Universal Behavior in Nonlinear Systems

by Mitchell J. Feigenbaum

Universal numbers, $\delta = 4.6692016...$ and $\alpha = 2.502907875...$, determine quantitatively the transition from smooth to turbulent or erratic behavior for a large class of nonlinear systems.

There exist in nature processes that can be described as complex or chaotic and processes that are simple or orderly. Technology attempts to create devices of the simple variety: an idea is to be implemented, and various parts executing orderly motions are assembled. For example, cars, airplanes, radios, and clocks are all constructed from a variety of elementary parts each of which, ideally, implements one ordered aspect of the device. Technology also tries to control or minimize the impact of seemingly disordered processes, such as the complex weather patterns of the atmosphere, the myriad whorls of turmoil in a turbulent fluid, the erratic noise in an electronic signal, and other such phenomena. It is the complex phenomena that interest us here.

When a signal is noisy, its behavior from moment to moment is irregular and has no simple pattern of prediction. However, if we analyze a sufficiently long record of the signal, we may find that signal amplitudes occur within narrow ranges a definite fraction of the time. Analysis of another record of the signal may reveal the same fraction. In this case, the noise can be given a *statistical* description. This means that while it is impossible to say what amplitude will appear next in succession, it is possible to estimate the probability or likelihood that the signal will attain some specified range of values. Indeed, for the last hundred years disorderly processes have been taken to be statistical (one has given up asking for a precise causal prediction), so that the goal of a description is to determine what the probabilities are, and from this information to determine various behaviors of interest—for example, how air turbulence modifies the drag on an airplane.

We know that perfectly definite causal and *simple* rules can have statistical (or random) behaviors. Thus, modern computers possess "random number generators" that provide the statistical ingredient in a simulation of an erratic process. However, this generator does nothing more than shift the decimal point in a rational number whose repeating block is suitably long. Accordingly, it is possible to predict what the nth generated number will be. Yet, in a list of successive generated numbers there is such a seeming lack of order that all statistical tests will confer upon the numbers a pedigree of randomness. Technically, the term "pseudorandom" is used to indicate this nature. One now may ask whether the various complex processes of nature themselves might not be merely pseudorandom, with the full import of randomness, which is untestable, a historic but misleading concept. Indeed our purpose here is to explore this possibility. What will prove altogether remarkable is that some very simple schemes to produce erratic numbers behave *identically* to some of the erratic aspects of natural phenomena. More specifically, there is now cogent evidence that the problem of how a fluid changes over from smooth to turbulent flow can be solved through its relation to the simple scheme described in this article. Other natural problems that can be treated in the same way are the behavior of a population from generation to generation and the noisiness of a large variety of mechanical, electrical, and chemical oscillators. Also, there is now evidence that various Hamiltonian systems—those subscribing to classical mechanics, such as the solar system—can come under this discipline.

The feature common to these phenomena is that, as some external parameter (temperature, for example) is varied, the behavior of the system changes from simple to erratic. More precisely, for some range of parameter values, the system exhibits an orderly *periodic* behavior; that is, the system's behavior reproduces itself every *period* of time T. Beyond this range, the behavior fails to reproduce itself after T seconds; it almost does so, but in fact it requires *two* intervals of T to repeat itself. That is, the period has *doubled* to 2T. This new periodicity remains over some range of parameter values until another critical parameter value is reached after which the behavior *almost* reproduces itself after 2T, but in fact, it now requires 4T for reproduction. This process of successive period doubling recurs continually (with the range of parameter values for which the period is $2^n T$ becoming successively smaller as n increases) until, at a certain value of the parameter, it has doubled *ad infinitum*, so that the behavior is no longer periodic. Period doubling is then a characteristic route for a system to follow as it changes over from simple periodic to complex aperiodic motion. All the phenomena mentioned above exhibit period doubling. In the limit of aperiodic behavior, there is a unique and hence *universal* solution common to all systems undergoing period doubling. This fact implies remarkable consequences. For a given system, if we denote by Λ_n the value of the parameter at which its period doubles for the nth time, we find that the values Λ_n converge to Λ_∞ (at which the motion is aperiodic) *geometrically* for large n. This means that

$$\Lambda_\infty - \Lambda_n \propto \delta^{-n} \tag{1}$$

for a fixed value of δ (the *rate* of onset of complex behavior) as n becomes large. Put differently, if we define

$$\delta_n \equiv \frac{\Lambda_{n+1} - \Lambda_n}{\Lambda_{n+2} - \Lambda_{n+1}}\ , \tag{2}$$

δ_n (quickly) approaches the constant value δ. (Typically, δ_n will agree with δ to several significant figures after just a few period doublings.) What is quite remarkable (beyond the fact that there is always a geometric convergence) is that, for all systems undergoing this period doubling, the value of δ is *predetermined* at the universal value

$$\delta = 4.6692016\ldots\ . \tag{3}$$

Thus, this definite number must appear as a natural rate in oscillators, populations, fluids, and all systems exhibiting a period-doubling route to turbulence! In fact, most measurable properties of *any* such system in this aperiodic limit now can be determined, in a way that essentially bypasses the details of the equations governing each specific system because the theory of this behavior is universal over such details. That is, so long as a system possesses certain *qualitative* properties that enable it to undergo this route to complexity, its *quantitative* properties are determined. (This result is analogous to the results of the modern theory of critical phenomena, where a few qualitative properties of the system undergoing a phase transition, notably the dimensionality, determine *universal* critical exponents. Indeed at a *formal* level the two theories are identical in that they are fixed-point theories, and the number δ, for example, can be viewed as a critical exponent.) Accordingly, it is sufficient to study the simplest system exhibiting this phenomenon to comprehend the general case.

Functional Iteration

A random number generator is an example of a simple iteration scheme that has complex behavior. Such a scheme generates the next pseudorandom number by a definite transformation upon the present pseudorandom number. In other words, a certain function is reevaluated successively to produce a sequence of such numbers. Thus, if f is the function and x_0 is a starting number (or "seed"), then $x_0, x_1, \ldots, x_n, \ldots$, where

$$x_1 = f(x_0)$$

$$x_2 = f(x_1)$$

$$\vdots$$

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

$$\vdots$$

is the sequence of generated pseudorandom numbers. That is, they are generated by *functional iteration*. The nth element in the sequence is

$$x_n = f(f(\ldots f(f(x_0))\ldots)) \equiv f^n(x_0)\ , \tag{5}$$

where n is the total number of applications of f. [$f^n(x)$ is not the nth power of f(x); it is the nth *iterate* of f.] A property of iterates worthy of mention is

$$f^n(f^m(x)) = f^m(f^n(x)) = f^{m+n}(x)\ , \tag{6}$$

since each expression is simply m + n applications of f. It is understood that

$$f^0(x) = x \,. \tag{7}$$

It is also useful to have a symbol, $\circ$, for functional iteration (or composition), so that

$$f^n \circ f^m = f^m \circ f^n = f^{m+n} \,. \tag{8}$$

Now f^n in Eq. (5) is itself a definite and computable function, so that x_n as a function of x_0 is known in principle.

If the function f is *linear* as, for example,

$$f(x) = ax \tag{9}$$

for some constant a, it is easy to see that

$$f^n(x) = a^n x \,, \tag{10}$$

so that, for this f,

$$x_n = a^n x_0 \tag{11}$$

is the solution of the *recurrence relation* defined in Eq. (4),

$$x_{n+1} = ax_n \,. \tag{12}$$

Should $|a| < 1$, then x_n geometrically converges to zero at the rate 1/a. This example is special in that the linearity of f allows for the explicit computation of f^n.

We must choose a *nonlinear* f to generate a pseudorandom sequence of numbers. If we choose for our nonlinear f

$$f(x) = a - x^2 \,, \tag{13}$$

then it turns out that f^n is a polynominal in x of order 2^n. This polynomial rapidly becomes unmanageably large; moreover, its coefficients are polynomials in a of order up to 2^{n-1} and become equally difficult to compute. Thus even if $x_0 = 0$, x_n is a polynomial in a of order 2^{n-1}. These polynomials are nontrivial as can be surmised from the fact that for certain values of a, the sequence of numbers generated for almost all starting points in the range $(a - a^2, a)$ possess *all* the mathematical properties of a random sequence. To illustrate this, the figure on the cover depicts the iterates of a similar system in two dimensions:

$$x_{n+1} = y_n - x_n^2$$

$$y_{n+1} = a - x_n \,. \tag{14}$$

Analogous to Eq. (4), a starting coordinate pair (x_0,y_0) is used in Eq. (14) to determine the next coordinate (x_1,y_1). Equation (14) is reapplied to determine (x_2,y_2) and so on. For some initial points, all iterates lie along a definite elliptic curve, whereas for others the iterates are distributed "randomly" over a certain region. It should be obvious that no explicit formula will account for the vastly rich behavior shown in the figure. That is, while the iteration scheme of Eq. (14) is trivial to specify, its nth iterate as a function of (x_0,y_0) is unavailable. Put differently, applying the simplest of *nonlinear* iteration schemes to itself sufficiently many times can create vastly complex behavior. Yet, precisely because the same operation is reapplied, it is conceivable that only a select few self-consistent patterns might emerge where the consistency is determined by the key notion of iteration and *not* by the particular function performing the iterates. These self-consistent patterns do occur in the limit of infinite period doubling and in a well-defined intricate organization that can be determined *a priori* amidst the immense complexity depicted in the cover figure.

The Fixed-Point Behavior of Functional Iterations

Let us now make a direct onslaught against Eq. (13) to see what it possesses. We want to know the behavior of the system after many iterations. As we already know, high iterates of f rapidly become very complicated. One way this growth can be prevented is to have the first iterate of x_0 be precisely x_0 itself. Generally, this is impossible. Rather this condition *determines* possible x_0's. Such a self-reproducing point is called a *fixed point* of f. The sequence of iterates is then $x_0, x_0, x_0, \ldots$ so that the behavior is *static*, or if viewed as periodic, it has period 1.

It is elementary to determine the fixed points of Eq. (13). For future convenience we shall use a modified form of Eq. (13) obtained by a translation in x and some redefinitions:

$$f(x) = 4\lambda x(1 - x) \,, \tag{15}$$

so that as λ is varied, $x = 0$ is always a fixed point. Indeed, the fixed-point condition for Eq. (15),

$$x^* = f(x^*) = 4\lambda x^*(1 - x^*) \,, \tag{16}$$

gives as the two fixed points

$$x^* = 0, \; x_0^* = 1 - 1/4\lambda \,. \tag{17}$$

The maximum value of f(x) in Eq. (15) is attained at $x = \frac{1}{2}$ and is equal to λ. Also, for $\lambda > 0$ and x in the interval (0,1), f(x) is always positive. Thus, if λ is anywhere in the range [0,1], then any iterate of any x in (0,1) is also always in (0,1). Accordingly, in all that follows we shall consider only values of x and λ lying between 0 and 1. By Eq. (16) for $0 \leq \lambda < \frac{1}{4}$, only $x^* = 0$ is within range, whereas for $\frac{1}{4} \leq \lambda \leq 1$, both fixed points are within the range. For example, if we set $\lambda = \frac{1}{2}$ and we start at the fixed point $x_0^* = \frac{1}{2}$ (that is, we set $x_0 = \frac{1}{2}$), then $x_1 = x_2 = \ldots = \frac{1}{2}$; similarly if $x_0 = 0$, $x_1 = x_2 = \ldots = 0$, and the problem of computing the nth iterate is obviously trivial.

What if we choose an x_0 *not* at a fixed point? The easiest way to see what happens is to perform a graphical analysis. We graph $y = f(x)$ together with $y = x$.

Where the lines intersect we have $x = y = f(x)$, so that the intersections are precisely the fixed points. Now, if we choose an x_0 and plot it on the x-axis, the ordinate of $f(x)$ at x_0 is x_1. To obtain x_2, we must transfer x_1 to the x-axis before reapplying f. Reflection through the straight line $y = x$ accomplishes precisely this operation. Altogether, to iterate an initial x_0 successively,

1. move *vertically* to the graph of $f(x)$,
2. move *horizontally* to the graph of $y = x$, and
3. repeat steps 1, 2, etc.

Figure 1 depicts this process for $\lambda = \frac{1}{2}$. The two fixed points are circled, and the first several iterates of an arbitrarily chosen point x_0 are shown. What should be obvious is that if we start from any x_0 in (0,1) ($x = 0$ and $x = 1$ excluded), upon continued iteration x_n will converge to the fixed point at $x = \frac{1}{2}$. No matter how close x_0 is to the fixed point at $x = 0$, the iterates diverge away from it. Such a fixed point is termed *unstable*. Alternatively, for almost all x_0 near enough to $x = \frac{1}{2}$ [in this case, all x_0 in (0,1)], the iterates converge towards $x = \frac{1}{2}$. Such a fixed point is termed *stable* or is referred to as an *attractor* of period 1.

Now, if we don't care about the *transient* behavior of the iterates of x_0, but only about some regular behavior that will emerge eventually, then knowledge of the stable fixed point at $x = \frac{1}{2}$ satisfies our concern for the *eventual* behavior of the iterates. In this restricted sense of eventual behavior, the existence of an attractor determines the solution *independently* of the initial condition x_0 provided that x_0 is within the *basin of attraction* of the attractor; that is, that it *is* attracted. The attractor satisfies Eq. (16), which is explicitly independent of x_0. This condition is the basic theme of universal behavior: if an attractor exists, the eventual behavior is independent of the starting point.

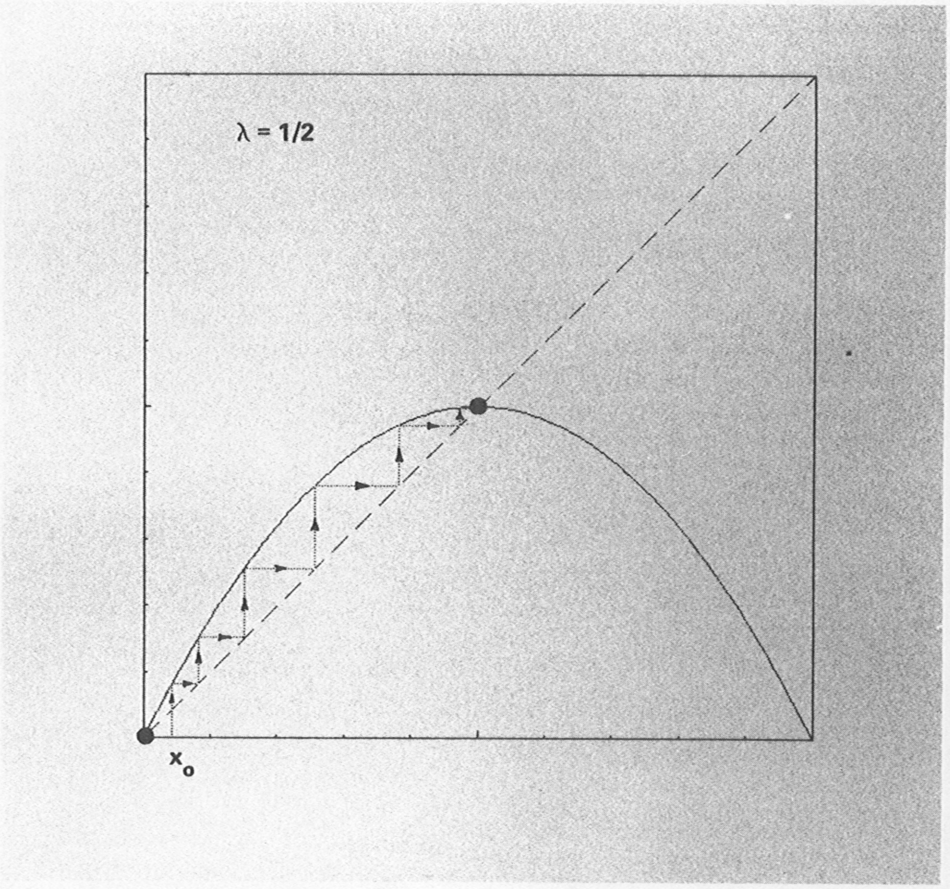

Fig. 1. Iterates of x_0 at $\lambda = 0.5$.

What makes $x = 0$ unstable, but $x = \frac{1}{2}$ stable? The reader should be able to convince himself that $x = 0$ is unstable because *the slope of $f(x)$ at $x = 0$ is greater than 1*. Indeed, if x^* is a fixed point of f and the derivative of f at x^*, $f'(x^*)$, is smaller than 1 in absolute value, then x^* is stable. If $|f'(x^*)|$ is greater than 1, then x^* is unstable. Also, only *stable* fixed points can account for the eventual behavior of the iterates of an arbitrary point.

We now must ask, "For what values of λ are the fixed points attracting?" By Eq. (15), $f'(x) = 4\lambda(1 - 2x)$ so that

$$f'(0) = 4\lambda \tag{18}$$

and

$$f'(x_0^*) = 2 - 4\lambda. \tag{19}$$

For $0 < \lambda < \frac{1}{4}$, only $x^* = 0$ is stable. At $\lambda = \frac{1}{4}$, $x_0^* = 0$ and $f'(x_0^*) = 1$. For $\frac{1}{4} < \lambda < \frac{3}{4}$, x^* is unstable and x_0^* is stable, while at $\lambda = \frac{3}{4}$, $f'(x^*_0) = -1$ and x_0^* also has become unstable. Thus, for $0 < \lambda < \frac{3}{4}$, the eventual behavior is known.

Period 2 from the Fixed Point

What happens to the system when λ is in the range $3/4 < \lambda < 1$, where there are no attracting fixed points? We will see that as λ increases slightly beyond $\lambda = 3/4$, f undergoes period doubling. That is, instead of having a stable cycle of period 1 corresponding to one fixed point, the system has a stable cycle of period 2; that is, the cycle contains two points. Since these two points are fixed points of the function f^2 (f applied twice) and since stability is determined by the slope of a function at its *fixed* points, we must now focus on f^2. First, we examine a graph of f^2 at λ just below $3/4$. Figures 2a and b show f and f^2, respectively, at $\lambda = 0.7$.

To understand Fig. 2b, observe first that, since f is symmetric about its maximum at $x = 1/2$, f^2 is also symmetric about $x = 1/2$. Also, f^2 must have a fixed point whenever f does because the second iterate of a fixed point is still that same point. The main ingredient that determines the period-doubling behavior of f as λ increases is the relationship of the slope of f^2 to the slope of f. This relationship is a consequence of the chain rule. By definition

$$x_2 = f^2(x_0) ,$$

where

$$x_1 = f(x_0), \; x_2 = f(x_1) .$$

We leave it to the reader to verify by the chain rule that

$$f^{2\prime}(x_0) = f'(x_0)f'(x_1) \tag{20}$$

and

$$f^{n\prime}(x_0) = f'(x_0)f'(x_1) \ldots f'(x_{n-1}) , \tag{21}$$

an elementary result that determines period doubling. If we start at a fixed point of f and apply Eq. (20) to $x_0 = x^*$, so that $x_2 = x_1 = x^*$, then

$$f^{2\prime}(x^*) = f'(x^*)f'(x^*) = [f'(x^*)]^2 . \tag{22}$$

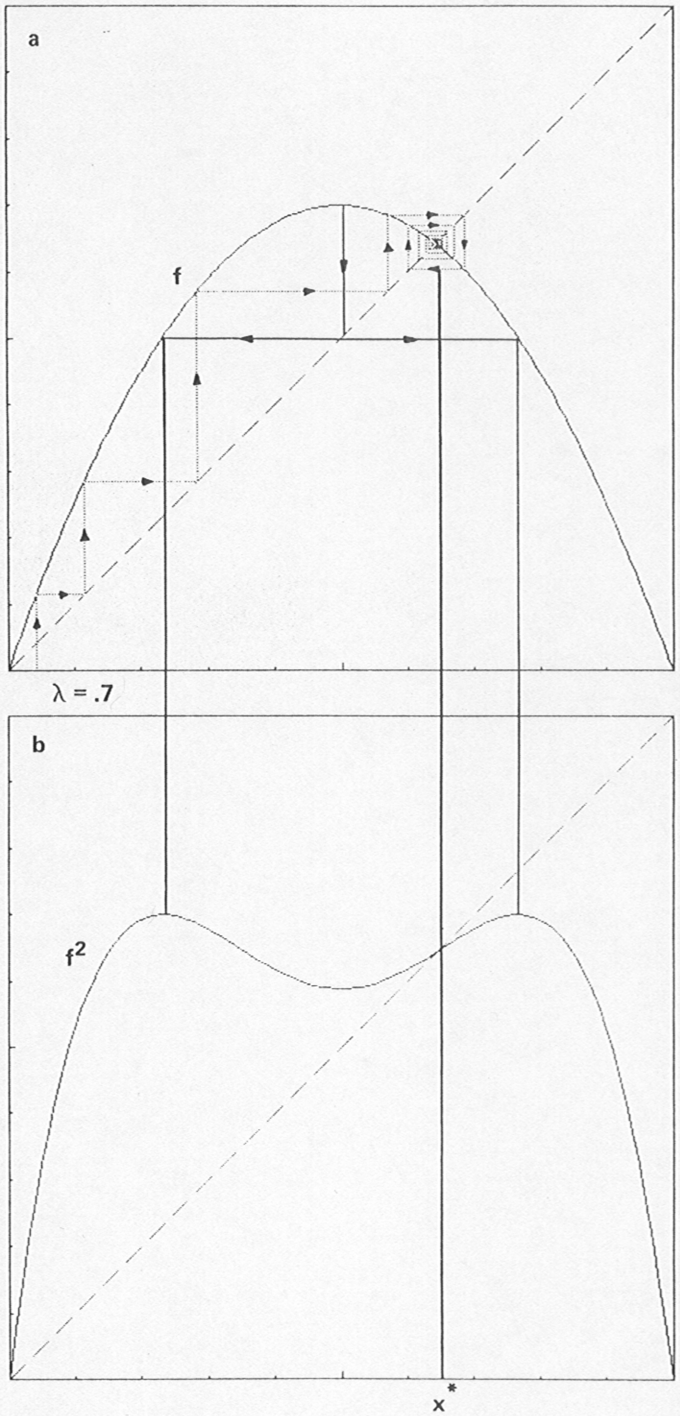

Fig. 2. $\lambda = 0.7$. x^* is the stable fixed point. The extrema of f^2 are located in (a) by constructing the inverse iterates of $x = 0.5$.

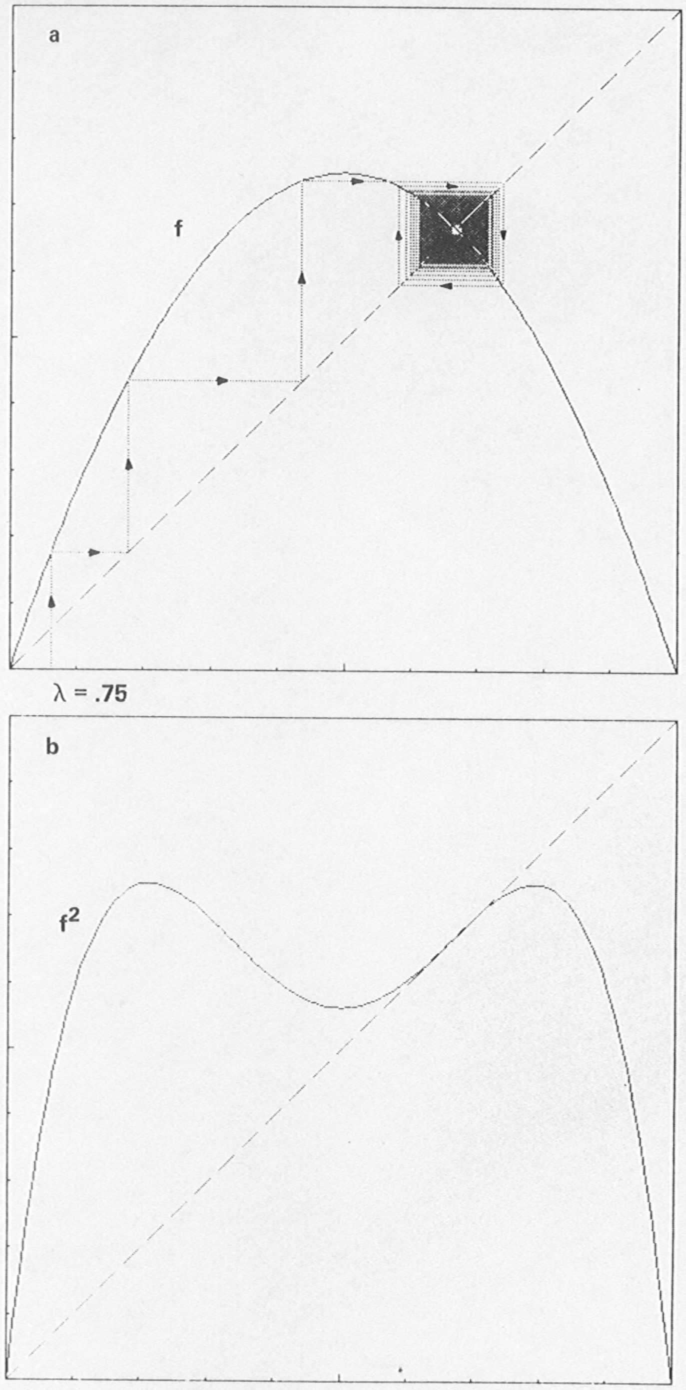

Fig. 3. λ = 0.75. (a) depicts the slow convergence to the fixed point. f^2 osculates about the fixed point.

Since at $\lambda = 0.7$, $|f'(x^*)| < 1$, it follows from Eq. (22) that

$$0 < f^{2\prime}(x^*) < 1 \; .$$

Also, if we start at the extremum of f, so that $x_0 = \frac{1}{2}$ and $f'(x_0) = 0$, it follows from Eq. (21) that

$$f^{n\prime}(\tfrac{1}{2}) = 0 \qquad (23)$$

for all n. In particular, f^2 is extreme (and a minimum) at $\frac{1}{2}$. Also, by Eq. (20), f^2 will be extreme (and a maximum) at the x_0 that will iterate under f to $x = \frac{1}{2}$, since then $x_1 = \frac{1}{2}$ and $f'(x_1) = 0$. These points, the *inverses* of $x = \frac{1}{2}$, are found by going *vertically* down along $x = \frac{1}{2}$ to $y = x$ and then *horizontally* to $y = f(x)$. (Reverse the arrows in Fig. 1, and see Fig. 2a.) Since f has a maximum, there are *two* horizontal intersections and, hence, the two maxima of Fig. 2b. *The ability of f to have complex behaviors is precisely the consequence of its double-valued inverse*, which is in turn a reflection of its possession of an extremum. A monotone f, one that always increases, *always* has simple behaviors, whether or not the behaviors are easy to compute. A *linear* f is always monotone. The f's we care about always fold over and so are *strongly* nonlinear. This folding nonlinearity gives rise to universality. Just as linearity in any system implies a definite method of solution, folding nonlinearity in any system also implies a definite method of solution. In fact folding nonlinearity in the aperiodic limit of period doubling in any system is solvable, and many systems, such as various coupled nonlinear differential equations, possess this nonlinearity.

To return to Fig. 2b, as $\lambda \rightarrow \frac{3}{4}$ and the maximum value of f increases to $\frac{3}{4}$, $f'(x^*) \rightarrow -1$ and $f^{2\prime}(x^*) \rightarrow +1$. As λ increases beyond $\frac{3}{4}$, $|f'(x^*)| > 1$ and $f^{2\prime}(x^*) > 1$, so that f^2 must develop two new fixed points beyond those of f; that is, f^2 will cross $y = x$ at two more points. This transition is depicted in Figs. 3a and b for f and f^2, respectively, at $\lambda =$

0.75, and similarly in Figs. 4a and b at $\lambda = 0.785$. (Observe the exceptionally slow convergence to x^* at $\lambda = 0.75$, where iterates approach the fixed point not geometrically, but rather with deviations from x^* inversely proportional to the square root of the number of iterations.) Since x_1^* and x_2^*, the new fixed points of f^2, are *not* fixed points of f, it must be that f sends one into the other:

$$x_1^* = f(x_2^*)$$

and

$$x_2^* = f(x_1^*) \ .$$

Such a *pair of points*, termed a *2-cycle*, is depicted by the limiting unwinding circulating square in Fig. 4a. Observe in Fig. 4b that the slope of f^2 is in excess of 1 at the fixed point of f and so is an unstable fixed point of f^2, while the two new fixed points have slopes smaller than 1, and so are *stable*; that is, every two iterates of f will have a point attracted toward x_1^* if it is sufficiently close to x_1^* or toward x_2^* if it is sufficiently close to x_2^*. This means that the sequence under f,

$$x_0, x_1, x_2, x_3, \ldots ,$$

eventually becomes arbitrarily close to the sequence

$$x_1^*, x_2^*, x_1^*, x_2^*, \ldots ,$$

so that this is a stable 2-cycle, or an *at-*

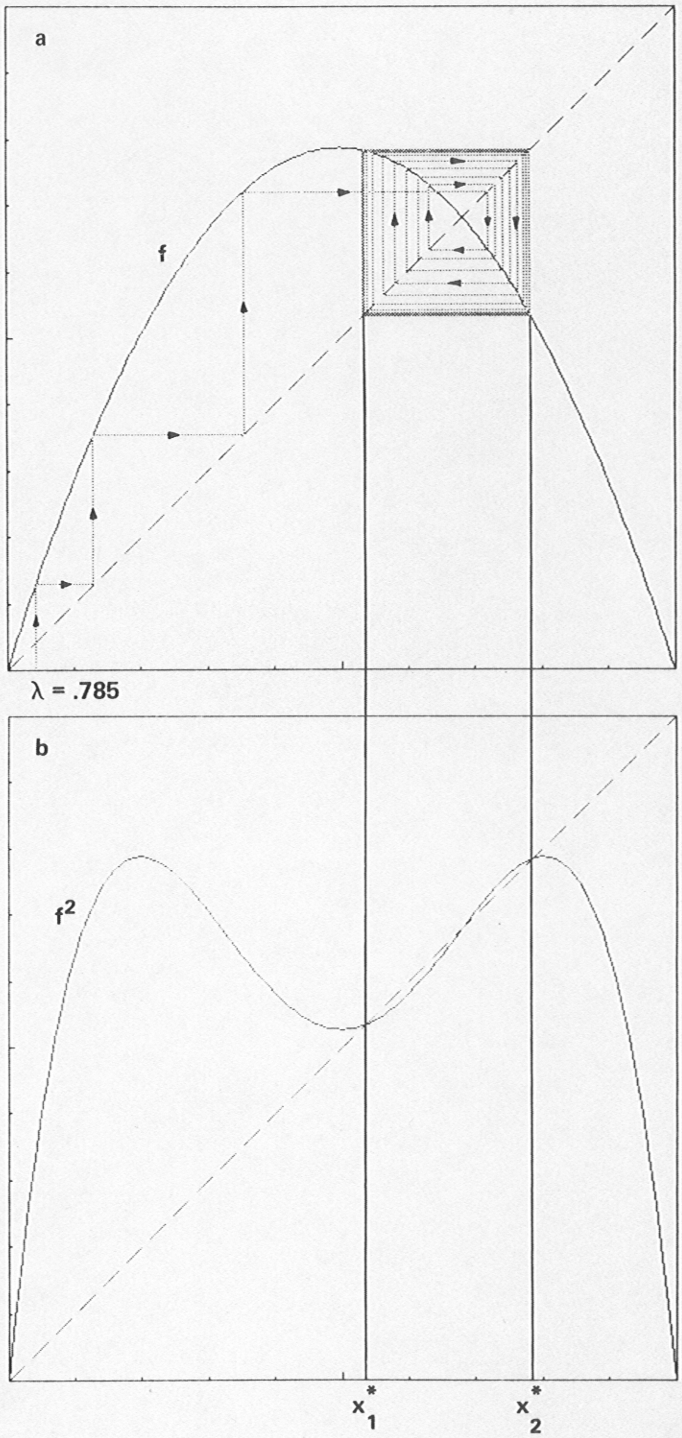

Fig. 4. ***$\lambda = 0.785$. (a) shows the outward spiralling to a stable 2-cycle. The elements of the 2-cycle, x_1^* and x_2^*, are located as fixed points in (b).***

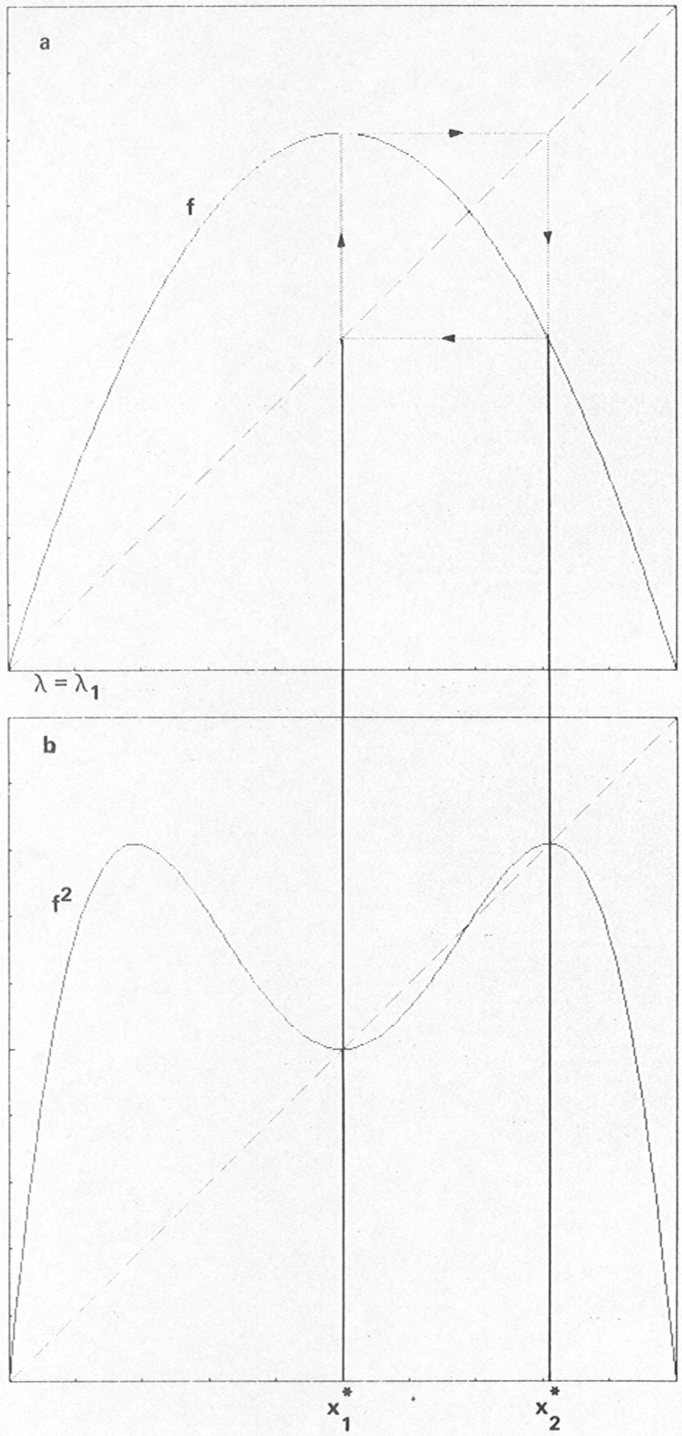

Fig. 5. $\lambda = \lambda_1$. A superstable 2-cycle. x_1^* and x_2^* are at extrema of f^2.

tractor of period 2. Thus, we have observed for Eq. (15) the first period doubling as the parameter λ has increased.

There is a point of paramount importance to be observed; namely, f^2 has the same slope at x_1^* and at x_2^*. This point is a direct consequence of Eq. (20), since if $x_0 = x_1^*$, then $x_1 = x_2^*$, and vice versa, so that the product of the slopes is the same. More generally, if $x_1^*, x_2^*, ..., x_n^*$ is an n-cycle so that

$$x_{r+1}^* = f(x_r^*) \qquad r = 1, 2, ..., n - 1$$

and

$$x_1^* = f(x_n^*)\,, \tag{24}$$

then *each* is a fixed point of f^n with identical slopes:

$$x_r^* = f^n(x_r^*) \qquad r = 1, 2, ..., n \tag{25}$$

and

$$f^{n\prime}(x_r^*) = f'(x_1^*) \ldots f'(x_n^*)\,. \tag{26}$$

From this observation will follow period doubling *ad infinitum*.

As λ is increased further, the minimum at $x = \frac{1}{2}$ will drop as the slope of f^2 through the fixed point of f increases. At some value of λ, denoted by λ_1, $x = \frac{1}{2}$ will become a fixed point of f^2. Simultaneously, the right-hand maximum will also become a fixed point of f^2. [By Eq. (26), both elements of the 2-cycle have slope 0.] Figures 5a and b depict the situation that occurs at $\lambda = \lambda_1$.

Period Doubling *Ad Infinitum*

We are now close to the end of this story. As we increase λ further, the minimum drops still lower, so that both x_1^* and x_2^* have negative slopes. At some parameter value, denoted by Λ_2, the slope at *both* x_1^* and x_2^* becomes equal to −1. Thus at Λ_2 the same situation has developed for f^2 as developed for f at $\Lambda_1 = 3/4$. This transitional case is depicted in Figs. 6a and b. Accordingly, just as the fixed point of f at Λ_1 issued into being a 2-cycle, so too does *each* fixed point of f^2 at Λ_2 create a 2-cycle, which in turn is a 4-cycle of f. That is, we have now encountered the second period doubling.

The manner in which we were able to follow the creation of the 2-cycle at Λ_1 was to anticipate the presence of period 2, and so to consider f^2, which would resolve the cycle into a pair of fixed points. Similarly, to resolve period 4 into fixed points we now should consider f^4. Beyond being the fourth iterate of f, Eq. (8) tells us that f^4 can be computed from f^2:

$$f^4 = f^2 \circ f^2 .$$

From this point, we can abandon f itself, and take f^2 as the "fundamental" function. Then, just as f^2 was constructed by iterating f with itself we now iterate f^2 with itself. The manner in which f^2 reveals itself as being an iterate of f is the slope equality at the fixed points of f^2, which we saw imposed by the chain rule. Since the operation of the chain rule is "automatic," we actually needed to consider only the fixed point of f^2 nearest to $x = 1/2$; the behavior of the other fixed point is slaved to it. Thus, at the level of f^4, we again need to focus on only the fixed point of f^4 nearest to $x = 1/2$: the other *three* fixed points are similarly slaved to it. Thus, a recursive scheme has been unearthed. We now increase λ to λ_2, so that the fixed point of f^4 nearest to $x = 1/2$ is again at $x = 1/2$ with slope 0.

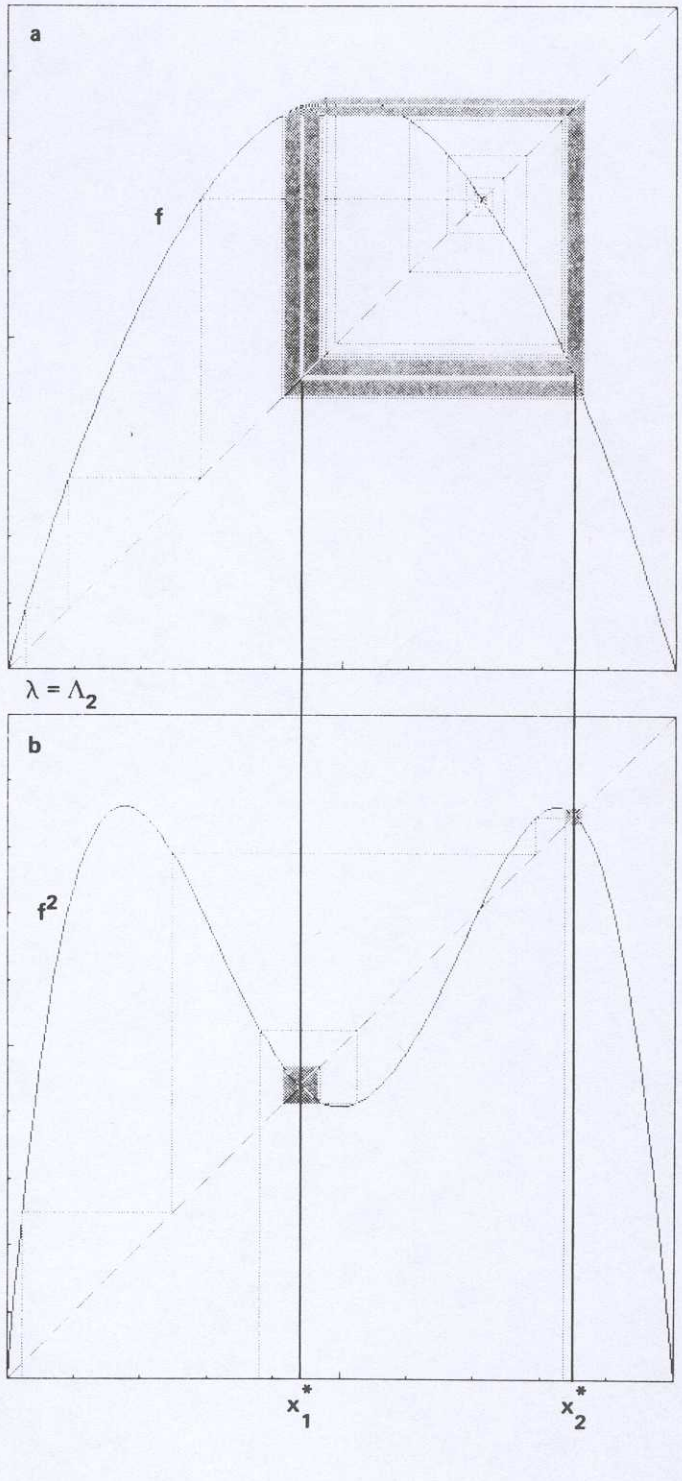

***Fig. 6.** $\lambda = \Lambda_2$. x_1^* and x_2^* in (b) have the same slow convergence as the fixed point in Fig. 3a.*

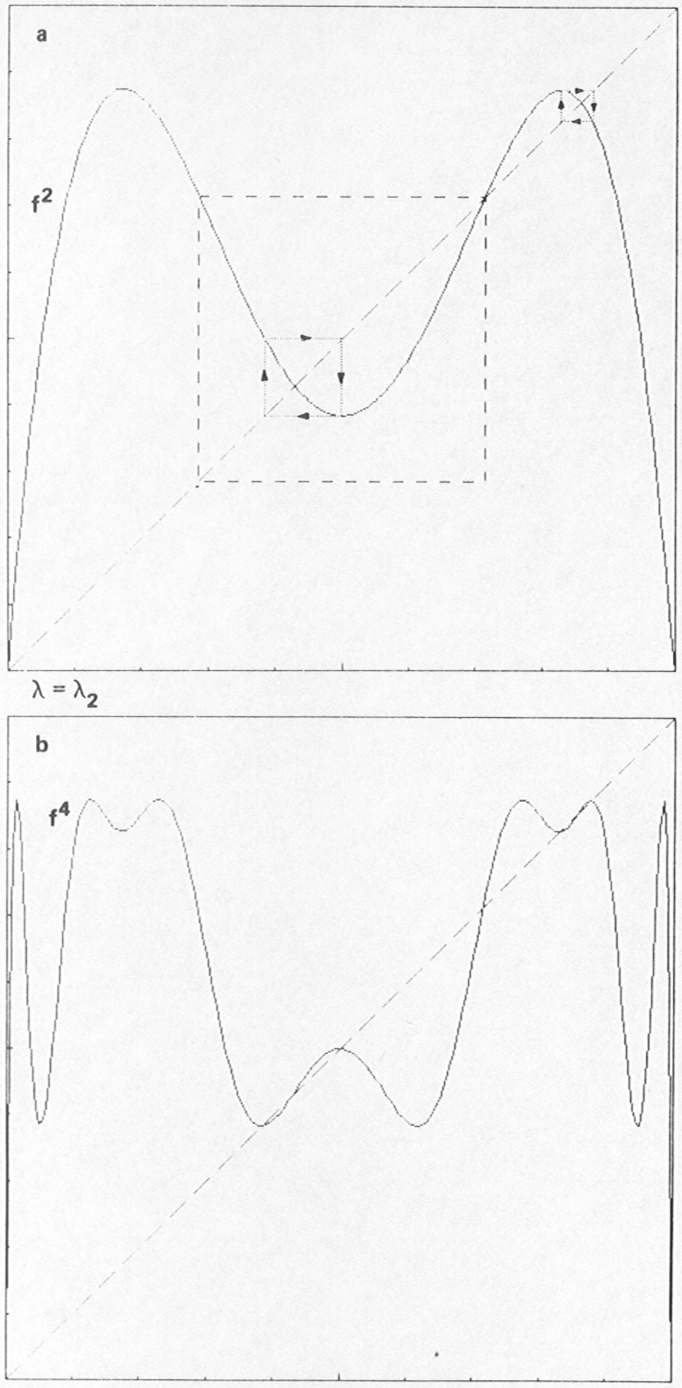

Fig. 7. $\lambda = \lambda_2$. A superstable 4-cycle. The region within the dashed square in (a) should be compared with all of Fig. 5a.

Figures 7a and b depict this situation for f^2 and f^4, respectively. When λ increases further, the maximum of f^4 at $x = \frac{1}{2}$ now moves up, developing a fixed point with negative slope. Finally, at Λ_3 when the slope of this fixed point (as well as the other three) is again -1, each fixed point will split into a pair giving rise to an 8-cycle, which is now stable. Again, $f^8 = f^4 \circ f^4$, and f^4 can be viewed as fundamental. We define λ_3 so that $x = \frac{1}{2}$ again is a fixed point, this time of f^8. Then at Λ_4 the slopes are -1, and another period doubling occurs. Always,

$$f^{2^{n+1}} = f^{2^n} \circ f^{2^n}. \tag{27}$$

Provided that a constraint on the range of λ does not prevent it from decreasing the slope at the appropriate fixed point past -1, this doubling must recur *ad infinitum.*

Basically, the mechanism that f^{2^n} uses to period double at Λ_{n+1} is the same mechanism that $f^{2^{n+1}}$ will use to double at Λ_{n+2}. The function $f^{2^{n+1}}$ is constructed from f^{2^n} by Eq. (27), and similarly $f^{2^{n+2}}$ will be constructed from $f^{2^{n+1}}$. Thus, there is a definite operation that, by acting on functions, creates functions; in particular, the operation acting on f^{2^n} at Λ_{n+1}, (or better, f^{2^n} at λ_n) will determine $f^{2^{n+1}}$ at λ_{n+1}. Also, since we need to keep track of f^{2^n} only in the interval including the fixed point of f^{2^n} closest to $x = \frac{1}{2}$ and since this interval becomes increasingly small as λ increases, the part of f that generates this region is also the restriction of f to an increasingly small interval about $x = \frac{1}{2}$. (Actually, slopes of f at points farther away also matter, but these merely set a "scale," which will be eliminated by a rescaling.) The behavior of f away from $x = \frac{1}{2}$ is immaterial to the period-doubling behavior, and in the limit of large n only the *nature of f's maximum* can matter. This means that in the infinite period-doubling limit, all functions with a quadratic extremum will have identical behavior. [$f''(\frac{1}{2}) \neq 0$ is the generic circumstance.] Therefore, the operation on functions will have a *stable fixed point* in the space of functions, which will be the common universal limit of high iterates of any specific function. To determine this universal limit we must enlarge our scope vastly, so that the role of the starting point, x_0, will be played by an arbitrary *function*; the attracting fixed point will become a universal function obeying an equation implicating only itself. The role of the function in the equation $x_0 = f(x_0)$ now must be played by an *operation* that yields a new function when it is performed upon a function. In fact, the heart of this operation is the functional composition of Eq. (27). If we can determine the exact operator and actually can solve *its* fixed-point problem, we shall understand why a special number, such as δ of Eq. (3), has emerged independently of the specific system (the starting function) we have considered.

The Universal Limit of High Iterates

In this section we sketch the solution to the fixed-point problem. In Fig. 7a, a dashed square encloses the part of f^2 that we must focus on for all further period doublings. This square should be compared with the unit square that comprises all of Fig. 5a. If the Fig. 7a square is reflected through $x = 1/2$, $y = 1/2$ and then *magnified* so that the circulation squares of Figs. 4a and 5a are of equal size, we will have in each square a piece of a function that has the same kind of maximum at $x = 1/2$ and falls to zero at the right-hand lower corner of the circulation square. Just as f produced this second curve of f^2 in the square as λ increased from λ_1 to λ_2, so too will f^2 produce another curve, which will be similar to the other two when it has been magnified suitably and reflected twice. Figure 8 shows this superposition for the first *five* such functions; at the resolution of the figure, observe that the last three

A DISCOVERY

The inspiration for the universality theory came from two sources. First, 1971 N. Metropolis, M. Stein, and P. Stein (all in the LASL Theoretical Di sion) discovered a curious property of iterations: as a parameter is varied, t behavior of iterates varies in a fashion independent of the particular functi iterated. In particular for a large class of functions, if at some value of t parameter a certain cycle is stable, then as the parameter increases, the cycle replaced successively by cycles of doubled periods. This period doubling co tinues until an infinite period, and hence erratic behavior, is attained.

Second, during the early 1970s, a scheme of mathematics called dynamic system theory was popularized, largely by D. Ruelle, with the notion of "strange attractor." The underlying questions addressed were (1) how could purely causal equation (for example, the Navier-Stokes equations that descri fluid flow) come to demonstrate highly erratic or statistical properties and (how could these statistical properties be computed. This line of thought merg with the iteration ideas, and the limiting infinite "cycles" of iteration syster came to be viewed as a possible means to comprehend turbulence. Indeed became inspired to study the iterates of functions by a talk on such matters S. Smale, one of the creators of dynamical system theory, at Aspen in the su mer of 1975.

My first effort at understanding this problem was through the compl analytic properties of the generating function of the iterates of the quadra map

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

This study clarified the mechanism of period doubling and led to a rather d ferent kind of equation to determine the values of λ at which the period do bling occurs. The new equations were intractable, although approximate sol tions seemed possible. Accordingly, when I returned from Aspen, numerically determined some parameter values with an eye toward discerni some patterns. At this time I had never used a large computer—in fact my s computing power resided in a programmable pocket calculator. Now, su machines are very slow. A particular parameter value is obtained iterativ (by Newton's method) with each step of iteration requiring 2^n iterates of t map. For a 64-cycle, this means 1 minute per step of Newton's method. At t same time as n increased, it became an increasingly more delicate matter locate the desired solution. However, I immediately perceived the λ_n's we converging geometrically. This enabled me to predict the next value with i creasing accuracy as n increased, and so required just one step of Newto method to obtain the desired value. To the best of my knowledge, this observ tion of geometric convergence has never been made independently, for the si ple reason that the solutions have always been performed automatically on large and fast computers!

That a geometric convergence occurred was already a surprise. I was interested in this for two reasons: first, to gain insight into my theoretical work, as already mentioned, and second, because a convergence rate is a number invariant under all smooth transformations, and so of mathematical interest. Accordingly, I spent a part of a day trying to fit the convergence rate value, 4.669, to the mathematical constants I knew. The task was fruitless, save for the fact that it made the number memorable.

At this point I was reminded by Paul Stein that period doubling isn't a unique property of the quadratic map, but also occurs, for example, in

$$x_{n+1} = \lambda \sin \pi x_n \ .$$

However, my generating function theory rested heavily on the fact that the nonlinearity was simply quadratic and not transcendental. Accordingly, my interest in the problem waned.

Perhaps a month later I decided to determine the λ's in the transcendental case numerically. This problem was even slower to compute than the quadratic one. Again, it became apparent that the λ's converged geometrically, and altogether amazingly, the convergence rate was the same 4.669 that I remembered by virtue of my efforts to fit it.

Recall that the work of Metropolis, Stein, and Stein showed that precise qualitative features are independent of the specific iterative scheme. Now I learned that precise quantitative features also are independent of the specific function. This discovery represents a complete inversion of accustomed ritual. Usually one relies on the fact that similar equations will have qualitatively similar behavior, but quantitative predictions depend on the details of the equations. The universality theory shows that qualitatively similar equations have the identical *quantitative* behavior. For example, a system of differential equations naturally determines certain maps. The computation of the actual analytic form of the map is generally well beyond present mathematical methods. However, should the map exhibit period doubling, then precise quantitative results are available from the universality theory because the theory applies independently of which map it happens to be. In particular, certain fluid flows have now been experimentally observed to become turbulent through period doubling (subharmonic bifurcations). From this one fact we know that the universality theory applies—and indeed correctly determines the precise way in which the flow becomes turbulent, without any reference to the underlying Navier-Stokes equations.

curves are coincident. Moreover, the scale reduction that f^2 will determine for f^4 is based solely on the functional composition, so that if these curves for f^{2^n}, $f^{2^{n+1}}$, *converge* (as they obviously do in Fig. 8), the scale reduction from level to level will *converge to a definite constant.* But the width of each circulation square is just the distance between $x = \frac{1}{2}$ when it is a fixed point of f^{2^n} and the fixed point of f^{2^n} next nearest to $x = \frac{1}{2}$ (Figs. 7a and b). That is, asymptotically, *the separation of adjacent elements of period-doubled attractors is reduced by a constant value from one doubling to the next.* Also from one doubling to the next, this next nearest element *alternates* from one side of $x = \frac{1}{2}$ to the other. Let d_n denote the algebraic distance from $x = \frac{1}{2}$ to the nearest element of the attractor cycle of period 2^n, in the 2^n-cycle at λ_n. A positive number α scales this distance down in the 2^{n+1}-cycle at λ_{n+1}:

$$\frac{d_n}{d_{n+1}} \sim -\alpha \ . \qquad (28)$$

But since rescaling is determined only by functional composition, there is some function that composed with itself will *reproduce* itself reduced in scale by $-\alpha$. The function has a quadratic maximum at $x = \frac{1}{2}$, is symmetric about $x = \frac{1}{2}$, and can be scaled by hand to equal 1 at $x = \frac{1}{2}$. Shifting coordinates so that $x = \frac{1}{2} \rightarrow x = 0$, we have

$$-\alpha g(g(x/\alpha)) = g(x) \ . \qquad (29)$$

Substituting $g(0) = 1$, we have

$$g(1) = -\frac{1}{\alpha} \ . \qquad (30)$$

Accordingly, Eq. (29) is a definite equation for a function g depending on x through x^2 and having a maximum of 1 at $x = 0$. There is a unique smooth solution to Eq. (29), which determines

$$\alpha = 2.502907875 \ldots \ . \qquad (31)$$

Knowing α, we can predict through Eq. (28) a definite scaling law binding on the iterates of any scheme possessing period doubling. The law has, indeed, been amply verified experimentally. By Eq. (29), we see that the relevant operation upon functions that underlies period doubling is functional composition followed by magnification, where the magnification is determined by the fixed-point condition of Eq. (29) with the function g the fixed point in this space of functions. However, Eq. (29) does not describe a stable fixed point because we have not incorporated in it the parameter increase from λ_n to λ_{n+1}. Thus, g is not the limiting function of the curves in the circulation squares, although it is intimately related to that function. The full theory is described in the next section. Here we merely state that we can determine the limiting function and thereby can *determine the location of the actual elements of limiting 2^n-cycles*. We also have established that g is an unstable fixed point of functional composition, where the rate of divergence away from g is precisely δ of Eq. (3) and so is computable. Accordingly, there is a full theory that determines, in a precise quantitative way, the aperiodic limit of functional iterations with an *unspecified* function f.

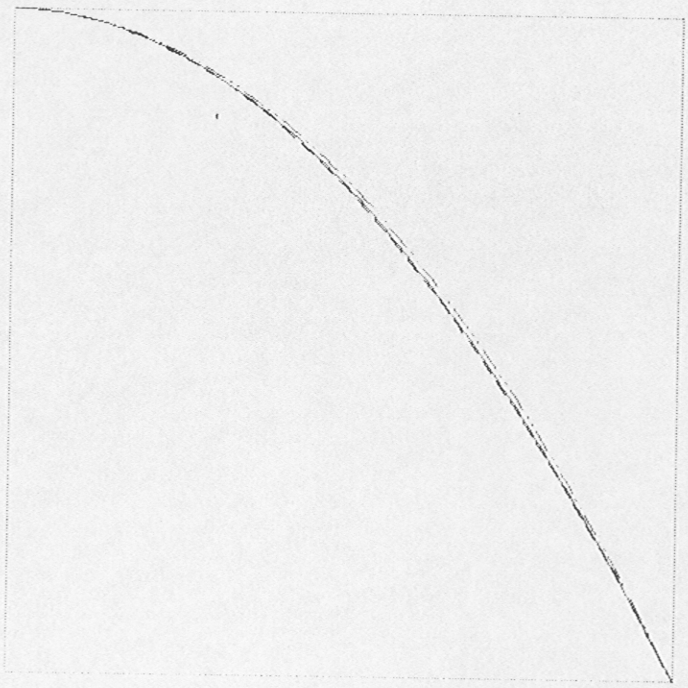

Fig. 8. The superposition of the suitably magnified dotted squares of $f^{2^{n-1}}$ at λ_n (as in Figs. 5a, 7a, ...).

Some Details of the Full Theory

Returning to Eq. (28), we are in a position to describe theoretically the universal scaling of high-order cycles and the convergence to a universal limit. Since d_n is the distance between $x = \frac{1}{2}$ and the element of the 2^n-cycle at λ_n nearest to $x = \frac{1}{2}$ and since this nearest element is the 2^{n-1} iterate of $x = \frac{1}{2}$ (which is true because these two points were coincident before the n^{th} period doubling began to split them apart), we have

$$d_n = f^{2^{n-1}}(\lambda_n, \tfrac{1}{2}) - \tfrac{1}{2}. \qquad (32)$$

For future work it is expedient to perform a coordinate translation that moves $x = \frac{1}{2}$ to $x = 0$. Thus, Eq. (32) becomes

$$d_n = f^{2^{n-1}}(\lambda_n, 0). \qquad (33)$$

Equation (28) now determines that the rescaled distances,

$$r_n \equiv (-\alpha)^n\, d_{n+1}.$$

will converge to a definite finite value as $n \to \infty$. That is,

$$\lim_{n\to\infty} (-\alpha)^n f^{2^n}(\lambda_{n+1}, 0) \qquad (34)$$

must exist if Eq. (28) holds.

However, from Fig. 8 we know something stronger than Eq. (34). When the n^{th} iterated function is *magnified* by $(-\alpha)^n$, it converges to a definite function. Equation (34) is the value of this function at $x = 0$. After the magnification, the convergent functions are given by

$$(-\alpha)^n f^{2^n}(\lambda_{n+1}, x/(-\alpha)^n).$$

Thus,

$$g_1(x) \equiv \lim_{n\to\infty} (-\alpha)^n f^{2^n}(\lambda_{n+1}, x/(-\alpha)^n) \qquad (35)$$