Journal of Statistical Physics, Vol. 19, No. 1, 1978

Quantitative Universality for a Class of Nonlinear Transformations

Mitchell J. Feigenbaum[1]

Received October 31, 1977

A large class of recursion relations $x_{n+1} = \lambda f(x_n)$ exhibiting infinite bifurcation is shown to possess a rich quantitative structure essentially independent of the recursion function. The functions considered all have a unique differentiable maximum $\bar{x}$. With $f(\bar{x}) - f(x) \sim |x - \bar{x}|^z$ (for $|x - \bar{x}|$ sufficiently small), $z > 1$, the universal details depend only upon z. In particular, the local structure of high-order stability sets is shown to approach universality, rescaling in successive bifurcations, asymptotically by the ratio α ($\alpha = 2.5029078750957...$ for $z = 2$). This structure is determined by a universal function $g^*(x)$, where the 2^nth iterate of f, $f^{(n)}$, converges locally to $\alpha^{-n}g^*(\alpha^n x)$ for large n. For the class of f's considered, there exists a λ_n such that a 2^n-point stable limit cycle including $\bar{x}$ exists; $\lambda_\infty - \lambda_n \sim \delta^{-n}$ ($\delta = 4.669201609103...$ for $z = 2$). The numbers α and δ have been computationally determined for a range of z through their definitions, for a variety of f's for each z. We present a recursive mechanism that explains these results by determining g^* as the fixed-point (function) of a transformation on the class of f's. At present our treatment is heuristic. In a sequel, an exact theory is formulated and specific problems of rigor isolated.

KEY WORDS: Recurrence; bifurcation; limit cycles; attractor; universality; scaling; population dynamics.

1. INTRODUCTION

Recursion equations $x_{n+1} = f(x_n)$ provide a description for a variety of problems. For example, a numerical computation of a zero of $h(x)$ is obtained recursively according to

$$x_{n+1} = x_n + \frac{\epsilon h(x_n)}{h(x_n - \epsilon) - h(x_n)} \equiv f(x_n)$$

Research performed under the auspices of the U.S. Energy Research and Development Administration.

[1] Theoretical Division, Los Alamos Scientific Laboratory, Los Alamos, New Mexico.

If $\bar{x} = \lim_{n\to\infty} x_n$ exists, then $h(\bar{x}) = 0$. As $\bar{x}$ satisfies

$$\bar{x} = f(\bar{x})$$

the desired zero of h is obtained as the "fixed point" of the transformation f. In a natural context, a (possibly fictitious) discrete population satisfies the formula $p_{n+1} = f(p_n)$, determining the population at one time in terms of its previous value. We mention these two examples purely for illustrative purposes. The results of this paper, of course, apply to any situation modeled by such a recursion equation. Nevertheless, we shall focus attention throughout this section on the population example, both for the intuitive appeal of so tangible a realization as well as for a definite viewpoint, rather different from the usual one toward this situation, that shall emerge in the discussion. It is to be emphasized, though, that our results are generally applicable.

If the population referred to is that of a dilute group of organisms, then

$$p_{n+1} = bp_n \tag{1}$$

accurately describes the population growth so long as it remains dilute, with the solution $p_n = p_0 b^n$. For a given species of organism in a fixed milieu, b is a constant—the static birth rate for the configuration. As the population grows, the dilute approximation will ultimately fail: sufficient organisms are present and mutually interfere (e.g., competition for nutrient supply). At this point, the next value of the population will be determined by a *dynamic* or effective birth rate:

$$p_{n+1} = b_{\text{eff}} p_n$$

with $b_{\text{eff}} < b$. Clearly b_{eff} is a function of p, with

$$\lim_{p\to 0} b_{\text{eff}}(p) = b$$

the only model-independent quantitative feature of b_{eff}. Since the volume and nutrient available to a population are limited, it is clear that $b_{\text{eff}} \simeq 0$ for p sufficiently large. Accordingly, the simplest form of $b_{\text{eff}}(p)$ to reproduce the qualitative dynamics of such a population should resemble Fig. 1, where $b_{\text{eff}}(0) = b$ is an adjustable parameter [say, the nutrient level of the milieu held fixed independent of $p(t)$, and measurable by observing very dilute populations in that milieu]. A simple specific form of b_{eff} is

$$b_{\text{eff}} = b - ap$$

so that

$$p_{n+1} = bp_n - ap_n^2$$

By defining $p_n \equiv (b/a)x_n$, we obtain the standard form

$$x_{n+1} = bx_n(1 - x_n) \tag{2}$$

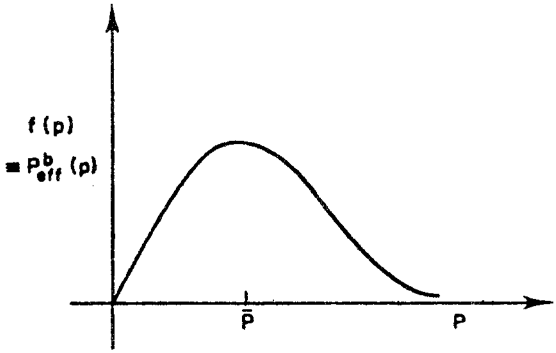

Fig. 1

In (2) the adjustable parameter b is purely multiplicative. With a different choice of b_{eff}, x_{n+1} would not in general depend upon b in so simple a fashion. Nevertheless, the internal b dependence may be (and often is) sufficiently mild in comparison to the multiplicative dependence that at least for qualitative purposes the internal dependence can be ignored. Thus, with $f(p) = pb_{\text{eff}}(p)$ any function like Fig. 1,

$$p_{n+1} = bf(p_n) \tag{3}$$

is compatible and representative of the population discussed. So long as $f'(0) = 1$ (so that the static birth rate is b and the dilute regime is correctly modeled) and f goes to zero for large p with a single central maximum, relation (3) correctly (at least qualitatively) models the situation. However, $f_2(p) = \sin(ap)$ affords an (a priori) equally good modeling as $f_1(p) = p - ap^2$. Thus only detailed quantitative results of (3) could determine which (if either) is empirically correct. One should then ask what the dynamical behavior of (3) is with f as in Fig. 1. It turns out that (3) enjoys a rich spectrum of excitations, with a universal behavior that would frustrate any attempt to discriminate among possible f's qualitatively. That is, providing (3) affords an honest model of a population's dynamics, so far as qualitative aspects are concerned, f is sufficiently specified by Fig. 1: the data could not qualitatively determine any more specific form [such as (2), say]. Conversely, *any* such choice of f—say Eq. (2)—is fully sufficient for study to comprehend all qualitative aspects of the dynamics. If the data should in any way disagree qualitatively with the predictions of (2), then (3) for any believable f must be an incorrect model.

The qualitative information available pertaining to (3) for any f of the form considered (see Appendix A for the exact requirements on f) is quite specific and detailed. In discussing the numerical solution to $h(x) = 0$ a fixed point was considered. In a population context, a fixed point

$$p^* = bf(p^*)$$

signifies zero population growth: $p_n = p^*$ for all n. However, p^* is "interesting" only so long as it is stable: if p fluctuates away from p^*, it should return to p^* in successive generations. For example, if $g(\bar{x})$ is finite, then

$$x_{n+1} = x_n + h(x_n)g(x_n) \tag{4}$$

will possess $\bar{x}$ as a fixed point if $h(\bar{x}) = 0$. However, unless

$$x_n \to \bar{x}$$

(4) is of no value to obtain $\bar{x}$; indeed, g is *chosen* so as to maximize the stability of $\bar{x}$. A stable fixed point is termed an "attractor," since points in its neighborhood approach it when iterated. An attractor is "global" if almost all points are eventually attracted to it. It is not necessary that an attractor be a unique isolated point. Thus, there might be n points $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n$ such that

$$\bar{x}_{i+1} = f(\bar{x}_i), \quad i = 1, \ldots, n-1; \qquad \bar{x}_1 = f(\bar{x}_n)$$

Such a set is called an "n-point limit cycle." Every n applications of f return an $\bar{x}_i$ to itself: each $\bar{x}_i$ is a fixed point of the nth iterate of f, $f^{(n)}$:

$$f^{(n)}(\bar{x}_i) = \bar{x}_i, \qquad i = 1, \ldots, n$$

Accordingly, $\{\bar{x}_1, \ldots, \bar{x}_n\}$ is a stable n-point limit cycle if each $\bar{x}_i$ is a stable fixed point of $f^{(n)}$. If it is a global attractor, then for almost every x_0, the sequence $x_n = f^{(n)}(x_0)$, $n = 1, 2, \ldots$, approaches the sequence

$$\bar{x}_1, \ldots, \bar{x}_n, \bar{x}_1, \ldots, \bar{x}_n, \ldots$$

Finally, there can be infinite stability sets $\{\bar{x}_i\}$ with

$$\bar{x}_{i+1} = f(\bar{x}_i)$$

such that the sequence $x_n = f^{(n)}(x_0)$ eventually becomes the sequence $\{\bar{x}_i\}$.

With this terminology, some of the detailed qualitative features of (3) can be stated as follows. (See Appendix A for more precise statements.) Depending upon the parameter value b, (3) possesses stable attractors of every order, with one attractor present and global for each fixed choice of b. As b is increased from a sufficiently small positive value, a fixed point $p^* > 0$ is stable until a value B_0 is reached when it becomes unstable. As b increases above B_0, a two-point cycle is stable, until at B_1 it becomes unstable, giving rise to a stable four-point cycle. As b is increased, this phenomenon recurs, with a 2^n-point cycle stable for

$$B_{n-1} < b < B_n$$

giving rise to a 2^{n+1}-point cycle above B_n until B_{n+1}, etc. The sequence of B_n is *bounded above* converging to a finite B_∞. This set of cycles (of order 2^n, $n = 1, 2, \ldots$) is termed the set of "harmonics" of the two-point cycle. For any value of $b > B_\infty$ (but not too large) some particular stable n-point cycle will be present. As b is increased, it becomes unstable, and is replaced with a stable $2n$-point cycle. Until the cycle has doubled ad infinitum, no new stability sets save for these appear. Moreover, the ordering (with respect to b) of the onset of new size stability sets (e.g., seven-point before five-point) is also independent of f. Thus, if b is the unique parameter governing a population, any deviation of the ordering of stability sets upon increase of b from that determined by (2), say, constitutes empirical proof that (3) for any believable f incorrectly models the population. On the other hand, if (3) is appropriate for some f, then (2), for all qualitative purposes, comprises the full theory of the population's evolution. The exact quantitative theory reduces to the problem of determining the particular f. Unfortunately, even if (3) might be applicable, the data of biological populations are too crude at present to significantly discriminate among f's.

With so much specific qualitative information about (3) independent of f available, we may ask if the form of (3) might not also imply some *quantitative* information independent of f. It is the content of the following to answer this inquiry in the affirmative. Thus, the local structure of high-order stability sets (the quantitative locations of all elements of a stability set nearby one another) is independent of f. The role of a specific f is to set a local scale size for each cluster of stability points and to set the spacing between them. If one plots the points of, say, a 2^8-point limit cycle of (2) (or any cycle highly bifurcated from some low-order one), then by unevenly stretching the axis, the same 2^8-point cycle of (3) for another f is produced. The points are distributed unevenly in clusters sufficiently small that the stretching is essentially a pure magnification over the scale of a cluster. Moreover, for a fixed f, if b is increased to produce a 2^9-point cycle, that cluster about $\bar{x}$ (the maximum point) reproduces itself on a scale approximately α times smaller, where

$$\alpha = 2.5029078750957\ldots$$

when f has a normal (i.e., quadratic) maximum. (This shall be assumed unless specifically stated otherwise.) The presence of the number α is a binding test on whether or not (3) is a correct model. α is a reflection of the infinitely bifurcative structure of (3), independent of any particular f. That is, the great bulk of the detailed quantitative aspect of solutions to (3) is independent of a specific choice of f: Eq. (3) and Fig. 1 comprise the bulk of the quantitative theory of such a population. Indeed, it is very difficult to extract the exact form of f from data, as so much quantitative information is determined purely by (3). In addition to α, another universal number determined by (3) should leave its mark on the data of a system described by (3). Thus, let b_0 be the value of b such that $\bar{x}$ (the abscissa of the maximum) is an element of a stable r-point cycle, and generally b_n the value of b such that $\bar{x}$ is an element of the stable $(r \times 2^n)$-point cycle n times bifurcated from the original. Then

$$\delta = \lim_{n \to \infty} \frac{b_{n+1} - b_n}{b_{n+2} - b_{n+1}}$$

is universal, with

$$\delta = 4.6692016091029...$$

It must be stressed that the numbers α and δ are *not* determined by, say, the set of all derivatives of (an analytic) f at same point. (Indeed, f need not be analytic.) Rather, universal functions exist that describe the local structure of stability sets, and these functions obey functional equations [independent of the f of (3)] implicating α and δ in a fundamental way.

2. QUALITATIVE ASPECTS OF BIFURCATION AND UNIVERSALITY

For definiteness (with no loss of generality), f is taken to map [0, 1] *onto* itself. At the unique differentiable maximum $\bar{x}$, $f(\bar{x}) = 1$,

$$x_{n+1} = \lambda f(x_n)$$

and λ lies in the interval [0, 1] to guarantee that if $x_0 \in [0, 1]$ then so, too, will all its iterates. When $\lambda = \bar{x}$,

$$\lambda f(\bar{x}) = \bar{x} f(\bar{x}) = \bar{x}$$

and $\bar{x}$ is a fixed point (Fig. 2). There is a simple graphical technique to determine the successive iterates of an initial point x_0:

(a) Draw a vertical segment along $x = x_0$ up to $\lambda f(x)$, intersecting at P.
(b) Draw a horizontal segment from P to $y = x$. The abscissa of the point of intersection is x_1.
(c) Repeat (a) and (b) to obtain x_{n+1} from x_n.

It is obvious from Fig. 2 that $\bar{x}$ is stable. Stability is locally analyzed by linear approximation about a fixed point. Setting

$$x_n = \bar{x} + \xi_n, \qquad \bar{x} f(x) \equiv g(x), \qquad g(\bar{x}) = \bar{x}$$
$$x_{n+1} = g(x_n) \Rightarrow \bar{x} + \xi_{n+1} = g(\bar{x} + \xi_n) = g(\bar{x}) + \xi_n g'(\bar{x}) + \cdots$$
$$\Rightarrow \xi_{n+1} = g'(\bar{x})\xi_n + O(\xi_n^2)$$

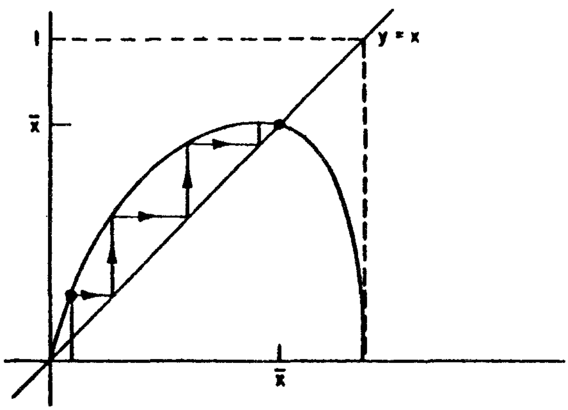

Fig. 2

Clearly $\xi_n \to 0$ if $|g'(\bar{x})| < 1$, the criterion for local stability. But $g'(\bar{x}) = \bar{x}f'(\bar{x}) = 0$, so that $\bar{x}$ is stable. With $r \equiv |g'(\bar{x})| < 1$,

$$\xi_n \propto r^n$$

so that convergence is geometric for $r \neq 0$. For $r = 0$, convergence is faster than geometric, and $\lambda = \bar{x}$ is that value of λ determining the most stable fixed point. We denote this value of λ by λ_0. Increasing λ just above λ_0 causes the fixed point x^* to move to the right with $g'(x^*) < 0$. At $\lambda = \Lambda_0$, $g'(x^*) = -1$ and x^* is marginally stable; for $\lambda > \Lambda_0$ it is unstable. According to Metropolis *et al.*,[1] a two-point cycle should now become stable. Stability of either of these points, say $x_1{}^*$, is determined by $|g^{(2)\prime}(x_1{}^*)|$, where

$$g^{(2)}(x) = g(g(x)); \qquad g^{(n+1)}(x) = g(g^{(n)}(x)) = g^{(n)}(g(x))$$

Accordingly, consider $g^{(2)}(x)$ when $g'(x^*) < -1$ (Fig. 3). Several details of Fig. 3 are especially important. First, $g^{(2)}$ has two maxima: this because $\bar{x}$ has two inverses for $\lambda > \lambda_0$. Each maximum is of identical character to that of g: a neighborhood of $x_m^{(1)}$ is mapped into a neighborhood about $\bar{x}$ by g; g has a nonvanishing derivative at $x_m^{(1)}$, so that the imaged neighborhood is the original simply translated and stretched; accordingly, g applied to this new neighborhood is simply a magnification of g about $\bar{x}$. Thus, if $g(x) \propto |x - \bar{x}|^z + g(\bar{x})$, $z > 1$ for $|x - \bar{x}|$ small, then $g^{(2)} \propto |x - x_m^{(1)}|^z + g(\bar{x})$ for $|x - x_m^{(1)}|$ small. Similarly, the minimum (located at $\bar{x}$) is of order z. This is, of course, the content of the chain rule: $g^{(n)\prime}(x_0) = \prod_{i=0}^{n-1} g'(x_i)$ with $x_i = g^{(i)}(x_0)$ $[g^{(0)}(x) \equiv x]$. In particular, observe that $\bar{x}$ is a point of extremum of $g^{(n)}$ for all n. Also, if $g(x^*) = x^*$, then $g^{(n)\prime}(x^*) = [g'(x^*)]^n$. With $g'(x^*) < -1$,

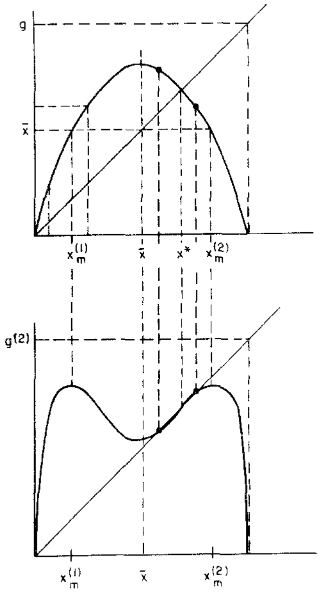

Fig. 3

$g^{(2)\prime}(x^*) > 1$, so that $g^{(2)}$ must develop two fixed points besides x^*: these two new fixed points are a two-point cycle of g itself, and for $\lambda - \Lambda_0$ sufficiently small, $0 < g^{(2)\prime} < 1$ at these points. Moreover, since $g(x_1^*) = x_2^*$ and $g(x_2^*) = x_1^*$, the chain rule implies that $g^{(2)\prime}(x_1^*) = g^{(2)\prime}(x_2^*)$, so that each element of the cycle enjoys identical stability. As λ is increased, the maxima of $g^{(2)}$ ($g^{(2)} = \lambda$ at maximum) also increase until a value λ_1 is reached when the abscissa of the rightmost maximum $x_m^{(2)} = \lambda_1$. By the chain rule, the other fixed point is now also at an extremum, and must be at $\bar{x}$ (Fig. 4).

As λ increases above λ_1, $g^{(2)}(\bar{x})$ decreases below $\bar{x}$, so that $g^{(2)\prime} < 0$ for the leftmost fixed point, and so, for the rightmost one. At $\lambda = \Lambda_1$, $g^{(2)\prime} = -1$ for both: otherwise the two-point cycle would always remain stable, in violation of the results of Metropolis *et al.* Thus, $g^{(2)\prime} < -1$ for $\lambda > \Lambda_1$, the two-point cycle is unstable, and we are now motivated to consider $g^{(4)}$, as a four-point cycle should now be stable. Alternatively, the region "a" of $g^{(2)}$ of Fig. 4 bears a distinct resemblance to g of Fig. 2 turned upside down and reduced in scale: the transition that led from Fig. 2 to Fig. 4 is now being reexperienced, with $g^{(2)}$ replacing g and $g^{(4)}$ replacing $g^{(2)}$. In particular, at $\lambda = \lambda_2 > \Lambda_1$ the fixed points of $g^{(4)}$ beyond those of $g^{(2)}$ will occur at extrema (Fig. 5). The region "a" of $g^{(4)}$ is again an upside-down, reduced version of that of $g^{(2)}$ in Fig. 4; the square box construction including $\bar{x}$ for $g^{(2)}$ of Fig. 5 is an upside-down, reduced version of that of g in Fig. 4. Since the boxes are *squares*, the Fig. 5 box is reduced by the *same* scale on both height

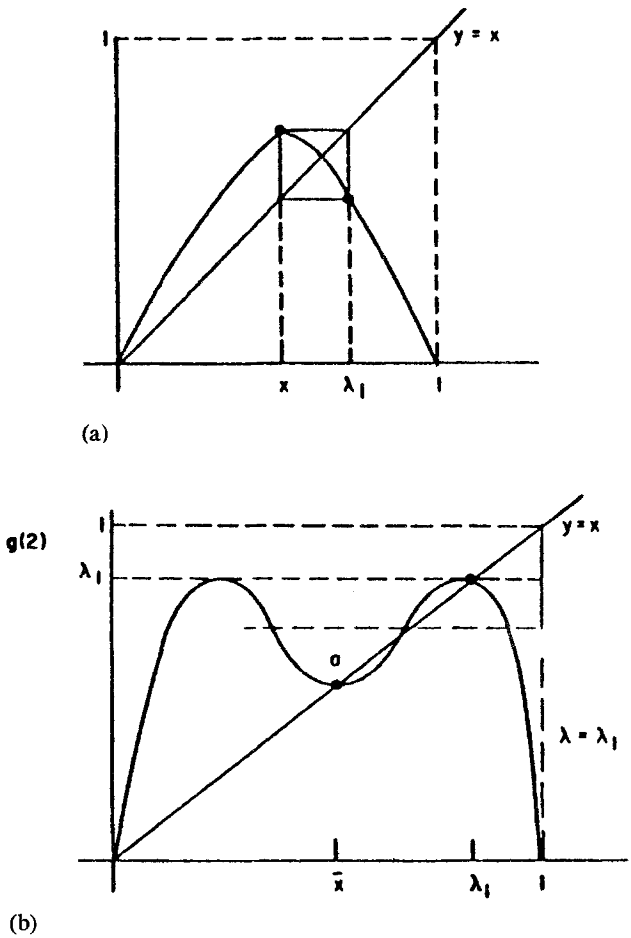

Fig. 4

and width from Fig. 4. Accordingly, the regions "a" are also rescaled identically on height and width.

It is very important to realize that in Fig. 5, g itself was not drawn since it is unnecessary: $g^{(2)}$ is sufficient to determine $g^{(4)}$:

$$g^{(4)}(x) = g(g(g(g(x)))) = g(g(g^{(2)}(x))) = g^{(2)}(g^{(2)}(x))$$

[and similarly, $g^{(2n+1)}(x) = g^{(2^n)}(g^{(2^n)}(x))$]. At the level of discussion of Fig. 5, $g^{(2)}$ has effectively replaced g as the fundamental function considered. $g^{(2)}$, though, is not simply proportional to λ, possessing internal λ dependence: the underlying role of g is exposed by $g^{(2)}$ in the simultaneous occurrence of the two box constructions. Similarly, by the nth bifurcation, only $g^{(2^{n-1})}$ and $g^{(2^n)}$ are important. If at λ_{n+1} (at $\lambda = \lambda_n$, $\bar{x}$ is an element of a 2^n-point cycle)

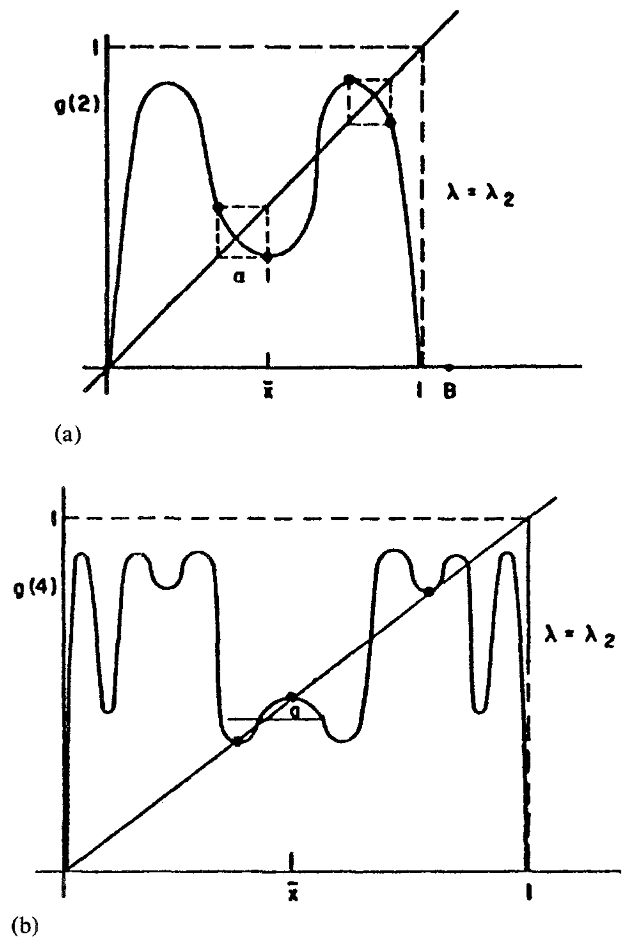

Fig. 5

we magnify the box containing $\bar{x}$ of $g^{(2^n)}$ and invert it to overlay that of $g^{(2^{n-1})}$ at $\lambda = \lambda_n$ (Fig. 6), we have two curves of identical order of maximum z, of identical height with identical zeros. Through a set of operations, $g^{(2^{n-1})}$ determines $g^{(2^n)}$, just as will $g^{(2^n)}$ determine $g^{(2^{n+1})}$. Referring back to Fig. 4, observe that the restriction of $g^{(2)}$ to the interval between maxima is determined entirely by the restriction of g itself to this same interval. The region "a" of $g^{(2)}$ is determined by g restricted a smaller interval plus essentially just the slope of g at λ_1 if g is sufficiently smooth. Analogously, the restriction of $g^{(2^n)}$ to its box part is determined through a similar restriction of $g^{(2^{n-1})}$. With the n scale reductions that have taken place by this level of iteration, $g^{(2^n)}$ is determined by g restricted to an increasingly small interval about $\bar{x}$ together with

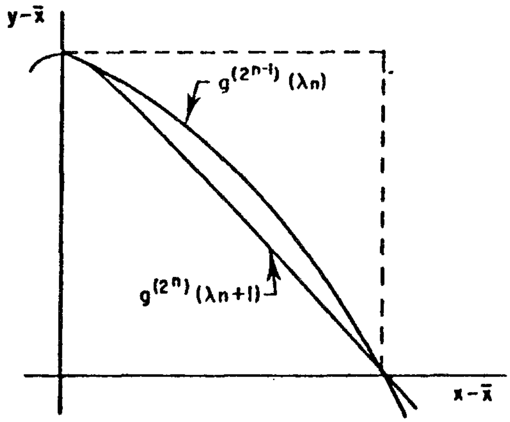

Fig. 6

the slope of g at n points. These slopes determine only the absolute scale of $g^{(2^n)}$: its shape is determined purely by the restriction of g to the immediate vicinity of $\bar{x}$. If we now set by hand the scale of a magnified $g^{(2^n)}$ so that the square is of unit length, then the role of the n slopes is eliminated. *Accordingly, we now conjecture that the rescaled $g^{(2^n)}$ about $\bar{x}$ approaches a function $g^*(x)$ independent of $f(x)$ for all f's of a fixed order of maximum z: g^* depends only on z.* It remains now to make this discussion formal, exactly defining the rescaling and the function g^*. The above heuristic argument for universality regrettably remains in want of a rigorous justification. However, we have carefully verified it, and all details to follow by computer experiment. In a sequel to this work we shall establish exact equations and isolate specific questions whose resolutions would establish the conjecture.

3. THE RECURSIVE NATURE OF SUCCESSIVE BIFURCATION

We have described a process that can be summarized as follows.

(0) We start at $\lambda = \lambda_n$, and look at $g^{(2^n)}$ near $x = \bar{x}$. Alternatively, we might look at $g^{(2^{n-1})}$ for the same λ and range of x, as depicted in Fig. 7.

(i) Form $g^{(2^n)}(x) = g^{(2^{n-1})}(g^{(2^{n-1})}(x))$, depicted in Fig. 8.

(ii) Increase λ from λ_n to λ_{n+1}, depicted in Fig. 9.

(iii) Rescale: $g^{(2^n)}(x) \to \alpha_n g^{(2^n)}(x/\alpha_n)$, depicted in Fig. 10 ($|\alpha| > 1$).

Calling the operations (i)–(iii) B_{n-1}, we have

$$\tilde{g}_n(x) = B_{n-1}[\tilde{g}_{n-1}(x)], \qquad n = 2, 3, \ldots$$

and are claiming $\tilde{g}_n(x) \to g^*(x)$ locally about $\bar{x}$.

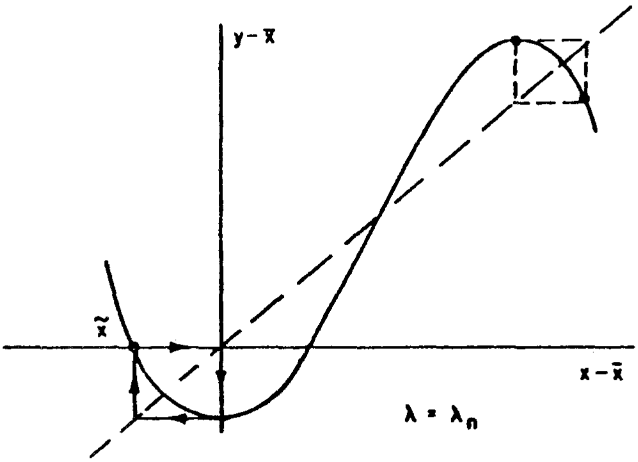

Fig. 7

Clearly (i) of B_n is recursive and n-independent; we call this part of B_n "doubling." We will motivate that (ii) becomes asymptotically n-independent; we term this part of B_n "λ-shifting." Also, with $\alpha_n \to \alpha$ essentially by (i), part (iii) of B_n becomes asymptotically n-independent; we term this part (obviously) "rescaling." Thus, $B_n \to B$. That is,

$$\lim_{r \to \infty} B^r[\tilde{g}_n(x)] = g^*(x)$$

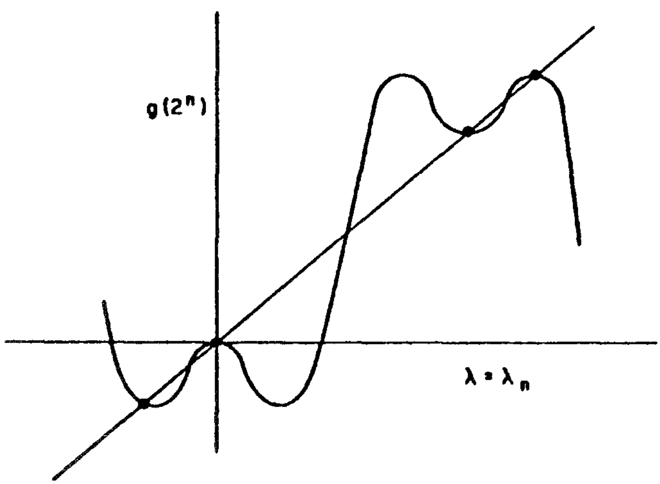

Fig. 8

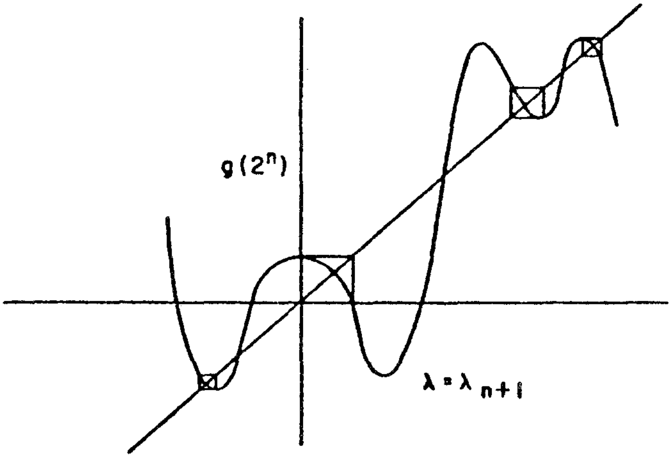

Fig. 9

Accordingly, g^* satisfies the equation

$$g^* = B[g^*] \tag{5}$$

Universality, thus, is the consequence of a recursion on the class of functions $f(x)$ considered. Under high-order bifurcation, the fixed point of B is approached—that fixed point being, within a certain domain, a property of B itself and not of the starting $f(x)$. Evidently, domains of the various fixed points of B are disjoint for different z. Also, each fixed-z domain clearly exceeds the class of f's specified by properties 1–4 of Appendix A, since $(f)^{(2^n)}$ for each n is also in the domain. At present we cannot specify just how

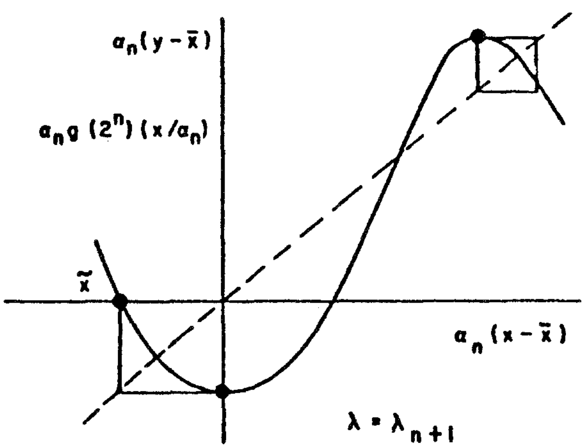

Fig. 10

large this domain is. The fixed-point equation (5) will certainly, for a given z, determine the rescaling ratio α as well as g^*. [For a variety of functions $f(x)$ with $z = 2$, we have determined g^*, with $\tilde{x}$ of Fig. 10 set to unit length.]

4. DETAILED FEATURES OF THE BIFURCATION RECURSION

We first indicate roughly how the parameters α and δ are interrelated and determined by g^*. At $\lambda = \lambda_n$, $\tilde{g}_{n-1}$ and $\tilde{g}_{n-1} \circ \tilde{g}_{n-1}$ appear as in Fig. 11. Increasing λ has $\tilde{g}_{n-1}(0)$ increase above 1, producing Fig. 12, where $\tilde{g}_{n-1}$ and $\tilde{g}_{n-1} \circ \tilde{g}_{n-1}$ at λ_n are shown dashed. By the definition of α_n, h_n of Fig. 12 satisfies

$$h_n = \alpha_n^{-1}$$

Clearly, though, in some rough sense

$$h_n \simeq (h_{n-1} - 1)|\tilde{g}'_{n-1}(1)| \equiv \delta h_{n-1}|\tilde{g}'_{n-1}(1)|$$

i.e.,

$$\delta h_{n-1} \simeq |\alpha_n \tilde{g}'_{n-1}(1)|^{-1} \tag{6}$$

Also,

$$\delta h_{n-2} \simeq |\tilde{g}'_{n-2}(1)|^{-1} \delta h_{n-1} \tag{7}$$

This is more nearly accurate than (6), since $\tilde{g}_{n-2}$ shifts less than $\tilde{g}_{n-1}$ for the same λ increase. Thus,

$$\delta h_{n-1} \simeq \prod_{2}^{n} |\tilde{g}'_{n-i}(1)| \; \delta h_0 \tag{8}$$

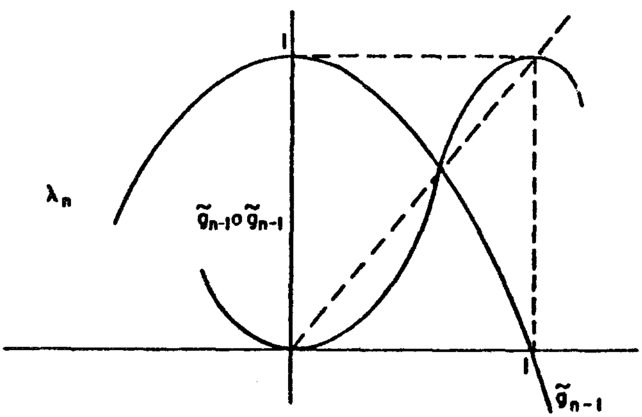

Fig. 11

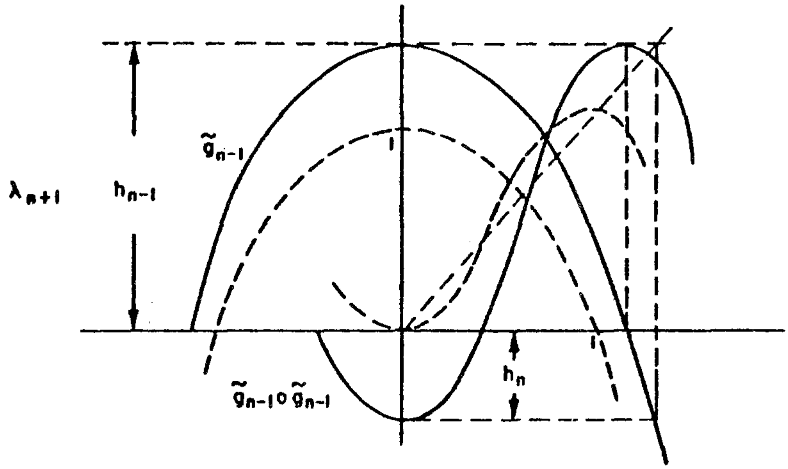

Fig. 12

However, $\delta h_0 = \delta\lambda_n = \lambda_{n+1} - \lambda_n$. Assuming $\tilde{g}_n \to g^*$ (this is not quite correct; see Section 5) one has, so far as n dependence is concerned,

$$\delta h_{n-1} \simeq \mu |g^{*\prime}(1)|^{n-1}\, \delta\lambda_n \tag{9}$$

with $\mu \sim 1$ an asymptotically n-independent factor. Substituting in (6),

$$\delta\lambda_n \simeq \mu^{-1}/|\alpha g^{*\prime}(1)|^n$$

with $\alpha = \lim \alpha_n$.

Accordingly, $\delta\lambda_n \propto \delta^{-n}$, with

$$\delta \simeq \alpha |g^{*\prime}(1)| \tag{10}$$

For $z = 2$, the computer-experimental value for $|g^{*\prime}(1)|$ is $+1.89$, to be compared with $\delta/\alpha = 1.87$.

With $f(x)$ real-analytic in an arbitrarily small domain about $\bar{x}$, the manner in which the $\tilde{g}_n$ are formed ensures for them a systematically larger domain of analyticity. With $\tilde{g}_n \to g^*$, an equivalent procedure for defining the α_n is to require (at least one-sided) agreement in $\tilde{g}_n^{(2)}(0)$. One has

$$\tilde{g}_n(x) = 1 - |x|^z(a + bx^2 + \cdots) \tag{11}$$

Then,

$$\begin{aligned}
\tilde{g}_n \circ \tilde{g}_n(x) &= 1 - a|\tilde{g}_n|^z - b|\tilde{g}_n|^{z+2} + \cdots \\
&= 1 - a|1 - a|x|^z \cdots|^z - b|1 - a|x|^z + \cdots|^{z+2} + \cdots \\
&= 1 - a - b + \cdots + a[az|x|^z + \cdots + b(z+2)|x|^z + \cdots] \\
&= \tilde{g}_n(1) + a|x|^z[1 - \tilde{g}_n{}'(1)] + \cdots \\
&= a|x|^z[1 - \tilde{g}_n{}'(1)] + \cdots
\end{aligned}$$

Next, the λ shift is performed:

$$\tilde{g}_n \to \tilde{g}_n \circ \tilde{g}_n \to -\nu + \mu a|x|^z[1 - \tilde{g}_n'(1)] + \cdots$$

and finally, α rescaling:

$$\tilde{g}_n \to -\{1 - \mu(a/\alpha^{z-1})[1 - \tilde{g}^{*\prime}(1)]|x|^z + \cdots\} \tag{12}$$

For (11) and (12) to agree, one has

$$\alpha^{z-1} \sim 1 - g^{*\prime}(1) \tag{13}$$

where $\mu \lesssim 1$ corresponds to λ-shifting being mostly a displacement in the immediate environs of $\bar{x}$. Again, for $z = 2$, one compares $\alpha = 2.50$ with $1 - g^{*\prime}(1) = 2.87$. Combining (10) and (13), one has

$$\delta \simeq |g^{*\prime}(1)|[1 - g^{*\prime}(1)]^{1/z-1} \tag{14}$$

While (13) and (14) are crude, they are roughly correct for $z \gtrsim 2$, but more important, indicate that g^* ultimately determines everything.

We now proceed to describe the situation more carefully, tacitly assuming convergence, and successively illustrating its details through consistency arguments.

By definition

$$g^*(x) = \lim(-1)^n\alpha^n g^{(2^n)}(x/\alpha^n, \lambda_{n+1}) \equiv \lim \tilde{g}_n(x) \tag{15}$$

where α^n is symbolic for α_n which becomes asymptotically a multiple of α^n: the multiple has been absorbed in $g^{(2^n)}$. For all n, $\tilde{g}_n$ satisfies

$$\tilde{g}_n(1) = 0, \qquad \tilde{g}_n(0) = 1, \qquad \tilde{g}_n'(0) = 0 \tag{16}$$

and near $x = 0$, $1 - \tilde{g}_n(x) \sim |x|^z$.

We now furnish an approximate equation for g^*:

$$\begin{aligned}(-1)^n\alpha^{n-1}g^{(2^n)}(x, \lambda_n) &= (-1)^n\alpha^{n-1}g^{(2^{n-1})}(g^{(2^{n-1})}(x, \lambda_n), \lambda_n)\\ &= (-1)^n\alpha^{n-1}g^{(2^{n-1})}\left(\frac{1}{\alpha^{n-1}}\,\alpha^{n-1}g^{(2^{n-1})}(x, \lambda_n), \lambda_n\right)\\ &= -\tilde{g}_n \circ \tilde{g}_n(x\alpha^{n-1}) \end{aligned}\tag{17}$$

or

$$(-1)^n\alpha^n g^{(2^n)}(x/\alpha^n, \lambda_n) = -\alpha\tilde{g}_n \circ \tilde{g}_n(x/\alpha) \tag{18}$$

or

$$-\alpha\tilde{g}_n \circ \tilde{g}_n(x/\alpha) = \tilde{g}_{n+1}(x) - (-1)^n\alpha^n(g^{(2^n)}(x/\alpha_n, \lambda_{n+1}) - g^{(2^n)}(x/\alpha_n, \lambda_n))$$

or

$$-\alpha\tilde{g}_n \circ \tilde{g}_n(x/\alpha) \simeq \tilde{g}_{n+1}(x) - (-1)^n\alpha^n(\lambda_{n+1} - \lambda_n)\,\partial_\lambda g^{(2^n)}(x/\alpha_n, \lambda_n) \tag{19}$$

assuming a "mild" λ-shifting.

Clearly, $\alpha^n \partial_\lambda g^{(2^n)}(x/\alpha_n, \lambda_n)$ diverges with n since $\lambda_{n+1} - \lambda_n \to 0$. Thus, a more careful analysis, like that used to treat Eq. (10), needs to be done. By (17),

$$\begin{aligned}\partial_\lambda g^{(2^n)}(x, \lambda_n) &= \partial_\lambda g^{(2^{n-1})}(g^{(2^{n-1})}(x, \lambda_n), \lambda_n) \\ &\quad + \partial_x g^{(2^{n-1})}(g^{(2^{n-1})}(x, \lambda_n), \lambda_n)\, \partial_\lambda g^{(2^{n-1})}(x, \lambda_n) \\ &= \partial_\lambda g^{(2^{n-1})}(g^{(2^{n-1})}(x, \lambda_n), \lambda_n) \\ &\quad + \partial_x \tilde{g}_{n-1}(\tilde{g}_{n-1}(x/\alpha^{n-1}))\partial_\lambda g^{(2^{n-1})}(x, \lambda_n)\end{aligned}$$

So,

$$\begin{aligned}\alpha^n \partial_\lambda g^{(2^n)}\left(\frac{x}{\alpha^n}, \lambda_n\right) &= \alpha^n \partial_\lambda g^{(2^{n-1})}\left(\frac{1}{\alpha^{n-1}} \tilde{g}_{n-1}\left(\frac{x}{\alpha}\right), \lambda_n\right) \\ &\quad + \tilde{g}'_{n-1} \circ \tilde{g}_{n-1}\left(\frac{x}{\alpha}\right) \alpha^n \partial_\lambda g^{(2^{n-1})}\left(\frac{x}{\alpha^n}, \lambda_n\right)\end{aligned} \tag{20}$$

At $x = 0$,

$$\alpha^n \partial_\lambda g^{(2^n)}(0, \lambda_n) = \alpha^n \partial_\lambda g^{(2^{n-1})}(1/\alpha^{n-1}, \lambda_n) + \tilde{g}'_{n-1}(1)\alpha^n \partial_\lambda g^{(2^{n-1})}(0, \lambda_n) \tag{21}$$

With

$$\alpha^n \partial_\lambda g^{(2^{n-1})}(1/\alpha^{n-1}, \lambda_n) = \mu\alpha^n \partial_\lambda g^{(2^{n-1})}(0, \lambda_n)$$

(such a μ exists if λ-shifting becomes n-independent), (21) becomes

$$\begin{aligned}\alpha^n \partial_\lambda g^{(2^n)}(0, \lambda_n) &= [\mu + \tilde{g}'_{n-1}(1)]\alpha^n \partial_\lambda g^{(2^{n-1})}(0, \lambda_n) \\ &\simeq [\mu + \tilde{g}'(1)]\alpha^n \partial_\lambda g^{(2^{n-1})}(0, \lambda_{n-1})\end{aligned} \tag{22}$$

($g^{(2^{n-1})}$ shifts more slowly than $g^{(2^n)}$: higher order λ derivatives have been neglected). Iterating (22), one has

$$\partial_\lambda g^{(2^n)}(0, \lambda_n) \simeq \rho[\mu + \tilde{g}'(1)]^n \tag{23}$$

with $\rho \sim 1$, n-independent. So,

$$(-1)^n(\lambda_{n+1} - \lambda_n)\alpha^n \partial_\lambda g^{(2^n)}(0, \lambda_n) \simeq \rho[\alpha(-\tilde{g}'(1) - \mu)]^n(\lambda_{n+1} - \lambda_n)$$

By (19) this is n-independent, and so,

$$\lambda_{n+1} - \lambda_n \sim \delta^{-n}$$

with

$$\delta \simeq \alpha(-\tilde{g}'(1) - \mu) \tag{24}$$

Defining $\tilde{h}_n(x) = (-1)^n\alpha^n(\lambda_{n+1} - \lambda_n)\, \partial_\lambda g^{(2^n)}(x/\alpha_n, \lambda_n)$, (19) reads

$$\tilde{g}_{n+1}(x) = \tilde{h}_n(x) - \alpha\tilde{g}_n \circ \tilde{g}_n(x/\alpha) \tag{25}$$

or

$$g^*(x) = h^*(x) - \alpha g^* \circ g^*(x/\alpha) \tag{26}$$

Returning to (20), multiplied by $\lambda_{n+1} - \lambda_n$, neglecting higher order derivatives,

$$\tilde{h}_n(x) \simeq -\omega(\tilde{h}_{n-1} \circ \tilde{g}_{n-1}(x/\alpha) + \tilde{h}_{n-1}(x/\alpha)\tilde{g}'_{n-1} \circ \tilde{g}_{n-1}(x/\alpha)) \tag{27}$$

with some $\omega \sim 1$, or, as $n \to \infty$, and repeating (26),

$$h^*(x) = -\omega(h^* \circ g^*(x/\alpha) + h^*(x/\alpha)g^{*\prime} \circ g^*(x/\alpha))$$

and

$$g^*(x) = h^*(x) - \alpha g^* \circ g^*(x/\alpha) \tag{28}$$

These constitute first-order (approximate) fixed-point equations, satisfying the boundary conditions

$$g^*(0) = 1, \qquad g^{*\prime}(0) = 0, \qquad g^*(1) = 0, \qquad h^*(0) = 1 \tag{29}$$

[We comment that (28) is recursively stable, and for $z = 2$ affords a 10% approximate solution.]

At this point, some remarks concerning convergence (say of $\tilde{g}_n \to g^*$) are in order. The function $g^*(x)$ describes the stability set for large n in the vicinity of $\bar{x}$: those x_i such that

$$g^* \circ g^*(x_i) = x_i$$

[and, of course, $g^{*\prime}(g^*(x_i)) \cdot g^{*\prime}(x_i) = 0$] are the stability set points near $\bar{x}$. Accordingly, all such x_i scale with α upon bifurcation: $|x_i - x_j| \to (1/\alpha)|x_i - x_j|$. For example, the distance between $\bar{x}$ and the nearest element to it of the stability set of order 2^n is α times greater than that distance in the stability set of order 2^{n+1}. (Also, if x_1 is the nearest point to $\bar{x}$ and x_2 the next nearest, then for all n large enough, $|\bar{x} - x_1|/|\bar{x} - x_2| \equiv \gamma$ is fixed.) This immediately leads to a difficulty: distances near $\bar{x}$ and those near λ_n (the furthest right element of a stability set) cannot possibly scale identically.

As is obvious from Fig. 13, with Δ_n the distance from $\bar{x}$ to x_1, and d_n the distance from λ_n to $\bar{x}$ (the next to rightmost point), $d_n \sim \Delta_n{}^z$, so that with $\Delta_n \propto \alpha^{-n}$,

$$d_n \propto (\alpha^z)^{-n} \neq \alpha^{-n} \tag{30}$$

Thus, convergence of $\tilde{g}_n(x)$ to $g^*(x)$ must be local in nature. The scale for which $g^*(0) = 1$ and $g^*(1) = 0$ is, of course, α^{-n} finer than usual measure on $[0, 1]$: for large n, $\sup|\tilde{g}_n - g^*| < \epsilon_{N_n}$ for $|x| < N_n$ is uniform convergence in "real" x of $|x| < N/\alpha^n$. To allow for a shifting rescaling of parts of g^*, $N_n \ll \alpha^n$. Thus, one anticipates that $\tilde{g}_n \to g^*$ (say in sup-norm) over *any* bounded part of R *but* with the $g^*(1) = 0$ measure. In any (small) interval about a given point in the stability set of order n, one sets the origin of $g^{(2^n)}$ at the point in question and forms $\tilde{g}_n$ with an appropriate (local) scale factor. As n increases, in the $\tilde{g}_n(1) = 0$ measure, any other point a finite distance away

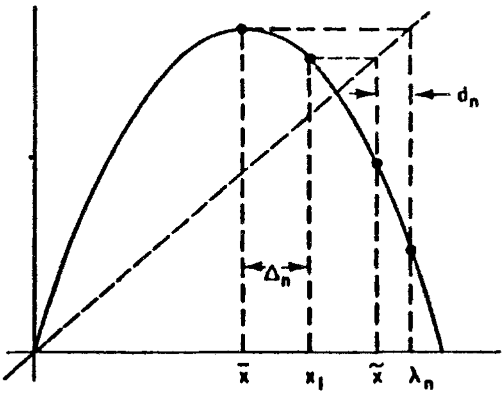

Fig. 13

in *usual* [0, 1] measure grows far remote from the chosen point: so far, that the local $\tilde{g}_n$ never converge to it. Thus, in effect, a class of g^* exist, each determining the large-n limiting stability set about a point.

For example, defining $\tilde{f}_n(x)$ about λ_{n+1} by Fig. 14 $[x_n \equiv g^{(2^n)}(\lambda_{n+1})$, $\lambda_{n+1} = g^{(2^n)}(x_n)]$, we have

$$\tilde{f}_n(x) = [g^{(2^n)}((\lambda_{n+1} - x_n)x + x_n) - x_n]/(\lambda_{n+1} - x_n) \to f^*(x) \qquad (31)$$

[so that $\tilde{f}_n(0) = 1, \tilde{f}_n'(0) = 0, \tilde{f}_n(1) = 0$]. In the notation of Fig. 13,

$$\tilde{f}_n(x) = [g^{(2^n)}(xd_n + x_n) - x_n]/d_n$$

and so $\tilde{f}_n$ scales by d^z rather than d. It is straightforward to relate f^* to g^*:

$$g^{(2^n)}(\lambda_{n+1}f(x)) = \lambda_{n+1}f(g^{(2^n)}(x)) \qquad (32)$$

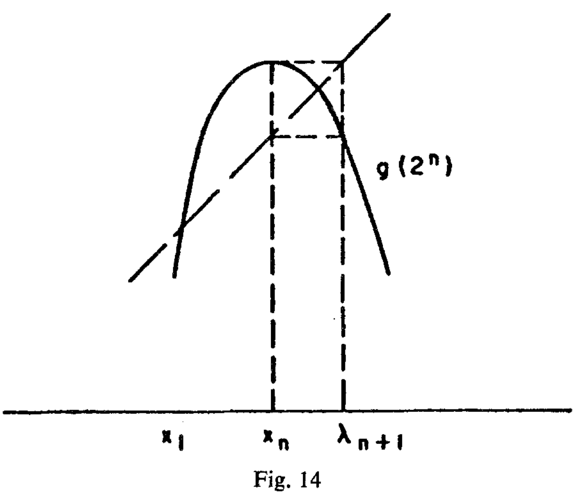

Fig. 14

so that for $x \sim 0$ (we have conveniently set $\bar{x} = 0$), $\lambda_{n+1} f(x)$ is near λ_{n+1} and (32) relates $g^{(2^n)}$ about λ_{n+1} to $g^{(2^n)}$ about 0. Thus

$$x \text{ small} \Rightarrow \lambda_{n+1} f(x) = \lambda_{n+1} - a\lambda_{n+1}|x|^z + O(|x|^{z+1})$$

$$\begin{aligned} g^{(2^n)}(\lambda_{n+1} f(x)) &= g^{(2^n)}((\lambda_{n+1} - x_n)(1 - a|x|^z) - ax_n|x|^z + x_n) \\ &= x_n + (\lambda_{n+1} - x_n)\tilde{f}_n\left(1 - a|x|^z - \frac{ax_n}{\lambda_{n+1} - x_n}|x|^z\right) \end{aligned}$$

Also,

$$\begin{aligned} \lambda_{n+1} f(g^{(2^n)}(x)) &= \lambda_{n+1} - a\lambda_{n+1}|g^{(2^n)}(x)|^z + \cdots \\ &= \lambda_{n+1} - a\lambda_{n+1}\Delta_n^z|\tilde{g}_n(x/\Delta_n)|^z \end{aligned}$$

Accordingly, (32) implies for small x that

$$(\lambda_{n+1} - x_n)\left[1 - \tilde{f}_n\left(1 - a|x|^z - \frac{ax_n}{\lambda_{n+1} - x_n}|x|^z\right)\right] = a\lambda_{n+1}\Delta_n^z\left|\tilde{g}_n\left(\frac{x}{\Delta_n}\right)\right|^z$$

By Fig. 14, $\lambda_{n+1} - x_n = d_n = a\Delta_n^z\lambda_{n+1}$, so that

$$1 - \tilde{f}_n\left(1 - \frac{\lambda_{n+1}a|x|^z}{\lambda_{n+1} - x_n}\right) = \left|\tilde{g}_n\left(\frac{x}{\Delta_n}\right)\right|^z$$

or

$$1 - \tilde{f}_n(1 - |x|^z/\Delta_n^z) = |\tilde{g}_n(x/\Delta_n)|^z$$

or

$$\tilde{f}_n(1 - |\xi|^z) = -|\tilde{g}_n(\xi)|^z + 1$$

or

$$\tilde{f}_n(x) = -|\tilde{g}_n((1 - x)^{1/z})|^z + 1 + \cdots \tag{33}$$

For large n, neglected terms are powers of $\Delta_n \to 0$. So,

$$f^*(x) = -|g^*((1 - x)^{1/z})|^z + 1 \tag{34}$$

and g^* determines f^*. (This has, of course, been computationally verified to full precision.) We are unsure of the size of the set of rescalings: clearly α and α^z belong to the set. However, about any point a fixed, finite number of iterates prior to $\bar{x}$, scaling goes by α, whereas any region a finite number of iterates after λ_{n+1} scales with α^z. On the other hand, points 2^{n-1} iterates from $\bar{x}$, as well as $2^{n-2}, \ldots, 2^{n-N}$ from $\bar{x}$, are just the N points nearest $\bar{x}$ that scale with α. That is, we are uncertain how to define a region which possesses a scaling intermediate between α and α^z; possibly the situation is simply an interspersal of regions scaling by either α or α^z. What is missing is some notion of ordered measure along a stability set.

At this point it is perhaps illuminating to indicate just what g^* looks like. Evidently Fig. 5 is a somewhat distorted version of g^* very near $x = 0$.

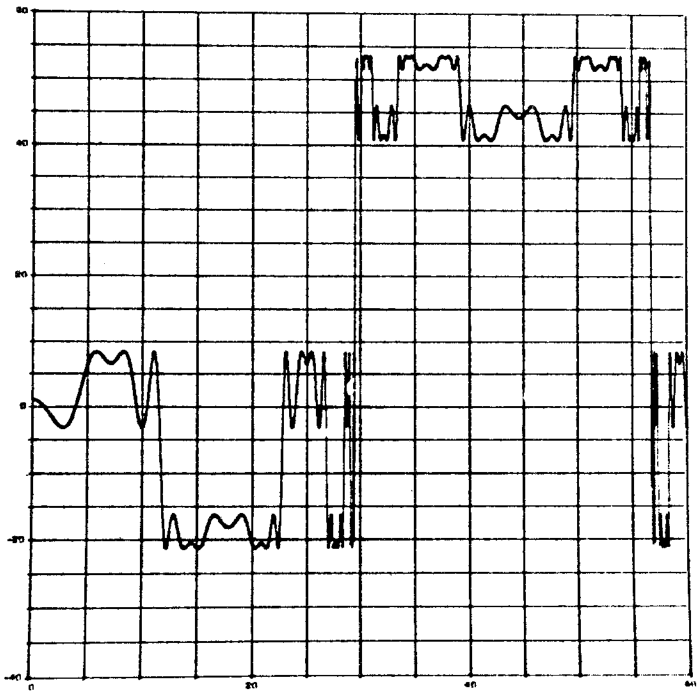

Fig. 15

Indeed, each large jump generates precursors about more mild features. At successive levels of $\tilde{g}_n$, more and more precursors are produced, whose oscillations grow narrower. Also, g^* grows with $|x|$, with a long string of features of roughly the same height as the end height of the string. Figure 15 shows g^* for $|x| < 50$ as computationally obtained. Evidently, convergence to such a function worsens with increasing $|x|$.

5. CONCLUSIONS AND A BRIEF SUMMARY OF THE EXACT THEORY

In this paper we have attempted to heuristically motivate our conjecture of universality, and indicate the form of an exact theory of highly bifurcated attractor sets. Our conclusion is that both of the numbers α and δ as well as the local structure of highly bifurcated attractors as determined by the universal function g^* are determined by functional equations. We have provided approximate such equations, but failed in establishing exact equations for an inability to exactly reflect an increase in the parameter λ. As described, the local structure determined by g^* pertains to values of λ asymptotically near λ_∞ at the specific values λ_n: we have not described that structure for values of

λ between λ_n and λ_{n+1}. However, at these choices of λ, the theory we describe holds regardless of the attractor from which the bifurcated attractors arise. (That is, λ_n refers to that attractor of order $m \cdot 2^n$, which includes the critical point $\bar{x}$ for any m.) Or, the local structure of all *infinite* attractors of $x_{n+1} = \lambda f(x_n)$ is described by g^*. Also, the rescaling parameter α and the convergence rate δ are common to all highly bifurcated attractors.

At this point we should like to briefly summarize some results of the exact theory; a full treatment of these results will shortly appear in a sequel.[(2)]

Figure 5a shows, near $\bar{x}$, what shall evolve into g^*, and represents the local structure at a "two-point" level. Figure 5b represents the graph of a function that shall evolve into $g^* \circ g^*$, and accordingly must also be universal: it represents the identical local structure as does g^*, but now at a "one-point" level of description. Evidently, one can view this local structure at a "2^r-point" level from the function

$$g_r(x) = \lim_{n \to \infty} (-\alpha)^n \tilde{g}^{(2^n)}(\lambda_{n+r}, x/(-\alpha)^n), \qquad r = 0, 1, 2,\ldots \tag{35}$$

where $g_1(x)$ is exactly $g^*(x)$. All g_r are of identical qualitative shape: each bump of g_1 aligned along $y = x$ contains two points of an attractor set, whereas each bump of g_r similarly situated now contains 2^r points of that set. Evidently, the lower r, the more magnified the local structure. Following immediately from the definition,

$$g_{r-1}(x) = -\alpha g_r(g_r(x/\alpha)) \tag{36}$$

(all the functions g_r are symmetric). The content of this equation is essentially the Cantor-set-like nature of highly bifurcated attractors: at each bifurcation, the rough locations of attractor points are unchanged, with a "microscopic" splitting of each such point; the scale of splitting is α below the previous level, so that rescaling by α after a bifurcation reveals the set at a next level of magnification. In fact, (36) provides the entire exact description, as we now synoptically elucidate. The central bump of g_r is effectively a $\lambda f(x)$ containing a 2^r-point cycle, which, as r increases, quickly approaches a λf containing the infinite attractor. That is, one expects

$$g(x) \equiv \lim_{r \to \infty} g_r(x) \tag{37}$$

to exist; g no longer affords the same description of attractor points as does g_r. Rather, g is the description at the level of infinite clusters of points, which is again a universal property. But g defined by (37) is simply a fixed point of (36). Accordingly,[2]

$$g(x) = -\alpha g(g(x/\alpha)) \tag{38}$$

[2] This exact equation was discovered by P. Cvitanović during discussion and in collaboration with the author.

The great virtue of g is that λ has been set at λ_∞ at the outset, and so the difficulty of modeling λ-shifting is totally bypassed. The price paid for this is that (38) defines no recursively stable equation like

$$\bar{g}_{n+1}(x) = -\alpha_n \bar{g}_n(\bar{g}_n(x/\alpha_n)) \tag{39}$$

[By (36), iterating produces g_r's for smaller values of r and hence diverging from g.] There are a variety of ways to solve (38). A method based on the fact that $\bar{g}_0(x) = f(\lambda_\infty(f)x)$ *must* cause (39) to converge [by (35), $g = \lim(-\alpha)^n \bar{g}^{(2^n)}(\lambda_\infty, x/(-\alpha)^n)$] together with the general recursive instability of (39) allows very fast, high-accuracy estimates of all λ_∞ for any chosen f. Alternatively, one can simply solve (38) by a numerical functional-Newton's method. (The result of the latter method is a 20-place determination of both α and g for $z = 2$.)

The g of (38) is a fixed point of (36). By setting

$$g = g_r + y_r$$

in (38), employing (36), and expanding to first order in y, one obtains

$$y_{r-1}(x) = -\alpha[y_r(g(x/\alpha)) + g'(g(x/\alpha))y_r(x/\alpha)] \tag{40}$$

(40) simply separates with the substitution

$$y_r = \lambda^{-r}\psi(x) \tag{41}$$

where ψ obeys

$$\mathscr{L}[\psi(x)] \equiv -\alpha[\psi(g(x/\alpha)) + g'(g(x/\alpha))\psi(x/\alpha)] = \lambda\psi(x) \tag{42}$$

The eigenvalue λ can clearly attain the value $+1$ corresponding to

$$\psi = g - xg'$$

reflecting the dilatation invariance of (38). In addition to a spectrum $|\lambda| \leqslant 1$, computationally there exists a unique alternate value δ strictly greater than $+1$.

It is possible to show that the eigenvalue δ is exactly that convergence rate discussed in this paper. Heuristically, if λ is held fixed at λ_n for $n \gg 1$, and $\lambda_n f$ iterated, it is indistinguishable from the iterates of $\lambda_\infty f$ that approximate g after an initial transient, until roughly n iterations have been performed to magnify the deviation of λ_n from λ_∞. Thus the argument about Eq. (8) can be made exact where the function $\tilde{g}_{n-i}$ there is essentially g for (logarithmically) all iterations.

One can next begin to investigate the nature of the n limit of (35). Defining

$$g_{r,n}(x) \equiv (-\alpha)^n \tilde{g}^{(2^n)}(\lambda_{n+r}, x/(-\alpha)^n)$$

it is immediate to verify that

$$g_{r-1,n+1}(x) = -\alpha g_{r,n}(g_{r,n}(-x/\alpha)) \tag{43}$$

But (36) is the large-n fixed point of (43), and so one can discuss in linear approximation the stability of (36). (We mention at this point that the antisymmetric parts of $g_{r,n}$ vanish in the large-n limit at the rate of $-\alpha$; this result is exactly observed computationally.)

With α and g obtained from (38), (42) determines both h and δ. (Again, both have been obtained to 20-place accuracy.) We stress that (38) and (42) are totally free of any reference to (1) and do produce the same α and δ to the 14-place accuracy of our best recursion data. Next,

$$g_r \sim g - \delta^{-r}h \tag{44}$$

so that a g_r for $r \gg 1$ is available. By successive application of (36) to this asymptotic g_r, g_1 can be obtained. (Regarding $r = 3$ as asymptotic produces a g_1 to six-place accuracy, to give an idea of the speed of onset of the asymptotic regime.) The approximate equations (28) are a high-z approximation to (38) and (42).

Since $\delta^{-r} \sim \lambda_\infty - \lambda_r$, (44) has an immediate continuation:

$$g_\lambda(x) \sim g(x) - (\lambda_\infty - \lambda)h(x)$$

which allows the determination of local structure for λ between λ_n and λ_{n+1} as well. Thus, the bifurcation points Λ_n also geometrically converge to λ_∞ at the rate δ, and logarithmically the behavior of bifurcation is periodic with period $\log \delta$. A demonstration that (36) is in fact a stable fixed point of (43) would constitute a proof of our universality conjecture: with exact (functional) equations at hand, it is possible to focus on the exact details requiring proof.[3]

APPENDIX A

In the formula $x_{n+1} = \lambda f(x_n)$ with $0 < \lambda < 1$, $f(x)$ satisfies the following conditions:

1. $f(x)$ is continuous, single-valued, piecewise $C^{(1)}$ on $[0, 1]$ possessing a unique, differentiable maximum at $\bar{x}$ with $f(\bar{x}) = 1$.
2. $f(x) > 0$ on $(0, 1)$, $f(0) = f(1) = 0$, and f is strictly decreasing on $(\bar{x}, 1)$ and strictly increasing on $(0, \bar{x})$.
3. For $\Lambda_0 < \lambda < 1$, $\lambda f(x)$ has two fixed points $[x^* = 0$, and some other $x^* \in (x, 1)]$ both of which are repellant (i.e., $|f'(x)| > 1/\lambda$).
4. In the interval N about $\bar{x}$ such that $|f'(x)| < 1$, f is concave downward.

[3] We are in possession of extensive high-precision data pertaining to all details discussed in this paper, as well as for the solutions to the functional equations discussed in the last sections. We will consider reasonable requests from individuals for copies of specific parts of this library.

Given these conditions, Metropolis *et al.*[1] have established among others the following universal, qualitative features:

(a) For $\Lambda_0 < \lambda < \lambda_\infty$, there exists stability sets of order 2^n, $n = 1, 2, \ldots$ only, with n increasing with λ.

(b) For $\Lambda_0 < \Lambda_{n-1} < \lambda < \Lambda_n < \lambda_\infty$ only 2^n-order stability sets exist. In particular, at λ_n, with $\Lambda_{n-1} < \lambda_n < \Lambda_n$, the 2^n-order stability set contains $\bar{x}$ as an element.

(c1) For $\lambda = \lambda_1$, under repeated application of λf, one has $\bar{x} \to x' \to \bar{x} \to \cdots$, where $x' > \bar{x}$. Calling an x "R" if $x > \bar{x}$ and "L" if $x < \bar{x}$, the "pattern" of motion through the stability set is abbreviated as R—meaning $\bar{x} \to R \to \bar{x}$.

(c2) The "harmonic" of a pattern P is a stability set of twice the order of P, with pattern PLP if P contains an odd number of Rs and PRP otherwise. The 2^n-order stability sets are exactly the successive harmonics of R (e.g., for λ_2, RLR; for λ_3, RLRRRLR; etc.).

(d) If P is a basic pattern (say, for an r-point cycle), then (a)–(c) hold with 2^n replaced by $r \cdot 2^n$.

APPENDIX B. COMPUTATIONAL RESULTS

The parameter values λ_n for a given recurrence function f are obtained by definition from

$$(\lambda_n f)^{(2^n m)}(\bar{x}) - \bar{x} = 0 \tag{B1}$$

(B1) possesses in general many roots. Accordingly, λ_0 is first obtained for a given fundamental pattern. λ is slowly increased to find the first new zero for $n = 1$; this λ by definition is λ_1. Next, λ_2 is similarly found as the next largest zero of (B1) for $n = 2$. At this point δ_1 is calculated

$$\delta_1 = (\lambda_1 - \lambda_0)/(\lambda_2 - \lambda_1)$$

and used to estimate–predict λ_3

$$\lambda_3 \simeq \lambda_2 + \delta_1^{-1}(\lambda_2 - \lambda_1) \tag{B2}$$

As n increases, $\delta_n \to \delta$ and the predicted value increases in precision, so that for large n, several Newton's-method iterations suffice to locate λ_n to full precision. Since the number of iterations increases geometrically and the number of zeros of (B1) similarly increases with collateral decrease of spacing between them, the prediction method is essential to locate high-n λ's. (For example, the set of all λ_n up to $n = 20$ for $f = x - x^2$ to 29-place precision requires just a few minutes of CDC 6600 time.)

Analogous to δ_n, one can also compute the rate of convergence of δ_n to δ through δ_n':

$$\delta_n' \equiv (\delta_{n+1} - \delta_n)/(\delta_{n+2} - \delta_{n+1})$$

With λ_n of 29-place accuracy, δ_n converges to δ to 13 places by $n = 20$ and δ_n' converges to three or four places. We quote some typical results in Tables I–III. Observe that for an f symmetric about its maximum, $\delta' \simeq \delta$ for $z \leqslant 2$, whereas $\delta' < \delta$ for $z > 2$. As related in the text, we shall explain these results in the sequel to this work.

With the λ_n determined, the parameter α is next obtained. We transform to variables in which $\bar{x} = 0$. The element of the limit cycle nearest to $\bar{x}$ is obtained as the 2^{n-1} iterate of $\bar{x}$,

$$z_n \equiv (\lambda_n f)^{(2^{n-1})}(\bar{x})$$

and the nth rescaling α_n, defined by

$$\alpha_n = -z_n/z_{n+1}$$

These α_n converge to α, also typically to 13 places.

Table I.[a] **Two-Cycle Data for f = 1 − 2x²**

N	λ	δ	δ'
1	0.70710678118654752440084436 21	4.74430946893705	—
2	0.80953772034934631684595410 18	4.67444782765301	2.7504
3	0.83112799388303047024828338 91	4.67079115022921	6.7888
4	0.83574677974388888508230093 95	4.66946164833746	3.7990
5	0.83673564559387058460370949 66	4.66926580979910	5.2553
6	0.83694741858280471080227218 46	4.66921427043589	4.3595
7	0.83699277324830473230907131 62	4.66920445137251	4.8560
8	0.83700248680244259434596829 76	4.66920220132661	4.5641
9	0.83700456714701499933137326 30	4.66920173797283	4.7307
10	0.83700501269305963494575502 66	4.66920163645133	4.6340
11	0.83700510811537583348518876 20	4.66920161499127	4.6896
12	0.83700512855191373187027246 60	4.66920161036023	4.6575
13	0.83700513292879431735839903 44	4.66920160937272	4.6759
14	0.83700513386618810557617655 11	4.66920160916069	4.6684
15	0.83700513406694914924927446 46	4.66920160911533	
16	0.83700513410994601691059292 49	4.66920160910564	
17	0.83700513411915462927322440 07		
18	0.83700513412112683204653650 15		

[a] In this and the following tables, the cycle of size 2^N for two-cycle data and $3 \times 2^{N-1}$ for three-cycle data is referenced by N. The parameter is denoted by λ; $\delta_N = (\lambda_{N+1} - \lambda_N)/(\lambda_{N+2} - \lambda_{N-1})$ and $\delta_N' = (\delta_{N+1} - \delta_N)/(\delta_{N+2} - \delta_{N+1})$.

Table II.[a] Three-Cycle Data for $f = 1 - 2x^2$

N	λ	δ	δ'
1	0.93671705072735084703311116496	3.36345171892599	5.6876
2	0.94151285264239059123560346 46	4.42501749338226	4.0902
3	0.94293870986164364886603081 30	4.61166338330503	4.9053
4	0.94326093620795424932356197 75	4.65729621052512	4.5349
5	0.94333080825184130135150375 59	4.66659897033153	4.7439
6	0.94334581095743025374165224 25	4.66865036240409	4.6264
7	0.94334902586955026529896963 76	4.66908278613357	4.6938
8	0.94334971448657455581687554 47	4.66917625427465	4.6550
9	0.94334986197100690205817374 54	4.66919616732989	4.6774
10	0.94334989355782742767401675 48	4.66920044506702	4.6774
11	0.94334990032276498729658660 91	4.66920135962593	
12	0.94334990177160781150424313 54		
13	0.94334990208190558320533568 25		

[a] See footnote to Table I.

Table III.[a] Two-Cycle Data for $f = x(1 - x^2)$

N	λ	δ	δ'
1	2.1213203435596425732025330 87	4.59165349403582	4.7073
2	2.2629896545363471897845547 81	4.65266937815338	4.6198
3	2.2938433134983575307521580 08	4.66563137254676	4.6704
4	2.3004747021602907471027097 71	4.66843710204515	4.6686
5	2.3018960293360623306427444 23	4.66903785250679	4.6680
6	2.3022004839414695502983032 34	4.66916653116789	4.6700
7	2.3022656910815752484787734 52	4.66919409734649	4.6687
8	2.3022796565590690203670882 64	4.66920000018325	4.6695
9	2.3022826475414960966948147 84	4.66920126453840	4.6690
10	2.3022832881185600388981314 27	4.66920153530566	4.6693
11	2.3022834253105609079220759 03	4.66920159329811	4.6691
12	2.3022834546928867363664720 37	3.66920160571803	4.6692
13	2.3022834609856811769873007 17	4.66920160837802	4.6831
14	2.3022834623334050432933708 93	4.66920160894770	4.2652
15	2.3022834626220462242369870 73	4.66920160906935	
16	2.3022834626838643261562712 19	4.66920160909788	
17	2.3022834626971036705355950 34	4.66920160909268	
18	2.3022834626999393755079069 36		
19	2.3022834627005466538414854 46		

[a] See footnote to Table I.

Finally, one computes the functions (where $\bar{x} = 0$)

$$g_n^*(x) \equiv (-1)^n \alpha_n^{-1} (\lambda_n f)^{(2^n - 1)}(\alpha_n x)$$

so normalized that

$$g_n^*(0) = 1, \qquad g_n^*(1) = 0$$

and observes convergence to g^* (in the interval [0, 1] also to 13 places).

ACKNOWLEDGMENTS

Initial thoughts on this work occurred during the author's stay as Aspen; accordingly, he thanks the Aspen Institute for Physics for their hospitality, and especially E. Kerner for early discussions. A crucial hint at the existence of δ followed from conversation with P. Stein. The author's rapid acquaintance with computational technique was strongly abetted by M. Bolsterli and L. Carruthers. Discussion with E. Larsen and R. Haymaker was especially useful during initial attempts at a comprehension of the existence of δ. To E. Lieb we owe strong thanks for his most profitable, critical remarks as well as for his high enthusiasm. Throughout this research, D. Campbell, F. Cooper, R. Menikoff, M. Nieto, D. Sharp, and P. Stein have offered continued critical interest. Finally, we thank P. Stein and P. Cvitanović for criticism of the manuscript. Beyond these acknowledgments, the author feels the strongest gratitude to Peter Carruthers, whose unflagging support of the author's research career has made this work possible.

REFERENCES

1. N. Metropolis, M. L. Stein, and P. R. Stein, On Finite Limit Sets for Transformations on the Unit Interval, *J. Combinatorial Theory* **15**(1):25 (1973).
2. M. Feigenbaum, The Formal Development of Recursive Universality, Los Alamos preprint LA-UR-78-1155.
3. B. Derrida, A. Gervois, Y. Pomeau, Iterations of Endomorphisms on the Real Axis and Representations of Numbers, Saclay preprint (1977).
4. J. Guckenheimer, *Inventiones Math.* **39**:165 (1977).
5. R. H. May, *Nature* **261**:459 (1976).
6. J. Milnor, W. Thurston, Warwick Dynamical Systems Conference, *Lecture Notes in Mathematics*, Springer Verlag (1974).
7. P. Stefan, *Comm. Math. Phys.* **54**:237 (1977).