A View on Chaotic Dynamical Systems

J. Palis
Instituto de Matemática Pura e Aplicada
Estrada Dona Castorina *110,* Rio de Janeiro, *Brazil*

Received November 20, 1994

Glancing at dynamics from the 60's to the 90's, we present a view on how the theory of chaotic dynamical systems *came* to be.

> For, *in* respect to the latter branch of the supposition, it should be considered that the most *trifling variation* in the facts of the two cases, might *give rise* to the most important miscalculations, *by* diverting thoroughly the *two courses* of events; very much as in arithmetic, an *error* which, in its own individuality, may be inappreciable, produces, at length, by dint of multiplication at *all* points, a result enormously at *variance* with truth.

This extraordinary sentence by Edgar Allan Poe in The Mistery of Marie *Rogêt*, embodies in a beautiful way the idea of sensitivity, and thus chaotic behavionr of a system, with respect to initial conditions. In the early 1960's, about a century and half after that, Lorenz in his remarkable work on the dynamical modelling of weather prediction, stressed the same idea, implying that some degree of uncertainty of such a prediction is unavoidable, no matter how accurate the initial data are. Incidentally, a curious discussion about priority: who first thought of sensitivity, i.e. chaotic systems? Americans, Russians (a great school in the area), the British (Maxwell) or the French (Poincaré)? Perhaps Poe's book is the answer... and, in present times, Lorenz's work...

The work of Lorenz, published in 1963, dealt with vector fields in three-dimensional space, i.e. dynamical systems given by assigning a vector to each point in the space. Each of the vector fields was deterministic (no random fluctuations in the coefficients of their equations) and, in fact, was algebraically very simple: the equations for the vector field at each point in space were just quadratic (second degree) in terms of the (usual) coordinates of the point. Yet, and that was a big novelty, the behaviour of most of the trajectories was, in the long run in the future, somewhat unpredictable: they got apart a certain constant amount even when starting at very nearby points (sensitivity property with respect to initial conditions), and they were accumulating everywhere in a complicated object, later called a strange attractor. Even more, this behaviour was still present when the coefficients of the equations of the vector field were slightly changed. Amazingly, the work became known to many of us only in the early 70's.

Actually, the topics on which most dynamicists focused in the 60's, although very important, were quite different from that in Lorenz's work. Indeed, two fundamental theories were constructed and took quite a full shape in that decade: the hyperbolic theory for general systems and the KAM (Kolmogorov, Arnold, Moser) theory for conservative systems.

Many of us were proud of the achievements in this area at the end of the 60's, thinking that we had been able to describe "very many" systems, or, at least, that we had set very solid foundations to that end. That is, that we had understood or even classified a great portion of the universe of dynamical systems. And all this, just to be rather shocked at the turn of the decade, when we learned about Lorenz's work. Then, in a crescendo of bewilderment in the 70's, we realized that we were unable to answer a question posed by the astronomer Hénon, concerning the possibility of having a somewhat persistent strange attractor (again, exhibiting the sensitivity property) even for a two-dimensional quadratic diffeomorphism and, finally, we learned by 1978 of the work of the physicists Feigenbaum and Coullet-Tresser on period-doubling bifurcations and, again, a strange attractor, now for a simple *one-dimensional* quadratic map. An interesting fact is that all these works on non-linear (actually quadratic) dynamics were based on computational experiments and no or not complete mathematical proofs were presented, although a program of a proof was in the (independent) papers by Feigenbaum and by Coullet-Tresser. It is also to be noted that the work of Ruelle-Takens challenging the previous ideas of Landau and others on models for the highly nonlinear phenomenon of turbulence of fluids, also played a role to create the "atmosphere" of the decade: they pledged for strange attractors instead of the previous non-sensitivity attractors as models. Again curiously, none of these colleagues, except for Takens, worked in dynamics at the time of their fundamental and indeed revolutionary contributions.

After this decade (of bewilderment), mathematicians started providing answers and perhaps now we can even suggest a scenario for the world of dynamical systems. Let us glance a little more closely at this great evolution of dynamics from the 60's to the 90's. The focus will be on general or dissipative systems.

Usually, a (discrete) dynamical system is given as a transformation (process) T on a space of events M. This space of events, or phase space, can be the real line, plane or the three- or higher- dimensional Euclidean space or more elaborate mathematical objects (manifolds) or even infinite-dimensional spaces or manifolds. One applies the transformation to an initial point (event) and then to the resulting point and so on, getting an ordered set of points, called the orbit or trajectory of that initial point under T. The main objective of dynamics is to describe the *w-limit set* of the system, i.e. the closure of the set of points where the orbits (or most of them) accumulate in the future when the number of iterates grows to infinity. *Most* here means total Lebesgue measure (assuming M is finite-dimensional). A fundamental concept in this sense is that of an attractor, which is a set of points, made of orbits, where all or most orbits starting at nearby points accumulate. A key topic in dynamics is to describe attractors, since most orbits tend, in the future, towards attractors.

If a transformation T is sensitive with respect to initial points near one of its attractors, then we say that T is *chaotic* near this attractor.

On the other hand, the system is *hyperbolic* near an attractor if distances between points increase and decrease exponentially in complementary directions when we apply the system. An important theory of the 60's, called hyperbolic theory, states that an *attractor persists with the same dynamics (continuous correspondente between the sets of orbits) when the system is slightly perturbed, if and only if it is hyperbolic.* In general, *a system is called hyperbolic if its w-limit set* is *hyperbolic*. Let us be more precise. A C^k differentiable transformation T on M, $k \geq 1$, is called C^k*-stable* if any C^k nearby transformation L on M has the same dynamics as T: there is a homeomorphism h on M such that $hT(x) = Lh(x)$ for all $x \in M$. In particular, the continuous transformation h with a continuous inverse sends trajectories of T onto trajectories of L. On the other hand, T restricted to an invariant compact set ***A***, i.e. made of T-trajectories, is *hyperbolic* if the tangent-bundle of M restricted to A splits into two subbundles E^s and E^u, invariant by the derivative dT, and dT on E^s is a contraction and dT on E^u is an expansion.

The time one map of gradient flows with nondegenerate critical values (nondegenerate gradients) and the map on the two-torus induced by the linear isomorphism $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$, or in general $\begin{pmatrix} 1 & b \\ c & d \end{pmatrix}$ with integer entries and determinant one, are important examples of hyperbolic diffeomorphisms. In the first case, the attractors are made of points (local minima) and in the second, the whole torus. When the whole manifold is hyperbolic for a certain diffeomorphism like above, the map is called an Anosov diffeomorphism. A remarkable fact, on the other hand, is that a *Cantor set* may also be a compact hyperbolic invariant set for a diffeomorphism, like in the famous Smale's *horseshoe* (see [1], [2]). *They behave like saddles not as attractors.* In these invariant Cantor sets, the periodic orbits are *dense*. Notice that we can construct hyperbolic diffeomorphisms having saddle-type (non-attracting) invariant sets which are Cantor sets.

After results of Anosov, Smale and myself, the latter two authors conjectured by 1967, that diffeomorphisms on attractors are C^k stable if and only if these attractors are hyperbolic (known as *The Stability Conjecture)*. Main contributions, showing one side of the conjecture, were given by Robbin and Robinson a few years later. About twenty years afterwards, Mañé in a remarkable paper completed the conjecture in the C^1 case. This result, together with a substantial work on the classification of hyperbolic diffeomorphisms (and flows), gives much strength to this theory. It explains, perhaps, why we thought we had *"understood very many sistems"* and felt very proud about it...

All these notions apply as well to flows or vector fields (non-discrete case). Of course, to define hyperbolicity, we have to consider, at non-singular points, three subbundles of the tangent-bundle to the manifold, one of them being one-dimensional and along the flow. Again, the nondegenerate gradient flows or the three-dimensional flow obtained by the suspension of an Anosov diffeomorphism (see [1]) constitute important classes of hyperbolic flows. And a flow is C^k *dynamically stable* if any C^k nearby flow is dynamically equivalent to it: there is a homeomorphism of the ambient manifold sending trajectories of the initial flow onto trajectories of the perturbed one.

Although persistent under small perturbations, the *Lorenz attractor is not hyperbolic:* the dynamics on it may change through small perturbations of the flow. Amazingly, the following challenge on this topic persists: we do not know yet if the original Lorenz's flow, given by quadratic equations in three variables

$$\begin{aligned} \dot{x} &= 10x + 10y \\ \dot{y} &= 28x - y - xz \\ \dot{z} &= -\frac{8}{3}z + xy \end{aligned}$$

has indeed, as his experimental computation showed, the sensitivity property with respect to initial conditions near the attractor. Notice that in Lorenz's case an attractor must exist since we can verify that there is a ball centered at the origin which is positively invariant by the flow. Despite this challenge, Guckenheimer and Williams, Robinson and Rychlik (see [2]), among others, have formally exhibited plenty of flows, even algebraic ones with cubic (but not quadratic) equations with the above properties: they are called Lorenz-like flows. *Notice that the attractor contains a singularity and infinitely many periodic orbits accumulating on it and this is why the flow is not hyperbolic.* The eigenvalues (Lyapunov exponents) at this singularity are, say, $\lambda_1 < 0$, $\lambda_2 > 0$, $\lambda_3 < 0$ such that $\lambda_1 + \lambda_2 < 0$ and $\lambda_2 + \lambda_3 > 0$. Thus, along the attractor we have some positive expansion and normal to it, a strong contraction. The Lorenz-like attractors are fully persistent: any small perturbation of the initial flow yields a flow with an attractor with the same properties as above, including *strangeness*, i.e. sensitivity with respect to inicial conditions.

Of course, the search for nonhyperbolic strange attractors is now intense. Very recently, Rovella, in his Ph.D. thesis at IMPA (Bull. Braz. Math. Soc., 1993), exhibited a much more subtle strange attractor for three-dimensional flow like in Lorenz's case. Again, the attractor contains a singularity with two negative and one positive Lyapunov exponents, but the sum of two of them is negative in all cases: there is some contraction along the attractor and a strong one normal to it. *Sensitivity is as before, but persistence of the attractor only occurs for a positive Lebesgue measure set in a two-dimensional parameter space.* There are, however, small perturbations that destroy the strange attractor, producing in turn a periodic one.

Actually, the result of Rovella was much based on a great achievement in dynamics in the late 80's: a remarkable, although still partial, answer to Hénon's question about a possible strange attractor for the quadratic diffeomorphism in the plane

$$\varphi_{a,b}(x, y) = (1 - ax^2 + y, bx)$$

for values of a and b near 1.4 and 0.3, respectively. The answer is yes and this important development went as follows. Before going into that, let us clarify that the word "partial" above is due to the fact that Hénon has asked for a "global" strange attractor and the answer we know *so far* corresponds to a strange attractor whose basin of attraction is not so big.

The main result is due to Benedicks-Carleson (Annals of Math., 1991), who proved that, for fixed very small b, there is a positive Lebesgue measure set of values of a near 2, for which $\varphi_{a,b}$ has a (nonhyperbolic) strange attractor. Immediately after that, again in a very remarkable paper, Mora and Viana (Acta Math., 1993) substantially extended the result to cover the much more general situation of bifurcating (unfolding) a *homoclinic tangency* through a one-parameter family of *dissipative (determinant of the Jacobian with norm less than one) surface diffeomorphisms.* The notion of a *homoclinic orbit* was introduced by Poincaré (see [1], [2]): *it corresponds to orbits of intersection of the* (invariant) *stable and unstable manifolds associated to a saddle periodic point; the orbits at which they are tangent are called homoclinic tangencies.*

Remarkably, when we *unfold a homoclinic tangency* through a one-parameter family of surface diffeomorphisms which are dissipative, we obtain as the parameter varies (like a in the quadratic family above), not only *Hénon-like strange attractors, but also cascades of period doubling bifurcations of sinks (periodic attractors) à la Feigenbaum and infinitely many coexisting sinks.* All these famous bifurcating phenomena are much discussed in [2].

We can now start to "see" a bit better the large (much larger than we thought in the 60's) world of nonhyperbolic dynamics: maybe homoclinic tangencies, strange attractors and cascades of period doubling bifurcations are to be seen densely in this part of the world of dynamical systems complementary to the hyperbolic part (this is a *conjecture*, see [2], that can also be formulated in higher dimension). Moreover, and this is a fact corresponding to very recent results of Ures and Catsigeras, *arbitrarily near, in the C^k sense, one of the above list of complicate bifurcating phenomena, we can obtain the other one.* The exception here is to know how to approximate a diffeomorphism with infinitely many sinks by one displaying a homoclinic tangency: this is a fascinating open problem, which, I conjecture, should have a positive answer.

I wish to point out that the unfolding of homoclinic tangencies does provide examples on *how one can pass directly from a simple nonchaotic system to the chaotic region, without going through cascades of period doubling bifurcations:* the cascades are "seen" in the parameter line only afterwards, i.e. the homoclinic tangency is a first bifurcating point in the parameter line (see [2]). These examples may be constructed to be robust in the sense that nearby parametrized families exhibit the same phenomenon of passing from simple to chaotic systems.

At this point, let us enrich our discussion with a main ingredient in bifurcations of dynamical systems, at least in the important case of the unfolding of a homoclinic tangency through one-parameter families. We restrict our discussion here to surface diffeomorphisms. Notice that a homoclinic tangency is associated with a saddle fixed or periodic point of a diffeomorphism. Now, this saddle may or not be part of a larger invariant set, say a horseshoe, i.e. a Cantor set. The *Hausdorff dimension* of such a Cantor set is a number strictly between zero and one: the Cantor set is called a *fractal* and its dimension a *fractal dimension.*

Results of Newhouse and myself (Astérisque, 1976), Takens and myself (Annals of Math., 1987) showed that *if this fractal dimension is smaller than one, then a homoclinic tangency, associated to the fractal, that generically unfolds yields many nonchaotic hyperbolic diffeomorphisms.* That is, the initial diffeomorphism is a point of *total Lebesgue density* in the parameter line for nonchaotic hyperbolic diffeomorphisms. Very recently, Yoccoz, a 1994 Fields Medal winner, and myself (Acta Math., 1994) proved a converse to that statement: *if the dimension is bigger than one, we have chaotic systems in abundance, i.e. the initial diffeomorphism is not a density point for nonchaotic hyperbolicity.*

These results convey a fundamental difference in the measure of the bifurcating set in the parameter line, depending on how large or how small is the fractal dimension of the hyperbolic set associated to the homoclinic bifurcation: small (smaller than one) dimension implies "few" dynamic bifurcations and large (larger than one) dimension is likely to imply "much" dynamic bifurcation.

Let us now turn briefly to conservative systems, or more properly to area-preserving surface diffeomorphisms. The classical KAM theorem states that generally there are many (their union is a set of positive Lebesgue measure) invariant curves near elliptic periodic points. Generally here means that the argument of the corresponding linear rotation given by the derivative of the map at the elliptic point satisfies certain generic diophantine inequalities. Often, we refer to such a region bounded by an invariant curve as an elliptic island.

After the initial work of Poincaré, Birkhoff and, somewhat later, Siegel, the KAM theorem was sketched in the analytical case by Kolmogorov in the late 50's and proved in the early 60's by Arnold in that case and by Moser in the smooth (i.e. C^{∞}) or highly differentiable case. From that, the KAM theory grew to higher and even infinite dimension, to constitute an important part of dynamics.

Recently, another beautiful theory by Aubry and Mather is being developed to detect *periodic orbits, invariant curves and Cantor sets, as well as orbits accumulating positively in one of these sets and negatzvely in a consecutive one* (consecutive along a radial line). This theory is much based on the notion of rotation number and rotation set (see [1]), again originally introduced by Poincaré for circle homeomorphisms.

I always thought that some of the techniques concerning homoclinic bifurcations and the like, discussed above for dissipative diffeomorphisms, should be applicable here. For that, there should be a little dictionary where (periodic) sinks should translate into elliptic (periodic) points or elliptic islands, saddles into saddles and the same for homoclinic tangencies. Indeed, very recently, Duarte obtained the following two results in the context of area-preserving diffeomorphisms.

The first one states that if we generically unfold a homoclinic tangency associated to a saddle point p, through one-parameter maps, we obtain *residually (Baire second category, see [1]) in intervals in the parameter line, infinitely many independent elliptic islands; in particular, the elliptic islands accumulate at p.*

His second and notable result concerns the famous *standard family of area preserving maps on the torus*

$$f_k(x, y) = (y + 2x + k \cos 2\pi x, x)/\mathrm{mod}\, 2\pi$$

where the parameter k is a real number. It is probably famous among physicists, or mathematical physicists, because it is expressed by very simple equations and yet we know very little about its robustness, meaning positive Lebesgue measure both in parameter and ambient space, concerning the existence of points in the torus for which the map exhibits positive Lyapunov exponent. In fact, it is *conjectured* that there is a set of positive measure K, in the parameter line, such that for each $k \in K$, the map f_k has positive Lyapunov exponents at points of a positive Lebesgue measure set. Some mathematicians even conjecture that the complement of the set with this property has zero measure. Actually, we do not know even how to prove that there is just one value of k for which the above is true. About the mathematical physicists' opinion on the standard map, Sinai, for instance, tells me that "if the conjecture is true then this means Arnold's diffusion is present in a meaningful way (positive Lebesgue measure)" and, which is amazing to me, "that may help to explain phase-transitions"!

The following result of Duarte shows how delicate the conjecture is, since, roughly speaking, elliptic islands "imply" zero Lyapunov exponents. Duarte proved that: *there exist $k_o > 0$ and a residual subset K^* of $[f_k, \infty)$ such that for each $k \in K^*$, the elliptic. islands accumulate in large, in the sense of Hausdorff metric, subsets of the torus and as k runs to infinity, the elliptic islands tend to fill in the whole torus, i.e. they tend to be dense in the torus.*

Books for Reference

1. J. Palis and W. de Melo, *Geometric theory of dynamical systems*, Springer-Verlag, 1982.
2. J. Palis and F. Takens, *Hyperbolicity and sensitive chaotic dynamics at homoclinic bifurcations*, Cambridge University Press, 1993.