ELSEVIER

Physica D 129 (1999) 35–56

PHYSICA D

Bifurcation rigidity

Brian R. Hunt [a,*], Jason A.C. Gallas[1] [a], Celso Grebogi [a], James A. Yorke [a], Hüseyin Koçak [b]

[a] *Institute for Physical Science and Technology, University of Maryland, College Park, MD 20742, USA*

[b] *Department of Mathematics and Computer Science, University of Miami, Coral Gables, FL 33124, USA*

Received 23 October 1997; received in revised form 12 June 1998; accepted 26 June 1998

Communicated by J.D. Meiss

Dedicated to Professor Junji Kato on his 60th birthday

Abstract

Bifurcation diagrams of periodic windows of scalar maps are often found to be not only topologically equivalent, but in fact to be related by a nearly linear change of parameter coordinates. This effect has been observed numerically for one-parameter families of maps, and we offer an analytical explanation for this phenomenon. We further present numerical evidence of the same phenomenon for two-parameter families, and give a mathematical explanation like that for the one-parameter case.

Keywords: Bifurcation diagrams; Period-doubling bifurcations; Saddle-node bifurcations

1. Introduction

Universal metric properties of bifurcation diagrams are some of the more striking phenomena in dynamical systems. A well-known example is the observation by Feigenbaum that for many one-parameter families of scalar maps, as one follows a cascade of period-doubling bifurcations in parameter space, the ratio of distances between successive period-doubling parameter values tends to approach a limiting ratio independent of the family of maps [3,4,12]. Our paper is concerned with a global scaling property for bifurcation diagrams of periodic orbits of smooth scalar maps with both one- and two-dimensional parameter spaces. Specifically, we examine "windows" of periodic behavior within chaotic regions of parameter space. In both the one-parameter and two-parameter cases we show that there is a canonical family of maps such that typically the bifurcations within a periodic window of a given scalar map are well approximated by a *linear* transformation of the bifurcation diagram of the canonical map.

The possibility of such a global linear scaling law was suggested for one-parameter families of maps by the numerical observations in [20]. More recently we have observed a similar phenomenon for two-parameter families (see also [13,17,18]). We will discuss this numerical evidence and exhibit the linear scaling law through several

* Corresponding author. Tel.: +1-301-405-4885; fax: +1-301-314-9363; e-mail: bhunt@ipst.umd.edu.

[1] Current address: Instituto de F´ısica, Universidade Federal do Rio Grande do Sul, 91501-970 Porto Alegre, Brazil.

0167-2789/99/$ – see front matter
PII: S0167-2789(98)00201-2

examples in Section 2. The topological similarities between the bifurcation diagrams of different windows is apparent from the figures in Section 2; we think of the further linear similarity as a sort of "rigidity" for bifurcation diagrams.

We find that a central feature of a region of periodic stability surrounded by chaotic behavior is a point in parameter space at which the map has a "superstable" orbit – a periodic orbit that includes a critical point of the map. Near a superstable period n orbit, the nth iterate of the map is generally well approximated by a pure quadratic map, and the canonical bifurcation diagram in the one-parameter case is found to be that of the quadratic family $x \mapsto x^2 - c$. The idea that iterated one-dimensional maps are often, under renormalization, nearly quadratic is by now well known – in particular, the scaling we describe here is shown (in various contexts) in [1,2,8,10] to hold asymptotically for certain sequences of windows approaching special parameter values (Misieurewicz points). However, with this approach the relationship between the closeness of the linear approximation and the period of the window depends on the particular sequence, and thus the renormalization results do not fully address the question of how well the linear scaling law applies to a typical window of a given period. Our aim is to demonstrate that the linear scaling property is global in parameter space, in the sense that a period n window chosen at random from all period n windows is likely to exhibit the linear scaling property, with the likelihood increasing as n increases. In Section 3 we formulate precise conditions under which we can prove a close linear correspondence between a given periodic window and the canonical diagram, and discuss the consequences of this result.

In the two-parameter case, the orbit will be superstable along a curve in parameter space. In general we expect that along lines transverse to the curve of superstability, the bifurcation diagram will resemble a one-parameter diagram to which the above scaling result applies. Indeed this should always be the case if the map has only one critical point. However, if the map has more than one critical point, then somewhere along the curve of superstability we can expect the orbit to become "doubly superstable" – to include a second critical point. Near such a point, we find that the parameter region of stability for the orbit is at its thickest, and the bifurcation diagram is not well described by the one-parameter scaling result.

Near a doubly superstable orbit of period n, the nth iterate of the map is well approximated by the composition of two quadratic maps, each of which depends linearly on the parameters. Typically each will depend on a different linear combination of the parameters, and a linear change of coordinates leads to the canonical two-parameter family $x \mapsto (x^2 - a)^2 - b$. We formulate a result for this case in Section 4, and discuss its ramifications.

2. Examples

The bifurcation diagram for the quadratic family of maps $x \mapsto x^2 - c$, shown in Fig. 1, is a familiar picture to dynamicists. For each value of the parameter c on the horizontal axis, the limiting behavior of a bounded trajectory

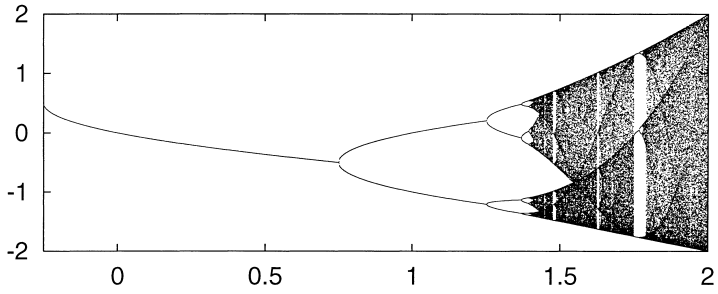


Fig. 1. Bifurcation diagram for the quadratic map $x \mapsto x^2 - c$. For each of 2000 parameter values along the (horizontal) c-axis, the map is iterated 1000 times starting with $x = 0$, then for the next 100 iterates a dot is plotted at coordinates (c, x). A vertical slice of the diagram at a parameter value c thus represents the asymptotic behavior found numerically for c.

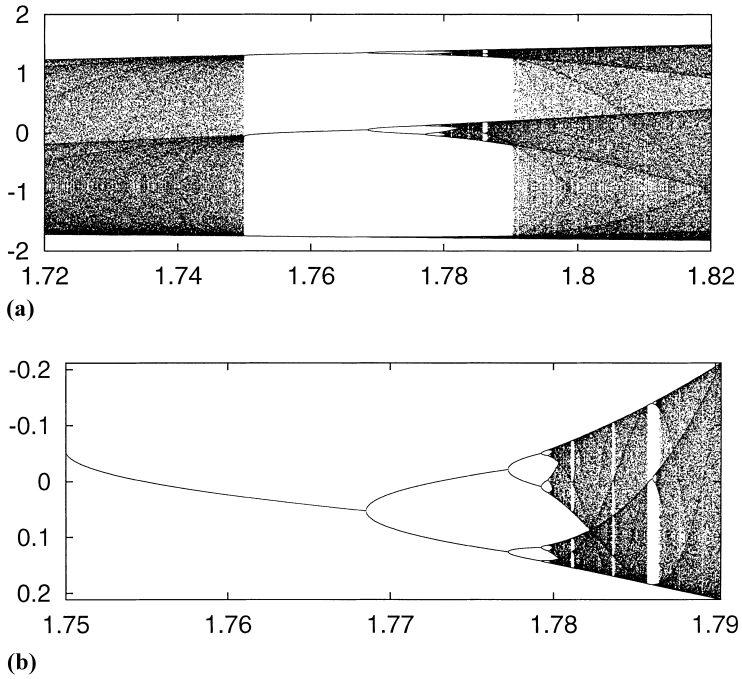


Fig. 2. (a) Enlargement of the interval $1.72 \leq c \leq 1.82$ in the bifurcation diagram in Fig. 1, created in the same manner. (b) Further enlargement of (a) near $x = 0$, with vertical axis flipped for comparison with Fig. 1.

is computed and plotted on the vertical axis. As c is increased, we see at first a stable fixed point, followed by a stable period 2 orbit, then period 4, 8, etc. For larger values of c the trajectories are often chaotic, filling one or more intervals on the x-axis.

Within the parameter interval in which chaotic behavior is observed, there are smaller intervals in which there is periodic behavior. To be precise, we say a map exhibit "period n behavior" if there is a cycle of n disjoint intervals that the map permutes. We call an interval of parameters for which the corresponding maps exhibit period n behavior a "period n window" (provided that the interval is maximal in this regard and that n is the minimal period for the given interval). We call a window "primary" if it is not contained in a window of another (smaller) period.

An example of a primary window is the period 3 window in Fig. 1 near $c = 1.75$, which is shown in more detail in Fig. 2(a) . As c is increased, a stable period 3 orbit is created by a saddle-node bifurcation, then a series of period-doubling bifurcations produce stable orbits of periods 6, 12, etc. What follows is a parameter region in which the dynamics are chaotic but still follow a period 3 pattern; the chaotic attractor consists of three intervals that are permuted by the map. Eventually this three-piece attractor is destroyed and a larger chaotic interval returns. The creation of the stable period 3 orbit at the saddle-node bifurcation parameter and the destruction of the three-piece attractor at the crisis parameter form the boundary of this window in parameter space, and between these two parameters the bifurcation diagram is qualitatively very similar to the global bifurcation diagram.

In [20] it was observed further that there is a quantitative correspondence between most primary periodic windows of one-parameter families and the "canonical" bifurcation diagram for the quadratic map $x \mapsto x^2 - c$. If the parameter interval for the window is linearly transformed to match the parameter interval for the canonical diagram, as in Fig. 2(b), the intermediate bifurcations occur at similar parameter values, with the correspondence becoming closer for smaller and smaller windows.

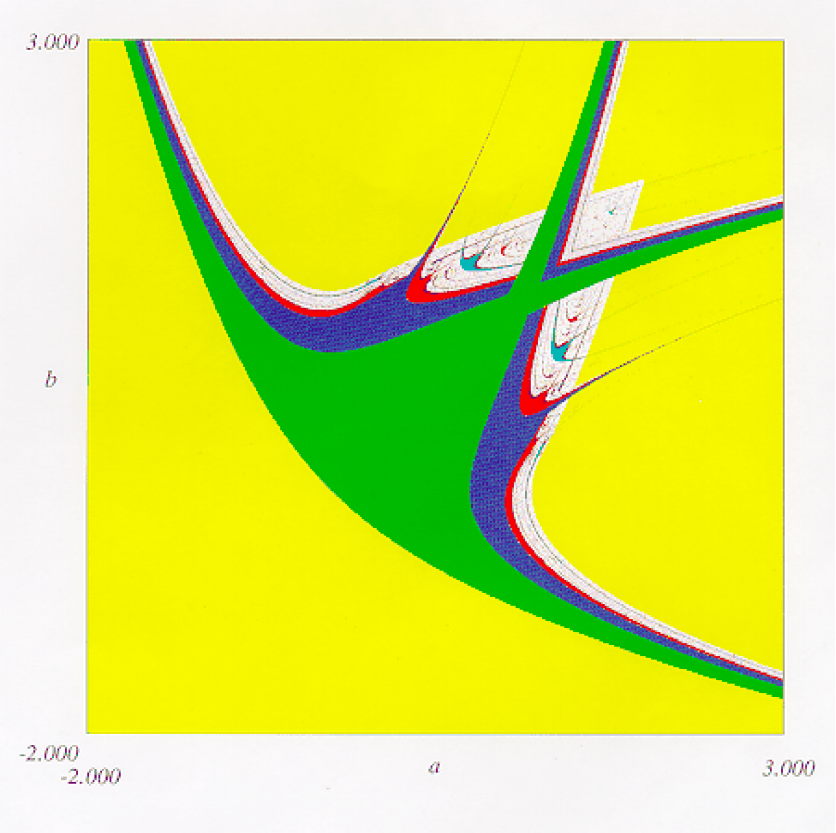
Table 1
Numerical data for all primary windows of period at most 8, sorted by window length. (The ratio which is compared with 9 in the last column is the length of the window divided by the distance from its left endpoint to the superstable parameter value in the window.)

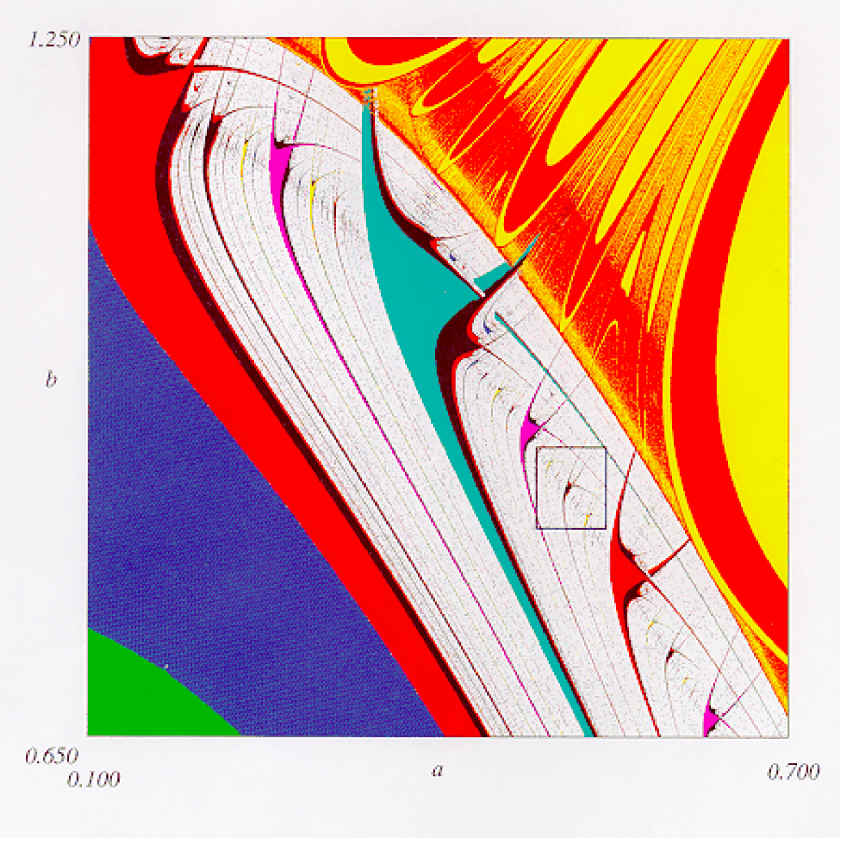
Period	Left endpoint	Window length	9−Ratio
3	1.750000	0.04032749	0.7322
6	1.474695	0.01090425	0.7346
5	1.624397	0.00896171	0.1858
4	1.940551	0.00221122	0.1202
5	1.860587	0.00174412	0.0809
7	1.574716	0.00155813	0.0161
7	1.673954	0.00100256	0.0550
8	1.521706	0.00098832	0.1174
8	1.711036	0.00038896	−0.0098
6	1.907251	0.00026439	0.0112
7	1.832291	0.00022141	0.0129
5	1.985410	0.00013069	0.0254
8	1.809989	0.00011436	0.0192
7	1.884793	0.00009851	0.0075
8	1.851721	0.00008554	−0.0192
6	1.966764	0.00008007	0.0086
7	1.927141	0.00006387	−0.0011
8	1.869997	0.00006016	0.0255
7	1.953702	0.00003529	0.0069
8	1.896916	0.00002210	0.0005
8	1.917096	0.00001721	0.0014
7	1.977178	0.00001345	0.0018
6	1.996375	0.00000801	0.0060
8	1.960758	0.00000660	0.0007
7	1.991814	0.00000461	0.0016
8	1.972199	0.00000427	0.0006
8	1.981655	0.00000282	0.0003
8	1.988793	0.00000166	0.0006
8	1.994333	0.00000079	0.0004
7	1.999096	0.00000050	0.0015
8	1.997963	0.00000028	0.0004
8	1.999774	0.00000003	0.0004

The accuracy of the linear scaling was tested in [20] by computing for various windows a particular ratio of distances between bifurcations in parameter space. The ratio chosen was the length of the window divided by the distance from the saddle-node bifurcation to the first period-doubling bifurcation. This ratio is of course preserved by linear transformations, so a necessary condition for close linear correspondence between a period n window and the canonical bifurcation diagram is that the ratio be close to $\frac{9}{4}$ – the ratio for the canonical diagram. Indeed the

Fig. 3. Bifurcation diagram for the canonical family $x \mapsto (x^2 - a)^2 - b$. This and the following figures depict a 1500×1500 grid in parameter space. For each parameter pair, the map is iterated 800 times, and then the trajectory is examined for periodic behavior. If the trajectory is becoming unbounded, a yellow dot is plotted. If the trajectory remains bounded but is not discernibly periodic, or has period greater than 32, a white dot is plotted. If the trajectory is found to approach an orbit with low period, the dot is colored according to the period, for instance 1 is green, 2 is blue, 3 is cyan (light blue), 4 is red, 5 is magenta, and 6 is black. The same set of colors is used for periods 7–32 as well, with most of these periods being colored black to enhance the visibility of smaller shrimp. Due to the existence of small parameter regions for which the map has two attractors, in this figure the orbits of both critical points were checked and the color was chosen according to the lower period observed. In each of the subsequent figures, a single arbitrarily chosen initial condition was used. The fact that only one attractor is observed for each parameter pair results in small regions of asymmetry in Figs. 6 and 9.

Fig. 4. Bifurcation diagram for the cubic family $x \mapsto -x^3 + 3ax + b$. See also the caption for Fig. 3. (In this figure only, initial conditions with unbounded trajectories are plotted either yellow or red.)


3.000
b
-2.000
-2.000
a
3.000


1.250
b
0.650
0.100
a
0.700

ratio is found numerically to rapidly approach $\frac{9}{4}$ for successively smaller primary windows of the quadratic map. In Table 1 we show numerical results for a similar ratio, namely the length of the window divided by the distance from the saddle-node bifurcation to the superstable parameter value; for the canonical diagram this ratio is 9.

For a one-parameter cross-section of the Hénon family, the ratio is also found in [20] to be close to $\frac{9}{4}$ for most of the windows computed, indicating that this phenomenon is not restricted to one-dimensional maps. In [19] it is shown that under some circumstances rescalings of one-parameter families of two-dimensional maps approach the (one-dimensional) quadratic family (see also [15]). Yet we feel that the phenomena discussed in this paper apply only to sufficiently dissipative maps with just one positive Lyapunov exponent – roughly speaking, maps that are only one-dimensionally expanding. We present some numerical results for the full two-parameter Hénon family below.

Recently there has been a great deal of interest in the structure of parameter space for two-parameter families of maps. Instead of a bifurcation diagram that includes both parameter and phase space variables, one generally draws the diagram entirely in parameter space, assigning colors or shading to regions according to the dynamics of the corresponding map; see for example [5–7,11]. A common shape has been observed, for instance in [16–18], and has been referred to as a "swallow" [13,14] and "shrimp" [5–7]. This shape is seen in Fig. 3, which shows the bifurcation diagram for the canonical family of maps $x \mapsto (x^2 - a)^2 - b$. This canonical family has been identified in both [17,18] and [13]. The former papers give an example of a linear transformation that brings the stability region of a period 7 orbit of a two-parameter circle map very close to the stability region for the fixed point in Fig. 3. In [13,14], two-parameter families of polynomial maps of the complex plane are considered, and the "swallow" is one of several canonical topological shapes found in parameter space for these complex maps. In [11], a two-parameter family of scalar maps is used to study the scaling of these stability regions as certain points on the boundary of the chaotic parameter set are approached; they observe linear scaling to a different canonical shape than we observe in the interior of the chaotic parameter set.

For the cubic family on the real line, which we parametrize as $x \mapsto -x^3 + 3ax + b$, we find the bifurcation diagram shown in Fig. 4. We see a number of shrimps, many of which look very different in shape from the one in Fig. 3. Nonetheless most of them are very close to a linear, though heavily sheared, image of Fig. 3. In Fig. 5 the square surrounding the period 6 shrimp in Fig. 4 is blown up. Next, we have transformed the parallelogram shown in Fig. 5 linearly onto the square that makes up Fig. 6.

As in the one-parameter case, we find numerically that this scaling property is not restricted to one-dimensional maps. In Fig. 7 we show a bifurcation diagram for the Hénon family

$$x \mapsto a - x^2 + by, \quad y \mapsto x.$$

As before, we blow up a square around a period 9 shrimp in Fig. 7 to get Fig. 8. The parallelogram in Fig. 8 is transformed linearly to a square to obtain Fig. 9.

In both of these cases, the linear change of coordinates was determined from the data by requiring it to transform three reference points on the bifurcation diagram to the corresponding points on the canonical bifurcation diagram. For one-dimensional maps, one can also find the canonical change of coordinates analytically, as described in Section 4.

3. One-parameter families

We begin this section with an explanation of our assertion that a central feature of a typical periodic window is a superstable orbit. Following that, we will discuss our main result, that the nth iterate of a smooth family of maps

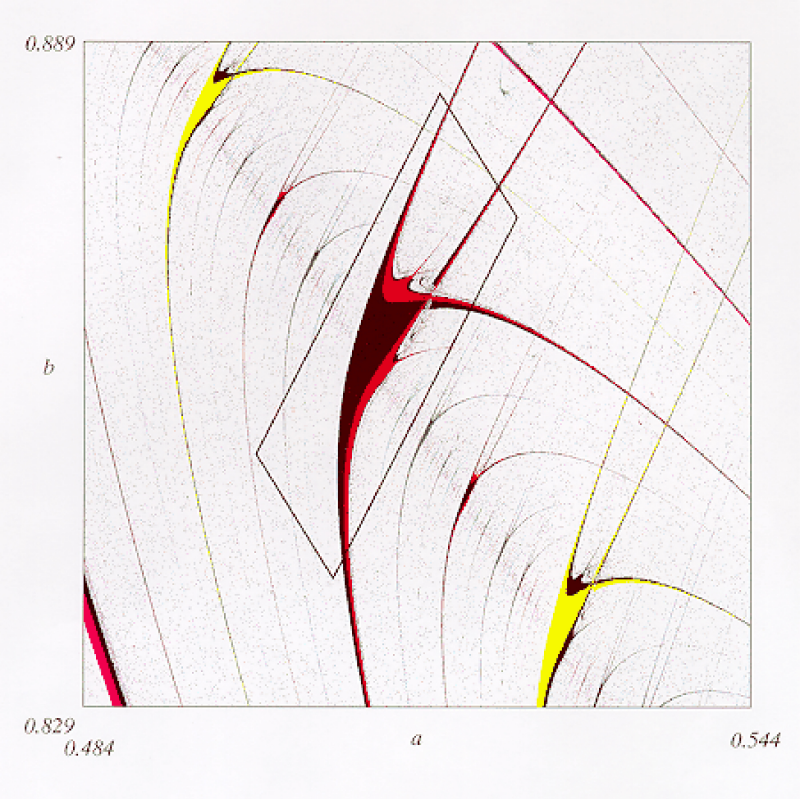


Fig. 5. Blow-up of square in Fig. 4.

Fig. 6. Linear transformation of parallelogram in Fig. 5.

near a superstable period n point is usually well approximated under a linear change of coordinates by the quadratic family $x \mapsto x^2 - c$, and explore its ramifications.

Though the superstable orbit is not a visible feature of a typical bifurcation diagram of a periodic window, there must be one between every saddle-node and a period-doubling bifurcation. This is because when a period n orbit is created by a saddle-node bifurcation, the nth iterate of the map has derivative 1 on the orbit, and at the period-doubling bifurcation the map has derivative -1 on the orbit, so at some intermediate parameter value the nth iterate has derivative zero, i.e., the orbit is superstable.

One way to understand the importance of superstability in our analysis of periodic windows is as follows. Chaotic maps are characterized by a positive Lyapunov exponent, i.e. exponential divergence of nearby trajectories. For a periodic orbit to be stable at nearby parameter values, some point of the orbit must move close to a critical point of the map, so that the rapid local contraction of the map near that point is enough to counteract the local expansion of the map near the other points of the orbit. Typically the orbit will then pass through the critical point, becoming superstable at that parameter value, somewhere in the region of stability.

Consider now the dynamics of a map $x \mapsto f(x, \lambda)$ near a period n superstable orbit. Let λ_0 be the superstable parameter value, and let $x_0, x_1, \ldots, x_{n-1}$ be the orbit with x_0 a critical point of f. Near x_0, the map will typically be quadratic (to lowest order), and a neighborhood of x_0 will be folded over by f and then stretched by the next $n-1$ iterates of f before returning close to x_0. Thus the nth iterate of f, which we write as f^n, will be like a quadratic function as well, but with large derivatives due to the stretching. By rescaling both x and λ near (x_0, λ_0), we can make the lowest order terms in f^n match those of the quadratic family, but in order to say that the rescaled map is close to the quadratic family, we must ensure that the higher order terms are made small by the rescaling. We next offer a heuristic argument that the higher order terms become insignificant when the amount of stretching is large, then we give a more formal explanation with Principle 1.

Let

$$S = \frac{\partial f^{n-1}}{\partial x}(x_1, \lambda_0),$$

i.e., S is the stretching rate of f^{n-1} near x_1, which we assume is large. The main technical point, which is addressed in Appendix A, is that typically the kth order partial derivatives of f^{n-1} near x_1 are of order S^k, so that for some p_1 of order 1,

$$f^{n-1}(x, \lambda) - x_0 = S(x - x_1) + p_1 S(\lambda - \lambda_0) + \mathrm{O}(S^2(|x - x_1|^2 + |\lambda - \lambda_0|^2)),$$

provided $|x - x_1|$ and $|\lambda - \lambda_0|$ are small compared with S^{-1}. Also, for some q and p_2 of order 1,

$$f(x, \lambda) - x_1 = q(x - x_0)^2 + p_2(\lambda - \lambda_0) + \mathrm{O}(|x - x_0|^3 + |x - x_0||\lambda - \lambda_0| + |\lambda - \lambda_0|^2)$$

for (x, λ) near (x_0, λ_0). It then follows by composing the above expansions that

$$\begin{aligned} f^n(x, \lambda) - x_0 = {} & qS(x - x_0)^2 + pS(\lambda - \lambda_0) + \mathrm{O}(S^2(|x - x_0|^4 + |\lambda - \lambda_0|^2)) \\ & + \mathrm{O}(S(|x - x_0|^3 + |x - x_0||\lambda - \lambda_0|)), \end{aligned}$$

where $p = p_1 + p_2$.

To be precise,

$$q = \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(x_0, \lambda_0), \quad p = S^{-1}\frac{\partial f^n}{\partial \lambda}(x_0, \lambda_0).$$

The linear change of coordinates we use is

$$y = qS(x - x_0), \quad c = -pqS^2(\lambda - \lambda_0).$$

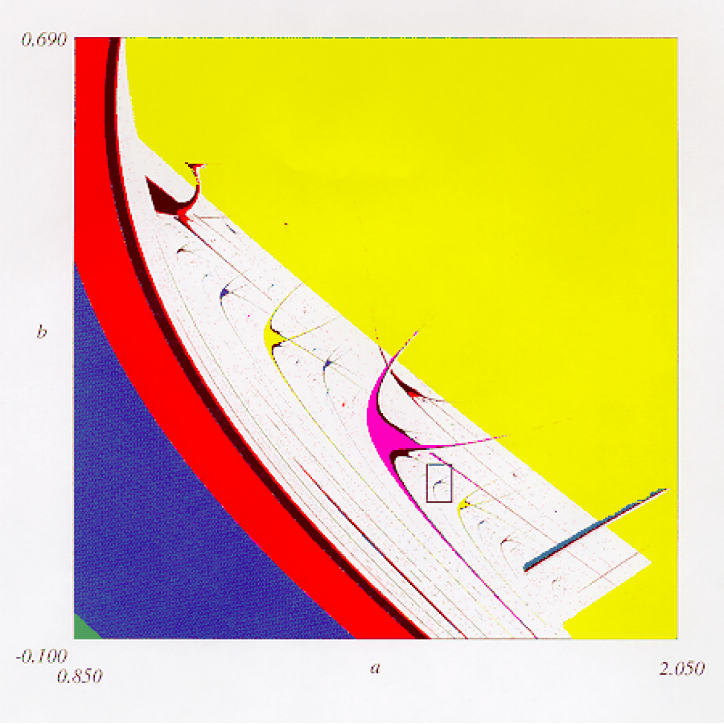


Fig. 7. Bifurcation diagram for the Henon family $(x, y) \mapsto (a - x^2 + by, x)$. See also the caption for Fig. 3.

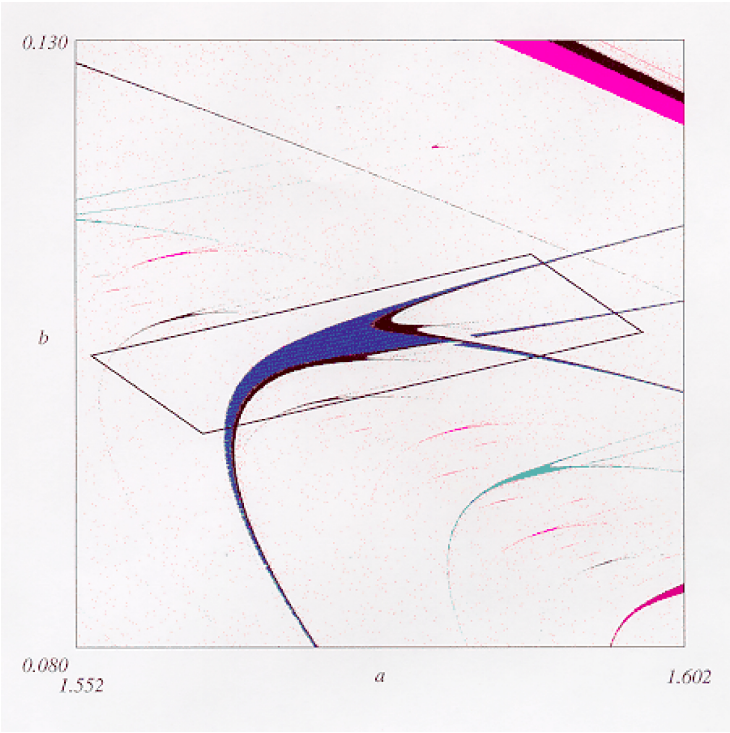


Fig. 8. Blow-up of square in Fig. 7.

Fig. 9. Linear transformation of parallelogram in Fig. 8.

Let $y \mapsto g(y, c)$ be the family of maps conjugate to $x \mapsto f^n(x, \lambda)$ under this change of coordinates. Then from the above expansion for f^n one finds that

$$g(y, c) = y^2 - c + \mathrm{O}(S^{-1}(|y|^3 + |y||c| + |c|^2)).$$

The following principle formalizes the above argument. Its goal is to provide a basis for saying not only that g is typically close to the quadratic family $y \mapsto y^2 - c$, but that certain features of the bifurcation diagrams of the two maps are close. Since the bifurcation diagram of the quadratic family spans the rectangle $(y, c) \in [-2, 2] \times [-0.25, 2]$, we need the maps to be close in a neighborhood of this rectangle, say $[-2.5, 2.5] \times [-2.5, 2.5]$. Further, we need not only that $|g(y, c) - (y^2 - c)|$ be small in this region, but that the partial derivatives (up to a certain order) of these two functions be close to each other as well. We choose to show that the maps are close in the C^2 norm, though some additional information about the bifurcation diagrams could be obtained from a C^3 estimate. At the end of this section, we explain how to conclude from this estimate that similar bifurcations of the two maps occur at similar parameter values.

Principle 1. *Let X and Λ be compact intervals. For a typical C^3 family of maps $f : X \times \Lambda \to X$ and a typical superstable period n orbit with critical point (x_0, λ_0), there is a linear change of coordinates that conjugates f^n near (x_0, λ_0) to the quadratic family $y \mapsto y^2 - c$ in the square $[-2.5, 2.5] \times [-2.5, 2.5]$ to within an error ε in the C^2 norm, where $\varepsilon \to 0$ as n increases.*

Remark. We state this and the main result of Section 4 as "principle" and "argument" rather than "theorem" and "proof" because our use of the word "typical" does not easily translate into a precise mathematical definition. In the course of the argument we will formulate precise conditions for f and (x_0, λ_0) to be "typical"; in brief, these

conditions may be understood as follows. First, f must not have an unusually small second derivative at the critical point x_0 (to ensure that the quadratic term of f^n is not dominated by higher order terms). Second, f^n must not have an unusually small parameter derivative at (x_0, λ_0) (to ensure that the higher order parameter terms of f^n do not dominate the linear parameter term). Third, the other points $x_1, x_2, \ldots, x_n$ of the orbit must not lie too close to the critical points of f (to ensure that the derivative of f^{n-1} near x_1 is not too small compared with its higher order derivatives).

We remark also that the C^3 hypothesis on f can be weakened to C^2 with $\partial^2 f/\partial x^2$ Lipschitz.

Argument for Principle 1. Recall that

$$S = \frac{\partial f^{n-1}}{\partial x}(x_1, \lambda_0), \qquad q = \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(x_0, \lambda_0), \qquad p = S^{-1}\frac{\partial f^n}{\partial \lambda}(x_0, \lambda_0).$$

The constant S represents the linear stretching factor of the $n-1$ iterates of f from x_1 to x_0, which should typically be large like an exponential function of n if f is chaotic at nearby parameter values. The constant q, the second derivative of f at the critical point, will typically be nonzero except perhaps at a finite number of parameter values; we assume λ_0 does not happen to lie close to one of those values. Next, p is S^{-1} times the parameter derivative of f^n at (x_0, λ_0); this derivative will typically be of the same order as S, as is illustrated by the expansion

$$\frac{\partial f^n}{\partial \lambda}(x_0, \lambda_0) = \frac{\partial f}{\partial \lambda}(x_{n-1}, \lambda_0) + \frac{\partial f}{\partial \lambda}(x_{n-2}, \lambda_0)\frac{\partial f}{\partial x}(x_{n-1}, \lambda_0) + \cdots + \frac{\partial f}{\partial \lambda}(x_0, \lambda_0)\frac{\partial f^{n-1}}{\partial x}(x_1, \lambda_0).$$

Notice that the last term is S multiplied by $(\partial f/\partial \lambda)(x_0, \lambda_0)$, and thus is typically of the same order as S. There are other terms in the sum of order S as well, but there is no reason to expect them to cancel out in general. Thus both p and q typically do not get closer to zero as n increases.

Recall that the linear change of coordinates we use is

$$y = qS(x - x_0), \quad c = -pqS^2(\lambda - \lambda_0).$$

Let g be the map conjugate to f^n under this change of coordinates. One then finds that

$$\frac{\partial g}{\partial y}(0, 0) = 0, \quad \frac{\partial^2 g}{\partial y^2}(0, 0) = 2, \quad \frac{\partial g}{\partial c}(0, 0) = -1,$$

so that to lowest order in y and c,

$$g(y, c) \approx y^2 - c.$$

We must prove that typically the higher order terms are small when n, and hence S, is large.

Let M be a number that bounds the partial derivatives of f up to order 3 on $X \times \Lambda$ as follows:

$$\left|\frac{\partial^{i+j} f}{\partial x^i \partial \lambda^j}\right| \le M|p^j q^{i+j-1}|, \quad 1 \le i + j \le 3.$$

(The reason that we choose this form for the bounds is so that M is dimensionless with regard to any units given to x and λ.)

Let R be the rectangle around (x_0, λ_0) that is mapped onto the square $[-2.5, 2.5] \times [-2.5, 2.5]$ under the above change of coordinates, i.e.

$$R = \left\{(x, \lambda) : |x - x_0| \le \frac{2.5}{|qS|}, |\lambda - \lambda_0| \le \frac{2.5}{|pqS^2|}\right\},$$

and let

$$\rho = |S|^{-1}\sup\left\{\left|\frac{\partial f^{k-j}}{\partial x}(f^j(x,\lambda),\lambda)\right| : 1 \le j < k \le n; (x,\lambda) \in R\right\}.$$

$$\sigma = \sup\left\{1, \left|\frac{\partial f}{\partial x}(f^j(x,\lambda),\lambda)\right|^{-1} : 1 \le j < n; (x,\lambda) \in R\right\}.$$

We then have the following lemma, which is proved in Appendix A.

Lemma 1. *Assume* $|S| \ge 1$. *The distance* ε *between* $g(y,c)$ *and* $y^2 - c$ *in the* C^2 *norm for* $(y,c) \in [-2.5, 2.5] \times [-2.5, 2.5]$ *can be bounded as follows:*

$$\varepsilon \le \tfrac{10}{3}(n+7)^3\rho^2\sigma M^3|S|^{-1}.$$

For the bound on ε to be small, we must have not only that p and q are not too small, but also that ρ and σ are not too large. We claim that this will be the case as long as the orbit $x_1, x_2, \ldots, x_{n-1}$ does not pass unusually close to a critical point of f, and does not spend unusually long periods of time in contracting regions of phase space.

The constant σ is simply one over the minimum of $\partial f/\partial x$ over the iterates from 1 to $n-1$ of an initial point $(x,\lambda) \in R$. Since

$$f(x,\lambda) \approx x_1 + \frac{1}{2}q(x - x_0)^2 + \frac{\partial f}{\partial \lambda}(x_0,\lambda_0)(\lambda - \lambda_0)$$

for (x,λ) near (x_0,λ_0), then for $(x,\lambda) \in R$, the distance from $f(x,\lambda)$ to x_1 should be of order $|S|^{-2}$. Further iterates of f will stretch distances by up to a cumulative factor of $|S|$, but then the trajectory of (x,λ) should remain within order $|S|^{-1}$ of $x_1, x_2, \ldots, x_{n-1}$. Now as n increases, we should expect the closest approach of an orbit $x_1, x_2, \ldots, x_{n-1}$ to a critical point of f to decrease, but typically it will be at least of order $1/n$. Since $|S|^{-1}$ is typically exponentially small as a function of n, we conclude that for large n, the constant σ will typically be at most of order n.

Next, notice that for $1 \le j < k \le n$, the effect on $|(\partial f^{k-j}/\partial x)(f^j(x,\lambda),\lambda)|$ of varying $(x,\lambda) \in R$ is small compared with S because of the proximity of $f(x,\lambda), \ldots, f^{n-1}(x,\lambda)$ to $x_1, \ldots, x_{n-1}$, as described in the above paragraph. Thus the only way for ρ to be large is for some $|(\partial f^{k-j}/\partial x)(x_j,\lambda_0)|$ to greatly exceed $|S|$, which can only happen if f is greatly contracting along the iterates $x_1, \ldots, x_{j-1}$ or $x_k, \ldots, x_{n-1}$. This is the situation we ruled out above, though as in the above paragraph we anticipate the possibility that f contracts by a factor of order $1/n$ occasionally. We conclude that for large n, the constant ρ will typically be at the very most polynomial in n, and therefore the upper bound on ε will decrease to zero as n increases. □

We conclude this section by arguing that when $g(y,c)$ is sufficiently close to $y^2 - c$ in the C^2 norm, then the bifurcations of the map $y \mapsto g(y,c)$ that determine the ratios that were considered in Section 2 occur at parameter values near to those for the quadratic family. That is, g has a saddle-node bifurcation creating a fixed point near $c = -0.25$, a period-doubling bifurcation for this fixed point near $c = 0.75$, and a crisis causing the escape of the orbit of the critical point from the interval $[-2.5, 2.5]$ near $c = 2$.

For a generic saddle-node bifurcation, g must have a fixed point at which

$$\frac{\partial g}{\partial y}(y,c) = 1, \quad \frac{\partial g}{\partial c}(y,c) \ne 0, \quad \frac{\partial^2 g}{\partial y^2}(y,c) \ne 0$$

(see, for instance, [9]). For the quadratic family $y \mapsto y^2 - c$, these conditions are met at $y = 0.5, c = -0.25$.

Consider the function $F : C^2(\mathbb{R}^2) \times \mathbb{R}^2 \to \mathbb{R}^2$ defined by

$$F(g, y, c) = \left(g(y, c) - y, \frac{\partial g}{\partial y}(y, c) - 1 \right).$$

The Jacobian matrix of F with respect to y and c is nonsingular at $g = y^2 - c$, $y = 0.5$, $c = -0.25$, so by the implicit function theorem, if g is sufficiently C^2 close to $y^2 - c$, there is a unique point (y, c) near $(0.5, -0.25)$ for which $F(g, y, c) = 0$. Further, since g is C^2 close to a function with second derivative 2 and parameter derivative -1, these derivatives of g are nonzero at (y, c), and the conditions for a generic saddle-node bifurcation at (y, c) are met.

Similar reasoning can be used to argue that any bifurcation of the quadratic family whose location is determined by conditions on the derivatives of the map at a periodic point (or conditions on the itinerary of the critical point) occurs at a nearby parameter value for g (and vice versa) when ε is sufficiently small. The period-doubling bifurcation requires conditions on the third derivative of the map, so to duplicate the above argument in that case would require a C^3 estimate on the distance between g and $y^2 - c$. Though there is no obstacle to such an estimate, we have omitted it for the sake of brevity. Even with only a C^2 estimate we can show that a transition from a stable fixed point to a stable period 2 orbit occurs somewhere in the vicinity of $c = 0.75$ (we just cannot say that a *generic* period-doubling bifurcation occurs). With this distinction in mind, we offer the following result in support of the principle that the bifurcation diagram of g is close to a linear image of the bifurcation diagram of the quadratic family.

Proposition 1. *For all $\delta \in (0, 0.5)$ there is an $\varepsilon > 0$ such that if $g(y, c)$ is within ε of $y^2 - c$ in $C^2([-2.5, 2.5] \times [-2.5, 2.5])$, then:*

(i) *For $c \in [-2.5, -0.25 - \delta]$, all orbits of g escape from $[-2.5, 2.5]$.*
(ii) *For $c \in [-0.25 + \delta, 2 - \delta]$, the orbit of the critical point of g approaches an attractor in $[-2.5, 2.5]$.*
(iii) *For $c \in [2 + \delta, 2.5]$, the orbit of the critical point of g escapes from $[-2.5, 2.5]$.*
(iv) *For $c \in [-0.25 + \delta, 0.75 - \delta]$, g has a stable fixed point.*
(v) *For $c \in [0.75 + \delta, 2.5]$, g has no stable fixed point.*
(vi) *For $c \in [0.75 + \delta, 1.25 - \delta]$, g has a stable period 2 orbit.*

Proof (Sketch). For $c \in [-2.5, -0.25 - \delta]$, we have $y^2 - c \geq y + \delta$, so if g is within $\delta/2$ of $y^2 - c$, then (i) follows. For $c \geq -0.25 + \delta$, notice that $y^2 - c$ has a fixed point $y_0 = (1 + \sqrt{1 + 4c})/2$ at which it has derivative at least $1 + 2\sqrt{\delta}$. It follows that if $y > y_0$, then $y^2 - c > y$, while if y is between y_0 and its other preimage $y_1 = -y_0$, then $y^2 - c \leq y_0$. Thus if the image of the critical point, $-c$, is greater than y_1, then the interval $[y_1, y_0]$ is mapped into itself. On the other hand, if $-c < y_1$, then the second iterate of the critical point is greater than y_0, and the orbit of the critical point increases thereafter.

One can check that if $c \in [-0.25 + \delta, 2 - \delta]$, then $-c \geq y_1 + \delta/2$, while if $c \in [2 + \delta, 2.5]$, then $-c \leq y_1 - \delta/2$. Thus if g is sufficiently C^2 close to $y^2 - c$ and c is in one of the above two intervals, then g also has a fixed point y_0 at which it has derivative greater than 1, that fixed point also has a second preimage y_1, and the image of the critical point of g is greater than or less than y_1, depending on which of the two intervals c is in. This implies (ii) and (iii).

For $c \in [-0.25 + \delta, 0.75 - \delta]$, one can check that $y^2 - c$ has a stable fixed point at which its derivative has absolute value at most $1 - \delta$. Then if g is sufficiently C^2 close to $y^2 - c$, it will have a nearby fixed point at which $|\partial g/\partial y| < 1$, and (iv) is verified. A similar argument establishes (v) and (vi). □

An application of Proposition 1 is that if two bifurcation points of a period n window of the original map f can be determined, say the saddle-node and superstable parameter values, then other bifurcation points, such as the period-doubling and crisis values, can typically be predicted to a high degree of accuracy by linear extrapolation.

We have found that this observation can be quite valuable in efficiently computing the endpoints of small periodic windows, because a very close initial approximation to the crisis value may be needed in order for Newton's method to converge.

4. Two-parameter families

As we argued in Section 1, for a two-parameter family of maps with more than one critical point, the regions of period n stability in parameter space are organized around curves that have superstable period n orbits, and are thickest when two such curves cross to form a doubly superstable orbit. We now build on the results of the previous section to assert that near such an orbit, the dynamics are well approximated by a composition of quadratic maps $x \mapsto (x^2 - a)^2 - b$, again under a linear change of coordinates. As before we call the points of the periodic orbit $x_0, x_1, \ldots, x_{n-1}$, and we assume that both x_0 and x_m are critical points of the map for some m.

Principle 2. *Let X be a compact interval and let Λ be a compact set in the parameter plane. For a typical C^3 family of maps $f : X \times \Lambda \to X$ with a doubly superstable period n point $(x_0, \boldsymbol{\lambda}_0)$, there is a linear change of coordinates that conjugates f^n near $(x_0, \boldsymbol{\lambda}_0)$ to the family $y \mapsto (y^2-a)^2-b$ in the box $[-2.5, 2.5]\times[-2.5, 2.5]\times[-2.5, 2.5]$ to within an error ε in the C^2 norm, where $\varepsilon \to 0$ as n increases.*

Remark. The conditions for f and $(x_0, \boldsymbol{\lambda}_0)$ to be "typical" are similar to before, with one additional constraint: as well as not being unusually small, the parameter gradients of f^m at $(x_0, \boldsymbol{\lambda}_0)$ and f^{n-m} at $(x_m, \boldsymbol{\lambda}_0)$ must point in substantially different directions. Again we will be more specific about the conditions in the course of the following argument. And again, it suffices for f to be C^2, with $\partial^2 f/\partial x^2$ Lipschitz, instead of C^3.

Argument for Principle 2. Let

$$S_1 = \frac{\partial f^{m-1}}{\partial x}(x_1, \boldsymbol{\lambda}_0), \quad S_2 = \frac{\partial f^{n-m-1}}{\partial x}(x_{m+1}, \boldsymbol{\lambda}_0),$$

$$q_1 = \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(x_0, \boldsymbol{\lambda}_0), \quad q_2 = \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(x_m, \boldsymbol{\lambda}_0),$$

$$\boldsymbol{p}_1 = S_1^{-1}\nabla_\lambda f^m(x_0, \boldsymbol{\lambda}_0), \quad \boldsymbol{p}_2 = S_2^{-1}\nabla_\lambda f^{n-m}(x_m, \boldsymbol{\lambda}_0).$$

As in the previous section, S_1 and S_2 will be exponentially large as functions of m and $n - m$ respectively, while $\boldsymbol{p}_1, \boldsymbol{p}_2, q_1, q_2$ will typically stay away from zero as n increases. Furthermore, there is no reason to expect that the directions of $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$ are correlated in any way, so for a typical orbit the angle between these two vectors will not be unreasonably small.

The linear change of coordinates this time is

$$y = (q_1 S_1)^{2/3}(q_2 S_2)^{1/3}(x - x_0), \quad z = (q_1 S_1)^{1/3}(q_2 S_2)^{2/3}(x - x_m),$$

$$a = -q_1^{1/3} S_1^{4/3}(q_2 S_2)^{2/3}\boldsymbol{p}_1 \cdot (\boldsymbol{\lambda} - \boldsymbol{\lambda}_0), \quad b = -(q_1 S_1)^{2/3} q_2^{1/3} S_2^{4/3}\boldsymbol{p}_2 \cdot (\boldsymbol{\lambda} - \boldsymbol{\lambda}_0).$$

Let g and h be the maps conjugate to f^m and f^{n-m} under this change of coordinates. One can check that

$$\frac{\partial g}{\partial y}(0,0) = 0, \quad \frac{\partial^2 g}{\partial y^2}(0,0) = 2, \quad \frac{\partial g}{\partial a}(0,0) = -1, \quad \frac{\partial g}{\partial b}(0,0) = 0,$$

$$\frac{\partial h}{\partial z}(0,0) = 0, \quad \frac{\partial^2 h}{\partial z^2}(0,0) = 2, \quad \frac{\partial h}{\partial a}(0,0) = 0, \quad \frac{\partial h}{\partial b}(0,0) = -1,$$

and thus taking into account lowest order terms,

$$h \circ g(y, a, b) \approx (y^2 - a)^2 - b.$$

We must again show that typically the higher order terms are small when n is large.

Let us make the preliminary parameter change of coordinates

$$\mu_1 = \boldsymbol{p}_1 \cdot (\boldsymbol{\lambda} - \boldsymbol{\lambda}_0), \qquad \mu_2 = \boldsymbol{p}_2 \cdot (\boldsymbol{\lambda} - \boldsymbol{\lambda}_0).$$

Let $q = \min(|q_1|, |q_2|)$, and let M bound the partial derivatives of f in the new parameter coordinates as follows:

$$\left| \frac{\partial^{i+j+k} f}{\partial x^i \partial \mu_1^j \partial \mu_2^k} \right| \leq M q^{i+j+k-1}, \quad 1 \leq i + j + k \leq 2.$$

Notice that M will not be inordinately large unless $\boldsymbol{p}_1$, $\boldsymbol{p}_2$, or q is unusually small or the directions of $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$ are unusually close.

Let R be the rectangle (in $\boldsymbol{\mu}$ coordinates) that is mapped to the square $(a, b) \in [-2.5, 2.5] \times [-2.5, 2.5]$ under the rescaling of parameter coordinates, i.e.

$$R = \left\{ \boldsymbol{\mu} : |\mu_1| \leq \frac{2.5}{|q_1|^{1/3}|S_1|^{4/3}|q_2 S_2|^{2/3}}, |\mu_2| \leq \frac{2.5}{|q_1 S_1|^{2/3}|q_2|^{1/3}|S_2|^{4/3}} \right\}.$$

Notice that in $\boldsymbol{\lambda}$ coordinates, R is a parallelogram centered at $\boldsymbol{\lambda}_0$. Let

$$R_1 = \left\{ (x, \boldsymbol{\mu}) : |x - x_0| \leq \frac{2.5}{|q_1 S_1|^{2/3}|q_2 S_2|^{1/3}}, \boldsymbol{\mu} \in R \right\}$$

and

$$R_2 = \left\{ (x, \boldsymbol{\mu}) : |x - x_m| \leq \frac{2.5}{|q_1 S_1|^{1/3}|q_2 S_2|^{2/3}}, \boldsymbol{\mu} \in R \right\}.$$

Let $\rho, \sigma \geq 1$ be constants such that for all $(x, \boldsymbol{\mu}) \in R_1$ and $1 \leq j < k \leq m$ or $(x, \boldsymbol{\mu}) \in R_2$ and $m+1 \leq j < k \leq n$,

$$\prod_{i=j}^{k-1} \sup_{(x,\boldsymbol{\mu})} \left| \frac{\partial f}{\partial x}(f^i(x, \boldsymbol{\mu}), \boldsymbol{\mu}) \right| \leq \rho |S|, \qquad \left| \frac{\partial f}{\partial x}(f^j(x, \boldsymbol{\mu}), \boldsymbol{\mu}) \right|^{-1} \leq \sigma,$$

where $S = S_1$ if $j < m$ and $S = S_2$ if $j > m$. The definition of ρ is stronger than in the previous section, but is related inasmuch as the right-hand side of the inequality still reflects a maximal stretching factor from the jth iterate of f to the kth iterate. One can argue as in the previous section that ρ and σ are typically at most polynomial in n.

Finally, let τ be a constant such that for all $(x, \boldsymbol{\mu}) \in R_1$ and $1 \leq j \leq m-1$ or $(x, \boldsymbol{\mu}) \in R_2$ and $m+1 \leq j \leq n-1$,

$$|\boldsymbol{\nabla}_\mu f^{\ell - j}(x_j, 0)| \leq (\tau - 1) \left| \frac{\partial f^{l-j}}{\partial x}(x_j, 0) \right|$$

(where $l = m$ if $j < m$ and $l = n$ if $j > m$). Since

$$\boldsymbol{\nabla}_\mu f^{l-j}(x_j, 0) = \sum_{i=j+1}^{l} \frac{\partial f^{l-i}}{\partial x}(x_i, 0) \boldsymbol{\nabla}_\mu f(x_{i-1}, 0),$$

and the terms on the right-hand side will typically be of the same order of magnitude as $\partial f^{l-j}/\partial x(x_j, 0)$ for i near j and decay exponentially as i increases, we conclude that τ typically need not be large. Principle 2 then follows from the following lemma, which is proved in Appendix A.

Lemma 2. *Assume* $|S_1|, |S_2| \geq 1$. *The distance* ε *between* $g(y, a, b)$ *and* $y^2 - a$ *in the* C^2 *norm for* $(y, a, b) \in [-2.5, 2.5] \times [-2.5, 2.5] \times [-2.5, 2.5]$ *can be bounded as follows:*

$$\varepsilon \leq \tfrac{25}{2}(m+5)^5 \rho^2 \sigma^3 \tau^2 M^5 (|S_1|^{-2/3}|S_2|^{-1/3} + |S_2|^{-1}).$$

The C^2 *distance between* $h(z, a, b)$ *and* $z^2 - b$ *can of course be bounded similarly* (*with* m *replaced by* $n - m$ *and* S_1 *and* S_2 *exchanged*). □

Since the parameter region R is a parallelogram whose sides may have very different lengths, Lemma 2 turns out not to be a simple extension of Lemma 1. The stretching factors S_1 and S_2, which are like exponential functions of m and $n - m$, respectively, will be of different orders of magnitude unless m and $n - m$ happen to be close. Some effort is required to avoid error terms like $|S_1||S_2|^{-2}$, which need not be small. Lemma 2 implies that the C^2 errors associated with both g and h are small as long as both S_1 and S_2 are large.

For a given two-parameter bifurcation of the canonical family $y \mapsto (y^2 - a)^2 - b$, as in Proposition 1 it is possible to show that when $g \circ h$ is close to the canonical family it has a corresponding bifurcation near the same parameter pair. Thus for the unscaled period n window corresponding to f, finding the parameter values for three noncollinear bifurcation points determines a linear transformation to the corresponding points in the canonical bifurcation diagram, and allows other features of the window to be well approximated by linear extrapolation.

Finally, we comment that we are doubtful whether the results of this paper extend to the case of maps with three or more critical points and three or more parameters. Applying the ideas in this section and Appendix A to the case of three parameters, with stretching factors S_1, S_2, and S_3, we encounter error terms that involve, for example, negative powers of S_1 and S_2 but a positive power of S_3. Though a substantial proportion of windows will have the three stretching factors close enough to each other to imply a small error, we anticipate that a substantial proportion of windows will have, say, S_3 enough larger than S_1 and S_2 so as to prevent a close linear correspondence with the canonical map $x \mapsto ((x^2 - a)^2 - b)^2 - c$.

Acknowledgements

We thank J.-P. Eckmann, J. Milnor, and J. Ringland for helpful discussions, and the referees for several excellent suggestions. This research was supported by the National Science Foundation (Divisions of Mathematical and Physical Sciences) and by the US Department of Energy (Offices of Scientific Computing and Energy Research).

Appendix A

Here we offer proofs of the central rigorous results of this paper, Lemmas 1 and 2 (stated in Sections 3 and 4, respectively). For the maps f and g we will use subscripts to denote partial differentiation and superscripts to denote iteration; if both a superscript and subscripts are present then iteration is to be done before differentiation (for instance, f_x^i means the partial derivative with respect to x of the ith iterate of f).

Proof of Lemma 1. As we mentioned in Section 3, the values of g, g_y, g_c, and g_{yy} are equal to those for $y^2 - c$ at the origin. We first show that $g_{yy} - 2$, g_{yc}, and g_{cc} are small on all of R; then it will follow by integration that $g_y - 2y$, $g_c + 1$, and $g - (y^2 - c)$ are also small on R. Thus our main task is to obtain reasonable upper bounds on the second order partial derivatives of f^n. In the estimates below we will frequently make use of the fact that $M \geq 2$ and $\rho, \sigma, |S| \geq 1$ (true by definition or, in the case of $|S|$, by hypothesis) to bound lower order terms by higher order terms.

We begin with some estimates which will be useful in the following paragraphs; all estimates apply to $(x, \lambda) \in R$. Since f_x is zero at the center of R,

$$|f_x| \le \frac{2.5}{|qS|}M|q| + \frac{2.5}{|pqS^2|}M|pq| \le \frac{5M}{|S|},$$

and hence for $1 \le k \le n$ we have

$$|f_x^k| = |f_x^{k-1} \circ f||f_x| \le \rho|S|\frac{5M}{|S|} = 5\rho M.$$

Also,

$$|f_\lambda^k| = \left|\sum_{i=1}^{k}(f_x^{k-i} \circ f^i) f_\lambda \circ f^{i-1}\right| \le k\rho M|pS|.$$

Next,

$$(f_x^{n-1} \circ f)_x = (f_x^{n-1} \circ f)\sum_{i=1}^{n-1}\frac{f_{xx} \circ f^i}{f_x \circ f^i} f_x^i,$$

and thus

$$|(f_x^{n-1} \circ f)_x| \le \rho|S|(n-1)M|q|\sigma 5\rho M = 5(n-1)\rho^2\sigma M^2|qS|.$$

Also for $1 \le k \le n$,

$$(f_x^{n-k} \circ f^k)_\lambda = (f_x^{n-k} \circ f^k)\sum_{i=k}^{n-1}\frac{1}{f_x \circ f^i}(f_{x\lambda} \circ f^i + f_{xx} \circ f^i f_\lambda^i),$$

so

$$|(f_x^{n-k} \circ f^k)_\lambda| \le \rho|S|\sum_{i=k}^{n-1}\sigma(M|pq| + M|q|i\rho M|pS|) \le \rho^2\sigma M^2|pqS^2|\sum_{i=k}^{n-1}\left(\frac{1}{2} + i\right)$$
$$= \frac{n^2-k^2}{2}\rho^2\sigma M^2|pqS^2|.$$

Now $f_x^n = (f_x^{n-1} \circ f) f_x$, so

$$f_{xx}^n = (f_x^{n-1} \circ f)_x f_x + (f_x^{n-1} \circ f) f_{xx},$$

and hence

$$|f_{xx}^n - 2qS| \le |(f_x^{n-1} \circ f)_x f_x| + |(f_x^{n-1} \circ f) f_{xx} - 2qS|.$$

Since $(f_x^{n-1} \circ f) = S$ at the center of R,

$$|(f_x^{n-1} \circ f) - S| \le 5(n-1)\rho^2\sigma M^2|qS|\frac{2.5}{|qS|} + \frac{n^2-1}{2}\rho^2\sigma M^2|pqS^2|\frac{2.5}{|pqS^2|} \le \frac{5(n^2+10n)}{4}\rho^2\sigma M^2.$$

Likewise,

$$|f_{xx} - 2q| \le M|q^2|\frac{2.5}{|qS|} + M|pq^2|\frac{2.5}{|pqS^2|} \le \frac{5M|q|}{|S|}.$$

Thus

$$|f^n_{xx} - 2qS| \le 5(n-1)\rho^2\sigma M^2|qS|\frac{5M}{|S|} + M|q|\frac{5(n^2+10n)}{4}\rho^2\sigma M^2 + \rho|S|\frac{5M|q|}{|S|}$$
$$\le \frac{5(n^2+30n)}{4}\rho^2\sigma M^3|q|.$$

Next,

$$f^n_{x\lambda} = (f^{n-1}_x \circ f)_\lambda f_x + (f^{n-1}_x \circ f) f_{x\lambda},$$

so

$$|f^n_{x\lambda}| \le \frac{n^2-1}{2}\rho^2\sigma M^2|pqS^2|\frac{5M}{|S|} + \rho|S|M|pq| \le \frac{5n^2}{2}\rho^2M^3|pqS|.$$

Finally,

$$f^n_\lambda = \sum_{i=1}^{n}(f^{n-i}_x \circ f^i) f_\lambda \circ f^{i-1},$$

so

$$f^n_{\lambda\lambda} = \sum_{i=1}^{n}((f^{n-i}_x \circ f^i)_\lambda f_\lambda \circ f^{i-1} + (f^{n-i}_x \circ f^i)(f_{\lambda\lambda} \circ f^{i-1} + (f_{x\lambda} \circ f^{i-1}) f^{i-1}_\lambda)),$$

and thus

$$|f^n_{\lambda\lambda}| \le \sum_{i=1}^{n}\left(\frac{n^2-i^2}{2}\rho^2\sigma M^2|pqS^2|M|p| + \rho|S|M|p^2q| + M|pq|(i-1)\rho M|pS|\right) \le \frac{n^3}{3}\rho^2\sigma M^3|p^2qS^2|.$$

It follows that

$$|g_{yy} - 2| \le \frac{5(n^2+30n)}{4}\rho^2\sigma M^3|S|^{-1},$$

$$|g_{yc}| \le \frac{5n^2}{2}\rho^2\sigma M^3|S|^{-1},$$

and

$$|g_{cc}| \le \frac{n^3}{3}\rho^2\sigma M^3|S|^{-1}.$$

Therefore, by integration we have that the C^2 distance ε between g and $y^2 - c$ satisfies

$$\varepsilon = \sup_R|g - (y^2-c)| + \sup_R|g_y - 2y| + \sup_R|g_c + 1| + \sup_R|g_{yy} - 2| + 2\sup_R|g_{yc}| + \sup_R|g_{cc}|$$
$$\le (2.5^2 + 2.5 + 1)\left(\sup_R|g_{yy} - 2| + 2\sup_R|g_{yc}| + \sup_R|g_{cc}|\right) \le 10\frac{(n+7)^3}{3}\rho^2\sigma M^3|S|^{-1}$$

as claimed. □

Proof of Lemma 2. The estimates below will be far from sharp with regard to powers of M, ρ, σ, and τ and other terms which are polynomial in the period n. The emphasis will be on getting the right powers of $|S_1|$ and $|S_2|$ in

the bounds. To keep the formulas concise, we use the notation $S = \max(|S_1|, |S_2|)$, so that expressions such as $|S_1| + |S_2|$ will be bounded by $2S$. Also, for $1 \le k \le m$, let

$$T_k = \max_{1\le j\le k} \prod_{i=j}^{k-1} \max_{(x,\mu)\in R_1} |f_x \circ f^i(x,\mu)|.$$

By the product from k to $k-1$, we mean simply the quantity 1; in particular, $T_1 = 1$. Notice that $T_k \le \rho|S_1|$ for all k.

We proceed as in the proof of Lemma 1 with some preliminary estimates which apply to $(x, \boldsymbol{\mu}) \in R_1$. When the derivation is very similar to the corresponding step in the previous proof, we omit the intermediate inequalities. First, we find that

$$|f_x| \le 7.5M|S_1|^{-2/3}|S_2|^{-1/3},$$

and for $1 \le k \le m$, it follows that

$$|f_x^k| \le 7.5M|S_1|^{-2/3}|S_2|^{-1/3}T_k.$$

Also,

$$|f_{\mu_1}^k| = \left|\sum_{j=1}^{k} (f_x^{k-j} \circ f^j) f_{\mu_1} \circ f^{j-1}\right| \le kMT_k,$$

and likewise for $|f_{\mu_2}^k|$.

The above bounds imply that

$$\begin{aligned} |f^k(x,\mu) - x_k| &\le 2.5(k+7.5)M|S_1|^{-4/3}|S_2|^{-2/3}T_k + 2.5kM|S_1|^{-2/3}|S_2|^{-4/3}T_k \\ &\le 5(k+4)MS^{2/3}|S_1|^{-4/3}|S_2|^{-4/3}T_k \end{aligned}$$

again for $1 \le k \le m$. Since T_k is of the same order of magnitude as $|S_1|$ for k near m, this estimate does not always imply that the trajectory of (x,μ) stays close to the trajectory of $(x_0, 0)$ – only when $S_1 \ll S_2^4$ is the right-hand side small for all k. The estimate does, however, allow us to obtain an improved bound on $|f_{\mu_2}^k|$ which will in turn imply a better estimate on $|f^k(x,\mu) - x_k|$.

First, observe that since

$$f_{\mu_2}^{m-k}(x_0,0) + f_x^{m-k} \circ f^k(x_0,0) f_{\mu_2}^k(x_0,0) = f_{\mu_2}^m(x_0,0) = 0, \qquad |f_{\mu_2}^k(x_0,0)| \le \tau - 1$$

for $1 \le k \le m-1$. Next, notice that

$$f_{\mu_2}^{k+1} = f_{\mu_2} \circ f^k + (f_x \circ f^k) f_{\mu_2}^k.$$

Using the notation $\Delta\varphi$ to denote $|\varphi(x,\mu) - \varphi(x_0,0)|$, we then have

$$\begin{aligned} \Delta f_{\mu_2}^{k+1} &\le \Delta f_{\mu_2} \circ f^k + (\Delta f_x \circ f^k)|f_{\mu_2}^k(x_0,0)| + |f_x \circ f^k(x,\lambda)|\Delta f_{\mu_2}^k \\ &\le \tau M 5(k+4)MS^{2/3}|S_1|^{-4/3}|S_2|^{-4/3}T_k + |f_x \circ f^k(x,\lambda)|\Delta f_{\mu_2}^k \\ &\le 5(k+4)\sigma\tau M^2 S^{2/3}|S_1|^{-4/3}|S_2|^{-4/3}T_{k+1} + |f_x \circ f^k(x,\lambda)|\Delta f_{\mu_2}^k. \end{aligned}$$

Now as in the estimate for $|f_x|$,

$$\Delta f_{\mu_2} \le 7.5M|S_1|^{-2/3}|S_2|^{-1/3},$$

and thus it follows by induction that

$$\Delta f_{\mu_2}^k \le \left(7.5M|S_1|^{-2/3}|S_2|^{-1/3} + \sum_{j=1}^{k-1} 5(j+4)\sigma\tau M^2 S^{2/3}|S_1|^{-4/3}|S_2|^{-4/3}\right) T_k$$
$$\le \frac{5(k^2+7k-5)}{2}\sigma\tau M^2|S_1|^{-2/3}|S_2|^{-1/3}T_k$$

for $1 \le k \le m$. Combining the estimates on $|f_{\mu_2}^k(x_0,0)|$ and $\Delta f_{\mu_2}^k$ and then bounding T_k by $\rho|S_1|$, we have

$$|f_{\mu_2}^k| \le \frac{5(k^2+7k-5)}{2}\sigma\tau M^2|S_1|^{-2/3}|S_2|^{-1/3}T_k + \tau - 1 \le \frac{5(k^2+7k)}{2}\rho\sigma\tau M^2 S^{1/3}|S_2|^{-1/3}.$$

Notice that we can now make the improved estimate

$$|f^k(x,\mu) - x_m| \le \frac{25(k+4)^2}{4}\rho\sigma\tau M^2|S_1|^{-1/3}|S_2|^{-2/3}.$$

Thus the trajectories of initial points in R_1 do in fact stay close to the trajectory of $(x_0, 0)$ for m iterates, provided $|S_1|$ or $|S_2|$ is sufficiently large.

Now we can proceed again like in the proof of Lemma 1. First,

$$|(f_x^{m-1} \circ f)_x| \le 7.5(m-1)\rho^2\sigma M^2 q|S_1|^{4/3}|S_2|^{-1/3},$$

and for $1 \le k \le m$,

$$|(f_x^{m-k} \circ f^k)_{\mu_1}| \le \frac{m^2-k^2}{2}\rho^2\sigma M^2 q|S_1^2|.$$

Similarly,

$$|(f_x^{m-k} \circ f^k)_{\mu_2}| \le \frac{5(m^3+9m)}{6}\rho^2\sigma^2\tau M^3 q S^{1/3}|S_1||S_2|^{-1/3}.$$

Next,

$$|(f_x^{m-1} \circ f) - S_1| \le \frac{25(m+4)^3}{12}\rho^2\sigma^2\tau M^3|S_1|^{2/3}|S_2|^{-2/3},$$

and

$$|f_{xx} - 2q_1| \le 7.5Mq|S_1|^{-2/3}|S_2|^{-1/3},$$

so

$$|f_{xx}^m - 2q_1 S_1| \le \frac{25(m+9)^3}{12}\rho^2\sigma^2\tau M^4 q S^{1/3}|S_1|^{1/3}|S_2|^{-2/3}.$$

Also,

$$|f_{x\mu_1}^m| \le \frac{15m^2}{4}\rho^2\sigma M^3 q S^{1/3}|S_1||S_2|^{-1/3},$$

and

$$|f_{x\mu_2}^m| \le \frac{25(m+3)^3}{12}\rho^2\sigma^2\tau M^4 q|S_1|.$$

Finally,

$$|f^m_{\mu_1\mu_1}| \le \frac{m^3}{3}\rho^2\sigma M^3 q|S_1|^2,$$

and

$$|f^m_{\mu_1\mu_2}| \le \frac{5(m+3)^4}{6}\rho^2\sigma^2\tau M^4 q S^{1/3}|S_1||S_2|^{-1/3}.$$

The above estimates are based on the expansion of $f^n_{\lambda\lambda}$ in the proof of Lemma 1; to estimate $|f^m_{\mu_2\mu_2}|$ we use a different expansion:

$$f^m_{\mu_2\mu_2} = \sum_{i=1}^{m}((f_x^{m-i}\circ f^i)(f_{\mu_2\mu_2}\circ f^{i-1} + 2(f_{x\mu_2}\circ f^{i-1})f^{i-1}_{\mu_2}) + (f_{xx}\circ f^{i-1})(f^{i-1}_{\mu_2})^2).$$

We also use our previous bound on $|f^k_{\mu_2}|$ involving T_k, together with the fact that

$$|f_x^{m-i}\circ f^i|T_{i-1} \le \rho\sigma|S_1|,$$

to obtain

$$\begin{aligned}|f^m_{\mu_2\mu_2}| &\le \sum_{i=1}^{m}|f_x^{m-i}\circ f^i|Mq\left(\frac{5(i^2+5i-11)}{2}\sigma\tau M^2|S_1|^{-2/3}|S_2|^{-1/3}T_{i-1}+\tau\right)^2\\ &\le \frac{5(m^3+9m^2)}{6}\rho^2\sigma^3\tau^2M^5q|S_1|.\end{aligned}$$

From the above inequalities and the change of coordinates described in Section 4, we find that

$$|g_{yy}-2| \le \frac{25(m+9)^3}{12}\rho^2\sigma^2\tau M^4 S^{1/3}|S_1|^{-2/3}|S_2|^{-2/3},$$

$$|g_{ya}| \le \frac{15m^2}{4}\rho^2\sigma M^3 S^{1/3}|S_1|^{-2/3}|S_2|^{-2/3},$$

$$|g_{yb}| \le \frac{25(m+3)^3}{12}\rho^2\sigma^2\tau M^4|S_2|^{-1},$$

$$|g_{aa}| \le \frac{m^3}{3}\rho^2\sigma M^3|S_1|^{-1/3}|S_2|^{-2/3},$$

$$|g_{ab}| \le \frac{5(m+3)^4}{6}\rho^2\sigma^2\tau M^4 S^{1/3}|S_1|^{-2/3}|S_2|^{-5/3},$$

and

$$|g_{bb}| \le \frac{5(m+3)^5}{4}\rho^2\sigma^3\tau^2M^5|S_2|^{-2}.$$

Once again, by integration it follows that the C^2 distance ε between g and y^2-a is bounded as claimed:

$$\varepsilon \le \frac{25}{2}(m+5)^5\rho^2\sigma^3\tau^2M^5S^{2/3}|S_1|^{-2/3}|S_2|^{-1}.$$

□

References

[1] A. Douady, J. Hubbard, Étude dynamique des polynômes complexes, Partie II, Publ. Math. d'Orsay 85 (4) (1985).
[2] J.-P. Eckmann, H. Epstein, Scaling of Mandelbrot sets generated by critical point preperiodicity, Comm. Math. Phys. 101 (1985) 283–289.
[3] M.J. Feigenbaum, Quantitative universality for a class of nonlinear transformations, J. Stat. Phys. 19 (1978) 25–52.
[4] M.J. Feigenbaum, The universal metric properties of nonlinear transformations, J. Stat. Phys. 21 (1979) 669–706.
[5] J.A.C. Gallas, The role of parameters in dynamical systems, Int. J. Modern Phys. C 3 (1992) 1295–1321 .
[6] J.A.C. Gallas, Dissecting shrimps: results for some one-dimensional physical models, Physica A 202 (1994) 196–223.
[7] J.A.C. Gallas, Structure of the parameter space of a ring cavity, Appl. Phys. B 60 (1995) S203–S213.
[8] J. Guckenheimer, Renormalization of one dimensional mappings, Contemp. Math. 58, pt. III (1987) 143–160.
[9] J. Hale, H. Koçak, Dynamics and Bifurcations, Springer, New York, 1991.
[10] H. Hurwitz Jr., M. Frame, D. Peak, Scaling symmetries in nonlinear dynamics: a view from parameter space, Physica D 81 (1995) 23.
[11] A.P. Kuznetov, S.P. Kuztenov, I.R. Sataev, L.O. Chua, Two-parameter study of transition of chaos in Chua's circuit: renormalization group, universality, and scaling, Int. J. Bifure Chaos 3 (1993) 943–962.
[12] O. Lanford, A computer-assisted proof of the Feigenbaum conjecture, Bull. Amer. Math. Soc. 6 (1982) 427–434.
[13] J. Milnor, Remarks on iterated cubic maps, Exp. Math. 1 (1992) 5–24.
[14] J. Milnor, Hyperbolic components in spaces of polynomial maps, SUNY Stony Brook IMS Preprint 1992/3.
[15] H.E. Nusse, L. Tedeschini-Lalli, Wild hyperbolic sets, yet no chance for the coexistence of infinitely many KLUS-simple Newhouse attracting sets, Comm. Math. Phys 144 (1992) 429–442.
[16] J. Ringland, N. Issa, M. Schell, From U sequence to Farey sequence: a unification of one-parameter scenarios, Phys. Rev. A 41 (1990) 4223–4235.
[17] J. Ringland, M. Schell, Universal geometry in the parameter plane of maps of the interval, Europhys. Lett. 12 (1990) 595–601.
[18] J. Ringland, M. Schell, Universal geometry in the parameter plane of maps of the interval, Europhys. Lett. 26 (1994) 637–638.
[19] L. Tedeschini-Lalli, J.A. Yorke, How often do simple dynamical processes have infinitely many coexisting sinks?, Comm. Math. Phys. 106 (1986) 635–657.
[20] J.A. Yorke, C. Grebogi, E. Ott, L. Tedeschini-Lalli, Scaling behavior of windows in dissipative dynamical systems, Phys. Rev. Lett. 54 (1985) 1095–1098.