Commun. Math. Phys. 68, 187–194 (1979)

Communications in Mathematical Physics

Chaotic Behavior in the Hénon Mapping

Frederick R. Marotto

Department of Mathematics, Drexel University, Philadelphia, PA 19104, USA

Abstract. In a previous work Hénon investigated a two-dimensional difference equation which was motivated by a hydrodynamical system of Lorenz. Numerically solving this equation indicated for certain parameter values the existence of a "strange attractor", i.e., a region in the plane which attracts bounded solutions and in which solutions wander erratically. In the present work it is shown that this behavior is related to the mathematical concept of "chaos". Using general methods previously developed, it is proven analytically that for some parameter values the mapping has a transversal homoclinic orbit, which implies the existence of the chaotic behavior observed by Hénon.

1. Introduction: The Hénon Mapping

In a recent work Hénon [2] investigated the dynamics of the mapping of the plane into itself defined by the difference equation:

$$\begin{aligned} x_{k+1} &= y_k + 1 - a x_k^2, \\ y_{k+1} &= b x_k, \end{aligned} \tag{1}$$

where $a, b \in \mathbb{R}$. Numerically solving (1) for a variety of initial values, he found this system to exhibit a very complex type of behavior. In particular for certain values of a and b Hénon found the existence of a "strange attractor" in (x_k, y_k) phase-space, that is, a region in the plane which attracts bounded solutions from outside under iteration of (1), and in which trajectories of (1) exhibit essentially random behavior. The implications of such behavior are significant. Once a strange attractor is observed, very unpredictable behavior of solutions will result. This is due to a lack of global stability of any solution, and more importantly, an extreme sensitivity to initial conditions. Ruelle and Takens [8] have suggested that such behavior is related to turbulence in the flow of fluids.

The principal motivation for consideration of (1) was an analysis conducted by Lorenz [4] upon a system of partial differential equations describing finite am-

0010-3616/79/0068/0187/$01.60

plitude convection in a fluid heated from below. In this work Lorenz converted the problem into a system of three ordinary differential equations and then solved this system numerically. For certain parameter values the resulting trajectories appeared to oscillate seemingly at random around either of two equilibrium points, alternating between them. The Lorenz system thus appeared to exhibit a random type of behavior.

With Lorenz's work in mind, Hénon attempted to develop a mathematical model which exhibits the same qualitative features as the Lorenz system, but which is more tractable to analysis. Such a model was constructed by consideration of a Poincaré map which can reduce a three-dimensional continuous problem to a two-dimensional discrete mapping. Hénon thus posed the model (1) as a qualitative approximation to the Lorenz system.

It is important to note that Hénon's observations concerning the complex behavior of (1) were based upon numerical studies of the system, not upon exact analytical methods. The purpose of this work is to provide just such a mathematical proof of the behavior observed by Hénon. In particular, we shall show analytically that (1) satisfies sufficient conditions for the system to be (what has been termed) chaotic.

2. The Concept of Chaos

The phenomenon of chaos is relatively new and not very well understood. Although chaotic forms of behavior had previously been observed in a variety of different settings, the first to use the term "chaos" were Li and Yorke [3], who considered the general scalar difference equation:

$$x_{k+1} = f(x_k) \quad f: \mathbb{R} \to \mathbb{R}. \tag{2}$$

A precise definition of chaos is presented in their work, but the essential implications of chaos are the following: (i) there exist an infinite number of periodic solutions of different periods; (ii) there exists an uncountably infinite set of points which exhibit random behavior when iterated under (2); and (iii) there is an extreme sensitivity to initial conditions. Hence there appears to be an intimate connection between the concept of chaos and the type of behavior exhibited by (1). Indeed, Ruelle and Takens have also proposed that chaos is the mathematical analogue of turbulence in the flow of fluids. It is also interesting to note that studies of the Lorenz system were the primary motivation for the investigations that led to the results of Li and Yorke.

In addition to defining chaos, Li and Yorke present a theorem giving sufficient conditions for its existence in problems of the form (2). Although this theorem has proven useful in explaining the complex behavior exhibited by many scalar equations, it is not extendable to multidimensional problems of the form:

$$X_{k+1} = F(X_k) \quad F: \mathbb{R}^n \to \mathbb{R}^n \tag{3}$$

(although the concept of chaos is extendable).

A theorem which does in fact provide conditions for chaos of (3) was previously proven by Smale [9] (although he did not refer to it by that name). Suppose Z is a conditionally stable fixed point of (3) where F is a diffeomorphism. That is, $F(Z)=Z$ and some eigenvalues of the jacobian of F at Z exceed 1 in norm, and the rest are less than 1 in norm. Under these conditions there exist stable and unstable manifolds of F at Z. The stable manifold is comprized of those points X_0 whose positive limit set is the point Z, i.e., $X_k \to Z$ as $k \to \infty$. Also, X_0 is in the unstable manifold if the negative limit set of X_0 is Z, i.e., $X_{-k} \to Z$ as $k \to \infty$. [Note that X_{-k} can be evaluated by iterating the inverse of (3) since F is a one-to-one mapping.] Suppose these manifolds intersect transversally (i.e., non-tangentially) at some point X_0 other than Z. The resulting trajectory $\{X_k\}_{k=-\infty}^{+\infty}$ has the properties $X_k \to Z$ and $X_{-k} \to Z$ as $k \to \infty$, and is called a *transversal homoclinic orbit* of F. If we let F^M represent the composition of F with itself M times, Smale has proven the following.

Theorem 1. *If F is a diffeomorphism and has a transversal homoclinic orbit, then there exists a Cantor set $\Lambda \subset \mathbb{R}^n$ in which F^M is topologically equivalent to the shift automorphism for some M.*

The existence of such a shift automorphism implies that within the set Λ there exists a dense collection of periodic solutions of different periods of (3) and an uncountably infinite collection of points which are asymptotically aperiodic. (There is also a sensitivity to initial conditions.) Thus a transversal homoclinic orbit implies a form of chaotic behavior similar to that defined by Li and Yorke and similar to that described by Hénon for the problem (1). In the next section we shall in fact show that (1) has a transversal homoclinic orbit for some parameter values, thus proving analytically the behavior observed by Hénon.

Although Theorem 1 is of theoretical interest, it is extremely difficult to apply directly to any particular problem. In order to check the hypotheses of this theorem, the stable and unstable manifolds must first be computed, and then shown to intersect transversally. In most cases these manifolds cannot be computed exactly. At best the eigenspaces tangent to them at the fixed point Z can be used as an approximation. In order to show a transversal intersection, these approximations to the manifolds can be discretized and then iterated under (3) hopefully producing a non-tangential intersection. This must usually be done visually with the aid of computer graphing devices. This was the approach used by Curry [1] in numerical studies of (1). In this work very careful numerical techniques were employed to demonstrate what is apparently a transversal homoclinic orbit. It is clear, however, that, although these findings strongly suggest the existence of such an orbit, they do not constitute an analytic proof, which is what we desire.

Another theorem which provides conditions for chaos of multidimensional problems of the form (3) was previously proven in [7]. Suppose Z is an unstable fixed point of F such that all eigenvalues of the jacobian of F at Z exceed 1 in norm. That is, no stable manifold exists. In this case the unstable manifold locally contains all points inside $B_r(Z)$, the ball of radius r around Z, for some $r>0$. Also, suppose there exists an initial point X_0 of (3) with $X_0 \in B_r(Z), X_0 \neq Z, X_M = Z$ for some $M>0$, and non-zero jacobian of F at each X_k for $0 \leq k < M$. Such a fixed point Z may be called a *snap-back repeller*. Note that F cannot be one-to-one if a snap-back repeller exists. In [7] the following is proven.

Theorem 2. *If F is differentiable and has a snap-back repeller, then* (3) *is chaotic.*

Here the definition of chaos is the same as that characterized by Li and Yorke. In fact, it is proven in [5] that Theorem 2 is (roughly) a generalization of their theorem. Theorem 2 is also closely related to Smale's conditions for chaos. That is, it is the complementary result for the case when no stable manifold exists. This can be easily seen by investigating the positive and negative limit sets of the point X_0 in the definition of a snap-back repeller. Since $X_M = Z$ and Z is a fixed point, we must have $X_k = Z$ for all $k \geqq M$. Also, since $X_0 \in B_r(Z)$ and $B_r(Z)$ is in the unstable manifold, we can find a sequence of X_k's for $k < 0$ where $X_k \to Z$ as $k \to -\infty$. The orbit $\{X_k\}_{k=-\infty}^{+\infty}$ is analogous to a homoclinic orbit of F. In addition it is apparent that the jacobian of $F(X_k)$ will be non-zero for all k. This is analogous to transversality. The existence of such a sequence $\{X_k\}_{k=-\infty}^{+\infty}$ may be taken as an equivalent characterization of a snap-back repeller Z.

There are, however, important practical differences between transversal homoclinic orbits and snap-back repellers. The former, as we have seen, is extremely difficult to compute. Snap-back repellers, on the other hand, are relatively easy to find, often requiring only finite iteration processes. The investigations in [7] show the simplicity with which snap-back repellers were computed for a number of problems of the form (3). (In the following section a simple method of computing them will be demonstrated.)

It would therefore be convenient if chaotic behavior could be proven with the use of Theorem 2 rather than Theorem 1. Unfortunately in the problem at hand, Theorem 2 is not applicable since both stable and unstable manifolds exist. We shall therefore require another result previously developed in [6]. In certain circumstances the dynamics of a problem of the form (3) can be determined by reducing the problem to one of lower dimension. This will be true when the higher dimensional problem is a small perturbation of the reduced equation. In particular, the two-dimensional problem:

$$\begin{aligned} u_{k+1} &= f(u_k, bv_k), \\ v_{k+1} &= u_k, \end{aligned} \tag{4}$$

where $f: \mathbb{R}^2 \to \mathbb{R}$ is differentiable, can be reduced to the one-dimensional equation:

$$u_{k+1} = f(u_k, 0) \tag{5}$$

when b is close to 0. The following is proven in [6].

Theorem 3. *Suppose* (5) *has a snap-back repeller. Then* (4) *has a transversal homoclinic orbit for all* $|b| < \varepsilon$ *for some* $\varepsilon > 0$.

Since, for each fixed value of a, (1) can be written in the form (4), we can show the existence of a transversal homoclinic orbit of (1) by first reducing it to a corresponding one-dimensional problem of the form (5), and applying Theorem 3. This is the method we shall use.

Note that Theorem 3 does not provide an explicit range of b values for which a transversal homoclinic orbit exists, i.e., no estimate for ε is given. This is of minor consequence to our analysis of (1) since we are primarily interested in the qualitative features of the model. The precise parameter values for which Hénon investigated the problem were chosen arbitrarily – to aid in a visual investigation of the strange attractor.

3. A Proof of Chaos for the Hénon Mapping

In order to apply Theorem 3 first note that (1) can be equivalently written:

$$u_{k+1}=bv_k+1-au_k^2,\qquad v_{k+1}=u_k, \tag{6}$$

where $x_k=u_k$ and $y_k=bv_k$. The dynamics of (6) must be identical to those of (1), i.e., if (6) has a transversal homoclinic orbit so does (1). For each fixed value of a, (6) is now of the form (4) where $f(u,v)=v+1-au^2$. According to Theorem 3, (6) will have a transversal homoclinic orbit for all $|b|<\varepsilon$ for some $\varepsilon>0$, if the problem:

$$u_{k+1}=1-au_k^2=g(u_k) \tag{7}$$

has a snap-back repeller. We shall show this for appropriate values of a, in particular for $a>1.55$.

First observe that $u^*=[-1+(1+4a)^{1/2}]/2a$ is an unstable fixed point of (7), i.e., $g(u^*)=u^*$ and $g'(u^*)=1-(1+4a)^{1/2}<-1$ for all $a>0.75$. Now if we can find a solution $\{u_k\}_{k=-\infty}^{+\infty}$ with not all $u_k=u^*$ satisfying: (i) $u_k=u^*$ for all $k\geqq M$ for some M; (ii) $u_k\to u^*$ as $k\to-\infty$; and (iii) $g'(u_k)\neq 0$ for all k, then u^* is a snap-back repeller. [The condition $g'(u_k)\neq 0$ is the jacobian condition required of the X_k's in the definition of a snap-back repeller.] Such a sequence can be generated in the following manner. If we let $u_0=u^*$ then, since u^* is a fixed point of (3), $u_k=u^*$ for all $k\geqq 0$. In addition u_k for $k<0$ can be constructed by iterating the multi-valued inverse of (7):

$$u_{k-1}=\pm[(1-u_k)/a]^{1/2}=g_{\pm}^{-1}(u_k) \tag{8}$$

provided $u_k\leqq 1$. With $u_0=u^*$ we have two choices for u_{-1} according to (8). Choosing the plus sign in (8), however, will not yield an appropriate sequence, since we would have $u_{-1}=g_+^{-1}(u_0)=g_+^{-1}(u^*)=u^*$. Therefore define $u_{-1}=g_-^{-1}(u_0)$. Note that $u_{-1}=-u^*$. Thereafter let $u_{k-1}=g_+^{-1}(u_k)$ in (8) for all $k\leqq -1$.

We shall show that this sequence $\{u_k\}_{k=-\infty}^{0}$ satisfies $u_k\to u^*$ as $k\to-\infty$ (for appropriate values of a). Note that since $u_{-1}=-u^*<u^*$:

$$u_{-2}=g_+^{-1}(u_{-1})\in g_+^{-1}[(-\infty,u^*)]\subset(u^*,+\infty).$$

Let us find those values of a for which $u_{-2}<1$. Since $u_{-1}=-u^*=[1-(1+4a)^{1/2}]/2a$, then by (8) $u_{-2}=g_+^{-1}(u_{-1})=[(1+u^*)/a]^{1/2}$. So, $u_{-2}<1$ implies $u^*<a-1$, or $[-1+(1+4a)^{1/2}]/2a<a-1$. This is equivalent to $a^3-2a^2+2a-2>0$. It can be shown that all values of $a>1.55$ satisfy this equation, and thus $u_{-2}\in(u^*,1)$ for these values of a. We restrict the remainder of the discussion to the problem when $a>1.55$.

Now because $u_{-2}\in(u^*,1)$ for these a values:

$$u_{-3}=g_+^{-1}(u_{-2})\in g_+^{-1}[(u^*,1)]\subset(0,u^*)$$

and consequently $u_{-3}\in(0,u^*)$. Also:

$$u_{-4}=g_+^{-1}(u_{-3})\in g_+^{-1}[(0,u^*)]\subset g_+^{-1}[(u_{-1},u^*)]\subset(u^*,u_{-2})$$

and so $u_{-4}\in(u^*,u_{-2})$. This implies:

$$u_{-5}=g_+^{-1}(u_{-4})\in g_+^{-1}[(u^*,u_{-2})]\subset(u_{-3},u^*)$$

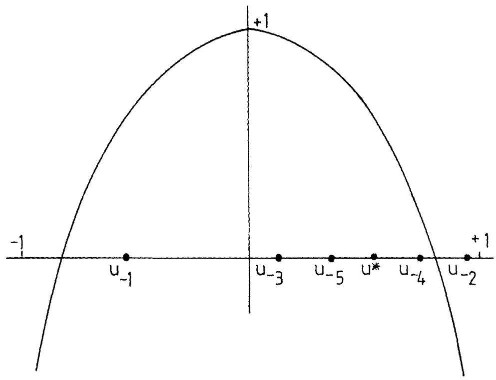

Fig. 1. Inverse iterates of $u_{k+1} = 1 - au_k^2$ for $a > 1.55$

and hence $u_{-5} \in (u_{-3}, u^*)$. Continuing in this manner, it is apparent that the sequence $\{u_k\}_{k=-\infty}^{0}$ thus constructed satisfies the following: u_{-2k} is a decreasing sequence bounded below by u^*, and u_{-2k-1} is an increasing sequence bounded above by u^* (Fig. 1). There must therefore exist a point $\alpha \in (0, u^*]$ which is the limit of u_{-2k-1}, and a point $\beta \in [u^*, 1)$ which is the limit of u_{-2k} as $k \to \infty$.

We shall show $\alpha = \beta = u^*$. Since $g(u_{-2k-1}) = u_{-2k}$ and $g(u_{-2k}) = u_{-2k+1}$, it must be that $g(\alpha) = \beta$ and $g(\beta) = \alpha$. Consequently, $g(g(\alpha)) = \alpha$, and α is thus a fixed point of the function $g \circ g$. But for $a > 1.55$ there are precisely four fixed points of $g \circ g$ (since this function is a quartic polynomial) each of which can be computed exactly:

$$[-1 \pm (1+4a)^{1/2}]/2a \quad \text{and} \quad [1 \pm (4a-3)^{1/2}]/2a.$$

It is clear that, for $a > 1.55$, $[-1-(1+4a)^{1/2}]/2a$ and $[1-(4a-3)^{1/2}]/2a$ are both negative, and thus neither of these can equal $\alpha \in (0, u^*]$. Also, suppose

$$[1+(4a-3)^{1/2}]/2a \leqq u^* = [-1+(1+4a)^{1/2}]/2a.$$

Then, $(4a+1)^{1/2} \geqq 2 + (4a-3)^{1/2}$. Squaring each side of this inequality implies $4(4a-3)^{1/2} \leqq 0$ which is a contradiction for $a > 1.55$. This implies $[1+(4a-3)^{1/2}]/2a > u^*$. Therefore it must be that $\alpha = [-1+(1+4a)^{1/2}]/2a = u^*$, and $\beta = g(\alpha) = u^*$. Hence $u_k \to u^*$ as $k \to -\infty$.

So far we have shown that the sequence $\{u_k\}_{k=-\infty}^{+\infty}$ satisfies (i) and (ii) from above. It can easily be shown that (iii) is also satisfied. Since $g'(u) = -2au$, the only possible way for $g'(u_k) = 0$ is if $u_k = 0$ for some k. But from the manner in which the sequence was constructed: $u_k = u^*$ for $k \geqq 0$; $u_{-1} = -u^* < 0$; $u_{-2} > u^* > 0$; $u_{-3} > 0$; and $u_k \in (u_{-3}, u_{-2}) \subset (0, 1)$ for all $k < -3$. Thus the sequence also satisfies (iii), and u^* is therefore a snap-back repeller of (7). Now applying Theorem 3, we see that for each $a > 1.55$ Eq. (6), and hence Eq. (1), have transversal homoclinic orbits for all $|b| < \varepsilon$ for some $\varepsilon > 0$.

4. Extending the Analysis

For each $a > 1.55$ we have proven the existence of a transversal homoclinic orbit, and hence chaotic behavior, of (1) for all $|b| < \varepsilon$. The region of a values for which this

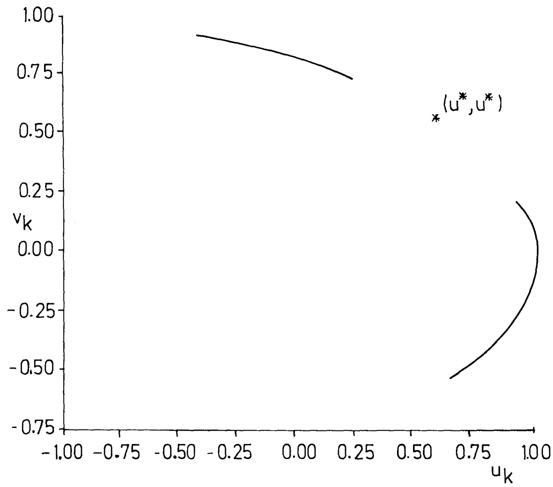

Fig. 2. For $b=0$ and $1.44<a<1.55$ the strange attractor of (6) is broken into two pieces around the fixed point (u^*, u^*). Each is a segment of $u_k=1-av_k^2$

type of behavior can be proven, however, can be extended below 1.55 in the following manner. A point $Z \in \mathbb{R}^n$ is a periodic point of period p if $F^p(Z)=Z$ but $F^k(Z) \neq Z$ for $1 \leqq k<p$, where F^k represents the composition of $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$ with itself k times. Since such a Z may be viewed as a fixed point of F^p, we may have the existence of stable and unstable manifolds that intersect transversally at some point other than Z. This also implies the existence of the type of chaotic behavior described in Theorem 1. This is precisely what occurs for Eqs. (6) and (1) for smaller values of a (and some $b>0$). We can show this by proving that the fixed points of g^p, where g is defined in (7), are snap-back repellers. In particular this can be done for the fixed points of $g^2=g \circ g$ which were computed in the previous section: $[1 \pm(4a-3)^{1/2}]/2a$. Following the analysis in Sect. 3, it can be shown that each of these is a snap-back repeller of $g \circ g$ for all $a>1.44$. Thus by Theorem 3 we can conclude the existence of a transversal homoclinic orbit (for a periodic point) of (6) or (1) for some $b>0$. For each such a value this proves the existence of chaotic behavior of (1) for all $|b|<\varepsilon$ for some $\varepsilon>0$. Continuing in this manner, the same can be shown for even smaller values of a.

5. Conclusions

In the previous analysis we have analytically proven the existence of a transversal homoclinic orbit of (1) for small values of b and appropriate values of a, by treating this equation as a small perturbation of the simplified problem with $b=0$. Although this verifies the existence of a chaotic Cantor set Λ described by Theorem 1, this does not prove that Λ attracts points under (1). It may be possible, however, to conclude such behavior by again reducing the analysis to the problem with $b=0$. For example, for $0 \leqq a \leqq 2$ and $b=0$, the set of initial points $(u_0, v_0) \in I \times I$, where $I=[-1, 1]$, leads to solutions contained within the one-dimensional invarient curve: $u_k=1-av_k^2$ for $v_k \in I$. In fact (x_1, y_1) lies on this curve. Hence the region

$I \times I$ is a "trapping region". We should thus expect the phase-plane portrait of solutions to be "close" to this curve for all $|b| < \varepsilon$. Hénon's numerical studies seem to corroborate this conjecture.

It is also possible to show that for those values of a for which (7) has a fixed point which is a snap-back repeller, the chaotic set of attraction contains the fixed point. For the problem (6) with $b=0$ the same must be true of the chaotic attracting set within the invarient curve: $u_k = 1 - av_k^2$. We may therefore expect this to be the case when b is close to 0 as well. On the other hand, for those a values for which chaos occurs but for which the fixed point is not a snap-back repeller, the chaotic attracting set of (7) appears to be broken into pieces around the fixed point. In particular for $1.44 < a < 1.55$ the two periodic points of period 2 are each snap-back repellers, but the fixed point is not. In this case the chaotic set of attraction for (7) is broken into at least two pieces each of which contains one of the periodic points, and neither of which contains the fixed point. Here points are mapped in a chaotic manner alternating from one piece to the other. This must again be the case for the two-dimensional problem (6) with $b=0$ inside the chaotic attracting curve: $u_k = 1 - av_k^2$ (Fig. 2). It is therefore likely that the strange attractor of (1) for these values of the parameter a and for small values of the parameter b is also broken into at least two pieces. Numerical studies of (1) seem to indicate that this does in fact occur. This behavior has also been observed in several other numerical investigations of discrete mappings modelling a variety of natural phenomena, but few analytical results have been obtained,

References

1. Curry, J.H.: On the Hénon transformation. Preprint National Center for Atmospheric Research, Boulder, Colorado (1978)
2. Hénon, M.: A two-dimensional mapping with a strange attractor. Comm. Math. Phys. **50**, 69–77 (1976)
3. Li, T.-Y., Yorke, J.A.: Period three implies chaos. Am. Math. Monthly **82**, 985–992 (1975)
4. Lorenz, E.N.: Deterministic nonperiodic flow. J. Atmos. Sci. **20**, 130–141 (1963)
5. Marotto, F.R.: Chaotic behavior in discrete dynamical systems with application to ecology. Doctoral Thesis, Boston University, Boston, Mass. (1977)
6. Marotto, F.R.: Perturbations of stable and chaotic difference equations. J. Math. Anal. Appl. to appear (1979)
7. Marotto, F.R.: Snap-back repellers imply chaos in $\mathbb{R}^n$. J. Math. Anal. Appl. **63**, 199–223 (1978)
8. Ruelle, D., Takens, F.: On the nature of turbulence. Comm. Math. Phys. **20**, 167–192 (1971)
9. Smale, S.: Differentiable dynamical systems. Bull. Am. Math. Soc. **73**, 747–817 (1967)

Communicated by J. Moser

Received January 16, 1979