) = \alpha^*$.

A second notion of fractal dimension is the capacity dimension. This is defined by

$$\mathrm{cap}(\Lambda) = \lim_{\epsilon \to 0^+} \frac{\log N(\epsilon)}{-\log \epsilon} \quad \text{(when the limit exists)},$$

where $N(\epsilon) = \min_{\|\mathcal{Q}\| \leqslant \epsilon} \mathrm{card}(\mathcal{Q})$ is the minimum number of sets with diameter ϵ needed to cover Λ.

It is easy to see that $\mathrm{HD}(\Lambda) \leqslant \mathrm{cap}(\Lambda)$, and equally easy to give examples where this inequality is strict, for example, $\Lambda = \{0\} \cup \{1/n\}_{n=1}^{\infty}$. There are also examples where $\mathrm{cap}(\Lambda)$ is not defined.

Our purpose here is to show that if Λ is invariant under an expanding map (that is, satisfies some (nonlinear) self-similarity law), then $\mathrm{cap}(\Lambda) = \mathrm{HD}(\Lambda)$. To be specific,

<u>Theorem 1</u>. Suppose that Λ is invariant under a map $f:[0,1] \to \mathbb{R}$, where f is $C^{1+\alpha}$ for some $\alpha > 0$, and expanding on $f^{-1}([0,1])$. Then $\mathrm{cap}(\Lambda) = \mathrm{HD}(\Lambda)$. (Expanding means $|f'(x)| \geqslant \lambda > 1$ for some number λ.)

We denote by $\{C_j^n\}$ the components of $f^{-n}([0,1])$. These sets form a natural cover of Λ. The proof of Theorem 1 depends on the following lemma, which says that the $\{C_j^n\}$'s approximate Λ so well that they can be used to compute the Hausdorff dimension of Λ.

<u>Lemma 1</u>. Let $d > \mathrm{HD}(\Lambda)$. Then

$$\lim_n \sum_j |C_j^n|^d = 0.$$

A second, well-known, fact will be required concerning expanding functions f:

<u>Lemma 2</u>. There exists $K \geqslant 1$ independent of n such that

$$\left| \frac{\frac{d}{dx} f^n(x)}{\frac{d}{dx} f^n(y)} \right| \leqslant K$$

whenever x, y lie in the same set C_j^n. ■

Proof of Theorem 1. Let $d > HD(\Lambda)$. Using Lemma 1, select N so that $n > N$ implies that $\sum_j |C_j^n|^d < K^{-d}$, where K is the constant in Lemma 2. Let $\epsilon_0 = \min_{j \neq k} \operatorname{dist}(C_j^n, C_k^n)$. For $\epsilon < \epsilon_0$, denote by $\mathcal{D}_\epsilon = \{D_i\}$ a cover of Λ by $N(\epsilon)$ sets of diameter ϵ. The collection $\{f^n(D_i \cap C_j^n)\}$ is then a cover of Λ by sets with diameters in the interval $\left[\epsilon K^{-1} |C_j^n|^{-1}, \epsilon K |C_j^n|^{-1}\right]$, and so

$$\text{(1)} \qquad \operatorname{card}\left\{D_i \in \mathcal{D}_\epsilon : D_i \cap C_j^n \neq \emptyset\right\} \leqslant N(\epsilon K^{-1} |C_j^n|^{-1}) .$$

Summing (1) over j we obtain

$$\text{(2)} \qquad N(\epsilon) \leqslant \sum_j N(\epsilon K^{-1} |C_j^n|^{-1}) .$$

Define $\phi(\epsilon)$ by $N(\epsilon) = \epsilon^{-d}\phi(\epsilon)$. Substituting this expression for $N(\epsilon)$ in (2) yields

$$\text{(3)} \qquad \phi(\epsilon) \leqslant K^d \sum_j |C_j^n|^d \phi(\epsilon K^{-1} |C_j^n|^{-1}) .$$

This expresses a bound for $\phi(\epsilon)$ as a convex combination of the values of ϕ at larger arguments, and so $\phi(\epsilon)$ is bounded. Consequently,

$$\overline{\lim_{\epsilon \to 0}} \frac{\log N(\epsilon)}{-\log \epsilon} \leqslant d \qquad \text{for each } d > HD(\Lambda)$$

and so $HD(\Lambda) = \operatorname{cap}(\Lambda)$. ∎

Proof of Lemma 1. First, note that it suffices to prove that $\lim_n \sum_j |C_j^n|^d$ is bounded for every $d > HD(\Lambda)$ since if $\sum_j |C_j^n|^{d-\epsilon} \leqslant M$ is bounded then $\sum_j |C_j^n|^d = \sum |C_j^n|^{d-\epsilon} |C_j^n|^\epsilon \leqslant M \max_j |C_j^n|^\epsilon$. But $\max_j |C_j^n| < \lambda^{-n}$ because f is expanding.

For $d > HD(\Lambda)$ define

$$a_d(\epsilon) = \inf\left\{\sum |A_i|^d : \{A_i\} \text{ is a cover of } \Lambda \text{ and } |A_i| \geqslant \epsilon\right\} .$$

Then $a_d(\epsilon)$ is a nondecreasing function of ϵ and $a_d(\epsilon) \to 0$ as $\epsilon \to 0$. For each $\epsilon > 0$ there is a cover of Λ, $\mathcal{A}_\epsilon$, with $|A_i| \geqslant \epsilon$ and $\sum_{\mathcal{A}_\epsilon} |A_i|^d < 2a_d(\epsilon)$. We then have that $\|\mathcal{A}_\epsilon\|^d = \max |A_i|^d < 2a_d(\epsilon)$ and so

$$\|\mathcal{A}_\epsilon\| < \left[2a_d(\epsilon)\right]^{1/d} .$$

For each n, find ϵ_n so that $\|\mathcal{A}_{\epsilon_n}\| < \frac{1}{2} \min_{j \neq k} \operatorname{dist}(C_j^n, C_k^n)$. Then, for each $A_i \in \mathcal{A}_{\epsilon_n}$, there is a unique C_j^n with $A_i \cap C_j^n \neq \emptyset$. We extend $f^n | C_j^n$ linearly with slope $|C_j^n|^{-1}$. Having done this, $f^n(A_i)$ is a well-defined set for each A_i which has nonempty intersection with C_j^n. In fact

$$\left\{ f^n(A_i) : A_i \cap C_j^n \neq \emptyset \right\} \text{ is a cover of } \Lambda$$

by sets having diameters at least $\epsilon_n K^{-1} |C_j^n|^{-1}$. Consequently,

$$\sum_{A_i \cap C_j^n \neq \emptyset} |A_i|^d K^d |C_j^n|^{-d} \geqslant a_d(\epsilon_n K^{-1} |C_j^n|^{-1})$$

and so

$$(4) \qquad \sum_i |A_i|^d = \sum_j \sum_{A_i \cap C_j^n \neq \emptyset} |A_i|^d$$

$$\geqslant \sum_j K^{-d} |C_j^n|^d a_d(\epsilon_n K^{-1} |C_j^n|^{-1}) .$$

Now $a_d(\epsilon_n K^{-1} |C_j^n|^{-1}) \geqslant a_d(\epsilon_n K^{-1} \lambda^n)$ for each j, where $\lambda = \min_{x \in f^{-1}[0,1]} |f'(x)|$. Then (4) becomes

$$2a_d(\epsilon_n) \geqslant a_d(\epsilon_n K^{-1} \lambda^n) K^{-d} \sum_j |C_j^n|^d ,$$

or for n such that $\lambda^n > K$,

$$2K^d > 2K^d \frac{a_d(\epsilon_n)}{a_d(\epsilon_n K^{-1} \lambda^n)} \geqslant \sum_j |C_j^n|^d ,$$

by the monotonicity of $a_d(\epsilon)$. This implies $\lim_n \sum_j |C_j^n|^d \leqslant 2K$. ∎

II. Final state sensitivity

We begin by considering maps $f: \mathbb{R} \to \mathbb{R}$ which have the essential features of the f shown in figure 1. We assume:

i) f is $C^{1+\alpha}$

ii) $|f'(x)| > 1$ if x and $f(x)$ are in $[0,1]$

iii) $[0,1] \subset \text{interior}(f([0,1]))$ and $f^{-1}[0,1] \subset [0,1]$

iv) f has two attractors, one attracting each point in $(1,\infty)$, the other attracting each point in $(-\infty, 0)$.

We wish to consider the case where the basin boundary of the attractors is a Cantor set. Thus we assume the $f \big| [0,1]$ has at least two laps. (That is, $f^{-1}([0,1])$ consists of at least two intervals.) Let r denote the number of laps of f. In case $r = 1$, the basin boundary is simply a point and Theorem 2 is trivially true. Note that conditions iii) and iv) imply that if $r \geqslant 2$ then $r \geqslant 3$ and r is odd. As mentioned earlier, we define p_ϵ to be the probability that two points x, y (chosen at random from $[0,1]$ according to the uniform distribution and subject to the condition $|x - y| < \epsilon$) tend to different attractors. That is, let $I_\epsilon(x) = [x - \epsilon, x + \epsilon]$ and set

$$p_\epsilon(x) = m\left(\left\{y \in I_\epsilon(x) \cap [0,1] : \lim_n f^n(x) \neq \lim_n f^n(y)\right\}\right) \Big/ m\left(I_\epsilon(x) \cap [0,1]\right)$$

and then set $p_\epsilon = \int_0^1 p_\epsilon(x)\,dx$. Then we have

Theorem 2:

$$\lim_{\epsilon \to 0} \frac{\log p_\epsilon}{\log \epsilon} = 1 - \alpha \quad ,$$

where $\alpha = \text{cap}(\Lambda) = \text{HD}(\Lambda)$.

Remark: As mentioned above, Theorem 2 is true for q_ϵ also. The proof is essentially the one given below.

Proof of Theorem 2: Let

$$A(\epsilon) = \{(x, y) \in [0, 1] \times [0, 1] : |x - y| < \epsilon\}$$

and

$$D(\epsilon) = \{(x, y) \in A(\epsilon) : \lim_n f^n(x) \neq \lim_n f^n(y)\} .$$

Then, denoting Lebesgue measure on $[0, 1]^2$ by m, we have that p_ϵ has the same behavior (as $\epsilon \to 0$) as $m(D(\epsilon)) / m(A(\epsilon))$. Note that $m(A(\epsilon)) = 2\epsilon - \epsilon^2$. Thus, to prove the theorem, it's enough to show that $\lim_{\epsilon \to 0} \log m\, D(\epsilon) / \log \epsilon = 2 - \alpha$.

We define a mapping $F : [0, 1]^2 \to \mathbb{R}^2$ by $F(x, y) = (f(x), f(y))$. It is immediate that $\left| \frac{\det DF^n(x, y)}{\det DF^n(w, z)} \right| \leqslant K^2$ whenever (x, y) and (x, z) belong to the same set $C_j^n \times C_j^n$. Also note that for each n and j we have that

(5) $$D(K^{-1}\epsilon\, |C_j^n|^{-1}) \subset F^n(D(\epsilon) \cap (C_j^n \times C_j^n)) \subset D(K\epsilon\, |C_j^n|^{-1})$$

for some k. Our goal is to derive a scaling law for $D(\epsilon)$ analogous to equation (2). In order to do this we will estimate $m(D(\epsilon))$ in terms of $m(D(\epsilon) \cap (\bigcup_j C_j^n \times C_j^n))$, and then apply (5).

Begin by considering $D(\epsilon)$ with

$$\epsilon \ll \min\left\{\min_{j \neq k} \text{dist}(C_j^{n+2}, C_k^{n+2}),\ \min_j |C_j^n|\right\} .$$

Then $(x, y) \in D(\epsilon)$ implies that one of x, y lies in some C_j^n. In this case (x, y) must lie in the region $*_\epsilon C_j^n$ shown in figure 2.

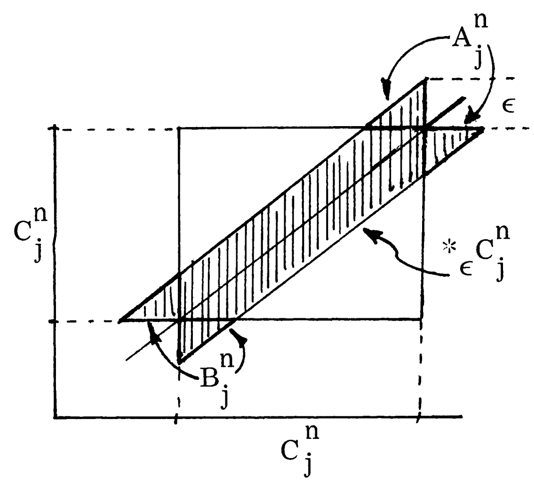

Figure 2

Claim: There is a subset of $\bigcup_j C_j^n \times C_j^n$ mapped in a 1-1 manner onto $\bigcup_j A_j^n \cup B_j^n$ by F^2. To establish the claim it is enough to show the following: If $C_j^n = [a,b]$ then at least one of the points in $f^{-2}(a)$ (and at least one of the points in $f^{-2}(b)$) lie in the interior of some C_ℓ^n. Such a point is an endpoint of some C_k^{n+2}. The corresponding set A_k^{n+2} (or B_k^{n+2}) determined by $*_\epsilon C_k^{n+2}$ is contained in some C_ℓ^n and is mapped onto A_j^n (or B_j^n) by F^2.

Proof of claim: Suppose $C_j^n = [a,b]$ and that all of the points $f^{-2}(a)$ are endpoints of sets C_ℓ^n. Apply f^{n-1}. We obtain a set $C_j^1 = [a',b']$ such that all the points $f^{-2}(a')$ are endpoints of other intervals C_k^1. But there are r intervals C_k^1 and r^2 points in $f^{-2}(a)$. Since $r^2 > 2r$ (when $r > 2$, as we have assumed), this is a contradiction.

Thus $D(\epsilon) \cap (\bigcup_j A_j^n \cup B_j^n)$ is contained in the image (under F^2) of a subset of $D(\epsilon) \cap (\bigcup_j C_j^n \times C_j^n)$. Since F^2 expands areas by at most a factor $B' \geqslant 1$, we have shown that (set $B = B'+1$)

$$m\big(D(\epsilon)\big) \leqslant B\, m\Big(\bigcup_j D(\epsilon) \cap C_j^n \times C_j^n\Big) .$$

We can now apply (5), sum over j and obtain

$$(6) \qquad K^{-2} \sum_j m\Big(D(\epsilon K^{-1} |C_j^n|^{-1})\Big) |C_j^n|^2 \leqslant m(D_\epsilon) \leqslant BK^2 \sum_j m\Big(D(K\epsilon |C_j^n|^{-1})\Big) |C_j^n|^2$$

Finally, suppose $2-d > \text{cap}(\Lambda)$ and write $m(D_\epsilon) = \epsilon^d \phi(\epsilon)$. Select n so large that $\sum_j |C_j^n|^{2-d} < (BK^2)^{-1}$. When ϵ is sufficiently small, equation (6) yields

$$\phi(\epsilon) \leqslant BK^2 \sum_j \left|C_j^n\right|^{2-d} \phi(K\epsilon \left|C_j^n\right|^{-1})$$

which implies that ϕ is bounded as $\epsilon \to 0$. Hence

$$\underline{\lim}_{\epsilon \to 0} \frac{\log mD(\epsilon)}{\log \epsilon} = \underline{\lim}_{\epsilon \to 0}\, d + \frac{\log \phi(\epsilon)}{\log \epsilon} \geqslant d$$

for every $d < 2 - \text{cap}(\Lambda)$.

Then let $2 - d < \text{cap}(\Lambda)$ and, selecting n so large that $\sum_j \left|C_j^n\right|^{2-d} > K^2$, write $m(D_\epsilon) = \epsilon^d \phi(\epsilon)$. The left hand inequality in (6) then yields

$$\phi(\epsilon) \geqslant \sum_j K^2 \left|C_j^n\right|^{2-d} \phi(K^{-1}\epsilon \left|C_j^n\right|^{-1}) .$$

This means $\phi(\epsilon)$ is bounded away from zero and so

$$\overline{\lim}_{\epsilon \to 0} \frac{\log m\, D(\epsilon)}{\log \epsilon} = d + \overline{\lim}_{\epsilon \to 0} \frac{\log \phi(\epsilon)}{\log \epsilon} \leqslant d$$

for every $d > 2 - \text{cap}(\Lambda)$. ∎

Remark: The quantity p_ϵ was interpreted as the probability of making an error when predicting the long term behavior of the system in the presence of a uniform error (of size ϵ) in determining the initial conditions of the system. The assumption that the errors are uniform is not, however, necessary. Let μ_ϵ be any probability measure on $[-\epsilon, \epsilon]$ which is equivalent to normalized Lebesgue measure on $[-\epsilon, \epsilon]$. (That is, both $\frac{d\mu_\epsilon}{dx}/2\epsilon$ and $2\epsilon/\frac{d\mu_\epsilon}{dx}$ are bounded independent of ϵ.)

We suppose that μ_ϵ describes the errors in observing the states of the system, and define a quantity $\overline{p_\epsilon}$ analogous to p_ϵ. Set

$$\overline{p_\epsilon}(x) = \mu_\epsilon \left\{ y \in [-\epsilon, \epsilon] : x + y \in [0,1] \text{ and } \lim_n f^n(x) \neq f^n(x+y) \right\}$$

and $\overline{p_\epsilon} = \int \overline{p_\epsilon}(x)dx$. Then $\overline{p_\epsilon}$ has the same behavior (as $\epsilon \to 0$) as p_ϵ. This is because the quantities $m\left(D(\epsilon) \cap C_j^n \times C_j^n\right)$ in the proof of Theorem 2 change by at most a bounded factor when Lebesgue measure is replaced by μ_ϵ.

III. Two-dimensional examples

A. Rational maps.

We begin with an explicit example: $f(z) = z^2 + c$ with $c \in \mathbb{C}$ of sufficiently small modulus. Considered as a map of the Riemann sphere $\mathbb{C} \cup \{\infty\}$, the map f has two attracting fixed points, one near 0 and the other at ∞. The common boundary of the basins of these attractors is the Julia set of f, frequently denoted $J(f)$.

More generally, we suppose that f is a rational map of the sphere $\mathbb{C} \cup \{\infty\}$ which has two attracting fixed points (say 0 and ∞). In addition, suppose that the forward orbit of each critical point of f tends to one of the attractors. Then (see [3]) the common boundary of the basins of the attractors of f is the set $J(f)$ and some power of f, say f^r, is uniformly expanding on a neighborhood of $J(f)$.

Let N be a neighborhood of $J(f)$ on which f^r is uniformly expanding. (That is, $\left|\frac{d}{dz} f^n(z)\right| \geqslant \gamma > 1$ for $z \in N$.) Denote by p_ϵ the probability that x, y chosen at random from N (according to the uniform distribution and subject only to $|x-y| < \epsilon$) tend to different attractors under the action of f. Then, corresponding to Theorems 1 and 2 we have:

(A1) $$\mathrm{HD}\left(J(f)\right) = \mathrm{cap}\left(J(f)\right) \quad \text{and}$$

(A2) $$\lim_{\epsilon \to 0^+} \frac{\log p_\epsilon}{\log \epsilon} = 2 - \mathrm{HD}\left(J(f)\right) \quad .$$

The proofs of (A1) and (A2) are almost identifical with the proofs of the corresponding theorems. The one difference is that, in stead of using the sets $\{C_j^n\}$ we use sets of the form

$$\left\{ z : f^j(z) \in A_{k_j} \ , \quad j = 0, 1, \ldots n-1 \ , \quad k_j \in \{1, \ldots, m\} \right\}$$

where $\{A_1, \ldots, A_m\}$ is some cover of $J(f)$ by open sets.

It should be noted that this (A2) provides a method for computing the Hausdorff dimension of some Julia sets which is particularly well-suited to the use of parallel or vector processors.

B. Horseshoes.

Consider a horseshoe diffeomorphism which acts on a disk in the plane as indicated in figure 3.

There are two attracting fixed points for f. Lebesgue almost every trajectory tends to one of these attractors under the action of f. The boundary of the basins of the attractors is the stable manifold W^s of the horseshoe. If we assume that f ristricted to $f^{-1}(R)$ is an affine transformation which preserves vertical lines (so that W^s is the product of a Cantor set and an interval) we obtain, as an application of Theorem 2, that

$$\lim_{\epsilon \to 0} \frac{\log p_\epsilon}{\log \epsilon} = 2 - HD(W^s) \quad . \tag{B1}$$

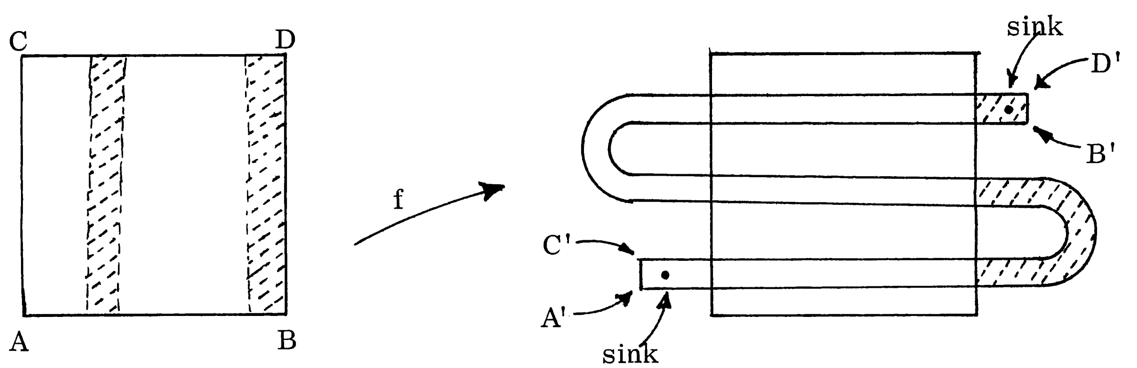

Figure 3. This illustrates the action of the "horseshoe" diffeomorphism on the square ABCD. The image of the square (relative to itself) is shown on the right. The shaded region indicates part of the basin of attraction of one of the sinks.

References

[1] Grebogi, C., S. McDonald, E. Ott, and J. Yorke. Final state sensitivity: an obstruction to predictability, Phys. Lett., 99A (1983), 415-418.

[2] Farmer, D., E. Ott, and J. Yorke. The dimension of chaotic attractors, Physica, 7D (1983), 153-180.

[3] Blanchard, P. Complex analytic dynamics on the Riemann sphere. Preprint.