What's new on Lorenz strange attractors ?

Marcelo Viana

Besides its philosophical implications on the ideas of determinism and (un)predictability of phenomena in Nature, E. Lorenz' famous article *Deterministic nonperiodic flow* [17], published nearly four decades ago in the Journal of Atmospheric Sciences, raised a number of mathematical questions that are among the leitmotivs for the extraordinary development the field of Dynamical Systems has been going through. This work is about those and related questions, and some remarkable recent results answering them. The first part is a general overview, mostly in chronological order. The four remaining sections contain more detailed expositions of some key topics.

Modeling the weather. Lorenz, a meteorologist at the MIT, was interested in the foundations of long-range weather forecast. With the advent of computers, it had become popular to try to predict the weather by numerical analysis of equations governing the atmosphere's evolution. The results were, nevertheless, rather poor. A statistical approach looked promising, but Lorenz was convinced that statistical methods in use at the time, specially prediction by linear regression, were essentially flawed because the evolution equations are very far from being linear.

To test his ideas, he decided to compare different methods applied to some simplified non-linear model for the weather. The size of the model (number and complexity of the equations) was a critical issue because of the limited computing power available in those days [1]. After experimenting with several examples, Lorenz learned from B. Saltzmann of recent work of his [35] concerning thermal fluid convection, itself a crucial element of the weather. A slight simplification

$$\begin{array}{ll} \dot{x} = -\sigma x + \sigma y & \sigma = 10 \\ \dot{y} = rx - y - xz & r = 28 \\ \dot{z} = xy - bz & b = 8/3 \end{array} \qquad (1)$$

of a system of equations studied by Saltzmann proved to be an ideal test model. I'll outline later how Saltzmann arrived at these equations, and why he picked these particular values of the parameters σ, r, b.

The key episode is recalled by Lorenz in [18]. At some stage during a computation he decided to take a closer look at a particular solution. For this, he restarted the integration using some intermediate value printed out by the computer as a new initial condition. To his surprise, the new calculation diverged gradually from the first one, to yield totally different results in about four "weather days" !

Lorenz even considered the possibility of hardware failure before he understood what was going on. To speed things up, he had instructed the computer program to print only three decimal digits, although the computations were carried out to six digits. So, the new initial condition entered into the program didn't quite match the value generated in the first integration. The small initial difference was augmented at each integration step, causing the two solutions to look completely different after a while. This phenomenon, first discovered in a somewhat more complicated system of equations, was reproduced in (1).

The consequences were far-reaching: assuming the weather does behave like these models, then long-range weather prediction is impossible: the unavoidable errors in determining the present state are amplified as time goes by, rendering the values obtained by numerical integration meaningless within a fairly short period of time.

[1] Lorenz' computer, a Royal McBee LGP-30, had 16Kb internal memory and could do 60 multiplications per second. Numerical integration of a system of a dozen differential equations required about a second per integration step.

Sensitivity and unpredictability. This observation was certainly not new. Almost a century before, J. C. Maxwell [22], one of the founders of the kinetic theory of gases, had warned that the basic postulate of Determinism *the same causes always yield the same effects* should not be confused with a presumption that *similar causes yield similar effects*, indeed there are cases in Physics where small initial variations may lead to a big difference in the final state[2]. A similar point was stressed early in the twentieth century by H. Poincaré [30], including the very setting of weather prediction:

Why have meteorologists such difficulty in predicting the weather with any certainty ? Why is it that showers and even storms seem to come by chance, so that many people think it is quite natural to pray for them, though they would consider it ridiculous to ask for an eclipse by prayer ? [...] a tenth of a degree more or less at any given point, and the cyclone will burst here and not there, and extend its ravages over districts that it would otherwise have spared. If they had been aware of this tenth of a degree, they could have known it beforehand, but the observations were neither sufficiently comprehensive nor sufficiently precise, and that is the reason why it all seems due to the intervention of chance.

On the other hand, gas environments and, particularly, the Earth's atmosphere, are very complicated systems, involving various types of interactions between a huge number of particles. Somehow, it is not surprising that their evolution be hard to predict. What was most striking about Lorenz's observations was the very simplicity of equations (1), combined with their arising in a natural way from a specific phenomenon like convection. That the solutions of such a simple set of equations, originating from a concrete important problem, could be sensitive with respect to the initial conditions, strongly suggested that *sensitivity is the rule in Nature rather than a particular feature of complicated systems.*

[2]Even before, the same idea appeared in E. A. Poe's *The mystery of Marie Roget*, in a context of crime investigations... So much for priorities on this matter.

Strange attractors. A few years later, in another audacious paper [33], D. Ruelle and F. Takens were questioning the mathematical interpretation of turbulent fluid motion. It had been suggested by L. Landau and E. Lifshitz [16], and by E. Hopf before them, that turbulence corresponds to quasi-periodic motions inside tori with very large dimension (large number of incommensurable frequencies) contained in the phase-space. However, Ruelle and Takens showed that such quasi-periodic tori are rare (non-generic) in energy dissipative systems, like viscous liquids. Instead, they sustained, turbulence should be interpreted as the presence of some *strange attractor*.

An *attractor* is a bounded region in phase-space, invariant under time evolution, such that the forward trajectories of most (positive probability) or, even, all nearby points converge to it. Ruelle and Takens did not really try to define what makes an attractor *strange*. Eventually, the notion came to mean that *trajectories converging to the attractor are sensitive with respect to the initial conditions.*

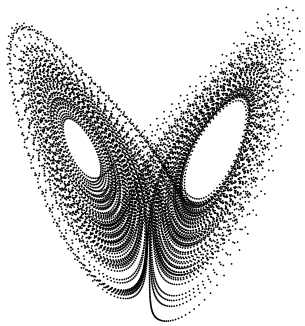

Figure 1: Lorenz strange attractor

Lorenz' system (1) provides a striking example of a strange attractor, and several others have been found in various models for experimental phenomena as well as in theoretical studies. However, not many examples were available at the time [33] was written. Still unaware of the work of Lorenz, which came slowly to

the attention of the mathematical community, Ruelle and Takens could only mention S. Smale's hyperbolic solenoids [38] which, although very important from a conceptual point of view, had no direct physical motivation.

Hyperbolic systems. Throughout the 60's, Smale was much interested in the concept of structurally stable dynamical system, introduced by A. Andronov and L. Pontryagin in [2]. The reader should be warned that the word stability is used in Dynamical Systems in two very different senses. One refers to trajectories of a system: a trajectory is stable (or attracting) if nearby ones get closer and closer to it as time increases. Another applies to systems as a whole: it means that the global dynamical behavior is not much affected if the laws of evolution are slightly modified [3]. *Structural stability* belongs to the second kind: basically, a system is structurally stable if small modifications of it leave the whole orbit structure unchanged, up to a continuous global change of coordinates.

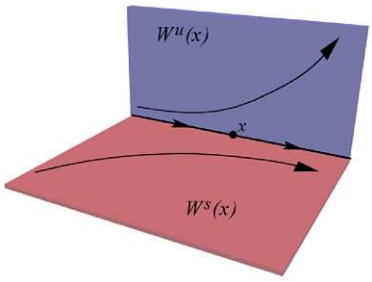

Figure 2: Hyperbolicity near a regular trajectory

In an insightful attempt to identify what the known stable systems had in common, Smale introduced the geometric notion of *hyperbolic dynamical system*. I will give a precise definition of hyperbolicity later, for now let me just refer to Figure 2, that describes its basic flavor: existence, at relevant points x in phase-space, of a pair of sub-manifolds that intersect transversely along the trajectory of x, such that points in one of them (the horizontal "plane") are forward asymptotic to x, whereas points in the other sub-manifold (the vertical "plane") are backward asymptotic to x. Most remarkably, hyperbolicity proved to be the crucial ingredient for stability: the hyperbolic systems are, essentially, the structurally stable ones. Moreover, a beautiful and rather complete theory of these systems was developed in the sixties and the seventies: hyperbolic systems and their attractors are nowadays well-understood, both from the geometric and the ergodic point of view. The reader may find precise statements and references to a number of authors, e.g., in the books [6, 28, 36]. Yet, not every system can be approximated by a hyperbolic one...

The flow described by equations (1) is not hyperbolic, nor structurally stable, so it doesn't fit into this theory. On the other hand, its dynamical behavior seems very *robust*. For instance: Figure 1, which represents a solution of (1) integrated over a long period of time [4], would have looked pretty much the same if one had taken slightly different values for the parameters σ, r, b. How can this be, if these systems are instable (and sensitive with respect to initial data!) ?

It was, probably, a fortunate thing that Smale and his students and colleagues did not know about this phenomenon at the time they were setting the foundations of the theory of hyperbolic systems, it might have convinced them that they were off in a wrong direction. In fact, a satisfactory theory of robust strange attractors of flows would come to existence only recently, building on several important advances obtained in the meantime. But I'm moving ahead of myself!

Lorenz-like flows. For now, let us go back to the mid-seventies, when the work of Lorenz was finally becoming widely known to dynamicists and, in fact, attracting a lot of attention. So much so, that by the end of that decade C. Sparrow could write a whole book [39] about the dynamics of equations (1) over different parameter ranges. Understanding and proving the observations of Lorenz in a rigorous fashion turned out to be no easy task, though.

[3] For instance, by altering the values of parameters appearing in the evolution equations.

[4] An initial stretch of the solution is discarded, so that the part that is plotted is already close to the strange attractor.

A very fruitful approach was undertaken, independently, by V. Afraimovich, V. Bykov, L. Shil'nikov [1], and by J. Guckenheimer, R. Williams [8, 44]. Based on the behavior observed in (1), they exhibited a list of geometric properties such that *any flow satisfying these properties must contain a strange attractor, with orbits converging to it being sensitive with respect to initial conditions.* Most important for the general theory, they proved that such flows do exist in any manifold with dimension 3. These examples came to be known as *geometric Lorenz models*.

The strange attractor has a complicated geometric structure like the "butterfly" in Figure 1 [5]. Sensitivity corresponds to the fact that trajectories starting at two nearby states, typically end up going around different "wings" of the butterfly. There are orbits inside the strange attractor that are dense in it. This means that the attractor is *dynamically indecomposable* (or transitive): it can not be split into smaller pieces closed and invariant under the flow. Another very important feature of these models is that the attractor contains an equilibrium point.

Now, one might expect that small modifications of the flow could cause such a complicated behavior to collapse. For instance, the attractor might break down into pieces displaying various kinds of behavior, or the different types of trajectories (regular ones and equilibria) might be set apart, if one changes the system only slightly. Surprisingly, this is not so: any flow close enough to one of these also has an attractor containing an equilibrium point and exhibiting all the properties I described before, including sensitivity and dynamical indecomposability.

A theory of robust strange attractors. As I'll explain later, the presence of equilibria accumulated by regular orbits of the flow implies that these systems can not be hyperbolic. On the other hand, they can not be disregarded as a pathology since, as we have just seen, this kind of behavior is robust. Indeed, these and other situations, often motivated by problems in the Natural Sciences, emphasized the need to enlarge the scope of hyperbolicity into a global theory of Dynamical Systems.

Profiting from the success in the study of specific classes of systems like the geometric Lorenz flows or the Hénon maps (M. Benedicks and L. Carleson [3], after pioneer work of M. Jakobson [14]), as well as from fundamental advances like the theory of bifurcations and Pesin's non-uniform hyperbolicity [29], a new point of view has been emerging on how such a global theory could be developed. In this direction, a comprehensive program was proposed a few years ago by J. Palis, built on the following core conjecture: *every smooth dynamical system (diffeomorphism or flow) on a compact manifold can be approximated by another that has only finitely many attractors, either periodic or strange.* I refer the reader to [27] for a detailed exposition.

In the context of flows, decisive progress has been obtained recently by C. Morales, M. J. Pacifico, E. Pujals, [25, 26], whose results provide a unified framework for robust strange attractors in dimension 3. While robust attractors without equilibria must be hyperbolic [11], they prove that *any robust attractor that contains some equilibrium point is Lorenz-like: it shares all the fundamental properties of the geometric Lorenz models.* A key ingredient is a weaker form of hyperbolicity, that Morales, Pacifico, Pujals call *singular hyperbolicity*. They prove that any robust attractor containing an equilibrium point is singular hyperbolic [26] [6]. This is a key step leading to a rather complete geometric and ergodic theory, applying to arbitrary robust attractors of 3-dimensional flows. More on this will come later.

Back to the original equations. While they were catalysing such fundamental developments in Dynamical Systems, equations (1) themselves kept resisting all attempts at proving that they do exhibit a sensitive attractor.

On the one hand, no mathematical tools could be devised to solve such a global problem for specific equations like (1). For instance, M. Rychlik [34] and C. Robinson [32] considered systems exhibiting

[5] The figure was produced by numerical integration of the original equations (1).

[6] See [4, 7] for related results about discrete-time systems, in any dimension. These and other important recent developments are surveyed in my article [43].

certain special configurations (codimension-2 bifurcations) and, using perturbation arguments, proved that nearby flows have strange attractors like the geometric Lorenz models. This enabled them to exhibit the first explicit examples (explicit equations) of systems with strange attractors of Lorenz type: those special configurations occur in some families of polynomial vector fields, with degree three, for appropriate choices of the parameters. However, it has not been possible to find parameter values σ, b, and r, for which (1) satisfy the assumptions of their theorems.

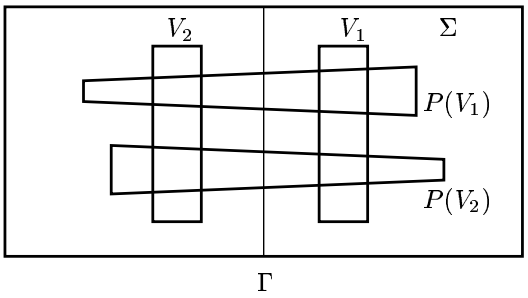

Figure 3: Suspended horseshoe

Another approach was through rigorous numerical calculations. Here a major difficulty arises from the presence of the equilibrium: solutions slow down as they pass near it, which means that a large number of integration steps are required, resulting in an increased accumulation of integration errors. This could be avoided in [9, 10, 23, 24], where all the relevant solutions remain far from the equilibrium point (error control remains delicate, nevertheless). In these works, the authors gave computer assisted proofs that the Lorenz equations have rich dynamical behavior, for certain parameters. More precisely, they used numerical integration with rigorous bounds on the integration errors, to identify regions $V = V_1 \cup V_2$ inside some cross-section of the flow, such that the image of V under the first-return map consists of two pieces that cross V as in Figure 3. By a classical result of Smale, see [38], this configuration ("suspended horseshoe") implies that there is an infinite set of periodic trajectories.

Still, the original question remained: do equations (1) really have a strange attractor ? These equations have no particular mathematical relevance, nor do the parameter values $(10, 28, 8/3)$: their great significance is to have pointed the possibility of a new and surprising kind of dynamical behavior that we now know to occur in many situations. Nevertheless, many of us felt that answering this question, for parameters near the original ones, was a great challenge and a matter of honor for mathematicians [7].

The Lorenz attractor exists! Remarkably, a positive solution was announced about a year ago, by W. Tucker, then a graduate student at the University of Uppsala, Sweden, working under L. Carleson's advice.

Theorem 1 (Tucker [41, 42]) *For the classical parameters, the Lorenz equations (1) support a robust strange attractor.*

Tucker's approach is a combination of two main ingredients. On the one hand, he uses *rigorous numerics* to find a cross-section Σ and a region N in Σ such that orbits starting in N always return to it in the future. After choosing reasonable candidates, Tucker covers N with small rectangles, as in Figure 4, and estimates the forward trajectories of these rectangles numerically, until they return to Σ. His computer program also provides rigorous bounds for the integration errors, good enough so that he can safely conclude that all of these rectangles return inside N. This proves that the equations do have some sort of attractor. A similar strategy is used to prove that the attractor is singular hyperbolic in the sense of [26].

The other main ingredient, *normal form theory* comes in to avoid the accumulation of integration errors when trajectories are close to the equilibrium sitting at the origin. Tucker finds coordinate systems near the equilibrium such that the expression of the

[7] See Smale's list of outstanding open problems for the next century in issue 20 of Mathematical Intelligencer, and his contribution to the book *Mathematics: Frontiers and Perspectives*, to be published under the aegis of the International Mathematical Union as part of the celebrations of the World Mathematical Year 2000.

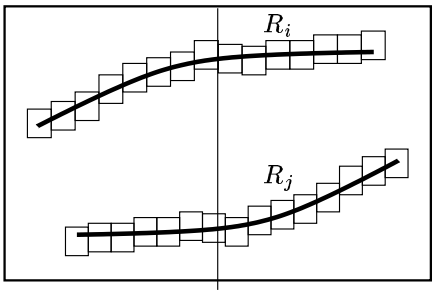

Figure 4: Covering a possibly invariant region with small rectangles

vector field in these coordinates is approximately linear. Thus, solutions of the linear flow (that are easily written down in analytical form) can be taken as approximations of the true trajectories, with efficient error estimates. Accordingly, Tucker instructs the computer program to switch the integration strategy when solutions hit some small neighborhood of the equilibrium: instead of step by step integration, it uses approximation by the linear flow to estimate the point where the solution will exit that neighborhood.

Verifying that a computer-assisted proof is correct involves both checking the algorithms for logical coherence, and making sure that the computer is indeed doing what it is supposed to do. The second aspect is, of course, less familiar to most mathematicians than the first one. In fact, as computer-assisted proofs are a rather new tool, there hasn't been much time to establish verification standards for computer programs. A basic procedure is to have the codes recompiled and rerun on different machine architectures. Preferably, beforehand the algorithm should be reprogrammed by different people. As far as I know, such a detailed independent verification of Tucker's computer programs has not yet been carried out. The first version did contain a couple of "bugs", which Tucker has fixed in the meantime [42], and which I mention briefly near the end. He has also made the text of his thesis and the computer codes, as well as the initial data used by his programs, available on his web page [41].

An outline of Tucker's arguments is given in the last section of this work. Right now, let us go back to where it all started, for a closer look.

1 From Thermal Convection to the Equations of Lorenz

Most of the motion in the Earth's atmosphere takes the form of convection, caused by warming of the planet by the Sun: heat absorbed by the surface of the Earth is transmitted to the lower layers of the atmosphere; warmer air being lighter, it rises, leaving room for downwards currents of cold air.

A mathematical model for thermal convection was proposed early in the twentieth century by the British physicist R. J. Stutt, better known as Lord Rayleigh. This model [31] describes thermal convection inside a fluid layer contained between two infinite horizontal plates that are kept at constant temperatures T_{top} and T_{bot}. It is assumed that the bottom plate is hotter than the top one, in other words, $T_{bot} > T_{top}$. If the temperature difference $\Delta T = T_{bot} - T_{top}$ is small, there is no fluid motion: heat is transmitted upwards by conduction only. In this case, the fluid temperature T_{steady} varies linearly with the vertical coordinate η. As ΔT increases, this *steady-state solution* eventually becomes unstable, and the system evolves into convective motion. Convection cells are formed, where hot fluid is cooled down as it rises, and then comes down to get heated again. See Figure 5.

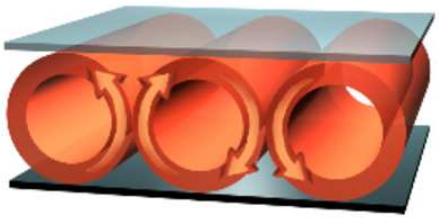

Figure 5: Convection motion

B. Saltzmann [35] analyzed a simplified version of Rayleigh's model. Firstly, he assumed that the system is invariant under translations along some fa-

vored direction, like the direction of convection rolls in Figure 5, so that the corresponding dimension in space may be disregarded. This brings the problem down to two spatial dimensions, the evolution equations reduce to

$$\frac{\partial \nabla^2 \Psi}{\partial t} = -\frac{\partial(\Psi, \nabla^2 \Psi)}{\partial(\xi, \eta)} + \nu \nabla^4 \Psi + g\alpha \frac{\partial \Theta}{\partial \xi} \qquad (2)$$

$$\frac{\partial \Theta}{\partial t} = -\frac{\partial(\Psi, \Theta)}{\partial(\xi, \eta)} + \frac{\Delta T}{H}\frac{\partial \Psi}{\partial \xi} + \kappa \nabla^2 \Theta \qquad (3)$$

where ξ and η are the spatial coordinates, t is time, and the dependent variables Ψ and Θ are interpreted as follows:

- $\Psi(\xi, \eta, t)$ is a stream function: the motion takes place along the level curves of Ψ, with velocity field
$$\left(-\frac{\partial \Psi}{\partial \eta}, \frac{\partial \Psi}{\partial \xi}\right).$$
- $\Theta(\xi, \eta, t) = T(\xi, \eta, t) - T_{steady}(\xi, \eta, t)$ is the temperature departure from the steady-state solution mentioned above.

The other letters represent physical parameters: H is the height of the fluid layer, g is the constant of gravity, α is the coefficient of thermal expansion, ν is the viscosity, and κ is the thermal conductivity.

If the *Rayleigh number*

$$R_a = g\alpha H^3 \frac{\Delta T}{\nu \kappa}$$

is small, the system remains in the steady-state equilibrium $\Psi \equiv 0$, $\Theta \equiv 0$. However, as observed by Rayleigh, as R_a crosses a threshold

$$R_c = \frac{\pi^4}{a^2}(1 + a^2)^3, \quad a > 0,$$

new solutions of (2)–(3) are created, of the form

$$\Psi(\xi, \eta, t) = X_0 \sin\left(\frac{\pi a}{H}\xi\right) \sin\left(\frac{\pi}{H}\eta\right) \qquad (4)$$

$$\Theta(\xi, \eta, t) = Y_0 \cos\left(\frac{\pi a}{H}\xi\right) \sin\left(\frac{\pi}{H}\eta\right), \qquad (5)$$

where X_0 and Y_0 are constants. They describe the motion in cylindrical convection cells in Figure 5, the parameter a being related to the eccentricity of the cylinders. They are *stationary solutions*, as time t does not appear explicitly on the right hand side of (4) and (5).

Nonperiodic behavior. Aiming to understand what happens when ΔT is further increased, Saltzman looked for more general solutions space-periodical in both dimensions. For this, he expanded Ψ and Θ as formal Fourier series in the variables ξ and η, with time-dependent coefficients. Replacing this formal expansion into (2), (3), one finds an infinite system of ordinary differential equations, with the Fourier coefficients as unknowns. Saltzmann truncated the system, keeping only a finite number of these equations.

He tested several possibilities, but one particular case, involving seven equations, was specially interesting. Numerical integration showed that, for convenient choices of the parameters, all but three of the dependent variables are transient: they go to zero as time increases to infinity. In other words, although the phase-space has dimension seven, many solutions seem to converge to some "attractor" contained in a three-dimensional subset of the phase-space. On the other hand, these three special non-transient modes seemed to have a rather complicated (nonperiodic) evolution with time.

Lorenz took the system of equations obtained in this manner, by truncating the original infinite system right from the start to only these three variables. This corresponded to looking for solutions of (2), (3), of the form

$$\Psi(\xi, \eta, t) = X(t) \sin\left(\frac{\pi a}{H}\xi\right) \sin\left(\frac{\pi}{H}\eta\right)$$

$$\Theta(\xi, \eta, t) = Y(t) \cos\left(\frac{\pi a}{H}\xi\right) \sin\left(\frac{\pi}{H}\eta\right) + Z(t) \sin\left(\frac{2\pi}{H}\eta\right).$$

Replacing these expressions in (2), (3), one obtains three ordinary differential equations on the coefficients $X(t)$, $Y(t)$, $Z(t)$ that are equivalent to equations (1). Actually, X, Y, Z are not exactly the same

as x, y, z in (1), but the two sets of variables are related to each other, simply, by rescaling:

$$X = A_0 x \qquad Y = B_0 y \qquad Z = C_0 z,$$

where A_0, B_0, C_0, are constants depending only on a and the physical quantities H, g, α, ν, κ, and ΔT. To obtain (1) one also re-scales time, by another constant D_0 depending on a and H. The parameters σ, r, b in (1) are given by

$$r = \frac{R_a}{R_c} \qquad \sigma = \frac{\nu}{\kappa} \qquad b = 4(1 + a^2).$$

Some simple facts about equations (1) are easy to check. For instance $(0, 0, 0)$ is an equilibrium point, for every value of r. This equilibrium is stable (attracting) when $r < 1$, corresponding to the stability of the steady-state solution. As r crosses the value 1, the origin becomes unstable, and two new stable equilibrium points P_1 and P_2 arise. They correspond to the stationary solutions given by (4) and (5). If one further increases r, these two solutions become unstable. This suggests that, for large Rayleigh number, the convection motion described by (4) and (5), is replaced by some different form of dynamics.

Let me also comment on the particular choice of parameter values in (1). Saltzmann took $a = 1/\sqrt{2}$, which is the value of a for which R_c is smallest. This gives $b = 8/3$. The *Prandtl number* $\sigma = 10$ is typical of liquids ($\sigma \approx 4.8$ for the water), for the air $\sigma \approx 1$. Finally, the *relative Rayleigh number* $r = 28$ is just slightly larger than the transition value $r \approx 24.7368$ at which the two equilibria P_1, P_2 become unstable.

2 Geometric Lorenz Models

Here is an outline of main facts about the geometric Lorenz models in [1, 8, 44], that are also relevant for the next two sections.

Equilibrium point. A first condition in the definition of the geometric models is that the flow should have an equilibrium point O. If X denotes the vector field associated to the flow, the derivative $DX(O)$ should have one positive eigenvalue $\lambda_1 > 0$ and two

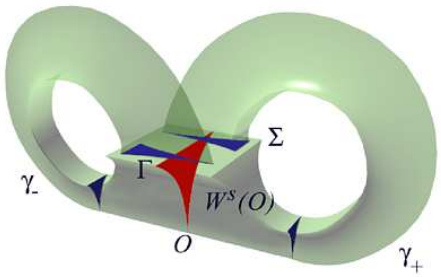

Figure 6: A geometric Lorenz flow

negative eigenvalues $-\lambda_2 < -\lambda_3 < 0$. As a consequence, there are two trajectories γ_+ and γ_- moving away from O in opposite directions as time increases, as shown in Figure 6. The union of γ_- and γ_+ with the equilibrium point O is called the *unstable manifold* of O, and denoted $W^u(O)$. Moreover, there is a two-dimensional surface containing the equilibrium point and formed by solutions that converge to O as time goes to $+\infty$. This is called the *stable manifold* of O, and denoted $W^s(O)$. Check also Figure 7.

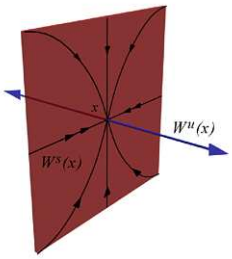

Figure 7: Stable and unstable manifolds at an equilibrium point

In addition, the expanding eigenvalue should *dominate* the weakest contracting one. In other words, one also asks that

$$\lambda_1 > \lambda_3 \,. \tag{6}$$

It is a simple exercise to check that the equilibrium point $O = (0, 0, 0)$ of (1) satisfies these eigenvalue

conditions for all values of the parameters r, σ, b close to the ones considered by Lorenz.

Cross-section. Next, one assumes that there exists some two-dimensional domain Σ that is cut transversely by the flow trajectories, and also intersects the stable manifold $W^s(O)$ along some curve Γ. In addition, both trajectories γ_- and γ_+ intersect Σ. Actually, all the trajectories starting in any of the two connected components of $\Sigma \setminus \Gamma$ should also intersect Σ in some future time. This means that Σ acts as a *trap*: solutions that hit it can not escape from returning to Σ indefinitely (unless they happen to hit Γ, in which case they simply converge to O and never come back to Σ).

The behavior of these trajectories can be understood by looking at the points where they successively meet Σ. In other words, under this assumption the evolution of the original flow can be reduced to that of the *first-return map* $z \mapsto P(z)$, that assigns to each z in $\Sigma \setminus \Gamma$ the point $P(z)$ where the trajectory of z first intersects Σ. Figure 8 contains a schematic representation of the image of $\Sigma \setminus \Gamma$ under P.

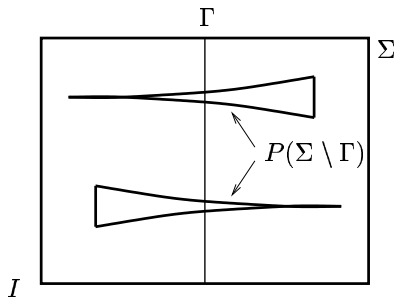

Figure 8: The image of the first-return map P

It is not difficult to find reasonable candidates for such a cross-section in equations (1). For instance, Sparrow [39] takes Σ contained in the horizontal plane $\{z = r - 1\}$. His computations of the first-return map, e.g. Figure 3.4(a) in [39, page 34], do suggest that P is well-defined, and its image has two connected components. The shape of these components is less clear from those pictures: they look more like curve arcs, compare Figure 4. The reason is that the return map is strongly area-dissipative, because the divergence of the flow

$$\operatorname{div}(-\sigma x + \sigma y, rx - y - xz, xy - bz) = -\sigma - 1 - b$$

is negative, for the parameter values we are interested in. Consequently, the image of P has very small area, which means that the cuspidal triangles represented in Figure 8 must be very thin.

Invariant contracting foliation. I move on to a more subtle condition in the definition of geometric Lorenz models, and definitely harder to check in specific situations. One assumes that there exists a *foliation*, that is, a decomposition of the cross-section Σ into roughly parallel curve segments, the *leaves* of the foliation, which is invariant under the first-return map P: if two points z_1 and z_2 are in a same leaf, then so are $P(z_1)$ and $P(z_2)$. Think of the leaves as vertical lines in Figure 8. Moreover, the foliation should be contracting: the distance from $P^n(z_1)$ to $P^n(z_2)$ goes to zero exponentially fast as n goes to $+\infty$, for any pair of points z_1 and z_2 in a same leaf.

The reason this property is very useful, is that it allows us to reduce the dimension of the problem even further [8]. Roughly speaking, this goes as follows: points in a same leaf of the foliation have roughly the same behavior in the future, since their trajectories get closer and closer; so, for understanding the dynamics of P it is enough to look at the trajectory of only one point in each leaf, for instance, the point where the leaf intersects a given horizontal segment.

Expanding map of the interval. Let me make the idea a bit more precise, with the aid of Figure 9. Let I be a horizontal segment in Σ, e.g., the bottom side of Σ. It is convenient to think of I as an interval in the real line, for instance $I = [0, 1]$. Given any point x in I, let γ_1 be the leaf containing it. By the invariance property, $P(\gamma_1)$ is contained in some leaf γ_2. Let $f(x)$ be the point where γ_2 crosses I. This is how the map f is defined. The graph of f is represented on the left hand side of Figure 10. Note

[8] Previously, existence of a cross-section permitted us to go from the 3-dimensional flow to the 2-dimensional map P.

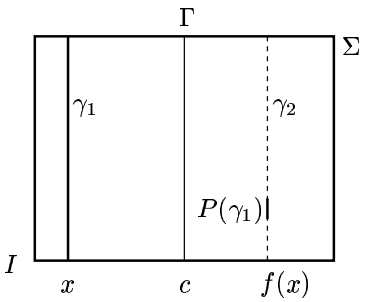

Figure 9: Definition of the interval map f

that the map has a discontinuity at the point c of I, corresponding to the special leaf Γ.

As a final condition, the map f must be expanding: there exists some constant $\tau > 1$ such that

$$\text{dist}(f(x), f(y)) \geq \tau \, \text{dist}(x, y) \qquad (7)$$

for any two points x, y located on the same side of the discontinuity point c.

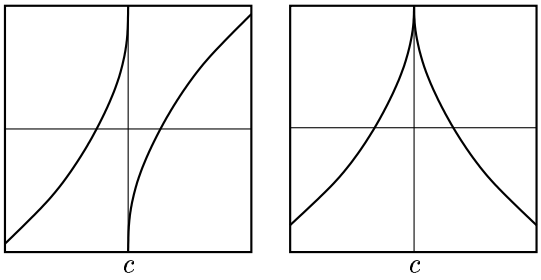

Figure 10: Interval maps related to Lorenz flows

It is interesting to point out that Lorenz computed such a map numerically for (1) in [17]. The cross-section he was considering (implicitly) is not quite the same as our Σ, so that he got a seemingly different picture, shown on the right hand half of Figure 10. Nevertheless, the information provided by either of the two maps is equivalent.

A sensitive attractor. Under these assumptions, [1] and [8, 44] prove that the flow exhibits a strange attractor Λ, that contains the equilibrium point O. The attractor is the closure of the set of trajectories that intersect the cross-section Σ infinitely many times in the past (as well as in the future). Any forward trajectory that cuts Σ accumulates in Λ, and these form a whole neighborhood of the attractor.

Moreover, these trajectories are sensitive with respect to initial conditions. Indeed, suppose you are given two nearby points z and w, whose forward flow trajectories intersect Σ. Typically, the intersection points will be in two different leaves γ_z and γ_w of the invariant foliation (corresponding to nearby points $x \neq y$ in I). The next time the two flow trajectories come back to Σ, the new intersection points will be in leaves γ'_z and γ'_w, corresponding to the points $f(x)$ and $f(y)$ in I. Now, because of the expansiveness property (7), the distance $\text{dist}(\gamma'_z, \gamma'_w)$ is larger than $\text{dist}(\gamma_z, \gamma_w)$. Thus, the distance between the two flow trajectories at successive intersections with Σ keeps increasing, one can check that it eventually exceeds some uniform lower bound that does not depend on how close the initial points z and w are [9].

Another important conclusion is that the attractor contains dense orbits (dynamical indecomposability): there exists some $z_0 \in \Lambda$ whose forward trajectory visits any neighborhood of any point of Λ. (For this one requires the constant τ in (7) to be larger than $\sqrt{2}$.) In particular, the trajectory of z_0 accumulates on the equilibrium point O.

Lorenz models without invariant foliations. Numerical investigations of the Lorenz equations carried out by M. Hénon and Y. Pomeau [12, 13], showed that when the relative Rayleigh number increases beyond $r \approx 30$ the image of the first-return map develops a "hooked" shape, described in Figure 11. This indicates that for such parameter values there is no longer an invariant foliation, as assumed in the geometric Lorenz models. As a simple, easy-to-iterate model of the behavior of the first-return map near the "bends", they introduced the family of maps of

[9] Of course, the growth must stop at some point, these distances can not exceed the order of magnitude of the attractor's diameter.

the plane

$$(x, y) \mapsto (1 - ax^2 + y, bx), \tag{8}$$

that is now named after Hénon.

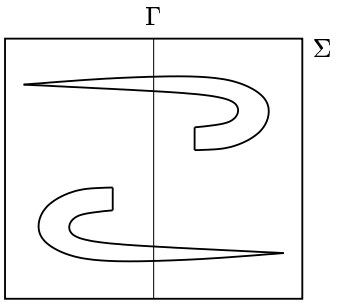

Figure 11: Hooked return maps in Lorenz equations

Based on their computations and, specially, on Chapter 5 of Sparrow's book [39], S. Luzzatto and I proposed an extended geometric model for Lorenz equations, including the creation of the hooks. The main result [21, 20] is that *the attractor survives the destruction of the invariant foliation, at the price of loosing its robustness:* after the hooks are formed, a strange attractor exists for a positive Lebesgue probability set of parameter values. This set is nowhere dense, which is related to the lack of robustness: near the parameter values for which the strange attractor exists there are others corresponding only to attracting periodic orbits. Moreover, the conclusions of [21], which treats a version of the problem for interval maps, have been further extended by Luzzatto and Tucker in [19].

3 Robust Strange Attractors

In order to state the results of Morales, Pacifico, Pujals mentioned before, I need to introduce the precise definition of a robust attractor. It makes sense also for discrete-time systems (diffeomorphisms, or just smooth transformations), but here I restrict myself to flows.

Robust invariant sets. Let $\varphi^t : M \to M$, $t \in \mathbb{R}$, be a flow on a manifold M, that is, a one-parameter group of diffeomorphisms satisfying

1. $\varphi^0(x) = x$ for every $x \in M$ and
2. $\varphi^t \circ \varphi^s(x) = \varphi^{t+s}(x)$ for all $x \in M$ and $s, t \in \mathbb{R}$.

Denote X the associated vector field on M, which is defined by

$$\frac{d\varphi^t}{dt}(x) = X(\varphi^t(x)).$$

We say that a subset Λ of M is *invariant* if trajectories starting in Λ remain there for all times: if $x \in \Lambda$ then $\varphi^t(x) \in \Lambda$ for all $t \in \mathbb{R}$. In what follows I always consider *compact* invariant sets. An invariant set Λ is *dynamically indecomposable* if there exists some point $x \in \Lambda$ whose forward orbit

$$\{\varphi^t(x) : t > 0\}$$

is dense in Λ. This property is often called *transitivity* in the specialized literature.

An invariant set Λ is called *robust* if it admits a neighborhood U such that the following two conditions are satisfied. Firstly, Λ consists of the points whose trajectories under φ^t never leave U:

$$\Lambda = \{x : \varphi^t(x) \in U \text{ for all } t \in \mathbb{R}\}.$$

Secondly, given any vector field Y close to the original one X the set

$$\Lambda_Y := \{x : \varphi_Y^t(x) \in U \text{ for all } t \in \mathbb{R}\}$$

of φ_Y^t-trajectories that never leave U is dynamically indecomposable for the flow φ_Y^t associated to Y. Here closeness means that the two vector fields X and Y, and their first derivatives, are uniformly close over M. Finally, Λ is *a robust attractor* if the neighborhood U can be chosen to be trapping (or forward invariant):

$$\varphi^t(U) \subset U \quad \text{for all } t > 0.$$

Hyperbolicity. Now I define when an invariant set Λ is *hyperbolic*. At each point x of Λ it should be possible to decompose the tangent space into three complementary directions (subspaces)

$$T_x M = E_x^s \oplus E_x^0 \oplus E_x^u$$

depending continuously on x, such that

- E_0^x is the one-dimensional subspace given by the direction of the vector field X at x;
- the linearized flow $D\varphi^t$ preserves the subbundles E^s and E^u (this is clear for E^0), moreover, it contracts E^s and expands E^u exponentially fast.

This last condition means that

$$D\varphi^t\, E^s_x = E^s_{\varphi^t(x)} \quad \text{and} \quad D\varphi^t\, E^u_x = E^u_{\varphi^t(x)}$$

and there are constants $C > 0$ and $\lambda < 1$ such that

$$\|D\varphi^t \mid E^s_x\| \le C\lambda^t \quad \text{and} \quad \|(D\varphi^t \mid E^u_x)^{-1}\| \le C\lambda^t$$

for all $t > 0$ and all $x \in \Lambda$.

I already mentioned the following important consequence, described in Figure 2. If an invariant set Λ is hyperbolic, in the sense of the previous paragraph, then every point x in it is contained in a pair of local sub-manifolds of M, the *stable manifold* $W^s(x)$ and the *unstable manifold* $W^u(x)$ such that

- if y is in $W^s(x)$ then $d(\varphi^t(x), \varphi^t(y)) \le Ce^{-\lambda t}$, for every $t > 0$,
- if z is in $W^u(x)$ then $d(\varphi^{-t}(x), \varphi^{-t}(z)) \le Ce^{-\lambda t}$, for every $t > 0$.

The existence of these manifolds also determines the local behavior of solutions close to the one passing through x.

So far, I have implicitly assumed that x is a regular point of the flow, that is, $X(x)$ is not zero. When, on the contrary, x is an equilibrium point then the definition has to be reformulated: in this case, the directions E^s_x and E^u_x alone must span the tangent space:

$$T_xM = E^s_x \oplus E^u_x\,.$$

Figure 7 describes stable and unstable manifolds through an equilibrium point.

Observe that if x is an equilibrium then

$$\dim E^s_x + \dim E^u_x = \dim M,$$

whereas in the regular case

$$\dim E^s_x + \dim E^u_x = \dim M - 1.$$

This has a simple, yet important, consequence: an invariant indecomposable set containing equilibria is never hyperbolic (except in the trivial case when it consists of a unique point). This is because the dimension of either E^s_x or E^u_x would have to have a jump at the equilibrium, contradicting the requirement of continuity. In particular, the geometric Lorenz attractors that I described above can not be hyperbolic.

Singular hyperbolicity. Morales, Pacifico, Pujals propose a notion of *singular hyperbolic set*, which plays a central role in their results. Let Λ be an invariant set for a flow φ^t. We say that Λ is *singular hyperbolic* if at every point $x \in \Lambda$ there exists a decomposition $T_xM = E_x \oplus F_x$ of the tangent space into two subspaces E_x and F_x such that the linear flow contracts E_x exponentially, and is exponentially *volume expanding* restricted to F_x:

$$\det\left(D\varphi^{-t}(x) \mid F_x\right) \le Ce^{-\lambda t}$$

for all $t > 0$. The decomposition must depend continuously on the point x. Note that the linear flow is allowed to either expand or contract F_x, or both. However, we also require that whatever contraction there is in this direction, it should be *dominated* by the one in the E_x direction:

$$\frac{\|D\varphi^t(x)e\|}{\|D\varphi^t(x)f\|} \le Ce^{-\lambda t}$$

for all norm 1 vectors $e \in E_x$ and $f \in F_x$. Finally, if there are equilibrium points in Λ they should all be hyperbolic (no eigenvalues with zero real part).

The strange attractors in the geometric Lorenz flows I mentioned before are singular hyperbolic. In fact,

Theorem 2 (Morales, Pacifico, Pujals [25, 26]) *Any robust attractor of a three-dimensional flow that contains an equilibrium point (and is not reduced to it) is singular hyperbolic.*

Actually, they prove an even stronger statement. Let Λ be any robust invariant *set*. If Λ contains points of equilibrium then they must all have the same stable and unstable dimensions. More precisely, the derivative DX has

(1) either two negative eigenvalues and one positive eigenvalue, at every equilibrium point;

(2) or two positive eigenvalues and one negative eigenvalue, at every equilibrium point.

These cases can be interchanged: replacing the vector field X by its symmetric $-X$ (in other words, reversing the orientation of trajectories) transforms Case 1 into Case 2, and vice-versa. So, from now on I'll speak only of Case 1. In this case, the eigenvalues satisfy the relation (6), just as the geometric Lorenz systems. Furthermore, the invariant set Λ must be an attractor for the flow. And this attractor is singular hyperbolic!

These results have several important implications. For one thing, they give that *every non-trivial robust attractor of a three-dimensional flow that contains an equilibrium point is sensitive with respect to initial conditions*. Moreover, the flow admits a contracting invariant foliation in a neighborhood of the attractor. Finally, there exists a cross-section Σ, with a finite number of connected components, and the first-return map induces an expanding one-dimensional map in the space of leaves restricted to Σ. The statements in the last sentence are being proved by a graduate student at the Federal University of Rio de Janeiro, as part of showing that these attractors are *stochastically stable*: time averages don't change much when small random noise is added to the system [10]. In fact, having reached this point the main geometric and ergodic properties of the classical Lorenz models extend to this whole class of robust attractors.

Higher dimensions. Not much is known about attractors of flows on manifolds of dimension larger than 3, apart from the hyperbolic case. Robust non-hyperbolic examples can be constructed in a sort of trivial way: the attractor lies inside a 3-dimensional submanifold that is invariant under the flow, and attracts all nearby trajectories. On the other hand, until recently one didn't know whether there are truly high-dimensional attractors of Lorenz type, that is, containing equilibria and yet robust. This was answered, affirmatively, by C. Bonatti, A. Pumariño, and myself in [5]: *for any $k \geq 1$ there exist smooth flows exhibiting robust strange attractors that contain equilibria with k expanding directions;* in particular, the topological dimension of the attractor is at least k. Needless to say, pictures of these multi-dimensional attractors are not easy to make...

[10] The study of systems with small random noise goes back to A. Andronov, L. Pontryagin, A. Kolmogorov and, specially, Ya. Sinai [37]. It was much developed by Yu. Kifer who showed, in particular, that hyperbolic attractors as well as the geometric Lorenz attractors are stochastically stable [15].

4 The Lorenz Attractor

In order to prove Theorem 1, Tucker begins by rewriting the equations in more convenient coordinates (x_1, x_2, x_3), related to the original ones (x, y, z) by a linear transformation, such that the linear part of the vector field at the origin now takes a diagonal form:

$$DX(0) = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & -\lambda_2 & 0 \\ 0 & 0 & -\lambda_3 \end{pmatrix}$$

For equations (1), with $\sigma = 10$, $b = 8/3$, $r = 28$, this roughly gives

$$\begin{aligned} \dot{x}_1 &= 11.8x_1 - 0.29(x_1 + x_2)x_3 \\ \dot{x}_2 &= -22.8x_2 + 0.29(x_1 + x_2)x_3 \\ \dot{x}_3 &= -2.67x_3 + (x_1 + x_2)(2.2x_1 - 1.3x_2). \end{aligned} \tag{9}$$

The next step is to choose a cross-section Σ for the flow: Tucker takes $\Sigma \subset \{x_3 = r - 1 = 27\}$, although this is fairly arbitrary. Now Theorem 1 can be restated in terms of the first-return map P of the flow to this cross-section. There are three essential facts to prove:

(A) There exists a region $N \subset \Sigma$ that is forward invariant under the first-return map, meaning that $P(N \setminus \Gamma)$ is contained in the interior of N.

(B) The return map P admits a forward invariant cone field. In other words, there exists a cone $\mathcal{C}(z)$ inside the tangent space of Σ at each point z of $N \setminus \Gamma$, such that $DP(z)\,\mathcal{C}(z)$ is strictly contained in $\mathcal{C}(P(z))$, for every z in $N \setminus \Gamma$.

(C) Vectors inside this invariant cone field are uniformly expanded by the derivative DP of the return map: there exist constants $c > 0$ and $\tau > 1$ such that

$$\|DP^n(z)v\| \geq c\tau^n\|v\|$$

for every $v \in \mathcal{C}(z)$ and $n \geq 1$.

Indeed, statements (A), (B), (C), together with some extra information on the value of the expansion constant τ, imply that the flow has a strange attractor. Since these three statements are robust, that is, if they hold for a given flow then they hold also for any nearby one (possibly with a slightly smaller τ), so is the attractor.

Let me begin by explaining how property (A) is obtained, I'll talk about the other two later on.

Existence of a forward invariant region. A first non-rigorous computation of the return map is used to guess what the region N could be. As I mentioned before, trajectories returning to Σ cut it along two "arcs". Tucker covers the approximate locations of these "arcs" by small rectangles R_i, of size $\delta_{max} = 0.03$. Then he takes N to be the union of these R_i. Recall Figure 4.

At this initial stage, Tucker has to deal with 700 rectangles, although he may take some advantage from the fact that the system (9) is symmetric with respect to the x_3 axis, to cut the number down to a half, that is, 350 rectangles. In any case, this number is soon going to increase, as we'll see. For simplicity, he always works with rectangles with sides parallel to coordinate axis. The ultimate goal is to prove that the *points in every one of the R_i will return to the cross-section Σ inside N.*

For this, he estimates the future trajectories of points in each of these rectangles in the way I now describe. First, the trajectory of the central point c_i of R_i is integrated (by the Euler method) from time $t = 0$ up to the moment where it intersects an intermediate horizontal plane Σ' placed at some small distance $h = 10^{-3}$ underneath Σ. Let c_i' be the intersection point. See Figure 12.

The place where the trajectories of the other points of R_i intersect Σ' is estimated from c_i' using Taylor

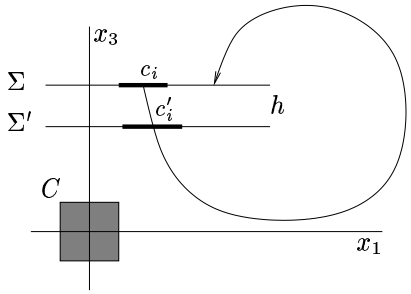

Figure 12: Step by step integration (x_2 direction not represented, for clearness)

expansion: since R_i is small, the distance from intersection points to c_i' can not exceed some small number ϵ_1, that depends on δ_{max} and h. One also has an upper bound ϵ_2 for the error committed in the integration of the trajectories of the central points c_i. This means that the whole image of R_i by the Poincaré map from Σ to Σ' must be contained in a rectangle R_i' of size not larger than $\epsilon_1 + \epsilon_2$ around c_i'. See Figure 13.

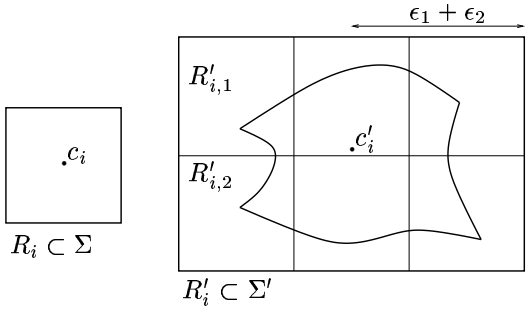

Figure 13: Accommodating all possible errors, then subdividing

The inductive step. Now the idea is to proceed for R_i' in the same way as for R_i. There are, however, several points one has to take into account. To start with, R_i' may be much larger than R_i. So, by repeating this step in a naive way, one is likely to end up with rectangles that are too big for the Tay-

lor expansion estimates to be of any help. To solve this, Tucker subdivides each R'_i into sub-rectangles $R'_{i,j}$ of size at most δ_{max}, and then treats each one of them individually. That is, the program goes on to integrate the trajectories of the central points of each $R'_{i,j}$ up to another intermediate horizontal plane Σ'', located at distance h from Σ', and so on.

Another problem is that, as one moves down, trajectories tend to approach the horizontal distance. See Figure 14. As a consequence, integration errors involved in estimating these Poincaré maps between horizontal planes increase dramatically. The way to avoid this is by switching from horizontal cross-sections to vertical ones, whenever trajectories are far from the vertical direction. More precisely, at each step the program checks whether the vertical component $\dot{x}_3$ is larger than the horizontal one $\dot{x}_1$. In the affirmative case, integration goes on as I was describing.

Otherwise, a Poincaré map is computed, from the horizontal cross-section that is currently under consideration, to a nearby vertical plane $\tilde{\Sigma}$. Bounds for the errors involved in this step are also computed. One ends up with another rectangle $\tilde{R}_i$ inside $\tilde{\Sigma}$, that contains the image of the previous one under this Poincaré map. Now, the algorithm proceeds just as before (apart from the fact that we are now dealing with vertical objects): $\tilde{R}_i$ is subdivided into sub-rectangles with size less that δ_{max}, the trajectories through the centers of these sub-rectangles are integrated up to intersecting another vertical plane $\tilde{\Sigma}'$ located at distance h to the side of $\tilde{\Sigma}$, and so on.

The question concerning the relative strength of the horizontal and the vertical components of the vector field is asked at each step. If the vertical component later becomes the stronger one once more, the program goes back to considering horizontal cross-sections. Switching back and forth from horizontal to vertical cross-sections may happen several times before one goes back to the original plane Σ. The whole process stops when the trajectories hit Σ again.

It is clear that, due to subdivision, the program has to deal with an increasing number of rectangles: the total number reaches some tenths of thousands, from the initial 350. The point is that, *if for every one of the rectangles that are created along the way, the algorithm finishes in finite time, and the return to Σ occurs inside N, then one is certain that N is indeed a forward invariant region for the flow.*

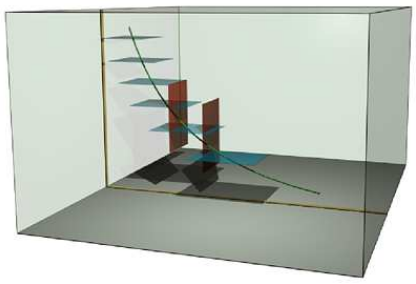

Figure 14: Switching from horizontal to vertical cross-sections

Passing close to the equilibrium. So far, I deliberately avoided talking about trajectories that go close to the origin. It is time to explain how this is dealt with.

For the computer program, "close to the origin" means "inside a cube C of size $1/5$ around the origin". If in the course of the integration, the trajectories do not hit the cube C, then the algorithm is precisely as I described. For trajectories that enter C at some step, the computer program calculates the exit point p_{exit} directly, as follows.

A key point is that the eigenvalues λ_1, $-\lambda_2$, $-\lambda_3$ of the vector field at the origin are sufficiently far from being *resonant*. What I mean by this is that linear combinations

$$n_1\lambda_1 - n_2\lambda_2 - n_3\lambda_3 \tag{10}$$

are not zero, nor too close to zero, for many positive integer values of n_1, n_2, n_3. The precise set $\mathcal{V}$ of values of n_1, n_2, n_3 for which this must be verified depends upon the goals one has in mind, and also upon specific properties of the system. In the present situation Tucker uses

$$\mathcal{V} = \{(n_1, n_2, n_3) \in \mathbb{N}^3 : n_1 + n_2 + n_3 \geq 2 \text{ and } \text{either } n_1 < 10 \text{ or } n_2 + n_3 < 10\}.$$

This set is actually infinite, so it may seem hopeless to check these conditions by means of actual computations. However, it is clear that (10) is far away from zero both when n_1 is much larger than $n_2 + n_3$ (in which case it is positive) and when $n_2 + n_3$ is much larger than n_1 (in which case it is negative). This observation leaves only a finite number of triples (n_1, n_2, n_3) for which non-resonance must be checked (almost 20,000 triples), which is readily carried out by an auxiliary computer program. The eigenvalues λ_1, λ_2, λ_3 can be computed with as much precision as required, of course.

Having checked this, a classical theory developed by H. Poincaré, S. Sternberg [40], and a number of other mathematicians, can be applied to conclude that there exist coordinates $y = (y_1, y_2, y_3)$ in C such that the expression of the vector field X in these coordinates is very close to being linear:

$$\begin{aligned} X(y) =& DX(0)y + \mathcal{R}(y) \\ |\mathcal{R}(y)| \leq& \text{const } |y_1|^{10} \left(|y_2| + |y_3|\right)^{10}. \end{aligned} \tag{11}$$

Why is this useful ? The trajectories of points under the linear flow $L(y) = DX(0)y$ can be expressed in a simple analytical form. Then, in view of (11), the trajectories of points in C for the actual flow X can be estimated with a good degree of accuracy. Note that we are talking of computations in the new y coordinates. One still has to go back to the x coordinates. This involves estimates for how far the transformation $x = \Psi(y)$ is from being the identity, that Tucker derives from explicit lower bounds for the norms of (10). According to his estimates, the errors involved in this step are bounded by 10^{-2}. So, he replaces the exit rectangle obtained for the linear flow L by another with size increased by 10^{-2} on each side. The latter rectangle is then treated in the same way as described above: subdivision into smaller rectangles, integration of small pieces of trajectories starting at the corresponding central points, and so on.

As explained before, this program should stop in finite time, with all the rectangles that were introduced in the course of the algorithm returning inside N. Tucker implemented this algorithm in C, with double precision floating point arithmetics (relative accuracy 10^{-15}). The program ran for about 30 hours on a Sparc Server station using both UltraSparc II 296 Mhz processors. And it stopped!

Invariant expanding cones. The steps I described so far show that equations (1) admit a trapping region, namely, the union of all the trajectories through the points of N. The set Λ of points that never leave the closure of this region, neither in the future nor in the past, is our candidate to being the attractor. However, at this point we still don't know much about Λ. For instance, it could be just an attracting periodic orbit. To rule out this possibility, and conclude that Λ is indeed a strange attractor, we still need properties (B) and (C). In fact, Tucker's program deals with all three statements simultaneously, as follows.

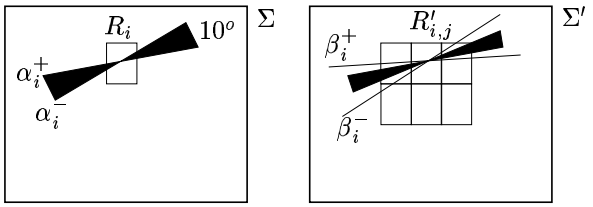

Figure 15: Invariant cones

Initially, besides feeding the program with a covering of N by small rectangles R_i, as explained before, Tucker also assigns to each R_i a cone C_i inside the tangent plane of R_i , represented by the slopes α_i^- and α_i^+ of its sides with respect to the x_1-direction. All these initial cones are chosen with total angle 10 degrees, but their inclination varies from one rectangle to another: their axis are more or less along the tangent directions to the two "arcs" that form the attractor, recall Figure 4, obtained from a preliminary (non-rigorous) computation [11].

[11] First, Tucker tried to use a constant cone field: axis everywhere horizontal, total angle 20 degrees. This turned out not to be invariant under the first-return map but the computer program failed to realize it, because cones were represented in terms of slopes only, so that the program was unable to tell a cone from its complement... Tucker modified the program to include orientation in the internal representation of cones, and he also added a safety device that would fire if cones were

Then, in parallel to finding a rigorous upper bound R_i' for the image of the rectangle R_i under the Poincaré map $\pi : \Sigma \to \Sigma'$, the program also looks for a corresponding upper bound for the image of C_i under the derivative of π. This derivative is computed through the formulae

$$\frac{\partial \pi_k}{\partial x_j}(x) = \frac{\partial \varphi_k^t}{\partial x_j}(x) - \frac{X_k}{X_3}(\pi(x)) \, \frac{\partial \varphi_3^t}{\partial x_j}(x),$$

where π_k, φ_k^t, X_k, are the components along the x_k-axis of π, the flow φ^t, and the vector field X. For this, it also needs to estimate the partial derivatives of the flow, which is done by numerical integration of the linearization

$$\frac{d}{dt}\left(D\varphi^t\right) = \left(DX \circ \varphi^t\right) \cdot D\varphi_k^t \quad \text{with } D\varphi^0 = \mathrm{Id} \tag{12}$$

of (1). Note that $D\varphi^t = (\partial \varphi_k^t / \partial x_j)$ is a 3×3 matrix, (12) corresponds to nine scalar differential equations.

These calculations, together with rigorous bounds on the integration errors, and on the total variation of $D\pi$ over the rectangle R_i, yield a cone C_i' inside the tangent plane of R_i' that contains $D\pi(x)C_i$ for every $x \in R_i$. See Figure 15. As before, this cone C_i' is described by the slopes β_i^- and β_i^+ of its sides. Finally, the program also computes a rigorous lower bound for the expansion inside C_i, that is, a positive number e_i such that

$$\|D\pi(x)v\| > e_i \|v\|$$

for all $x \in R_i$ and $v \in C_i$.

Now R_i' is subdivided into rectangles R_{ij}', as explained before, and the program proceeds the integration for each $(R_{ij}', \beta_i^-, \beta_i^+)$ in the place of the initial $(R_i, \alpha_i^-, \alpha_i^+)$. There is no subdivision of cones. Every integration step uses the same procedure as I just described for the first one, except that (i) the roles of x_1 and x_3 are interchanged when dealing with vertical cross-sections and (ii) linear approximations is used, instead, when solutions pass close to the origin.

ever to get larger than 20 degrees in the course of the iteration. This device remained silent when the program was run for the new initial cone field.

The program keeps track of successive expansion lower bounds, so that property (C) can be readily checked at the time of return to Σ: the product should be larger than some constant $\tau > 1$. In fact this is not quite so: somewhat surprisingly, vectors in the invariant cone field may even be contracted by the *first*-return map, if the starting point is close to the tips of the attractor [12]. However, this is compensated by expansion in subsequent returns: according to the computer output, the accumulated expansion exceeds 2 before the trajectory can cross Γ. This sort of estimate is sufficient for most purposes, including the proof that the attractor is indecomposable.

Moreover, the invariance property (B) translates into the following statement:

$$[\gamma_j^-, \gamma_j^+] \text{ is contained in } [\alpha_i^-, \alpha_i^+]$$

for every one of the initial $(R_i, \alpha_i^-, \alpha_i^+)$ and every returning $(S_j, \gamma_j^-, \gamma_j^+)$ such that S_j intersects R_i.

Successful verification of all these inequalities shows that Λ is a robust strange attractor for (1). Thus Λ is singular hyperbolic, and so it fits directly into the theory I described in the previous section. Its geometry and its ergodic properties can be well-understood, and they do correspond to those of the classical geometric Lorenz models. A happy conclusion to a beautiful story!

Acknowledgments. I'm most grateful to Maria José Pacifico and Warwick Tucker for explaining their work to me, and providing some of the bibliographical references. Many thanks also to Mattias Lindkvist, for caring to produce the 3-D pictures.

This work was partially supported by Faperj and Pronex-Dynamical Systems, Brazil.

References

[1] V. S. Afraimovich, V. V. Bykov, and L. P. Shil'nikov. On the appearence and structure of

[12] This is the reason why we need a constant c in the statement of condition (C). This fact was also not apparent when the program was first run, because the expansion was not correctly estimated.

the Lorenz attractor. *Dokl. Acad. Sci. USSR*, 234:336–339, 1977.

[2] A. Andronov and L. Pontryagin. Systèmes grossiers. *Dokl. Akad. Nauk. USSR*, 14:247–251, 1937.

[3] M. Benedicks and L. Carleson. The dynamics of the Hénon map. *Annals of Math.*, 133:73–169, 1991.

[4] C. Bonatti, L. J. Díaz, and E. Pujals. A C^1-generic dichotomy for diffeomorphisms: weak forms of hyperbolicity or infinitely many sinks or sources. Preprint, 1999.

[5] C. Bonatti, A. Pumariño, and M. Viana. Lorenz attractors with arbitrary expanding dimension. *C. R. Acad. Sci. Paris*, 325, Série I:883–888, 1997.

[6] R. Bowen. *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, volume 470 of *Lect. Notes in Math.* Springer Verlag, 1975.

[7] L. J. Díaz, E. Pujals, and R. Ures. Partial hyperbolicity and robust transitivity. *Acta Math.*, 1999. To appear.

[8] J. Guckenheimer and R. F. Williams. Structural stability of Lorenz attractors. *Publ. Math. IHES*, 50:59–72, 1979.

[9] B. Hassard, S. Hastings, W. Troy, and J. Zhang. A computer proof that the Lorenz equations have "chaotic" solutions. *Appl. Math. Lett.*, 7:79–83, 1994.

[10] S. Hastings and W. Troy. A shooting approach to the Lorenz equations. *Bull. Amer. Math. Soc.*, 27:298–303, 1992.

[11] S. Hayashi. Connecting invariant manifolds and the solution of the C^1 stability and Ω-stability conjectures for flows. *Annals of Math.*, 145:81–137, 1997.

[12] M. Hénon. A two dimensional mapping with a strange attractor. *Comm. Math. Phys.*, 50:69–77, 1976.

[13] M. Hénon and Y. Pomeau. Two strange attractors with a simple structure. In *Turbulence and Navier-Stokes equations*, volume 565, pages 29–68. Springer Verlag, 1976.

[14] M. Jakobson. Absolutely continuous invariant measures for one-parameter families of one-dimensional maps. *Comm. Math. Phys.*, 81:39–88, 1981.

[15] Yu. Kifer. *Random perturbations of dynamical systems.* Birkhäuser, 1988.

[16] L. D. Landau and E. M. Lifshitz. *Fluid mechanics.* Pergamon, 1959.

[17] E. N. Lorenz. Deterministic nonperiodic flow. *J. Atmosph. Sci.*, 20:130–141, 1963.

[18] E. N. Lorenz. On the prevalence of aperiodicity in simple systems. *Lect. Notes in Math.*, 755:53–75, 1979.

[19] S. Luzzatto and W. Tucker. Non-uniformly expanding dynamics in maps with singularities and criticalities. *Publ. Math. IHES.* To appear.

[20] S. Luzzatto and M. Viana. Lorenz-like attractors without invariant foliations. In preparation.

[21] S. Luzzatto and M. Viana. Positive Lyapunov exponents for Lorenz-like maps with criticalities. *Astérisque*, 1999.

[22] J. C. Maxwell. *Matter and motion.* Dover Publ., 1952. First edition in 1876.

[23] K. Mischaikow and M. Mrozek. Chaos in the Lorenz equations: a computer assisted proof (I). *Bull. Amer. Math. Soc.*, 32:66–72, 1995.

[24] K. Mischaikow and M. Mrozek. Chaos in the Lorenz equations: a computer assisted proof (II). *Math. Comp.*, 67:1023–1046, 1998.

[25] C. Morales, M. J. Pacifico, and E. Pujals. Partial hyperbolicity and persistence of singular attractors. Preprint 1999.

[26] C. Morales, M. J. Pacifico, and E. Pujals. On C^1 robust singular transitive sets for three-dimensional flows. *C. R. Acad. Sci. Paris*, 326, Série I:81–86, 1998.

[27] J. Palis. A global view of Dynamics and a conjecture on the denseness of finitude of attractors. *Astérisque*, 261:339–351, 1999.

[28] J. Palis and F. Takens. *Hyperbolicity and sensitive-chaotic dynamics at homoclinic bifurcations.* Cambridge University Press, 1993.

[29] Ya. Pesin. Families of invariant manifolds corresponding to non-zero characteristic exponents. *Math. USSR. Izv.*, 10:1261–1302, 1976.

[30] H. Poincaré. *Science and method.* Dover Publ., 1952. First edition in 1909.

[31] Lord Rayleigh. On convective currents in a horizontal layer of fluid when the higher temperature is on the under side. *Phil. Mag.*, 32:529–546, 1916.

[32] C. Robinson. Homoclinic bifurcation to a transitive attractor of Lorenz type. *Nonlinearity*, 2:495–518, 1989.

[33] D. Ruelle and F. Takens. On the nature of turbulence. *Comm. Math. Phys.*, 20:167–192, 1971.

[34] M. Rychlik. Lorenz attractors through Shil'nikov-type bifurcation. Part 1. *Erg. Th. & Dynam. Syst.*, 10:793–821, 1989.

[35] B. Saltzman. Finite amplitude free convection as an initial value problem. *J. Atmos. Sci.*, 19:329–341, 1962.

[36] M. Shub. *Global stability of dynamical systems.* Springer Verlag, 1987.

[37] Ya. Sinai. Gibbs measure in ergodic theory. *Russian Math. Surveys*, 27:21–69, 1972.

[38] S. Smale. Differentiable dynamical systems. *Bull. Am. Math. Soc.*, 73:747–817, 1967.

[39] C. Sparrow. *The Lorenz equations: bifurcations, chaos and strange attractors*, volume 41 of *Applied Mathematical Sciences.* Springer Verlag, 1982.

[40] E. Sternberg. On the structure of local homeomorphisms of euclidean n-space - II. *Amer. J. Math.*, 80:623–631, 1958.

[41] W. Tucker. *The Lorenz attractor exists.* PhD thesis, Univ. Uppsala, 1998. Text and program codes available at www.math.uu.se/∼warwick/.

[42] W. Tucker. The Lorenz attractor exists. *C. R. Acad. Sci. Paris*, 328, Série I:1197–1202, 1999.

[43] M. Viana. Dynamics: a probabilistic and geometric perspective. In *Procs. International Congress of Mathematicians ICM98-Berlin*, Documenta Mathematica, vol I, pages 557–578. DMV, 1998.

[44] R.F. Williams. The structure of the Lorenz attractor. *Publ. Math. IHES*, 50:73–99, 1979.

Marcelo Viana (viana@impa.br)
IMPA, Est. D. Castorina 110,
22460-320 Rio de Janeiro, Brazil
www.impa.br/∼viana/