10. Exploring Chaos on an Interval

T. Y. Li

Michigan State University

James A. Yorke

University of Maryland

In the late 1970s, many scientists learned of chaos. They found through computer studies that their favorite systems of equations behaved chaotically for some range of parameters. Until then they thought the simple deterministic physical systems they studied exhibited only two kinds of behavior (if we ignore transient motions), namely, steady state and quasi-periodic (or periodic). This third type was characterized by "sensitivity to initial data," a superb phrase that is apparently due to David Ruelle.

Much of our work can be characterized as trying to discover the pure mathematics the scientist needs; we have often aimed at reformulating ideas so they would be of interest and value to scientists.

When chaos theory seemed to suddenly spring to life full grown, scientists were unaware that the ideas had actually been developed by mathematicians over a period of many decades. Ideas established before 1900 include sensitivity to initial data (Maxwell's investigation of gas laws), Cantor sets, concepts of countable and uncountable infinities (Cantor), and the dynamic complexity of infinitely many coexisting periodic orbits of different periods (Poincaré, who also invented stable and unstable manifolds and investigated period doubling bifurcations).

By 1920 the following ideas were growing: fractal dimensions (beginning with Caratheodory and Hausdorff), symbolic dynamics (Hadamard), and there were topological investigations of connectedness, compactness, and continuity, and numerical methods including efficient differential equation solvers (Runge-Kutta).

The twenties and thirties brought topologically strange invariant attractors of Birkhoff, embedding (Whitney), and ergodicity theorems (Birkhoff and von Neumann—based on the new measure theory of Lebesgue). The generic dynamics theory of Smale and his collaborators had been prepared by Sard in the 1940s. The ideas of the Kolmogorov-Arnold-Moser theory had been prepared by Siegel's studies of "Siegel disks," also in the '40s. In the 1940s and '50s Cartwright and Littlewood explored the complicated topology of dynamics, topology that had been developed over a long period, including Poincaré's ideas that eventually became homology theory, including the Lefschetz Fixed Point theorem that is particularly useful for nonlinear dynamics.

The more one tries to touch upon the basic roots, the more one realizes how hopeless the task is, because nonlinear dynamics is also a description of how the world works, and almost every mathematical tool is at least occasionally needed, from probability theory and numerical analysis, to ordinary and partial differential equations. For each name mentioned above, several are unfairly omitted.

Our goal here is to describe the environment and personal interactions that drew us into this realm. We have often thought of our work as attempts to make connections between the mathematical theories and the needs of the scientists, showing the scientists how useful the mathematics can be at identifying concepts as well as at proving theorems.

One of our efforts began with a colleague, a fluid dynamicist Alan Faller, who was fascinated by several of Edward Lorenz's papers [2-5], at a time when few if any mathematical dynamicists knew of his work. In the early '70s, the faculty of the Institute of Fluid Dynamics and Applied Mathematics (IFDAM) at the University of Maryland, currently known as the Institute for Physical Science and Technology (IPST), researched diverse areas such as plasma physics, applied mathematics, meteorology and others. One of its faculty members, Prof. Alan Faller, an experimental fluid dynamicist and oceanographer, brought to our attention the four papers of Lorenz on fluid dynamics and weather prediction models. Faller had made many copies (this was before Xerox machines) and gave them to anyone he could interest. He confessed that those works were somewhat too theoretical for generic practical meteorologists even though they were all published in meteorology journals. They were beautifully written and Lorenz clearly described the structure of what is now called the "Lorenz Attractor."

What attracted us the most was the highly irregular asymptotic behavior of the trajectories of the set of equations:

$$x' = \sigma(y - x)$$
$$y' = \rho x - y - xz \qquad \text{(L)}$$
$$z' = xy - bz$$

for certain parameter values σ, ρ, and b. According to the Poincaré-Bendixon theory, the limit sets of bounded trajectories of two-dimensional smooth differential equations in the phase space are quite regular. While no one had been able to extend the theory in a useful way to higher dimensional equations, most mathematicians and scientists tended to believe that the asymptotic trajectories in higher dimensions were, more or less, quite regular.

Lorenz discovered a useful way of thinking of the solution of his system of equations. As he solved his system numerically he stored the local maxima of $z(t)$ as it oscillated. When he plotted a huge number of pairs of consecutive maxima

$$(z_n, z_{n+1})$$

they seemed to lie along a curve that was smooth except at one point; that is, they seemed to satisfy a relationship

$$z_{n+1} = F(z_n)$$

The dynamics of this process seemed to us much easier to explain than the original system.

The function was quite similar in spirit to the tent-map

$$F_a(x) := \begin{cases} ax & for\, 0 \le x \le \frac{1}{2} \\ a(1-x) & for\, \frac{1}{2} \le x \le 1 \end{cases}$$

where $1 < a < 2$.

One-dimensional Topology. We knew from Smale's horseshoe, the stretching property of a map can have infinitely many periodic points of different periods. Let f be a continuous map $f : R^1 \to R^1$ and write $f^2 = f(f)$ and $f^3 = f(f^2)$. If f has the property that for some point x_0

$$f^3(x_0) \le x_0 < f(x_0) < f^2(x_0) \quad \left(or, f^2(x_0) < f(x_0) < x_0 \le f^3(x_0)\right),$$

then f has periodic points of all periods when it is iterated. And, it happens that a continuous one-dimensional map having a periodic point of period three constitutes a special case of these properties. When $f^k(a) = a$, while $f^j(a) \ne a$ for $1 \le j < k$, then a is called a periodic point of period k. To explain Lorenz's results we discovered a theorem in early 1973 on sensitivity to initial data. For $f : R^1 \to R^1$, write $f^n = f\left(f^{n-1}\right)$.

Theorem 1 ("Period Three Implies Chaos"): If a continuous map $f : J \to J$ has a point of period three, then

(a) for each positive integer n, there exists a point of period n;

(b) there exists an uncountable set $S \subset J$ that is "scrambled," that is for any two points $x \ne y$ in S, we have

$$\lim_{n\to\infty} \inf\left|f^n(x) - f^n(y)\right| = 0$$

$$\lim_{n\to\infty} \sup\left|f^n(x) - f^n(y)\right| > 0$$

We completed the proof in April 1973, and submitted the first version of the paper [5] to MAA (Mathematic Association of America) *Monthly*. The first four *References* of that draft were [2-5]. Shortly after submission, we learned that the editorial opinion on the paper was not too favorable. The major criticism was the strong *research* tendency of the article, which might not be suitable for the readership of the magazine. However, if we insisted on resubmitting the paper back to the *Monthly*, we must revise the paper massively to make the paper more readable. This was a result we felt a lot of people should know about, so we persisted.

Our target for our simple ideas was the large readership of the magazine, so we never considered submitting it elsewhere. Nevertheless, occupied by other research projects, we put off the revisions for almost a year. Yorke had once submitted a paper to *Science* and got referee reports including assertions like, "This paper contains an identity. Identities are always true, so nothing can be learned from them." So going to an even larger readership journal seemed out of the question.

Each academic year, the Math Department of the University of Maryland routinely organized a special year program. The topic of the program for the academic year of 1973-4 was mathematical biology. Robert May, who was trained as a physicist but had become a professor of biology at Princeton University, was one of the distinguished invited speakers of the program.

During his visit in the first week of May, 1974, Professor May delivered five lectures, one per day. The subject of his fifth talk was the Logistic Model, commonly known as the iterations of the quadratic map.

$$T_a - (x) = ax(1-x), \qquad x \in [0,1], a \in [0,4].$$

He described his discovery, the now well-known *doubling period bifurcations* of the map as a varies. He did not know what happened in the *chaotic region*, the region beyond the main cascade of doubling period bifurcations, and we had not known of period doubling! Since he addressed the topic only at the end of his visit, we only had a chance to discuss the topic with him in the airport that afternoon. We revised our paper the very next week. It was accepted by the *Monthly* in August, 1974.

In the summer of 1974, Professor R. May was invited to give talks by many institutions in different countries in Europe. He adopted our use of "chaos" as a mathematical term, and *Period three implies chaos* therefore began to attract considerable worldwide attention by his strong advocacy in talks and papers [5,6]. A byproduct was the emerging popularity of the *strange attractor* of Lorenz. As use of the word "chaos" spread, it became a word people loved to hate: they didn't have a better word but didn't like chaos. We were quite surprised when use of the word spread to the Soviet Union, where there already was an established term: "stochasticity."

Worldwide attention led to our awareness of the existence of the Sharkovskii ordering:

$$\begin{gathered} 3 < 5 < 7 \ldots < 2 \bullet 3 < 2 \bullet 5 < \ldots \\ < 2^2 \bullet 3 < 2^2 \bullet 5 < 2^2 \bullet 7 \ldots < 2^3 \bullet 3 < 2^3 \bullet 5 < 2^3 \bullet 7 < \ldots \\ < 2^4 \bullet 3 < 2^4 \bullet 5 < \ldots < 2^3 < 2^2 < 2 < 1, \end{gathered}$$

here, if "$p < q$" then the existence of a period p point implies the existence of a period q point. Actually, when we were preparing the original version of the paper, a natural question was automatically raised: what about period five? The counter-example we constructed that period five does not imply period three essentially terminated our desire to pursue it any further.

In the fall of 1975, Yorke attended a large conference on nonlinear oscillations in East Berlin, meeting Sharkovskii on a tour boat on the river Elbe. Sharkovskii did not speak English, but A. Lasota and C. Mira translated. He said he had a better theorem. Despite the translator's efforts, he wouldn't say what it was, but promised to mail the paper to us. Mailing scientific papers out of the Soviet Union was discouraged and required permissions and explanations. We were quite amazed by the Sharkovskii Theorem in [7] when we received it from him in 1976, and disappointed that his result was so much sharper than part (a) of our theorem.

Of course it turned out that May was not the first discoverer of cascades of period doubling (see [9]), but his papers were so well written that the idea became so well known that he was the *last* discoverer of period doubling cascades!!

Chaos Implies Statistical Regularity. Yorke's first chaos paper was actually extremely different in approach. The distinguished Polish mathematician Andrezej (Andy) Lasota was a frequent visitor to IFDAM. On one visit he announced that he had an approach to understanding the theory of certain one-dimensional maps like the tent map F_a. He had a beautiful technique but could only apply it in very special cases. Yorke, who knew nothing of ergodic theory, was able to push the method through in considerable generality, yielding a result that we might call "chaos implies statistical regularity," [10]. The assumption used in [10] is

C: For some interval J, assume $F: J \to J$ is piecewise continuous (continuous except at a finite number of points), and piecewise C^1 and C^2, and for some constant $B > 1$

$$\left|\frac{dF}{dx}(x)\right| \geq B \quad \text{where } \frac{dF}{dx} \text{ is defined.}$$

This is a chaos assumption because it guarantees that every trajectory is unstable.

It is remarkable that we felt papers [1] and [10] were so different that we did not have them refer to each other. One was based on one-dimensional topology [1] and the other [10] was based on linear operators in a Banach space, but both applied to at least some of the tent maps F_a. This paper established that hypothesis **C** implied that there is an absolutely continuous invariant density g for F, a density that would be called today a Sinai-Ruelle-Bowen density or measure, even though it preceded the Bowen-Ruelle paper [11] by 2 years. Sinai had introduced this topic (See for example his survey [15]) by studying cases where the attractor is the entire space.

Our goal was to be able to "choose a point x at random" and describe how its trajectory points were distributed on J. "Choosing x at random" might mean choosing it from a uniform distribution or more generally from a probability distribution with some density g, so that the probability of choosing a point in an interval $(x, x + \Delta x)$ is approximately $g(x)\Delta x$ for small Δx. The goal was to tell people what they would typically see in a trajectory. Hence, it was important to study initial distributions that were absolutely continuous with respect to Lebesgue measure. Since the maps were piecewise expanding, there was a guarantee of sensitivity to initial data. There was no assumption that the map takes an interval onto itself nor of continuity. Hence sensitivity to initial data was the key hypothesis. Our approach was to choose an initial probability density g and see how it evolved in time. So if a point x is distributed with density g, we can ask how $F(x)$ is distributed. The point $F(x)$ will have a density $g_1(x)$ which we write as Pg where P is called the Frobenius-Perron operator. The second iterate $F^2(x)$ would have a distribution P^2g. If you compute the average up to time n of these densities, the resulting densities converge

$$h_n := \frac{1}{n}\sum_{j=1}^{n} P^j g \to h^* \qquad \text{as } n \to \infty$$

and the resulting limit density h^*, is invariant. That is $Ph^* = h^*$. We can summarize [10] by the following statement in which h^* is constructed as shown above.

Theorem 2: Condition C implies F has an invariant density h^*.

Lasota, who was on an extended visit to the United States visiting the University of Maryland, had seen the entire theoretical structure of this problem. He realized that densities of bounded variation were the key. What makes it possible to prove a theorem is the fact that if one starts with g which has bounded variation, then Pg does also, and there is a uniform bound $B > 0$ on the variation of all the densities $\{P^j g\}$. Hence the averages

$$\frac{1}{n}\sum_{j=1}^{n} P^j g$$

will also have the same bound on their variations, and the set of densities G_B with variation bounded by B is compact in the metric $d(g_1, g_2) = \int |g_1(x) - g_2(x)| dx$. Hence some limit point will exist.

Li wanted to make this approach more explicit to the scientist. He asked how you would find h^*. He felt it could be computed. So he began to write a Ph.D. thesis on this topic and discovered how to compute it [12], and in the process he solved a long-standing problem of Ulam. We followed up with [13,14]. It was a very exciting time for us since we were learning a great deal of established mathematics as we worked, from topology to operator theory to numerical methods!

References

[1] T.Y. Li and J.A. Yorke, Period three implies chaos, *Amer. Math. Monthly, 82* (1975), 985-992.

[2] E.N. Lorenz, The problem of deducing the climate from the governing equations, *Tellus, 16* (1964), 1-11.

[3] E.N. Lorenz, Deterministic nonperiodic flow, *J. Atmospheric Sci., 20* (1963), 130-141.

[4] E.N. Lorenz, The mechanics of vacillation, *J. Atmospheric Sci., 20* (1963), 448-464.

[5] E.N. Lorenz, The predictability of hydrodynamic flow, *Trans. N.Y. Acad. Sci., Ser. II, 25* (1963), 409-432.

[6] R.M. May, Simple mathematical models with very complicated dynamics, *Nature, 261* (1976), 459-467.

[7] R.M. May, Biological populations with non overlapping generations, stable points, and chaos, *Science, 186* (1974), 645-647.

[8] A.N. Sarkovskii, Coexistence of cycles of a continuous map of a line into itself, *Ukrainian Math. J., 16* (1964), 61-71.

[9] P.J. Myrberg, Sur l'iteration des polynômes réels quadratiques, *J. Math. Pures Appl., 41* (1962), 339-351.

[10] A. Lasota, and J.A. Yorke, On the existence of invariant measures for piecewise monotonic transformations, *Trans. Amer. Math. Soc., 186* (1973), 481-488.

[11] R. Bowen and D. Ruelle, The Ergodic Theory of Axiom-A Flows, *Inv. Math, 29* (1975), 181-202.

[12] T.Y. Li, Finite approximation for the Frobenius-Perron operator: A solution to Ulam's conjecture, *J. Approx. Theory, 17* (1976), 177-186.

[13] T.Y. Li and J.A. Yorke, Ergodic transformations from an interval into itself, *Trans. Amer. Math. Soc., 235* (1978), 183-192.

[14] T.Y. Li and J.A. Yorke, Ergodic maps on [0,1] and nonlinear pseudo-random number generators, *Nonlinear Anal., 2* (1978), 473-481.

[15] Sinai, Ya. Gibbs measure in ergodic theory, *Russ. Math. Surveys, 166* (1972), 21-69.