Physica 13D (1984) 90–104
North-Holland, Amsterdam

THE LOCAL STRUCTURE OF A CHAOTIC ATTRACTOR IN FOUR DIMENSIONS

Edward N. LORENZ
Center for Meteorology and Physical Oceanography, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Received 5 December 1983

A simple system of four ordinary differential equations exhibiting chaotic behavior is introduced. The Lyapunov exponents are approximately 3.34, 0.00, −1.79, and −5.55. Points on the attractor where two variables simultaneously assume prechosen values are located by a numerical integration procedure in which suitable antecedent conditions are obtained by successive approximations. The points are organized into curves, which appear more numerous under higher resolution. A backward integration procedure indicates that a line transverse to the curves intersects each curve rather than passing through gaps, while the set of intersections forms a Cantor set. Thus the attractor appears to be locally the product of three continua and one Cantor set. The continua are found to extend in the expanding, neutral, and less rapidly contracting local directions, while the Cantor set extends in the more rapidly contracting direction.

1. Introduction

Some time ago we introduced the dynamical system

$$\dot{x} = -\sigma x + \sigma y,$$
$$\dot{y} = -xz + rx - y,$$
$$\dot{z} = xy - bz, \tag{1}$$

where σ, b and r are positive constants and a dot denotes differentiation with respect to time t. These equations were derived as a model of forced dissipative fluid convection, although they do not describe intense convection very realistically. Wholly apart from their physical meaning, they define a steady flow in (x, y, z)-space, with each point moving along its orbit. We found that with $\sigma = 10$, $b = 8/3$, and $r = 28$, the numerically determined solutions were generally aperiodic, while the attractor appeared to be the product of two continua and a Cantor set [1].

In the same paper we identified a class of dynamical systems, some of which were existing models of forced dissipative processes, defined by the N equations

$$\dot{x}_i = \sum_{j,k=1}^{N} a_{ijk} x_j x_k - \sum_{j=1}^{N} b_{ij} x_j + c_i, \tag{2}$$

with the constants chosen so that $\Sigma a_{ijk} x_i x_j x_k$ vanishes identically and $\Sigma b_{ij} x_i x_j$ is positive definite. For such systems

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\sum_{i=1}^{N} x_i^2 = -\sum_{i,j=1}^{N} b_{ij} x_i x_j + \sum_{i=1}^{N} c_i x_i. \tag{3}$$

Because of the restrictions on b_{ij}, the right-hand side of eq. (3) vanishes on the surface of an ellipsoid E in $(x_1, \ldots, x_N)$-space, and is negative throughout the region exterior to E. As a consequence, if S(0) is any sphere centered at the origin and enclosing a region R(0) containing E, every orbit either is contained in R(0) in its entirety or else penetrates S(0) and remains in R(0) afterward. The attractor A is therefore contained in R(0).

Under the flow, S(0) is carried after any time interval δt into a surface S(δt), completely contained in R(0) and enclosing a region R(δt) which

0167-2789/84/$03.00

also contains the attractor. Likewise $S(2\delta t)$ is completely contained in $R(\delta t)$, etc., so that the attractor is contained in $R(\infty) = R(\delta t) \cap R(2\delta t) \cap \ldots$.

Of special interest is the subclass of systems where, in addition to the restrictions already imposed on the constants, a_{ijk} vanishes if $j = i$ or $k = i$. In this event an infinitesimal volume v is governed by

$$\dot{v}/v = -\sum_{i=1}^{N} b_{ii} < 0, \tag{4}$$

so that v diminishes exponentially. Because Σb_{ii} is constant, eq. (4) holds even when v is not infinitesimal, and, in particular, when v is the volume of $R(0)$. Hence $R(n\delta t)$ shrinks exponentially with n, and $R(\infty)$ and therefore A have zero volume.

We extend the definitions of the class and the subclass to include any system which satisfies the specified conditions after a linear transformation of the variables. The subclass thus includes eqs. (1), which acquire the proper form after $z' = z - r - \sigma$ is substituted for z.

The purpose of this study is to examine the nature of the attractor for a system of more than three equations which belongs to the subclass and possesses an aperiodic general solution. We are particularly interested in whether the attractor is locally the product of identifiable continua and Cantor sets, and, if so, how many continua and how many Cantor sets. We shall confine our attention to the local structure of A, leaving unanswered such involved questions as the manner in which separate sheets or hypersheets of the attractor are joined together.

2. Lyapunov exponents and fractal dimension

We shall restrict our attention to the subclass defined above. Let us rewrite eqs. (2) as

$$\dot{\boldsymbol{X}} = \boldsymbol{F}(\boldsymbol{X}), \tag{5}$$

where $\boldsymbol{X}$ and $\boldsymbol{F}$ are matrices of N rows and one column, or vectors, with elements x_i and f_i. An infinitesimal departure $\boldsymbol{Z}(t)$ from $\boldsymbol{X}(t)$ is then governed by

$$\dot{\boldsymbol{Z}} = G\boldsymbol{Z}, \tag{6}$$

where the elements g_{ij} of the square matrix G are the partial derivatives $\partial f_i/\partial x_j$. Integrating eq. (6) between times t_0 and t_1 we obtain

$$\boldsymbol{Z}(t_1) = H\boldsymbol{Z}(t_0), \tag{7}$$

where the square matrix H depends upon the values of $\boldsymbol{X}$ between t_0 and t_1. An easy way to evaluate H is to perturb $\boldsymbol{X}$ separately in each coordinate direction, and subtract the unperturbed numerical solution of eq. (5) from each perturbed solution to obtain each column.

An infinitesimal sphere S′, given by $\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{Z} = \varepsilon^2$ at time t_0, where T denotes a transpose, will be carried at time t_1 into an ellipsoid E′ given by $\boldsymbol{Z}^{\mathrm{T}}(HH^{\mathrm{T}})^{-1}\boldsymbol{Z} = \varepsilon^2$. The semi-axes of E′ are $\gamma_1\varepsilon, \ldots, \gamma_N\varepsilon$, where γ_j are the singular values of H, or the square roots of the eigenvalues of HH^{T}, arranged so that $\gamma_1 > \cdots > \gamma_N$. If the limits

$$\lambda_j = \lim_{t_1 \to \infty} \log \gamma_j/(t_1 - t_0) \tag{8}$$

exist, they are the characteristic exponents or Lyapunov exponents of the orbit passing through $\boldsymbol{X}(t_0)$. If these limits are the same for almost all orbits, they are the Lyapunov exponents of the system.

If $\boldsymbol{X}(t_0)$ and hence $\boldsymbol{X}(t_1)$ are on A, the limits in eq. (8) may be replaced by limits as $t_0 \to -\infty$ while t_1 remains fixed. Unit vectors $\boldsymbol{U}_j$ parallel to the axes of E′ may also approach limits as $t_0 \to -\infty$, in which case these limits will be called the Lyapunov vectors at $\boldsymbol{X}(t_1)$.

To evaluate λ_j and $\boldsymbol{U}_j$ we choose $t_1 - t_0$ sufficiently large, and factor H as

$$H = U\Gamma V, \tag{9}$$

where U and V are orthogonal and Γ is diagonal. The equation for E′ becomes $(U^{\mathrm{T}}\boldsymbol{Z})^{\mathrm{T}}\Gamma^{-2}(U^{\mathrm{T}}\boldsymbol{Z}) = \varepsilon^2$, so that the diagonal elements of Γ are γ_j,

while the vectors $\boldsymbol{U}_j$ are the rows of U^{T}, i.e., the columns of U.

To determine $\boldsymbol{U}_j$ at a prechosen point $\boldsymbol{X}(t_1)$ we can perturb $\boldsymbol{X}(t_1)$ separately in each coordinate direction and integrate backward to t_0, obtaining the separate columns of H^{-1}. We then factor H^{-1}, noting that

$$H^{-1} = V^{\mathrm{T}}\Gamma^{-1}U^{\mathrm{T}}. \tag{10}$$

It is algebraically simpler to factor H as

$$H = U'\Gamma'W, \tag{11}$$

where U' is orthonormal, Γ' is diagonal, and W has elements w_{ij} which vanish if $j < i$ and equal unity if $j = i$. Although U' and Γ' differ from U and Γ, they ordinarily lead to the same Lyapunov vectors and exponents as $t_1 \to \infty$. If $t_1 - t_0$ is not large, U and Γ appear to yield better approximations.

For a chaotic system, evaluations of H or H^{-1} are subject to sampling variations. In practice, if $t_1 - t_0$ is made large enough to render these variations unimportant, the diagonal elements of Γ (or Γ') are likely to differ by many orders of magnitude, and the roundoff errors may completely falsify all except the largest one. This difficulty may be circumvented by using special algorithms which factor H at frequent intervals between t_0 and t_1.

The sum $\Sigma\lambda_j$ is the exponential damping rate $-\Sigma b_{ii}$. If the orbit through $\boldsymbol{X}(t_0)$ approaches a fixed point asymptotically, $\lambda_1 < 0$; otherwise the points where this orbit intersects S′ will remain separated by a distance comparable to 2ε, and one exponent will vanish (see [2]). If $\lambda_1 = 0$ and $\lambda_2 < 0$, the orbit is tangent to $\boldsymbol{U}_1$; if $\lambda_K = 0$ and $\lambda_{K+1} < 0$ for some K, the orbit is tangent to the manifold formed by $\boldsymbol{U}_1, \ldots, \boldsymbol{U}_K$. If $\lambda_1 > 0$, the system behaves chaotically.

A number of recent studies [3, 4, 5] have related the fractal dimension of A to λ_j; a detailed review and bibliography appear in [5]. A formula suggested by Kaplan and Yorke [3] is

$$d = L + \sum_{j=1}^{L} \lambda_j / |\lambda_{L+1}|, \tag{12}$$

where L is the largest number of exponents which can be added to yield a non-negative sum. The various definitions of dimension are not equivalent; in particular, if the number of spheres of radius ρ needed to cover all points of A increases as $\rho^{-d'}$ as $\rho \to 0$, the Hausdorff dimension of A is d'. We shall treat d as simply another definition of dimension. For many systems d and d' are not far apart, but, if some points of A are visited much less frequently than others, d is generally smaller [5].

When $N = 3$ and the system is chaotic, $\lambda_1 > 0$, $\lambda_2 = 0$, and $\lambda_3 < 0$, so that $d = 2 + \lambda_1/|\lambda_3|$. Our original computations with eqs. (1) indicated that the sheets of the attractor formed two clusters, and that, along a line cutting across the sheets, the maximum distance between sheets in the same cluster was about 0.00007 times the distance between clusters. Thus d' would be about $2 + 1/\log_2(1/0.00007) = 2.073$. Our most recent estimates of d, with H factored as in eq. (11) and with $t_1 - t_0$ large, yield values of 0.92, 0.00, and -14.59, so that $d = 2.063$ (compare [4, 6]). The agreement is remarkable in view of the crudeness in estimating d'. Our intent in the present study is to treat a system where $d > 3$, and where the fractional part of d is considerably greater than 0.063.

3. The equations

Our choice of a dynamical system has been influenced by our desire to determine a very simple system of four equations, if possible the simplest one, capable of exhibiting chaotic behavior, belonging to the class and the subclass previously defined, and satisfying the additional condition that the equations be unaltered by a cyclic permutation of the variables. The last condition was

introduced with the hope of obtaining four points on the attractor for the price of one, but its chief effect has been to limit the choice so greatly as to lead almost immediately to the system

$$\begin{aligned}\dot{y}_1 &= y_2(y_3 - y_4) - y_1 + c,\\ \dot{y}_2 &= y_3(y_4 - y_1) - y_2 + c,\\ \dot{y}_3 &= y_4(y_1 - y_2) - y_3 + c,\\ \dot{y}_4 &= y_1(y_2 - y_3) - y_4 + c.\end{aligned} \tag{13}$$

Other arrangements of the quadratic terms are possible, but none appears to be simpler. Eqs. (13) have some obvious similarities to eqs. (1), but we have not succeeded in associating them with any particular physical system.

Through the transformation

$$\begin{aligned}x_1 &= (y_1 + y_2 + y_3 + y_4)/2,\\ x_2 &= (y_1 - y_2 + y_3 - y_4)/2,\\ x_3 &= (y_1 + y_2 - y_3 - y_4)/2,\\ x_4 &= (y_1 - y_2 - y_3 + y_4)/2,\end{aligned} \tag{14}$$

the equations become

$$\begin{aligned}\dot{x}_1 &= -x_2^2 + x_3^2/2 + x_4^2/2 - x_1 + 2c,\\ \dot{x}_2 &= x_1x_2 - x_3^2/2 + x_4^2/2 - x_3x_4 - x_2,\\ \dot{x}_3 &= (x_1 - x_2)(x_3 + x_4)/2 - x_3,\\ \dot{x}_4 &= (x_1 + x_2)(x_3 - x_4)/2 - x_4.\end{aligned} \tag{15}$$

The matrix of the transformation is self-inverse. The system retains four-way symmetry, since, if (x_1, x_2, x_3, x_4) is a time-dependent solution, so are $(x_1, -x_2, x_4, -x_3)$, $(x_1, x_2, -x_3, -x_4)$, and $(x_1, -x_2, -x_4, x_3)$. We shall study the attractor A of eqs. (15). However, since eqs. (13) are much simpler to integrate numerically, we shall use them in our computations, transforming between x_i and y_i whenever input or output is required.

The ellipsoid E becomes the sphere of radius c centered at $(c, 0, 0, 0)$. For the sphere S(0) centered at the origin we choose a radius infinitesimally greater than $2c$.

Eqs. (15) always possess the steady solution $(2c, 0, 0, 0)$; this is stable when c is small enough, but becomes unstable when $c > 1/2$. Two more stable steady solutions $(1, \pm\sqrt{2c-1}, 0, 0)$ then appear; these become unstable when $c > 3$. Four more stable steady solutions then appear; numerical integration indicates that these become unstable when $c > 3.85$ and give way to periodic behavior, which in turn gives way to chaotic behavior when $c > 60.7$. We shall study the supercritical case $c = 100$.

For negative c the first solution becomes unstable when $c < -1$. Chaotic behavior sets in when $c < -11.84$. Detailed consideration of this range of c is beyond the scope of this study.

In our numerical integrations with $c = 100$ we have chosen a time increment $\Delta t = 0.005$. At the beginning t_0 of each time step, we compute the first four derivatives of $\boldsymbol{X}(t_0)$, and evaluate the truncated Taylor series

$$\boldsymbol{X}(t_0 + \Delta t) = \sum_{k=0}^{4} (\mathrm{d}^k\boldsymbol{X}/\mathrm{d}t^k)\,\Delta t^k/k! \tag{16}$$

To evaluate $\boldsymbol{X}$ between time steps we substitute a fraction of Δt for Δt in eq. (16). Successive approximations allow us to determine $\boldsymbol{X}$ at some special time, such as a zero-crossing of one variable.

Fig. 1 shows 400 points at intervals of Δt along the projection of a typical orbit on the (x_1, x_2)-plane. Fig. 2 shows the projection of the same orbital segment on the (x_3, x_4)-plane. Evidently, each variable undergoes several oscillations during the two time units. Although the system is chaotic, each variable exhibits a distinctive behavior; figs. 1 and 2 could not possibly be mistaken for one another.

The sum of the Lyapunov exponents must be -4. To determine the individual exponents we have factored H according to eq. (11). We have made two estimates of λ_j with $t_1 - t_0 = 25.0$. The average values are 3.34, 0.00, -1.79, and -5.55, so that $d = 3.28$. The separate estimates of d are

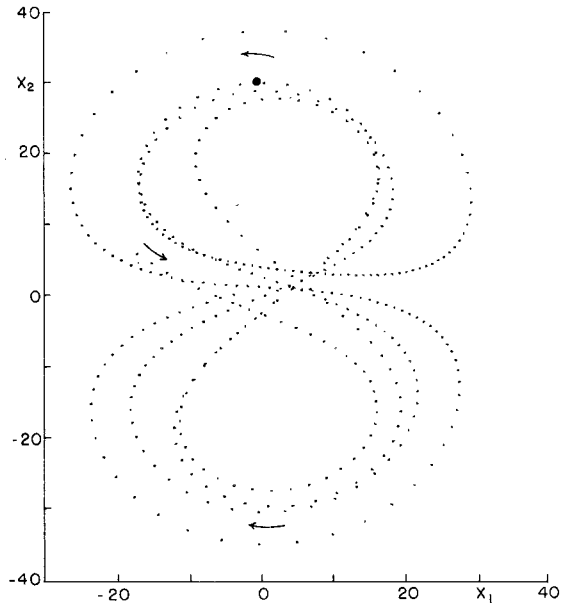

Fig. 1. Projection of particular numerical solution of eqs. (15), with $c = 100$, on (x_1, x_2)-plane, as represented by 400 points at time intervals $\Delta t = 0.005$. Arrows show direction of flow. Large dot at (0,30) indicates value of (x_1, x_2) chosen for detailed study.

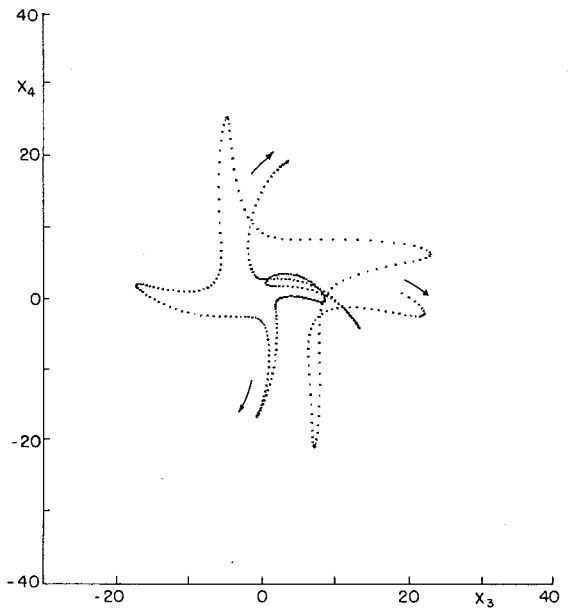

Fig. 2. Same as fig. 1, except projection of solution is on (x_3, x_4)-plane.

3.26 and 3.30, suggesting that the sampling variations are under control. We can therefore expect the Cantor-set structure of A to be easily visible, in contrast to eqs. (1).

4. The attractor

The attractor of eqs. (1) consists of a Cantor set of surfaces, but, because d is only slightly greater than 2.0, it is closely approximated by a pair of surfaces. From the numerically determined solution we were able to construct contours of x, on a diagram with y and z as coordinates.

In the four-dimensional space of eqs. (15) we cannot immediately use the same technique, but we can first eliminate one dimension by considering only the intersection A′ of A with a prechosen hypersurface, say the hyperplane Q* where x_1 assumed a fixed value x_1^*. An orbit in A will cross Q* at frequent intervals, and the values of the remaining variables at these crossings may be determined with high accuracy. If d were only slightly greater than 3.0, A′ might be approximated by just a few surfaces, and we might be able to draw contours of x_2 for each surface, with x_3 and x_4 as coordinates.

Because d is nearly 3.3, the Cantor set of surfaces comprising A′ requires too many surfaces to approximate it closely, and drawing contours is not feasible. We therefore attempt to display the structure of A through two-dimensional cross sections.

Fig. 3 shows such a section for eqs. (1). The curve is the intersection of the attractor with the plane $x = 4$, and is identical to the single contour $x = 4$ in the contour representation. The two apparent branches are actually Cantor sets of curves which have not been resolved.

For eqs. (15) we choose for our cross section the intersection A″ of A with the plane P* where x_1 and x_2 assume fixed values x_1^* and x_2^*. We choose a point $\boldsymbol{X}_0$ on Q*, with $\dot{x}_1 < 0$, and let $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots$ be the subsequent downward ($\dot{x}_1 < 0$) intersections of the orbit through $\boldsymbol{X}_0$ with Q*. We choose an

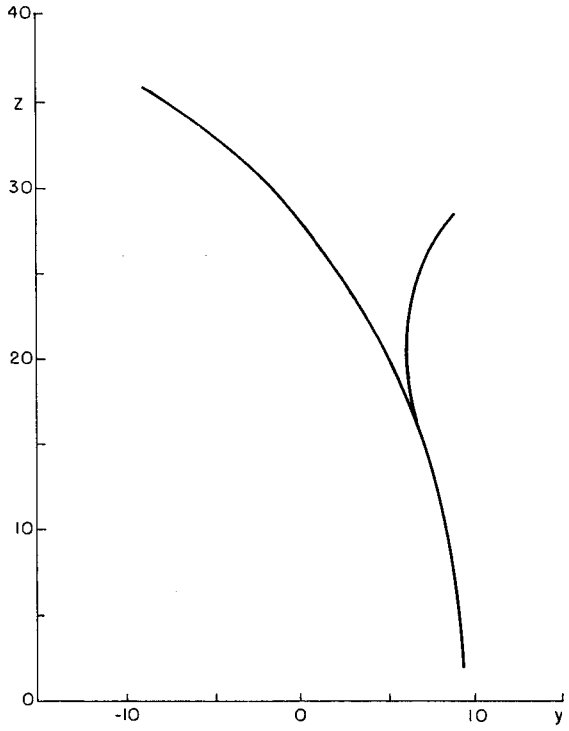

Fig. 3. Intersection of attractor of eqs. (1), for $\sigma = 10$, $b = 8/3$, and $r = 28$, with plane $x = 4$.

initial perturbation $\boldsymbol{Z}_0$ of magnitude ε, parallel to Q*. For $m = 1, 2, 3, \ldots$, we let $\boldsymbol{X}_m + \alpha_m \boldsymbol{Z}_m$ be the next downward intersection of the orbit through $\boldsymbol{X}_{m-1} + \boldsymbol{Z}_{m-1}$ with Q*, where $\boldsymbol{Z}_m$ is of magnitude ε and α_m is an amplification factor. We anticipate that, except for small m, $\boldsymbol{Z}_m$ will be parallel to the manifold formed by $\boldsymbol{U}_1$ and the orbit, which is the manifold formed by $\boldsymbol{U}_1$ and $\boldsymbol{U}_2$, and hence will be parallel to A′.

If $\boldsymbol{X}_m$ is reasonably close to the hyperplane $x_2 = x_2^*$ (or $x_2 = -x_2^*$, in view of the symmetry), we estimate, for $n = 1, 2, \ldots$, on the basis of $\boldsymbol{X}_m$, $\boldsymbol{X}_{m-n}$ and the amplification factors $\alpha_m, \alpha_{m-1}, \ldots$, the factor β_{mn} which will make the nth subsequent downward intersection of the orbit through $\boldsymbol{X}_{m-n} + \beta_{mn}\boldsymbol{Z}_{m-n}$ with Q* lie on P*. Presumably the intersection will prove not to lie on P*, but its proximity will allow us to make further estimates of β_{mn}. We continue until the intersection is within a distance ε of P*, and record the corresponding values of x_3 and x_4, except that, if the procedure has not converged after a few estimates, we quit and go to the next value of m.

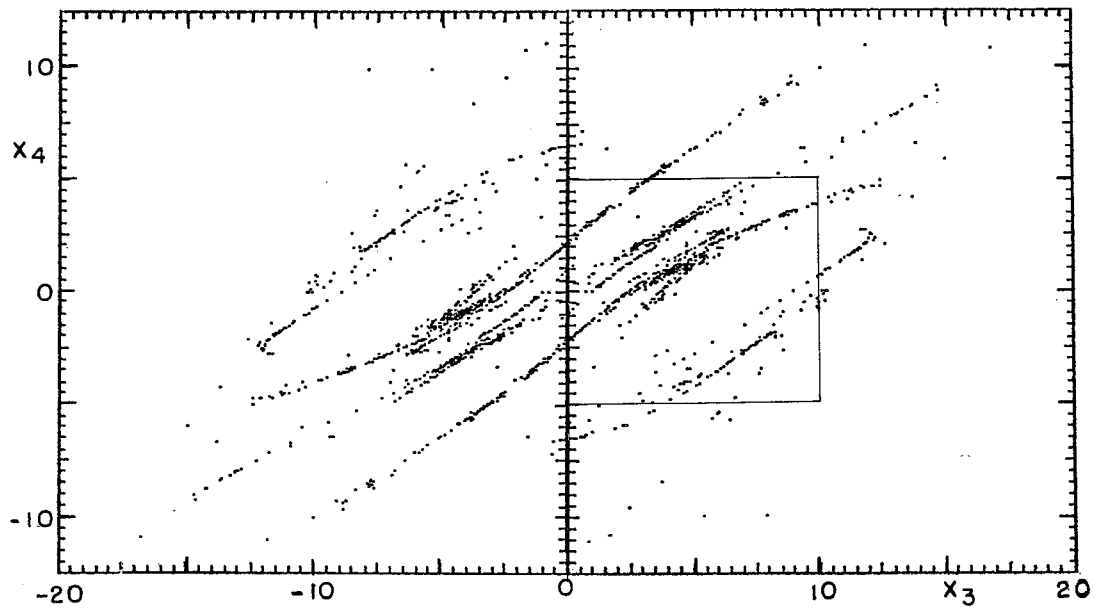

Fig. 4. Intersection of attractor of eqs. (15), for $c = 100$, with plane $x_1 = 0$, $x_2 = 30$. Square indicates region to be enlarged in fig. 5.

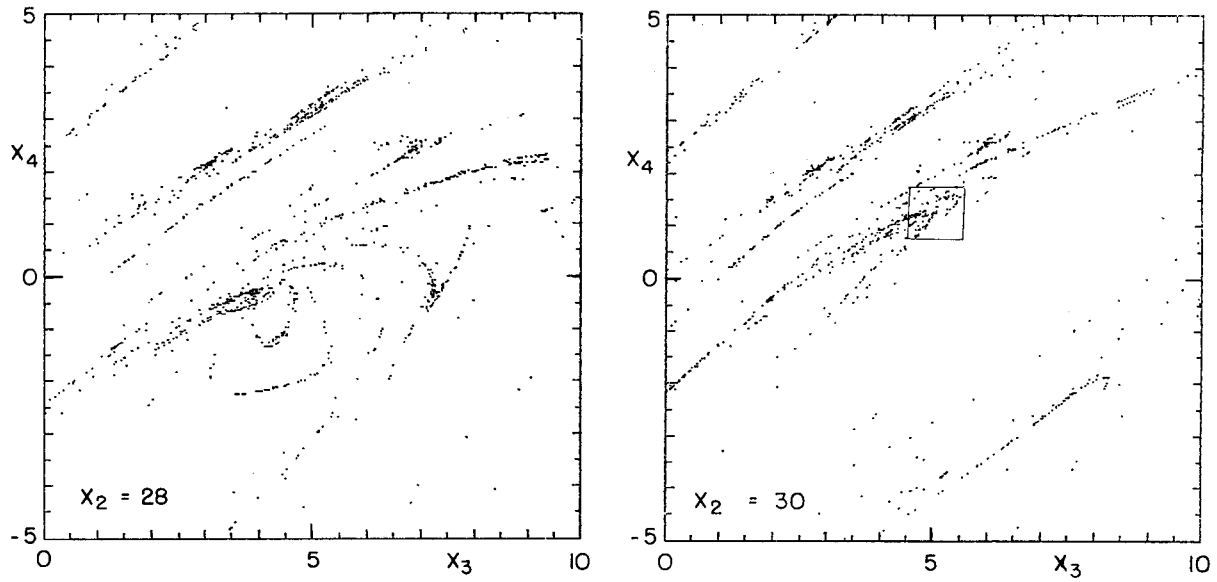

Fig. 5. Portion of intersection of attractor of eqs. (15) with plane $x_1 = 0$, $x_2 = 28$ (left), and plane $x_1 = 0$, $x_2 = 30$ (right). Right portion is enlargement of region contained in square in fig. 4. Small square indicates region to be enlarged in fig. 6.

With $x_1^* = 0$, $x_2^* = 30$, and $\varepsilon = 10^{-5}$, we obtain convergence for about half of the values of m. Finding a point on A″ requires about 50 times the computations needed to find one on A′. We have proceeded to determine 4000 points on A″, with $x_3 > 0$. All of these project onto the large dot in fig. 1.

Fig. 4 is the central result of this study, toward which the remainder of the study is addressed. It shows the first 1500 computed points on A″, together with their reflections in the origin; many are indistinguishable with this resolution. The points are organized into a number of curves, which appear to be continua. Locally they tend to be more or less parallel. In places they bend sharply and reverse their direction. If we had attempted to represent A′ by contours of x_2, every curve in fig. 4 would be an $x_2 = 30$ contour.

Fig. 5 contains an enlargement of the square region in fig. 4, and compares it with a similar cross section where $x_2^* = 28$ instead of 30, whose implications we shall consider later. Fig. 6 is a further enlargement of the small square in fig. 5. What looked like a few curves are now revealed as a great many. Reversals of direction are abundant.

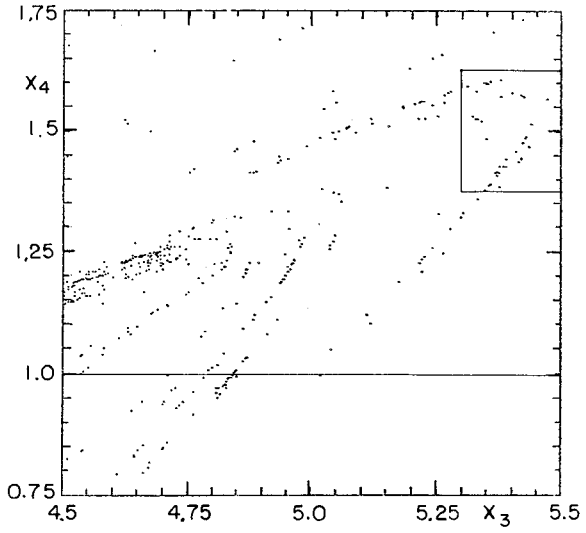

Fig. 6. Enlargement of region contained in small square in fig. 5. Rectangle indicates region to be examined in fig. 11. Horizontal line indicates segment studied with backward integration procedure (see text).

The appearance of additional curves under higher resolution is consistent with the presence of a Cantor set of curves; thus A″ seems to be the

product of a continuum and a Cantor set. Pending further results, we hypothesize that A is the product of three continua and one Cantor set.

5. Continua or Cantor sets?

Despite their apparent continuity, the curves as shown in figs. 4 and 5 possess prominent gaps, as they must when they are represented by a finite number of points. We therefore cannot rule out the possibility that the curves are Cantor sets. If they are indeed continua, further computation will eventually discover points in any present gap, but the required amount of computation may be prohibitively large. We therefore turn to another procedure.

If the curves are continua, any line or other smooth curve cutting across them will intersect each of them. If instead they are Cantor sets, a typical transverse line will pass through a gap in almost every curve. Possibly such a line will intersect the curves in a Cantor set of smaller fractal dimension than that of the Cantor set of curves, but there should be identifiable curves which it does not intersect.

Our new computations are designed to determine whether a chosen point $\boldsymbol{X}$ lies on the attractor, or, if $R(\infty)$ is larger than the attractor, whether it lies on $R(\infty)$. If $\boldsymbol{X}$ is in $R(0)$ but not in $R(\infty)$, it lies on $S(\tau)$ for some τ. In principle we can determine τ by integrating backwards from $\boldsymbol{X}$, and finding the time required for the orbit to intersect $S(0)$. Large values of τ presumably imply proximity to $R(\infty)$.

Accordingly, we choose a line segment cutting across a collection of curves, and seek τ at each of a set of points along the segment. If the curves are continua, the maximum values of τ should in-

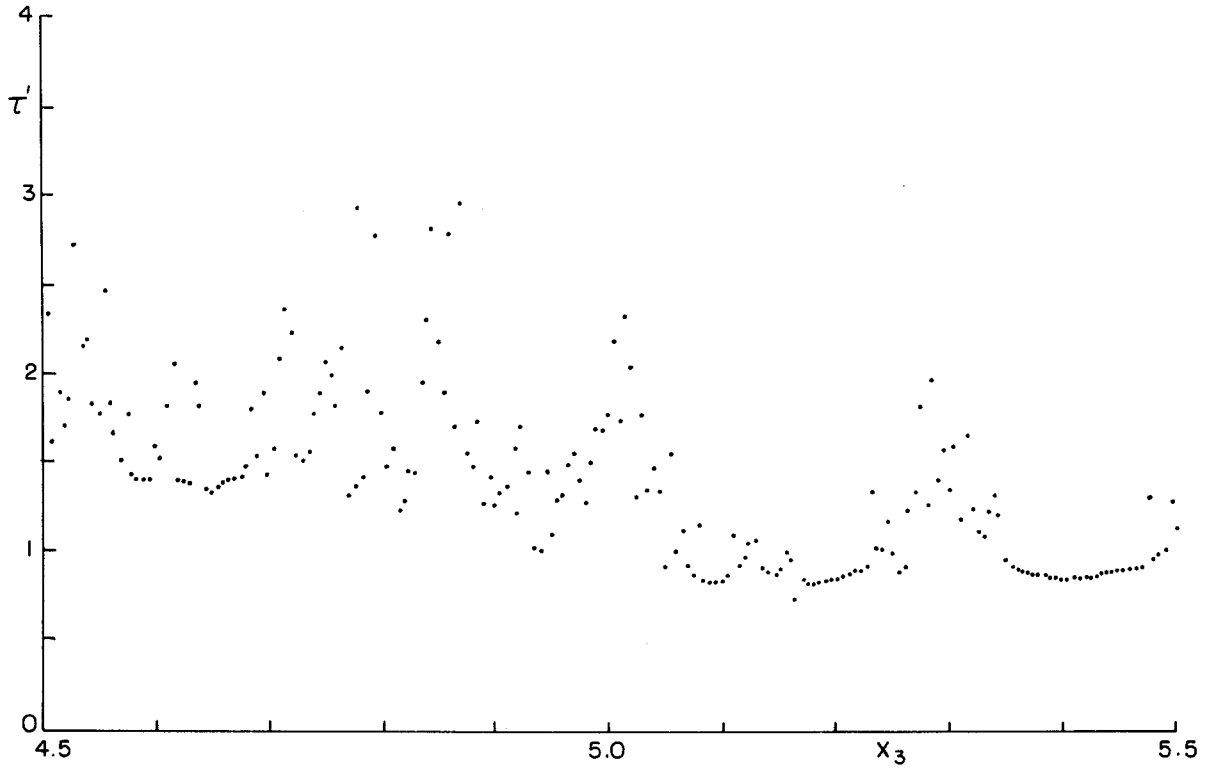

Fig. 7. Values of τ' (see text) at intervals $\delta x = 0.005$ along horizontal line in fig. 5.

crease without limit with increased resolution along the segment.

Our initial computations of τ yielded values which varied rather erratically, even with very high resolution. This may have occurred because points which are further from the origin tend to move much faster along their orbits (note the spacing of points in fig. 1), and, with a fixed Δt, the computations become less stable as S(0) is approached. Decreasing Δt sufficiently would increase the computation prohibitively. Examination of earlier computations revealed no instances when the distance r' from the origin of a point moving in A reached 100 units (the radius of S(0) is 200), and only brief intervals when r' exceeded 50. If there are close approaches of orbits in A to S(0), these are very rare events. Accordingly, we terminate each backward integration when r' reaches 100, and define τ' as the time when r' last equals 50. We then adopt τ' in place of τ as a measure of proximity to R(∞).

Backward integration cannot be the exact reverse of forward integration, nor can either one duplicate the exact behavior of the differential equations. We assume that the principal effect of the numerical approximations is to displace the values of the variables or of c at which key features of A occur, rather than destroying key features or creating new ones.

We have chosen the horizontal segment shown in fig. 6, where $x_4 = 1$ and x_3 extends from 4.5 to 5.5. Fig. 7 shows the behavior of τ' along the segment, as revealed with a resolution $\delta x = 0.005$. There are intervals where τ' is small and varies

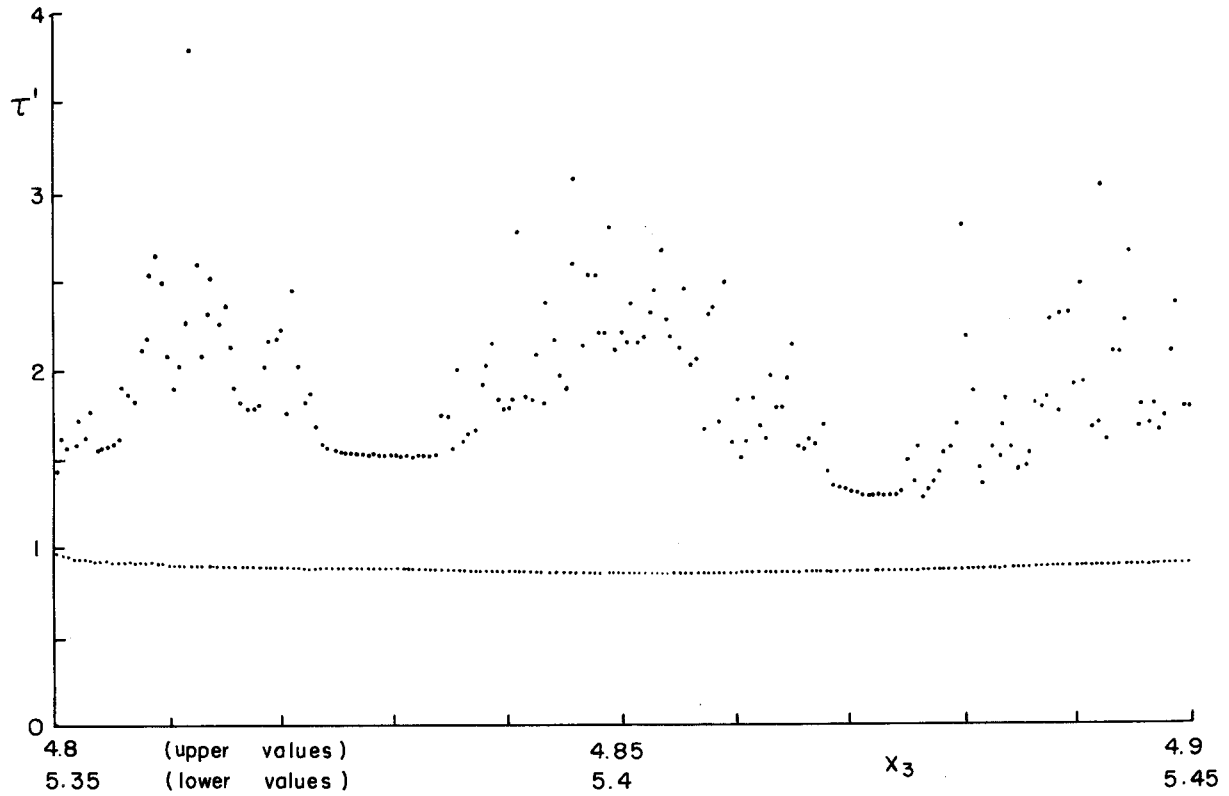

Fig. 8. Values of τ' at intervals $\delta x = 0.0005$ along segments of horizontal line in fig. 5, extending from $x_3 = 5.35$ to $x_3 = 5.45$ (lower values), and from 4.7 to 4.8 (upper values).

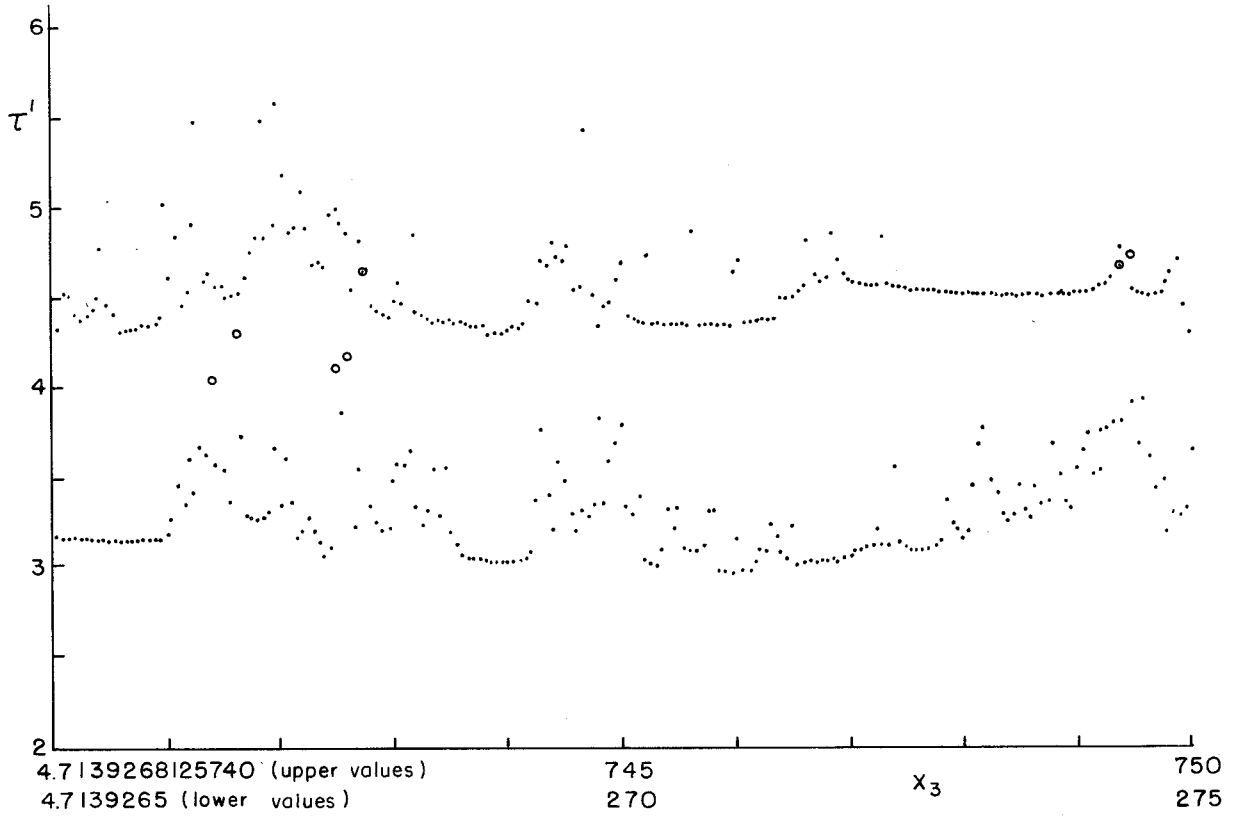

Fig. 9. Values of τ' at intervals $\delta x = 0.5 \times 10^{-8}$ along segment of horizontal line in fig. 5 extending from $x_3 = 4.7139265$ to $x_3 = 4.7139275$ (lower values), and at intervals $\delta x - 0.5 \times 10^{-14}$ along segment from 4.713926812574 to 4.713926812575 (upper values). Lower dots have been replaced by circles where they would otherwise be confused with upper dots.

smoothly; these correspond to locations in fig. 6 where no curves are evident, and presumably lie between curves. There are other intervals where τ' is generally larger and varies erratically. The clusters of high values showing peaks at $x_3 = 4.53$, 4.715, 4.78, 4.87, and 5.015 correspond to recognizable curves in fig. 6; the weaker peak at 5.285 presumably corresponds to a downward and leftward extension of a line of points just to the right of the square in fig. 5.

In fig. 8, we decrease δx to 0.0005 in the interval from 5.35 to 5.45, where the behavior is smooth, and from 4.7 to 4.8, where it is erratic. The smooth interval remains smooth, and shows no evidence of finer structure. The erratic interval proves to possess smooth and erratic subintervals, and, in fact, the upper curve in fig. 8 shows no qualitative difference from fig. 7, except for a general upward displacement. There is no suggestion that maximum values of τ' have been encountered.

With greater resolution, say with δt a million times smaller, we would expect nearly all intervals of length 200 δx to be smooth, with only a small fraction of them resembling fig. 7. Resolving the whole segment to this degree would require excessive computation, but we can obtain similar results by choosing intervals at random, or in any other manner not guided by the values of τ'. We have examined the ten intervals of length 10^{-6} beginning at $4.55, 4.65, \ldots, 5.45$, with $\delta x = 0.5 \times 10^{-8}$

Only the interval beginning at 4.75 shows any erratic behavior, and this occupies less than five percent of the interval.

The situation is different when we decrease δx stepwise by factors of ten, subjectively choosing at each step an interval which exhibited erratic behavior during the previous step. The lower curve in fig. 9 shows the behavior of τ' in an erratic interval of length 10^{-6}, with $\delta x = 0.5 \times 10^{-8}$. Again, the curve is much like fig. 7, although the smooth subintervals are somewhat more prominent. In the upper curve in fig. 9 the interval has been decreased stepwise to 10^{-12}, and $\delta x = 0.5 \times 10^{-14}$. The qualitative behavior is still unchanged.

Upon nearly exhausting the "double precision" employed, we find no sign that τ' possess an upper limit. Moreover, we have not found any smooth local maxima of τ', which we would expect if the segment passed through gaps in some of the curves, whereas smooth local minima abound. Whenever we have noted an exceptionally high value of τ', we have succeeded in finding a still higher value nearby, by increasing the resolution. There is therefore a strong suggestion that the curves are continua, and that A″ is locally the product of a continuum and a Cantor set.

We can, in fact, estimate the Hausdorff dimension of the Cantor set by noting what fraction of the segment remains unresolved into smooth intervals as δx is successively decreased. Since the end points of a smooth interval may be hard to identify, we choose the simpler procedure of asking what fraction of pairs of points on the segment with horizontal separation δx have values of τ' differing by $2\Delta t$ or more. We assume that as δx varies, this fraction is proportional to the fraction which remains unresolved into smooth intervals under some other definition.

We choose the 200 values of x_3 used in fig. 7 (omitting the last value) as reasonably representa-

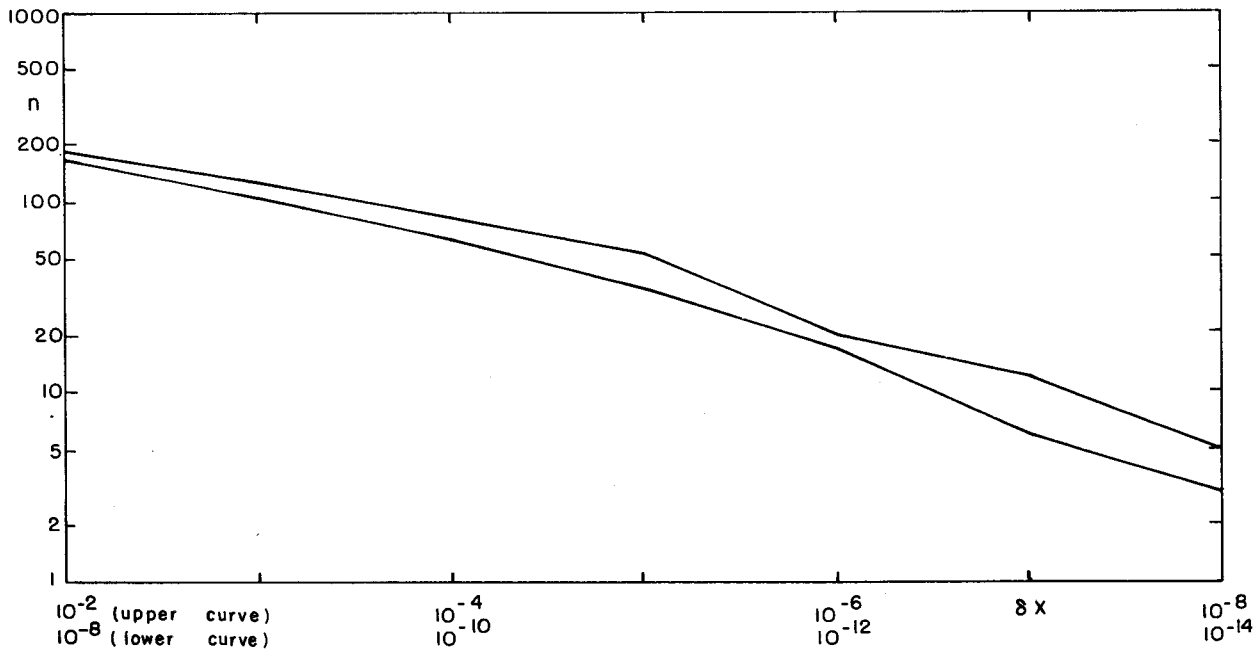

Fig. 10. Number n of unresolved pairs of points (see text), out of a total of 200 points, along horizontal line in fig. 5, with resolutions δx from 10^{-2} to 10^{-8} (upper curve), and along segment of horizontal line extending from $x_3 = 4.7139265$ to $x_3 = 4.7139275$, with resolutions δx from 10^{-8} to 10^{-14} (lower curve).

tive first members of a pair, and determine the number n of "unresolved" pairs as δx decreases by factors of ten from 10^{-2} to 10^{-8}. The results are given by the upper curve in fig. 10; δx and n are plotted logarithmically. The dimension should be $1+s$, where s is the slope of the curve. For good measure we have also chosen the values of x_3 used in the lower curve in fig. 9, letting δx decrease from 10^{-8} to 10^{-14}. The lower curve in fig. 10, which shows the results, is generally similar in slope to the upper curve.

Although the slopes of the curves are not uniform, they seem to be no steeper than -0.3, implying that the dimension is at least 0.7, and thus suggesting that d' is at least 3.7. This estimate is considerably higher than d. Possible R(0) has a higher dimension than A, but perhaps the difference occurs simply because d is ordinarily less that d' when different points of A are visited with different frequencies [5].

6. The directions of the curves

Whether or not the curves in figs. 4 and 6 are continua, it seems evident that locally they possess definite directions of orientation. These directions are among the most noticeable features of A″. To interpret them we note that the curves will retain their identity if we alter x_1 or x_2. If we should make fig. 4 three-dimensional by adding x_2 as a vertical coordinate, the curves would become surfaces. We return to fig. 5, which effectively shows the intersections of these surfaces with two horizontal planes. Evidently the surfaces are nearly vertical in the upper and left portions of the region, but in the central portion some surfaces intersecting $x_2 = 28$ fail to reach $x_2 = 30$, and perhaps are dome shaped.

If we could make fig. 4 four-dimensional by adding x_1 also, the curves would become hypersurfaces. We hypothesize that these hypersurfaces are tangent to the manifold formed by $\boldsymbol{U}_1$, $\boldsymbol{U}_2$, and $\boldsymbol{U}_3$. The directions of the curves should therefore

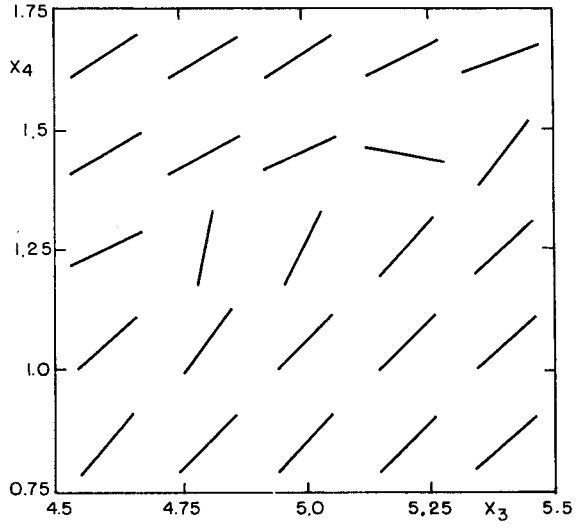

Fig. 11. Vectors parallel to intersection of manifold formed by first three Lyapunov vectors with plane $x_1 = 0$, $x_2 = 30$, in region covered by fig. 6.

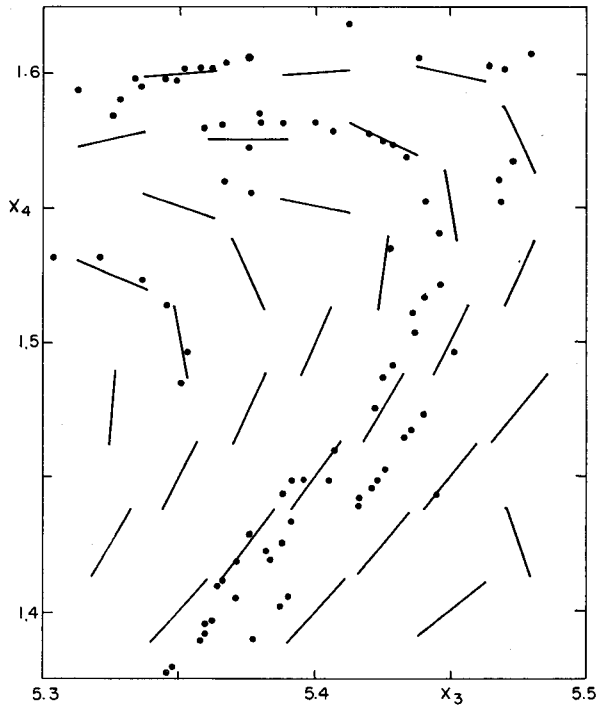

Fig. 12. Same as fig. 11, in region contained in rectangle in fig. 6. Dots indicating computed points on attractor are also included.

be the directions of the intersection of this manifold with P*.

Since U_1, U_2, and U_3 are the first three columns of the orthonormal matrix U in eq. (9), any vector in the hyperplane is orthogonal to the fourth column of U. If such a vector is also parallel to P*, its first two components vanish, and, aside from a scalar factor, it must be $(0, 0, u_{44}, -u_{34})$.

To evaluate U at a given point $X(t_1)$, we integrate backward to some time t_0, obtaining H^{-1}, which we then factor according to eq. (10). The procedure may be carried out even when $X(t_1)$ is not on the attractor, in which case it will appear to converge before finally diverging. We prefer this procedure to the computationally simpler factorization of H^{-1} because it appears to give a better approximation when $t_1 - t_0$ is small.

In fig. 11 we show tangent vectors at a grid of points, in the region covered by fig. 6. We obtained convergence with $t_0 = t_1 - 2.0$ at each point, with one exception, where one more time unit sufficed. The general positive slope, which is generally steeper in the lower portion of the figure, agrees well with the curves. Some odd directions appear where the curves are turning sharply, but the resolution is not sufficient to determine whether the vectors follow the turns. To see whether they do, we have, in fig. 12, enlarged the region indicated by the rectangle in fig. 6, where the curves are turning rapidly. We have superposed computed vectors on the available computed points, which include some points not appearing in fig. 6. It is evident that the vectors turn with the curves. The one odd direction, in the lower right portion where there are no available points, suggests that some structure of still finer scale may be present.

7. Concluding remarks

We have chosen a dynamical system where N is greater than in eq. (1) and d is closer to N than in eqs. (1). We have obtained numerical evidence that the attractor is the product of continua and Cantor sets. Locally, the continua extend in the directions of U_1, U_2, and U_3, while the Cantor set extends in the direction of U_4. Because λ_3 and λ_4 are both negative, we might have anticipated two Cantor sets, the sum of whose fractal dimensions could exceed unity.

The added variable merely complicates the computational procedure. The more complex appearance of the two-dimensional cross section, with many curves and direction reversals, is due entirely to the higher fractional part of d. The same features will appear with suitably chosen three-dimensional flows or two-dimensional mappings. An example of such a mapping is

$$x_{n+1} = \tfrac{1}{2}(1 + (ax_n - y_n)/(a+1))\cos(\pi a x_n),$$
$$y_{n+1} = \tfrac{1}{2}(1 + (ax_n - y_n)/(a+1))\sin(\pi a x_n), \quad (17)$$

defined for $a > 0$ and $x_n^2 + y_n^2 < 1$. Eqs. (17) resemble the simpler Hénon mapping [7]; they compress the unit circle vertically, stretch it horizontally, and then roll it up. Figs. 13 and 14 show portions of the attractor, for $a = 2.0$, when d is about 1.15, and $a = 2.1$, when d is about 1.35. The qualitative resemblance of fig. 3 to fig. 13, and fig. 4 to fig. 14, is apparent.

The Cantor sets of eqs. (15), and more generally of eqs. (2), seem to be associated with the same circumstances which are present under eqs. (1). Because A is bounded and the flow within it is aperiodic, any point moving along its orbit must eventually be approached asymptotically by a point moving in from some other portion of A. Each point carries with it, in A, a surface or hypersurface, extending in the direction of each Lyapunov vector with a non-negative exponent; there are at least two such directions. These surfaces or hypersurfaces must also approach each other asymptotically, and what appears to become a sheet or hypersheet must remain at least a pair. This pair is subsequently approached by another pair, after which the resulting pair of pairs is approached by another pair of pairs, etc., and the inevitable consequence is a Cantor set.

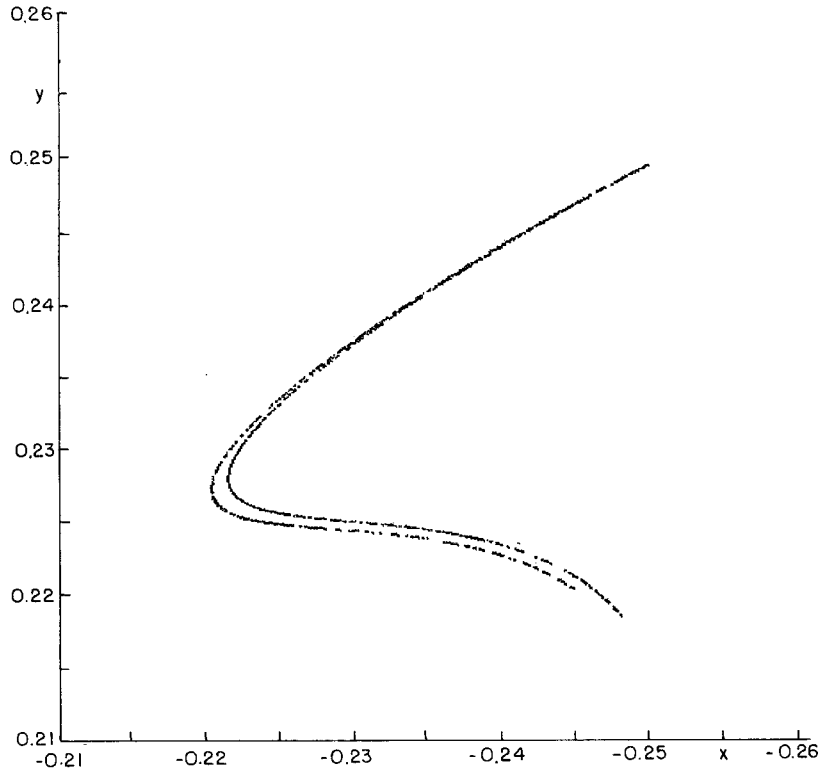

Fig. 13. Portion of attractor of eqs. (17), for $a = 2.0$.

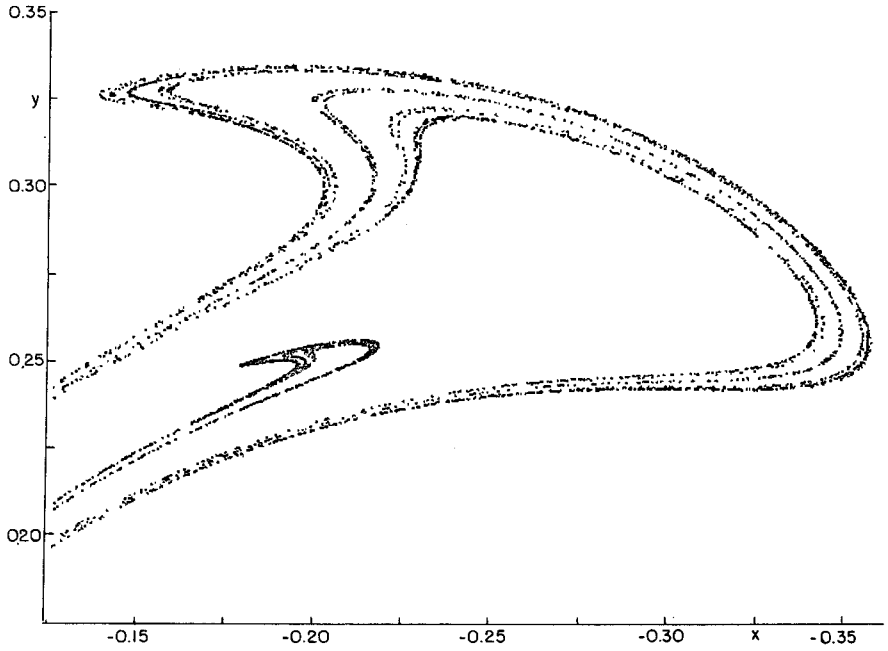

Fig. 14. Same as fig. 13, for $a = 2.1$.

Acknowledgements

A part of this research, including the portion described in sections 3 and 4, was performed while the author was a summer visitor at the National Center for Atmospheric Research, Boulder, Colorado. NCAR is supported by the National Science Foundation. The remainder of the work, including the portion described in sections 5 and 6, was performed at the author's permanent address, and was supported by the GARP Program of the Atmospheric Sciences Section, National Science Foundation, under Grant 82-14582 ATM.

References

[1] E.N. Lorenz, J. Atmos. Sci. 20 (1963) 130.
[2] H. Haken, Phys. Lett. 94A (1983) 71.
[3] J.L. Kaplan and J.A. Yorke, Lecture Notes in Math. 730, (Springer, Berlin, 1978) p. 228.
[4] H. Mori, Prog. Theor. Phys. 63 (1980) 1044.
[5] J.D. Farmer, E. Ott and J.A. Yorke, Physica 7D (1983) 153.
[6] I. Shimada and T. Nagashima, Prog. Theor. Phys. 61 (1979) 1605.
[7] M. Hénon, Comm. Math. Phys. 50 (1976) 69.