Journal of Statistical Physics, Vol. 26, No. 4, 1981

On Computing the Entropy of the Henon Attractor

James H. Curry[1]

Received February 27, 1980; revised April 14, 1981

In a recent article D. Ruelle [in *Lecture Notes in Physics*, No. 80 (Springer, Berlin, 1978)] has conjectured that for the Hénon attractor its measure theoretic entropy should be equal to its characteristic exponent. This result is known to be true for systems which satisfy Smale's Axiom A. In this article we report the results of our computations which suggest that Ruelle's conjecture may be true for the Hénon attractor. Further, in our study we are confronted with fundamental questions which suggest that certain existence theorems from ergodic theory are not sufficient from a computational point of view.

KEY WORDS: Characteristic exponent; entropy; partition.

1. INTRODUCTION

Two quantities which provide evidence for random behavior in a given dynamical system are the measure theoretic entropy and the characteristic exponents. If there is nontrivial recurrent behavior for a dynamical system, then one is likely to find that there is an exponential rate of separation of neighboring trajectories (therefore at least one positive characteristic exponent) and also positive entropy (see Section 2 for definition of entropy).

For more than a decade it has been known that there is a relationship between the entropy and characteristic exponents; precisely, it was known that in general the sum of the positive characteristic exponents was greater than or equal to the measure theoretic entropy. However, it has also been established that for systems which satisfy Smale's Axiom A[(2)] the measure theoretic entropy and sum of the positive characteristic exponents are equal. In Ref. 1 D. Ruelle has posed the question of whether these two quantities will be equal in a broader context than that allowed by systems

[1] Institut des Hautes Etudes Scientifiques, 35, route de Chartres, 91440 Bures-sur-Yvette, France. Permanent address: Mathematics Department, Box 426, University of Colorado, Boulder, Colorado 80309, USA.

0022-4715/81/1200-0683$03.00/0

considered in Ref. 2 (for the exact formulation of Ruelle's question, we refer the reader to Ref. 1).

More recently I. Shimada[3] has reported that for the Lorenz system of three coupled nonlinear ordinary differential equations the characteristic exponent and measure theoretic entropy are in good numerical agreement for the parameter values which he considers, which differ from those considered in Ref. 4. The Lorenz system does not satisfy Axiom A, hence the work of Shimada gives some evidence that Ruelle's question will have an affirmative answer for a wider class of systems than those considered in Ref. 2.

In Ref. 5 M. Hénon, motivated by a careful numerical study of the Lorenz system by Y. Pomeau, introduced a transformation of the plane into itself which seems to admit an attractor set which locally has the structure of the product of a Cantor set with an interval. Hénon's transformation has received much attention recently, both numerical[6-8] and analytic[9,10] and also does not seem to satisfy Axiom A. We remark that in Ref. 10 it has been proven that there is a transverse crossing of the stable and unstable manifolds of the fixed point for Hénon's parameter values and hence an explanation of the graphs in Ref. 5.

In Ref. 1 Ruelle also conjectures that the measure theoretic entropy of the Hénon attractor should be approximately 0.4. This conjecture is based in part on the question raised above and results of S. Feit. In Ref. 6 Feit has computed the characteristic exponent for the Hénon attractor and found that it has a value of approximately 0.42 to two decimal places.

In this article we report the results of our numerical experiments which were an attempt to verify the conjecture of Ruelle for the case of the Hénon attractor. In Section 2 we recall the definition of measure theoretic entropy.

In Section 3 we present the results of our numerical experiments for two well-known examples, and in Section 4 we report our findings for Hénon's transformation when $a = 1.40$, $b = 0.3$.

In Section 4.2 we mention results for other parameter values than those studied by Hénon. Finally, in Section 5, we discuss our findings and their possible significance.

2. ENTROPY

Let $(X, \mathcal{Q}, \mu)$ be a probability space. We define a partition of $(X, \mathcal{Q}, \mu)$ to be any disjoint collection of elements of $\mathcal{Q}$ whose union is X. Let β be a finite partition of X, $\beta = \{\beta_1, \beta_2, \ldots, \beta_k\}$, and suppose that $T: X \to X$. Given any $x \in X$ it is possible to identify x with a sequence of symbols (i_j) where $i_j = m$ if $T^j x \in \beta_m$ and $m \in \{1, 2, \ldots, k\}$. The action of T is to shift

$(i_j) \to (i_{j+1})$. Of course if we choose a different partition, β, the point $x \in X$ remains unchanged but the sequence of symbols which represent x could change.

The measure theoretic entropy of β is defined to be

$$H(\beta) = \sum_i \mu(B_i) \log \mu(B_i) \tag{2.1}$$

while the entropy of T with respect to β is defined to be

$$H(T, \beta) = \lim_{n\to\infty} \frac{1}{n} H(\beta \vee T\beta \vee \cdots \vee T^{(n-1)}\beta) \tag{2.2}$$

where, if β and α are two finite partitions, $\beta \vee \alpha = \{B_j \cap A_i\}$ and B_j and A_i are chosen from the elements of β and α, respectively. In what follows we shall refer to the quantity computed in (2.2) as the partition entropy for the partition β and transformation T.

The entropy of T, for the given measure μ, is then

$$h_\mu(T) = \sup_\beta h_\mu(T, \beta) \tag{2.3}$$

Two comments are necessary prior to reporting the results of our experiments. Formula (2.2) is essentially the one used in Ref. 1. An unfortunate fact about this formula is that it converges to the partition entropy very slowly. In order to speed up the convergence we have used the following formula, which also converges to the partition entropy provided the limit in (2.2) exists:

$$\lim_{n\to\infty} \left\{ H\left(\bigvee_0^n T^i\beta\right) - H\left(\bigvee_1^n T^i\beta\right)\right\} = \lim_{n\to\infty} H\left(\beta \mid \bigvee_1^n T^i\beta\right) \tag{2.4}$$

where $H(\alpha \mid \beta) = H(\alpha \bigvee \beta) - H(\beta)$ is the conditional entropy of α given β. We call the quantity (2.4) F_n^k, where k refers to the number of elements in the partition and n is given in (2.4). In order to compute the entropy of T we should compute the supremum over all partitions. This is not possible. Hence we note that if α is a generator for T and β is any partition, then $h(T,\alpha) \geqslant h(T, \beta)$. And in particular $h(T) = h(T,\alpha)$ for any generator α. This fact is the Kolmogorov–Sinai theorem and we shall make use of it by looking for a partition that is in some sense close to a generator. In general there is no algorithm for finding a generating partition. Therefore we try to find a partition which maximizes the entropy among those partitions we consider. Finally, note that (2.2) and (2.4) decrease to the partition entropy and the entropy of a refined partition is greater than the entropy of the original one. For a more complete introduction to measure theoretic entropy we refer the reader to Ref. 13 and the references cited there.

3. COMPUTATIONAL CONSIDERATIONS AND TWO EXAMPLES

Let us begin by giving a description of our underlying computational strategy for computing the measure theoretic entropy for a transformation T. The idea behind the strategy is simple: compute the frequency of occurrence of configurations having various lengths, then use formula (2.2). More explicitly, first partition the domain of the transformation into some number of disjoint pieces; then associate with each member of the partition a symbol, e.g., an integer between 1 and n inclusive, if there are n elements in the partition. In this way it is possible to represent symbolically the orbit of a point by noting in which element of the partition it lies, at each iteration of the transformation.

By a configuration of length k we shall mean a finite segment of the total orbit consisting of k symbols. In order to compute the partition entropy for T, formula (2.2) can then be used. There are three empirical considerations which must be kept in mind throughout the remainder of this article: (1) in general, it will be difficult to know a priori how many times a transformation must be iterated before the asymptotic frequency of a given configuration is achieved; (2) it is not possible to consider configurations of arbitrary length, and hence the limiting operations in (2.2), for example, cannot be performed; and (3) a question which is fundamental to both (1) and (2): *How is the domain of T to be subdivided*?

In view of the question and remarks of the last paragraph it will be useful to consider two illustrative examples, the first from an analytic point of view and the second from a computational point of view.

The first mapping we will consider is a transformation which carries the unit interval into itself and is given by

$$g(x) = \begin{cases} 2x, & 0 \leqslant x < \frac{1}{2} \\ 2(1-x), & \frac{1}{2} \leqslant x \leqslant 1 \end{cases} \tag{3.1}$$

This is the so-called "rooftop" map. The second, equally well known, map of the interval is

$$f(x) = 4x(1-x) \tag{3.2}$$

These two mappings are both orbit and measure theoretically conjugate with measure-theoretic entropy $\log 2$. A generating partition for both examples is the two-element partition given by $\{[0,\frac{1}{2}),[\frac{1}{2},1]\}$.

Since question (3) above is fundamental to all subsequent considerations, let us consider it and its relationship to (3.1). Suppose we divide the domain of this map into three subintervals having equal length. The partition is then given by $\alpha = \{[0,\frac{1}{3}),[\frac{1}{3},\frac{2}{3}),[\frac{2}{3},1]\}$. If we now label the sets with the symbols 1, 2, 3, then g has the same dynamics as a three state Markov shift whose initial vector is $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ and whose transition matrix is

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$$

(to see this it is sufficient to see what g does to the elements of α).

Now, given the initial vector and transition probabilities of a Markov shift it is possible to compute its entropy. In this case we find the number $\frac{2}{3}\log 2$.[14] If we now divide the unit interval into k (odd) subintervals all having equal length, then the dynamics of g for this partition is equivalent to a k-state Markov shift whose entropy is $[(k-1)/k]\log 2$. But, in order to arrive at the measure-theoretic entropy of g we must take the supremum over all partitions and in particular over all equally spaced partitions; when we do this, not surprisingly, $\log 2$ is achieved.

Recall that any partition into an even number of subintervals is a refinement of the generating partition $\{[0, \frac{1}{2}), [\frac{1}{2}, 1]\}$ and is therefore itself a generating partition. Hence the partition entropy must be $\log 2$. This can be checked by a straightforward computation.

We may summarize our findings for example (3.1) as follows:

(a) If we choose the wrong partition, the value we compute for the partition entropy will be less than the true entropy of the transformation.

(b) Refining a partition increases the corresponding partition entropy.

As we noted in the introduction to this article, I. Shimada[3] has found good agreement between the characteristic exponent and measure-theoretic entropy for the Lorenz system of equations by using a partition consisting of two elements. The choice of the partition used by Shimada was no doubt guided by the symmetries in the Lorenz equations. The comments made about the rooftop map and Shimada's computation suggest the following numerical experiment: Choose a dynamical system whose measure-theoretic entropy is known. Then using formula (2.2), compute the partition entropy when obvious symmetries are "overlooked." That is, what can we expect from a "bad" partition?

As one of the few examples of a dynamical system whose measure-theoretic entropy is known we consider (3.2). In computing the partition entropy we have used formula (2.2). For this example f was iterated 500,000 times and configurations of maximum length 8 were examined.

In Fig. 1, the results of our entropy computation for the partition consisting of two equal pieces is presented. Here it can be seen that there is

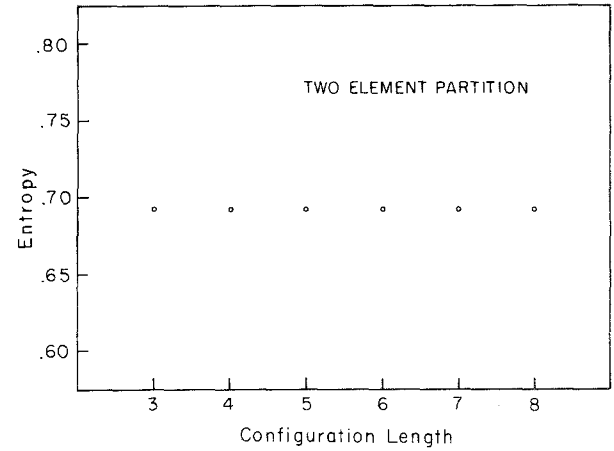

Fig. 1. Graph of the entropy of $4x(1-x)$ as a function of length for a two-element partition.

rapid convergence to the entropy of f when configurations having lengths between three and eight symbols are considered. Table I contains the data used to generate Fig. 1.

In Fig. 2 we have used a partition which contains three elements having equal length (the partition α considered before). Table II contains the data used to generate this graph. Notice that these data are not asymptotic to $\log 2$. Indeed, the fact that the configuration entropy is strictly less than $\log 2$ when configurations of length 8 are considered indicates that the partition entropy of f for this partition will be strictly less than its measure-theoretic entropy. There was no appreciable change in the results when f was iterated 1,000,000 times.

The above example indicates that the symmetries of a dynamical system can be a powerful aid in choosing a partition and achieving optimal convergence to the measure-theoretic entropy; while when the partition is

Table I. Two-Element Partition

Length of configuration	Entropy	Number of configuration
3	0.6931	8
4	0.6931	16
5	0.6931	32
6	0.6931	64
7	0.6931	128
8	0.6931	256

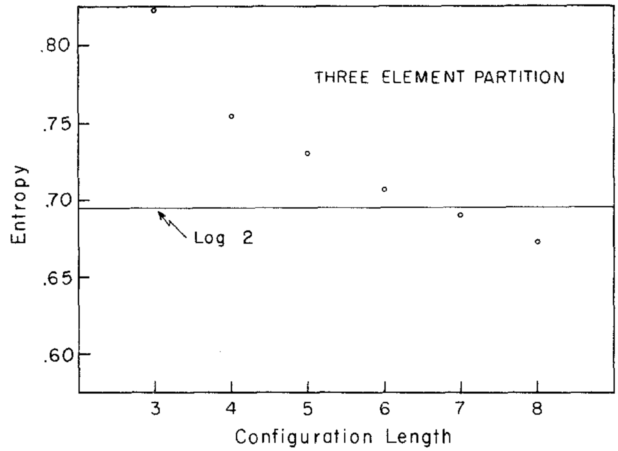

Fig. 2. Graph of the entropy of $4x(1-x)$ as a function of configuration length for a three-element partition.

arbitrary, convergence to a limiting value is slower (there is no way to choose a priori a generating partition).

4. THE HÉNON TRANSFORMATION

The Hénon transformation is an invertible mapping of the plane into itself whose definition is given by

$$T(x, y) = (1 + y - ax^2, bx)$$

In Ref. 5 M. Hénon considers the behavior of iterates of T when $a = 1.40$ and $b = 0.3$. In what follows we shall initially restrict our attention to these parameter values. For properties of T which have been discovered since Ref. 5 we refer to Refs. 6–10.

Table II. Three-Element Partition

Length of configuration	Entropy	Number of configuration
3	0.8243	14
4	0.76717	26
5	0.7323	48
6	0.70651	88
7	0.6880	161
8	0.6742	293

The goal of this article is to compare the characteristic exponent and measure theoretic entropy of the Hénon attractor. Computer experiments seem to indicate that there is an invariant measure associated with the Hénon attractor. This invariant measure has characteristic exponents and an entropy associated with it. We are interested in determining if these quantities are related in the manner proposed in Ref. 1 by Ruelle.

As was remarked earlier, given a dynamical system having no apparent symmetries it is not at all obvious how a good partition is to be chosen. What was done in the present case was to choose a partition, α, of the x-range of T which is contained in the interval $[-1.3, 1.3]$, i.e., partitions into vertical strips. The vertical partitions considered consisted of subdivisions of the x-range into two, three, five, and seven subintervals of equal length. We found that of all of the above partitions, the best results were achieved when the division into two equal subintervals was considered. Hence, the specific partition considered in many of the computations reported on here is the vertical partition $\alpha = \{[-1.3, 0), [0, 1.3]\}$ and its refinements which are obtained by dividing the x-range into four and eight subintervals having equal length. The "odd" partitions were not considered further.

A typical numerical experiment was performed as follows: Given α, symbolically represent the orbit of a given point on the attractor as a sequence of 0's and 1's. In order to compute the partition entropy of T given α we must tabulate the frequency of occurrence of configurations having lengths tending to infinity. What was done in practice was to consider configurations of maximum length 19, 16, and 15 for the two-, four-, and eight-element partitions, respectively, and to iterate T 1,000,000 times starting from a point on the attractor. Finally we recall the notation F_n^k which we use to denote the number computed from formula (2.4) when configurations of length n are considered and k elements are in the given partition.

4.1. Numerical Results

In Fig. 3 we have graphed F_n^k for the two-, four-, and eight-element partitions. Tables III, IV, and V were used to generate this figure. From the figure we can conclude that for configurations of lengths 19, 16, and 15 the partition entropy is approximately 0.40, which is within 5% of the characteristic exponent (value of 0.42) but is *less* than the characteristic exponent. It is straightforward to prove that F_n^k is a decreasing function of n, the configuration length. Hence for this partition we conclude that the characteristic exponent is strictly less than the measure theoretic entropy.

In Table VI we have tabulated the behavior of F_{19}^2 for various different vertical partitions. From the table it is clear that for all two-element vertical

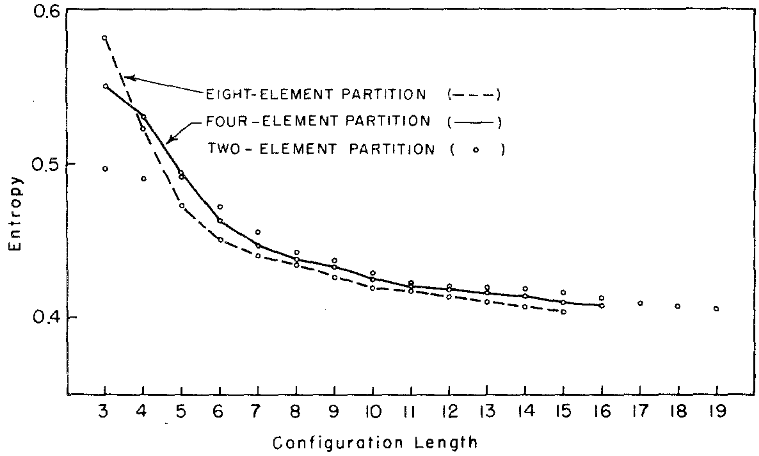

Fig. 3. Entropy of the Hénon attractor as a function of configuration length for a two-, four-, and eight-element partition, $a = 1.40$.

partitions considered the characteristic exponent strictly dominates the partition entropy. We call special attention to the last entry in the table. It was found that when the x-range was divided in this way the frequency of being in either the left or right element of the division was about equal. Note that $F_{19}^2 = 0.33650$, which is well below the characteristic exponent.

Table III. Two-Element Partition

n	F_n^2
3	0.49522
4	0.48886
5	0.48558
6	0.47051
7	0.45569
8	0.44111
9	0.43691
10	0.42794
11	0.42188
12	0.41954
13	0.41778
14	0.41685
15	0.41509
16	0.41262
17	0.41054
18	0.40888
19	0.40672

Table IV. Four-Element Partition

n	F_n^4
3	0.55578
4	0.53221
5	0.49499
6	0.46658
7	0.44610
8	0.43849
9	0.43473
10	0.42542
11	0.42196
12	0.41945
13	0.41471
14	0.41351
15	0.41061
16	0.40874

Table V. Eight-Element Partition

n	F_n^8
3	0.58591
4	0.52310
5	0.47268
6	0.44980
7	0.43970
8	0.43305
9	0.42287
10	0.41948
11	0.41612
12	0.41300
13	0.40936
14	0.40659
15	0.40353

Table VI. Two-Element Partition

	F_{19}^2
$[-1.3, -0.2), [-0.2, 1.3]$	0.37472
$[-1.3, -0.15), [-0.15, 1.3]$	0.38575
$[-1.3, -0.05), [-0.05, 1.3]$	0.39055
$[-1.3, 0), [0, 1.3]$	0.40672
$[-1.3, 0.05), [0.05, 1.3]$	0.40832
$[-1.3, 0.075), [0.075, 1.3]$	0.38307
$[-1.3, 0.2), [0.2, 1.3]$	0.35134
$[-1.3, 0.3), [0.3, 1.3]$	0.32355
$[-1.3, 0.419), [0.419, 1.3]$	0.33650

This demonstrates that the best possible partition is *not necessarily* a division into two equally probable pieces.

4.2. Other Parameter Values

We report here some results for other parameter values and confine our attention to the two-element partition of the previous section.

In the Introduction we noted that Misiurewicz and Szewc have shown that the stable and unstable manifolds of the fixed point near the Hénon attractor ($a = 1.40$, $b = 0.3$) has a transverse crossing. The first a value for which this occurs is approximately $a = 1.1538$.[8] The characteristic exponent for this parameter value is C.E. = 0.283. We have found that for this parameter value $F_{19}^2 = 0.257$ for the partition entropy given α. Another parameter value of interest is $a = 1.42690$, which is the value above which the unstable manifold of the fixed point lying in a neighborhood of the Hénon attractor tends to infinity. This parameter value "corresponds" to the value 4 for the mapping $f(x) = bx(1 - x)$ of the interval [0, 1] to itself.

In this case the C.E. = 0.424 while $F_{18}^2 = 0.412$. Here the partition entropy is within 3% of the characteristic exponent for α but is once again strictly less than the characteristic exponent. We have not, however, made any attempt to find the best possible two-element partition for these parameter values.

5. DISCUSSION

In this article we have reported on our attempts to verify the conjecture of Ruelle that the characteristic exponent and measure-theoretic entropy for the Hénon attractor should be equal. One of the principal problems encountered in this investigation was that of choosing a partition. In the case of the Hénon attractor, which has no apparent symmetries, we have considered only vertical partitions. Our choice was based on the hypothesis that elements in such a partition would mix sufficiently rapidly to achieve the desired limit.[7]

The numerical results of Section 4 indicate that two-element vertical partitions will give a value of F_n^2 which is strictly less than the characteristic exponent. A surprising result of our computations is that the partition, α, we considered seems to account for most of the entropy in the dynamical system. A somewhat more interesting fact is that one of the very worst partitions is the two-element vertical partition which makes being in either element equally probable.

In Section 3 we noted that when we chose a good partition there was almost immediate convergence to the entropy of the transformation, while when we chose the wrong partition the number computed was less than the entropy of the transformation and seemed to achieve its limit at a greatly reduced rate. The fact that we captured almost all the entropy of the Hénon attractor with a two-element vertical partition was "lucky." There are, however, a large number of "unlucky" partitions.

There are two points which must be made concerning the computations reported here. In order to compute the characteristic exponent of a dynamical system it is necessary to iterate the transformation and perform a minimal amount of tabulation. On the other hand, computing the measure-theoretic entropy for a given partition is a much larger task. Suppose we want to retain all configurations of length 20 and to iterate until the frequency of occurrence of each configuration has converged to its asymptotic value. Our computation suggests that for a two-element partition there are more than 20,000 such configurations. To ensure that each configuration has achieved its asymptotic frequency of occurrence our experiments suggest that the number of iterations required will be of the order of 100 times the number of configurations. If we extrapolate these empirical observations to a partition having eight elements, assuming the same asymptotic growth rate for new configurations as for two-element partitions, then we would expect about 8×10^{12} configurations of length 20.

If we now assume that we must iterate T 100 times in order for each configuration to achieve its asymptotic frequency of occurrence, then we need on the order of 10^{15} iterations of T. If each iteration requires 10^{-6} sec, then the total task would require about 10^9 sec or 10^4 days with current computing power.

In view of the impracticability of computations of this size, there are some mathematical questions which need to be addressed. Among them are the following: Is there a more efficient method for computing the partition entropy than formulas (2.2) or (2.3)? Is there any way to optimize the rate of convergence of the partition entropy?

Finally, the numerical results presented here are consistent with the conjecture of D. Ruelle,[(1)] but they are not definitive. I believe that if there were some guiding mathematical principal which would aid in choosing a partition, the results could be made definitively. Or if it were known for example that the unstable manifold of the Hénon attractor was the support of a smooth invariant measure, then the recent work of P. Walters [15] would allow us to conclude that Ruelle's question has an affirmative answer.

ACKNOWLEDGMENTS

Thanks to E. N. Lorenz for stimulating discussions concerning the computation of entropy and suggesting a closer look at the "rooftop map."

Thanks to P. Collet and S. K. Wong for several discussions. Special thanks to O. Lanford and D. Ruelle for incisive comments and remarks on an earlier version of this manuscript. Finally, all computations reported were performed at the National Center for Atmospheric Research. This work was supported in part by the Air Force Geophysics Laboratory under grant No. AF F 19628-78-C-0032 and by the Centre National de la Recherche Scientifique, while the author was a visitor at the IHES Institut des Havtes Études Scientifiques).

REFERENCES

1. D. Ruelle, Dynamical Systems with Turbulent Behavior, International Mathematics Physics Conference, Rome 1977, *Lecture Notes in Physics* No. 80 (Springer, Berlin, 1978).
2. D. Ruelle, A Measure Associated with Axiom A Attractors, *Am. J. Math.* **98**:619–654 (1976).
3. I. Shimada, Gibbsian Distribution on the Lorenz Attractor, *Prog. Theor. Phys.* **62**:61–69 (1979).
4. E. Lorenz, Deterministic Nonperiodic Flow, *J. Atmos. Sci.* **20**:130–141 (1963).
5. M. Hénon, A Two-Dimensional Mapping with a Strange Attractor, *Commun. Math. Phys.* **50**:69–77 (1976).
6. S. Feit, Characteristic Exponents and Strange Attractors, *Commun. Math. Phys.* **61**:249–260 (1978).
7. J. H. Curry, On the Hénon Transformation, *Commun. Math. Phys.* **68**:129–140 (1979).
8. C. Simó, On the Hénon-Pomeau Attractor, *J. Stat. Phys.* **21**:465–493 (1979).
9. F. Morroto, *Commun. Math. Phys.* **68**:187–194 (1979).
10. M. Misiurewicz and B. Szewc, Existence of a Homoclinic Point for the Hénon Map, *Commun. Math. Phys.* **75**:285–291 (1980).
11. R. Adler, A. Konheim, and M. McAndrew, Topological Entropy, *Trans. Am. Math. Soc.* **114**:309–319 (1965).
12. R. Bowen, On Axiom A Diffeomorphisms, *CBMS Lectures in Mathematics, Held at North Dakota State University* (Am. Math. Society, Providence, 1977).
13. P. Walters, Ergodic Theory—Introductory Lectures, *Lecture Notes in Mathematics* No. 458 (Springer, Berlin, 1975).
14. P. Billingsley, *Ergodic Theory and Information* (John Wiley and Sons, New York, 1965).
15. P. Walters, Characteristic Exponents and the Entropy of Diffeomorphisms, Preprint (1979).