Chapter 5: Scaling laws and self reference

Goals:
- To understand why the scaling exhibited by period-doubling bifurcation sequences of one-dimensional maps is universal, and to calculate the scaling exponents.
- To examine other natural processes which give rise to fractal structures.

This chapter continues our study of dilation-symmetric structures, or fractals. These structures are sometimes called “self-similar”. First, we will pursue our investigation of the period-doubling bifurcation sequence of the logistic map and get some understanding of why the scaling exponents $\alpha$ and $\delta$ which characterize it are universal. We hope that this will give you some sense of how understanding fractals can give new insight into natural phenomena. Then we will survey a few other systems in nature in which fractals arise. Some of these other systems are not nearly so well understood, so perhaps you will get some idea of the scope of the question of why fractals are seen in many different situations.

A. The universality of the period-doubling bifurcation sequence.

In the previous chapter and in Required Project 3 we investigate the scaling behavior of the period-doubling sequence of the logistic map. We saw that the sequence of $r$-values for the orbits of period $2^k$ obeys

$$
\lim_{k \to \infty} \frac{r_{k+1} - r_{k-2}}{r_k - r_{k-1}} = \delta,
$$

where $\delta = 4.7...$, and that the sequence of values $y_k \equiv f^{(2^k-1)}(x)\big|_{1/2}$ satisfies

$$
\lim_{k \to \infty} \frac{y_{k+1} - y_{k-2}}{y_k - y_{k-1}} = -\alpha,
$$

with $\alpha = 2.5...$. These results mean that the bifurcation diagram looks the same when it is magnified about the point $(x = 1/2, r = r_\infty = 3.57...)$ by a factor $-\alpha$ in the $x$-direction and a factor $\delta$ in the $r$-direction. This mapping of the bifurcation diagram onto itself under certain rescaling factors is known as dilation symmetry (because the diagram is invariant under dilation of co-ordinate scales). In Required Project 3 you show that $\alpha$ and $\delta$ are not noticeably different for some different choices of the map function.

In this chapter we address the question of why $\alpha$ and $\delta$ have the values that they do and why they are the same for many different mapping functions. We will write an equation that embodies the observation of dilation symmetry, and we will find that this equation determines the values of the exponents $\alpha$ and $\delta$. In other words, knowing that dilation symmetry exists is enough information to determine what the magnification factors must be. This result explains the observation of universality (different maps having the same exponents). The book by Cvitanovic in the reference list is a good resource for further reading about this universality. Several arguments in this chapter are based on this source.

A1: Why smooth functions lead to period-doubling

We first review how the iteration of a smooth function leads to $n$-cycles of ever longer length. We start with some smooth, positive function $f(x) = rf_o(x)$ that vanishes for $x = 0$ and $x = 1$. Such an $f$ must have at least one maximum. Such an $f$ must also cross the line $f = x$ when the “amplitude” $r$ exceeds some positive value that we denote $r_0$. When $r$ is just at this threshold value $r_0$, $f' = 1$, and as $r$ increases from there, the derivative at one of the crossing points becomes smaller than 1, so that the fixed point is stable. We follow this point as $r$ continues to increase, and see that $f'$ becomes more and more negative, eventually reaching $f' = -1$. At this point the fixed point is marginally stable. We denote it as $r_1$. Any further increase in $r$ will drive the fixed point unstable. To see what happens then, we look at the second iterate $f^{(2)}(x)$. Up to this value of $r$ it has had a stable fixed point at the same $x$ value as for $f(x)$. When $f'$ becomes -1 at the fixed point, $f^{(2)}$ becomes +1, i.e., $f' f''$. It becomes tangent to the $f = x$ line, making the $f^{(2)}$ fixed point marginally stable.

To understand what happens when $r$ increases above $r_1$, we have to look at the shape of the $f^{(2)}$ function when $r = r_1$. Specifically, we need to know whether $f^{(2)}$ is tangent to the $f = x$ line from below, with negative
curvature or from above, with positive curvature. The curvature, or second derivative \( d^2f(x)/dx^2 \) at \( x^* \) can readily be calculated by the chain rule, starting from \( df(x)/dx = f'(x)f''(f(x)) \):

\[
\frac{d^2f(x)}{dx^2} = f''(x)f'(f(x)) + f'(x)f''(f(x))f'(x)
\]

Now, at \( x^* \) \( f(x^*) = x^* \), and at \( r = r_1 \) \( f'(x^*) = -1 \). Thus

\[
\frac{d^2f(x)}{dx^2} = f''(x^*)(-1) + (-1)f''(x^*)(-1) = 0
\]

Evidently, \( f^{(2)} \) has an inflection point when \( x = x^* \) and \( r = r_1 \): it neither curves below the \( f = x \) line at its tangent point, nor does curve above the line; it crosses the \( f = x \) line.

As \( r \) increases from this \( r_1 \), the derivative \( f^{(2)}/r \) must increase above 1, so that \( x^* \) becomes unstable. A sketch of \( f^{(2)} \) near this \( x^* \), such as Figure 5.1 quickly convinces one that \( f^{(2)} \) must now cross \( f = x \) three times. To spell out the behavior of this sketch, our former inflection point now crosses the \( f = x \) line from below: \( f^{(2)}(x^*) \) \(< x^* \). But further away from \( x^* \) the inflection behavior dictates that \( f^{(2)}(x^*) \) \( > x^* \). This means that \( f^{(2)} \) must cross \( x \) for some \( x \) near \( x^* \). That is, if there is another fixed point just below the unstable one. We call it \( x_1^* \). At this \( x_1^* \), \( f^{(2)} \) must cross the \( f = x \) line from above, opposite to the crossing at \( x^* \). Thus the slope \( f^{(2)}/x_1^* \) must be less than 1, and this fixed point must be stable. The same reasoning applied to \( x \) slightly greater than \( x^* \) implies that there is a second stable fixed point there. We call it \( x_2^* \).

These two new fixed points of \( f^{(2)} \) cannot be fixed points of \( f \). Near \( r_1 \) the derivative \( f' \) is near \(-1 \), and there can only be one crossing point in the vicinity, viz. the one we have called \( x^* \). Since \( x_1^* \) and \( x_2^* \) are not fixed points of \( f \), each must be in a 2-cycle of \( f \). Clearly they are the same two-cycle. Thus \( r_1 \) represents the onset of a stable 2-cycle.

5.1 Problem: functional form of \( f^{(2)} \) At the point \( r_1 \) and for \( x \) near \( x^* \), we can expand \( f(x) \) in the form

\[
f(x) = x^* - (x - x^*) + b(x - x^*)^2 + c(x - x^*)^3 + \ldots
\]

The \( x^* \) and the coefficients \( b, c, \ldots \) must vary smoothly with the amplitude \( r \). Since \( x^* \) is an inflection point when \( r = r_1 \) the coefficient \( b = 0 \) there. Now for \( r \geq r_1 \), suppose that \( x^* \) is slightly unstable, so that the linear term \(- (x - x^*) \) becomes \(- (1 + \epsilon)(x - x^*) \), where the parameter \( \epsilon \) is arbitrarily small. The other coefficients \( b \) and \( c \) also have changes proportional to \( \epsilon \). Thus our expansion has the form

\[
f(x) = x^* - (1 + \epsilon)(x - x^*) + b'(x - x^*)^2 + \ldots,
\]

where \( b' \) is some other coefficient. The new crossing points \( x_1^* \) and \( x_2^* \) separate from \( x^* \) as \( \epsilon \) increases, as a power of \( \epsilon \). a) Find this power. b) Comment on the numerically observed change of \( x_1^* \) and \( x_2^* \) near \( r_1 \) in view of your finding in a).
We now look at the derivative of \( f^{(2)} \) at its upper fixed point \( x^*_2 \). This derivative is slightly less than 1. As \( r \) continues to increase from \( r_1 \), the derivative \( f^{(2)} \) continues to decrease, and ultimately decreases to -1. Thus \( x^*_2 \) becomes unstable, by the same process that happened with \( f \). But at this point, denoted \( r_2 \), \( f^{(4)} = f^{(2)}(f^{(2)}) \) is just tangent to the \( f = x \) line, as \( f^{(2)} \) was at \( r_1 \). Repeating the reasoning at \( r_1 \) we conclude that \( f^{(4)} \) has an inflection point at \( r_2 \). Slightly above \( r_2 \) and near \( x^*_2 \) \( f^{(4)} \) must look like Figure 5.1. Thus two stable fixed points \( x^*_{21} \) and \( x^*_{22} \) of \( f^{(4)} \) must appear near \( x^* \). Thus \( f^{(4)} \) repeats the process that \( f^{(2)} \) did before it. Meanwhile, what is happening at \( x^*_1 \)? Evidently, when \( x^*_2 \) ceases to be a stable fixed point of \( f^{(2)} \), \( f \) ceases to have a stable 2-cycle. That means that \( x^*_1 \) must become unstable at the same \( r_2 \) as above. Our reasoning at \( x^*_2 \) evidently applies equally to \( x^*_1 \). Thus at \( r_2 \), both points of the two cycle must split to form a 4-cycle of \( f \).

Each time the stable fixed points switch to a higher iterate of \( f \), it means that \( f \) has an \( n \)-cycle or period of twice the length as before. For each period doubling there is a corresponding \( r \) value, that we denote \( r_k \) for the \( k \)th doubling. Thus between \( r_k \) and \( r_{k+1} \) \( f^{(2^k)} \) has stable fixed points that aren’t stable for lower iterates. By continuing to increase \( r \) it appears that we may increase the length of these cycles as much as we wish, leading to behavior of arbitrarily great complexity. For the logistic map and many other \( f \) functions the limit cycles become infinite for some finite \( r_\infty \). This complexity as \( r \to r_\infty \) is what we want to probe.

When we are close to \( r_\infty \), an arbitrarily slight change in the original function \( f \) can evidently cause a qualitative change in behavior: a period doubling. We saw above that each period doubling can be understood by looking at a tiny range of \( x \). We expect the same to be true as we approach \( r_\infty \). A clever observation shows us what \( x \) to focus on. We saw that the derivative of our iterate \( f^{(2^k)} \) starts out at the marginally stable value of 1 at \( r = r_k \). Then it steadily decreases to -1 as we approach \( r_0 \). At some \( r \) value in between, the derivative must vanish. We denote this \( r \) as \( \tilde{r}_k \). When the derivative vanishes at a fixed point, this point is called superstable, as noted in the previous chapter and in Project II. We may determine a superstable fixed point \( x^* \) of any iterate by using the Floquet-multiplier expression for the derivative of \( f^{(2^k)} \): \( f^{(2^k)}(x) = f'(x_1)f'(x_2)...f'(x_{2^k}) \). At a superstable point, this multiplier must vanish. This means that one of the \( f' \)’s in the product must vanish. One point of the \( 2^k \)-cycle must be at a maximum or minimum of \( f(x) \). These maxima and minima don’t shift as we change \( r \); \( r \) is just a multiplicative factor. For example in the logistic map, there is a single maximum located at \( x = 1/2 \). From now on, we’ll focus on some chosen maximum or minimum of \( f \) located at \( x_m \) and consider the region around \( x_m \). At any \( \tilde{r}_k \), \( x_m \) is a fixed point of \( f^{(2^k)} \). This fixed point moves as \( r \) increases to the next period-doubling point \( r_k \). Since this change of \( r \) becomes arbitrarily small, we expect the motion of this \( x^* \) to be small as well.

A2: Limiting dependence on \( x \) To see how the instabilities pile up as \( r \to r_\infty \) we clearly need to know the functional form of \( f^{(2^k)}(x) \) near \( x = x_m \). This functional form cannot be very smooth. We know that each point in the current \( 2^k \)-cycle must be a fixed point of \( f^{(2^k)} \). Thus this function must cross the \( f = x \) line at least \( 2^k \) times. The number of wiggles must double each time \( k \) increases by one. We see this happening as we iterate the logistic map: each time \( k \) increases by one \( f^{(2^k)} \) becomes a polynomial of twice as high an order.

How can we possibly find any well-defined limiting behavior for such an impossibly complicated function? A limiting behavior would mean that \( f^{(2^k)} \) should somehow converge to a limiting smooth function that we can treat with our usual analytical tools. We cannot hope to have a this kind of manageable limit for \( f^{(2^k)} \) as \( k \to \infty \) unless we use a scale of \( x \) which expands to compensate for the increasing wiggliness of \( f^{(2^k)} \). Accordingly we adopt a rescaled co-ordinate for describing \( f^{(2^k)}(x) \) called \( z^{(k)} \) and defined by \( z^{(k)} = a(k)(x - x_m) \). Evidently \( a(k) \) must become huge, so that the tiny region of \( x \) where \( f^{(2^k)} \) looks smooth gets expanded to a finite, nonvanishing interval of \( z \). Now when we increase \( k \) we can hope to account for the increased wiggliness of \( f^{(2^k)}(x) \) near \( x_m \) by means of the scale factor \( a(k) \), so that the remaining \( k \) dependence is smooth. We can now express the iteration equation \( x_{i+2^k} = f^{(2^k)}(x_i) \) in the \( z^{(k)} \) variable using

\[
x = x_m + \frac{z^{(k)}}{a(k)}.
\]
We find
\[ x_m + \frac{z^{(k)}}{a(k)} = f(2^k)(x_m + \frac{z^{(k)}}{a(k)}) \]
or
\[ z^{(k)}_{i+2^k} = a(k)[f(2^k)(x_m + z^{(k)}/a(k)) - x_m] \]
This is an iterative map of the form \( z \rightarrow g_k(z) \), where
\[ g_k(z) \equiv a(k)[f(2^k)(x_m + z/a(k)) - x_m] \] (5.2)

Though the function \( f(2^k) \) has no smooth limit, the corresponding function \( g_k \) in the rescaled \( z^{(k)} \) variable may have a smooth limit. We shall suppose that the sequence of functions \( g_k \) does have a smooth limit called \( g(z) \), and deduce the properties it must have if it exists. Figure 5.2 illustrates the idea using \( f(2) \) and \( f(4) \).

Figure 5.2. Dark line: \( f(2) \) light line, \( f(4) \) near superstable points. A section of the \( f(2) \) is copied below it. This section is reduced by a factor of 2.74 and inverted and placed near the \( f(4) \) curve to show that the \( f(4) \) curve resembles the \( f(2) \) curve except for a change of horizontal and vertical scale.

We may use the known behavior of \( f(2^k) \) to infer the behavior of the hypothesized limiting function \( g \). What we know about \( f(2^k) \) is the way it is related to the next function in the series, \( f(2^{k+1}) \): \[ f(2^{k+1})(x) = f(2^k)(f(2^k)(x)). \] (5.3)

We may express these functions in terms of our reduced functions \( g_k \) and \( g_{k+1} \) using (5.2): \[ f(2^k)(x) = x_m + \frac{g_k(z^{(k)}(x))}{a(k)}. \]
Then Eq. (5.3) becomes
\[ x_m + \frac{g_{k+1}(z^{(k+1)}/a(k+1))}{a(k+1)} = f(2^k) \left( x_m + \frac{g_k(z^{(k)})}{a(k)} \right), \]
\[ = x_m + \frac{g_k(z^{(k)}(x_m + g_k(z^{(k)})/a(k)))}{a(k)}, \]

\[ \text{† This law relates } f(2^k) \text{ to its successor only at the } \text{same value of the parameter } r. \text{ Thus this law doesn’t apply in general to the sequence of superstable fixed points } r_{sk}. \text{ This sequence only becomes } k\text{-independent at the limit point } r_{\infty}. \text{ Thus our strategy to find the limiting behavior of } g_k \text{ works only for the sequence of } f_{r_{sk}} \text{ at } r = r_{\infty}. \text{ These } f’s \text{ that lead to the limiting } g \text{ found here are thus not at superstable fixed points, though we used the superstable fixed points to motivate the scaled variables } z^{(k)}. \text{ This explains why the limiting function } g \text{ found below is not at a superstable fixed point, for which } g(0) \text{ would be 0.} \]
Where \( z^{(k)}(x) \) is given by \( x = x_m + z^{(k)}/a(k) \). Thus

\[
x_m + \frac{g_{k+1}(z^{(k+1)})}{a(k + 1)} = x_m + \frac{g_k(a(k)g(z^{(k)}))}{a(k)}
\]

We remove the \( x_m \) from both sides and cancel the \( a(k) \) and \( 1/a(k) \) factors. We simplify further by defining \( \beta(k) = a(k + 1)/a(k) \) and noting that \( z^{(k+1)} \) is just \( \beta(k)z^{(k)} \). Then the simplified equation reads:

\[
g_{k+1}(\beta(k)z^{(k)}) = \beta(k)g_k(g_k(z^{(k)})).
\]

for a range of \( z^{(k)} \) around zero.

Now we take the limit as \( k \to \infty \). The \( g \) functions go to their assumed limiting form \( g(z) \). The equation can then only make sense if there is also a limiting value for \( \beta(k) \), which we’ll call \( \beta \). We are left with a simple but strange equation for \( g \) called the Feigenbaum equation [1]:

\[
g(\beta z) = \beta g(g(z)).
\]

(5.4)

Can there be functions \( g(z) \) with this property? If so, what are the possible values for the scale factor \( \beta \)? We can learn something about \( g \) by looking at its fixed points. By looking at the definition of \( g_k \) we see that every fixed point of \( f^{(2k)} \) at some \( x^* \) implies a fixed point of \( g_k \) at \( z^* = a(k)(x^* - x_m) \). Thus our limiting function \( g(z) \) must have at least one fixed point. Let’s consider some fixed point \( z^* \) and look at what the Feigenbaum equation (5.4) says: \( g(\beta z^*) = \beta g(g(z^*)) = \beta g(z^*) = \beta z^* \). This says that \( \beta z^* \) is also a fixed point. There must be another fixed point a factor \( \beta \) away from the original one. By the same argument there must be a third fixed point another factor of \( \beta \) away, and so forth. This behavior sets constraints on \( \beta \): it cannot have magnitude less than 1. If it did, our sequence of fixed points would march into the origin, ending all possibility of the smooth \( g(z) \) we have assumed.

The Feigenbaum equation doesn’t determine \( g \) uniquely. There is an arbitrariness in \( g \) that reflects an arbitrariness in our rescaling procedure. If we had expanded all the \( z^{(k)} \) scales by the same factor \( b \), it would not have affected our argument for finding \( g \). If we obtained convergence to a \( g \) function using the original \( z^{(k)} \) variables, we would have obtained it in the expanded \( z^{(k)} \) variables as well. This shift would have introduced an an extra factor \( b \) relating \( f^{(2k)}(x) \) to \( g_k \); thus \( g_k \) would have been divided by \( b \). Then we would have obtained not \( g(z) \) but \( \tilde{g}(z) \equiv g(bz)/b \). One can readily show that if \( g \) satisfies the Feigenbaum equation, then so does \( \tilde{g} \). It is conventional to settle this ambiguity by defining \( g(0) = 1 \). With this convention, one may readily relate \( g \) to \( \beta \) by applying the Feigenbaum equation at \( z = 0 \). This gives \( 1 = \beta g(1) \).

To get further information about \( \beta \), we take the assumed smoothness of \( g \) seriously and approximate it as a truncated Taylor expansion: \( g(z) \simeq g(0) + \frac{1}{2}g''z^2 \). We note that there is no linear term. Indeed, the first derivative of \( g \) must vanish at 0 because the first derivative of \( f^{(2k)}(x) \) vanishes at \( x_m \); we chose \( x_m \) because the derivative of \( f \), and therefore its iterates, vanishes at \( x_m \). Our \( g \) is merely a rescaled version of the \( f \)'s and thus its derivative vanishes also. Then the Feigenbaum equation reads

\[
g(0) + \frac{1}{2}g''\beta^2z^2 = \beta[g(0) + \frac{1}{2}g''(g(0) + \frac{1}{2}g''\beta z^2)^2]
\]

Equating the terms in \( z^2 \) gives \( \frac{1}{2}g''\beta^2 = \beta \frac{1}{2}g''g(0)g'' \), or \( \beta = g(0)g'' \). Equating the constant terms gives \( g(0) = \beta g(0) + \frac{1}{2}g''g(0)^2 \) or \( 1 = \beta[1 + \frac{1}{2}\beta] \). This quadratic equation implies \( \beta = -1 \pm \sqrt{3} \). (Reassuringly, \( \beta \) is insensitive to our convention \( g(0) = 1 \).) One of these roots has a magnitude smaller than 1. This one is incompatible with a smooth \( g(z) \), as explained above. Thus our argument leads us to the conclusion that \( \beta \simeq -2.73 \). It is negative! But a little reflection shows that we might have expected this negative value. The \( f^{(2k)}(x) \) at \( x_m \) alternates between being a maximum and a minimum. (The derivative of \( f^{(2)}(x) \) at the fixed point closest to \( x_m \) must be negative at the threshold of a period doubling; however, the derivative of the next iterate \( f^{(2k+1)}(x) \) must be positive there. If \( f^{(2k)}(x) \) and \( f^{(2k+1)}(x) \) are to have the same smooth form, this suggests that the two iterates should switch from a maximum at \( x_m \) to a minimum, or vice versa.) Thus the \( a(k) \to (constant)\beta^k \) factor relating \( f^{(2)}(x) \) to \( g_k \) must also alternate sign with every \( k \).
The Feigenbaum equation can be solved to successively better approximations by using refinements of our polynomial procedure. We start by assuming that \( g(z) \) is an even polynomial of order \( 2n \) with \( n-1 \) unknown coefficients (supposing that \( g(0) = 1 \)). Then we obtain \( g(g(z)) \) as a polynomial of order \( (2n)^2 \). Equating this with the order-\( 2n \) polynomial on the left side, we obtain \( n \) equations for the \( n-1 \) unknown coefficients, plus the scale factor \( \beta \). By this means one can obtain \( g \) and \( \beta \) to arbitrary accuracy: \( \beta = 2.502907875.... \) A graph of \( g \) is shown in Figure 5.3

\[ \beta = 2.502907875.... \]

Figure 5.3. Plot of iterates \( f^{(16)}, f^{(32)}, \) and \( f^{(64)} \) of the logistic map in the vicinity of \( x = 1/2 \) and \( f^{(N)} = 1/2 \), using rescaled co-ordinates. Specifically, vertical axis is \( \tilde{f} = (f^{(N)}(x) - 1/2)/a_k \), where \( a_k = (f^{(N)}(1/2) - 1/2) \). Thus \( \tilde{f}(1/2) = 1 \) by construction. Horizontal axis is \( \tilde{x} = (x - 1/2)/a_k \). Thus the scale factors for vertical and horizontal scales are the same. Plots span the range \(-30 < \tilde{x} < 30\). The \( r \) value is indicated on the legend and is equal to \( r_\infty \) to the precision given. Plots were made using a Java applet. Evidently, the rescaled functions agree in a range near \( \tilde{x} = 0 \). Convergence improves with higher \( N \).

The significance of this Feigenbaum equation and its solution is that it applies to a wide variety of situations. When we explained it, we didn’t use the fact that our iterated function \( f(x) \) was the logistic map or any other specific function. Thus we expect similar convergence equally much whenever we iterate a

---

\[ \text{It can be shown that there are no odd terms in the Taylor series for } g. \] For example, if one assumes that \( g(z) \) has a cubic, \( a_3z^3 \), term and examines the \( z^3 \) term on each side of Eq. 5.4 to find the coefficient \( a_3 \), one readily verifies that both sides of the equation have factors of \( a_3 \). Thus one is free to set \( a_3 = 0 \) The same thing occurs for each higher odd power. In any case we know that there must be \textit{some} even solution to Eq. 5.4, since logistic-map iterates \( f^{(2k)} \) are all even and converge to \( g \). Further, by applying the methods of the next section one can show that if a tiny odd part is added to \( g \), it becomes smaller under iteration.
function that behaves like \( f(x) \): it has a smooth maximum at some \( x_m \). The resulting function \( g \) satisfies an equation that makes no reference to the iterated function \( f \) that led to it. Thus neither \( g \) nor \( \beta \) can depend on this function. A whole class of \( f(x) \) functions yield one single behavior at the onset of chaos. We say that \( g \) and \( \beta \) are universal. This is the first example of universality we have encountered. The mere existence of dilation symmetry, together with certain qualitative constraints, led to a quantitative behavior that remains unchanged when we change our implementation. Since many forms of time development can be expressed as iterations of a function, these universal properties must show up in a broad class of physical situations as well as in mathematical examples.

The universal scale factor \( \beta \) explains the geometrically spaced fixed points observed as one approaches \( r_\infty \), described at the beginning of the chapter. One may readily verify that the geometrically-spaced fixed points \( z^* \) of \( g \) imply a spacing of fixed points \( x_k \) satisfying \((x_{k-1} - x_{k-2})/(x_k - x_{k-1}) = \beta \). We have thus explained the convergence behavior of the points \( y_k \) noted at the beginning of the chapter. The \( \alpha \) appearing there is simply the negative of the \( \beta \) in the Feigenbaum equation. No one has found a way to express \( \beta \) in terms of simpler constants. It seems to be a new fundamental constant, like Euler’s constant \( e \). Indeed, there is a certain resemblance [2] between the Feigenbaum equation and the defining equation for the exponential function \( \exp(x) \): \( \exp(x + y) = \exp(x) \exp(y) \); \( \exp(0) = 1 \). One may say that \( g \) is to functional composition...
what exp is to functional multiplication.

A3: Limiting dependence on $r$  Our original motivation for finding the Feigenbaum function $g$ was to understand the sequence of $r_k$ that led to the onset of chaos. We have now understood the $x$ dependence near this limit point, but not the $r$ dependence. The treatment above simply ignored any $r$ dependence; thus it is incomplete†. We expect to express the $r$ dependence of the $f^{(2^k)}$ functions using the same scaling ideas we used to infer the $x$ dependence. Evidently, the $f^{(2^k)}$ become increasingly sensitive to $r$ just as they do to $x$. We expect a well-behaved limit when we examine a small enough region around each $r_k$. We must expand the region progressively more as $k$ increases, since we know the behavior becomes arbitrarily sensitive to changes in $r$. Thus, to describe the $r$ dependence in terms of smooth functions, we must rescale. We define a scale factor $c(k)$ by $r_k = r_\infty - 1/c(k)$. These $c$’s must go to infinity with $k$; they play an analogous role to $a(k)$ used above. As we have argued above, $f^{(2^k)}(r_\infty, x)$, when expressed in terms of $g_k(z^{(k)})$ yields the limiting smooth function $g(z)$.

Now we want to know how the limiting function is altered if we make slight changes in $r$. The shift required to obtain corresponding behavior for different $k$ reflects the diverging sensitivity to $r$. Thus it must be expanded by the scale factor $c(k)$. Accordingly we define a rescaled quantity $R_k$ analogous to $z^{(k)}$. We try the representation: $R_k = c(k)(r - r_\infty)$. Then our scaling hypothesis becomes

$$f^{(2^k)}(r, x) = x_m + \frac{1}{a(k)}g_k(R_k, z^{(k)})$$

Translating the equation

$$f^{(2^{k+1})}(r, x) = f^{(2^k)}(r, f^{(2^k)}(r, x))$$

into rescaled language as before, we obtain

$$g_{k+1}(R_{k+1}, z^{(k+1)}) = \beta g_k(R_k, g_k(R_k, z^{(k)})).$$

As before, we express both sides in terms of common variables $R_k, z^{(k)}$ by using $z^{(k+1)} = (a(k+1)/a(k))z^{(k)} \to \beta z^{(k)}$ and $R_{k+1} = c(k+1)/c(k)R_k$. Then

$$g_{k+1}((c(k+1)/c(k))R_k, \beta z^{(k)}) = \beta g_k(R_k, g_k(R_k, z^{(k)}))$$

for any $R_k$ and $z^{(k)}$. We’re free to rename these variables $R$ and $z$. We also observe that for the $k \to \infty$ limit to make sense in this equation, the ratio $c(k+1)/c(k)$ must go to a finite limit, which we denote as $\gamma$. Then

$$g(\gamma R, \beta z) = \beta g(R, g(R, z)). \quad (5.5)$$

Evidently $g(0, z)$ is the Feigenbaum function $g(z)$. Anticipating that $g$ will have a smooth dependence on $R$, we write $g$ as a Taylor expansion for small $R$:

$$g(R, z) = g(0, z) + Rh(z) + \mathcal{O}(R^2). \quad (5.6)$$

When we insert this form into the Feigenbaum equation, we obtain an expression for $h$:

$$g(0, \beta z) + \gamma Rh(\beta z) = \beta \{g(0, [g(0, z) + Rh(z)]) + Rh([g(0, z) + Rh(z)])\}$$

$$= \beta \{g(0, g(0, z)) + g'(0, g(0, z))Rh(z) + Rh(g(0, z)) + \mathcal{O}(R^2)\}$$,

so that

$$g'(g(z))h(z) + h(g(z)) = (\gamma/\beta) h(\beta z). \quad (5.7)$$

† In fact the equation (5.3) that we used to find the Feigenbaum function $g(z)$ is only true for if $r$ is the same for all the $f$’s in the equation. Thus, as noted above, the $f^{(2^k)}$ only converge to $g(z)$ if $r = r_\infty$. 

Physics 251/CS 279/Math 292  Chapter 5  typeset October 26, 2012  page 8
Here \( g(z) \) means \( g(0, z) \) and \( g' \) means the derivative with respect to the \( z \) argument. Since \( g(z) \) and \( \beta \) are known, this gives an explicit, linear equation for the function \( h \). (Linearity means that if \( h_1 \) and \( h_2 \) were solutions, any linear combination of them would be a solution as well. This is clearly true of Eq. (5.7).) Since the equation is a linear functional of \( h \), it amounts to an eigenvalue equation, whose eigenvalue is \( \gamma/\beta \). By approximating \( g \) and the unknown \( h \) as polynomials, one can obtain successive approximations to \( \gamma \).

Cvitanovic’s Introduction gives some explicit prescriptions.

Eigenvalue equations have many solutions. For example if we approximate \( h \) by a linear combination of \( N \) basis functions such as \( z, z^2, \ldots, z^N \), there will be \( N \) eigenvalues \( \gamma_n \) and eigenvectors \( h_n \). We denote the largest of the \( \gamma_n \) as \( \gamma_1 \).

The \( \gamma \) factor tells us what happens if we iterate a function that is slightly different from the original function \( g(0, z) \). If our function differs only slightly from \( g \), we can express the difference as a sum of eigenfunctions \( h_n \) of Eq. (5.7). Typically there will be some contribution of \( h_1 \) to this sum. Let’s consider this contribution by itself. Then the iterated function has the same form as the original one, with \( R \) replaced by \( \gamma_1 R \). It is clearly important whether \( \gamma_1 \) has a magnitude smaller than or greater than one. If it is smaller than one, the departure \( Rh_1 \) diminishes as one iterates, and the initial perturbation from \( g(0, z) \) dies away. If it is greater than one, the slight departure from \( g(0, z) \) gets amplified as one iterates: our function gets farther and farther from \( g(0, z) \).

Now, when one determines the \( \gamma \)’s numerically, one finds that the largest \( \gamma \) is greater than 1. In fact \( \gamma_1 \) is 4.6692016... Thus, one cannot hope to converge to \( g(0, z) \) simply by plugging some initially guessed function into the right side of the Feigenbaum equation. In general, the departure from the true \( g \) involves \( h_1 \) (as well as the other \( h \)’s.) This piece involving \( h_1 \) gets multiplied by \( \gamma_1 \). It grows without bound. Even if the guess is almost perfect, successive iterations will drive you away from it. Since the \( \gamma_1 \) contribution grows faster than any of the others, the result after a few iterations is a departure from \( g \) that is dominated by \( h_1 \). Thus, the \( \gamma \) value that matters in the limit \( k \to \infty \) is \( \gamma_1 \). The relevant solution to Eq. (5.7) is \( \gamma = \gamma_1 \).

Now that we have determined \( \gamma \), we know how \( r \) must be rescaled to give finite limiting behavior as \( k \to \infty \). From this rescaling we can now see how the doubling points \( r_k \) must approach their limiting value. Since \( c(k+1)/c(k) \) approaches\(^\dagger \) the constant \( \gamma_1 \), we must have \( c(k) \to (\text{constant})\gamma_1^k \), so that \( r_{\infty} - r_k = (\text{constant})\gamma_1 - k \). This implies a simple scaling for the ratio of successive shifts: \( (r_{\infty} - r_{k-1})/(r_k - r_{k-1}) = \gamma_1 \). This is the behavior noted at the beginning of the chapter. We now see that observed ratio \( \delta \) is simply the constant \( \gamma_1 \) defined by our eigenvalue equation (5.7). Since \( \gamma_1 \) is defined by a prescription that is independent of the chosen iteration function \( f(x) \), it is universal, like \( \beta \) and \( g(z) \). Now that we’ve shown that \( \delta \) is our \( \gamma_1 \), we’ll use the name \( \delta \) for both henceforth. It is called the Feigenbaum number. Likewise, we’ll use the conventional name \(-\alpha\) in place of our \( \beta \).

**Menu Project. Calculating \( \alpha \) and \( \delta \).** This chapter sketched methods of obtaining \( \alpha \) and \( \delta \) by approximating \( g \) and \( h \) as polynomials. The aim of this project is to obtain successive approximations to \( \alpha \) and \( \delta \) using this analytical scheme or another one. The Introduction in Cvitanovic is a useful starting point. It points to detailed information in the rest of the book, which includes a reprint of Ref. 1. Hilborn, section 5.7 has a heuristic discussion.

a) Use the quadratic approximation for \( g(z) \) used in Chapter 5 (that gave \( \beta = -1 - \sqrt{3} \)) to solve the eigenvalue equation (5.7) for \( h(z) \), assuming \( h(z) \) has the form \( \sum_{i=1}^{n} A_i z^i \). We omit the \( z^0 \) term in order to maintain our regularization condition on \( g \): \( g(0) = 1 \). Thus \( 1 = g(R,0) = g(0) + Rh(0), \) so \( h(0) = 0. \) (If we don’t maintain the regularization condition, the resulting scaling indeterminacy of \( g(z) \to \tilde{g}(z) \) discussed below Eq. (5.4) leads to an additional neutral eigenmode with eigenvalue 1.) The left side of Eq. (5.7) contains powers of \( z \) up to \( 2n \). To obtain solutions, we require that the two sides be equal to order \( z^n \). We ignore \( z^{n+1}, z^{n+2}, \ldots \) terms. Likewise we ignore the \( z^0 \) term since we are not trying to determine \( A_0 \). This gives \( n \) equations for the remaining \( n \) coefficients \( A_i \). There are in general \( n \) independent solutions to this eigenvalue problem. Verify that the largest \( \gamma \) is roughly independent of the order \( n \) of the approximation using \( n = 2, 3 \). This is done most efficiently using an algebraic

\(^\dagger \) This amounts to saying that slight shifts in \( r \) introduce perturbations in \( f \) and \( g \) that involve \( h_1 \). While this is true in general, one could invent cases where it isn’t true. Then the observed \( \delta \) would be different and smaller.
B. Description of chaos in iterative maps

We have seen above that the approach to the point of infinite periodicity is regulated by a symmetry under dilation. Since the symmetry involves changes of scale needed to put the functions \( f(x^n) \) into a standard or normal form, this is often called renormalization symmetry, and the set of transformations that take \( g_k \) into \( g_{k+1} \) (or \( g_{k+2} \) etc.) is called the renormalization group. If the transformation leaves the function invariant, we say that the function is at a fixed point of the renormalization group. The limiting function \( g(z) \) is such a fixed point.

While this symmetry is a powerful and elegant feature of iterative maps, it doesn’t begin to describe the richness we have seen in our numerical calculations when our amplitude parameter \( r \) exceeds \( r_\infty \). In this section we’ll note a few of these features and suggest properties that can be proven. Two good sources of further information are the descriptive book by Hilborn and the readable math text by Devaney, on the course reference list.

For \( r < r_\infty \), all initial points \( x_m \) led to a single periodic sequence. We called the set of final values for a given region of \( x \) an attractor, and the region that ends up there as the basin of attraction. When \( r = r_\infty \) the attractor contains an infinite number of points on the 0-1 interval. Thus the sequence of \( x_i \) never repeats. When \( r \) exceeds \( r_\infty \) the attractor also contains infinitely many points. The dynamics has become complex. However, not all starting points lead to this complex behavior. There are many isolated fixed points and cycles, as we may readily verify by graphing various \( f(n)(x) \) and noting that they cross the fixed-point line in several places. Thus there are two kinds of attractors, those with a finite number of points and the chaotic one with an infinite number of points.

The attractor begins as a set of narrow bands that broaden as \( r \) increases. As we have seen, these bands grow to encompass the whole interval as \( r \) reaches 4. For that special case of \( r = 4 \) we saw that one can describe the map in an equivalent way using the \( \theta \) representation of Chapter 3. Here we saw that the map has three properties:

1) The map is sensitive to initial conditions. Neighboring points spread apart with successive iterations. (Since \( \theta_{i+1} = 2\theta_i \), the distance between neighboring points doubles at each iteration.)

2) The map mixes up the entire interval. Within any small interval of \( x \) values some \( x \)'s will get mapped into any other interval. Devaney calls maps with this property topologically transitive.

3) Despite this complexity, there are periodic orbits arbitrarily close to any \( x \). That is, the periodic orbits are dense in the interval.

These three properties when \( r = 4 \) embody the paradoxical coexistence of simplicity and complexity. A dynamical system with these three properties is called chaotic in Devaney’s definition.

When \( r \) is between \( r_\infty \) and 4, there are also regions where chaos occurs. In these regions of \( r \), as elsewhere, almost all initial \( x \)'s lead to the chaotic attractor. The fraction \( x \)'s that aren’t chaotic is vanishingly small. Any starting point \( x \) that is not part of an \( n \)-cycle comes arbitrarily close to every point in the attractor.

We can gain insight into the chaotic regions by following the images of the extremal point of \( f(x) \), namely \( x = 1/2 \). If \( f \) is chaotic, \( f(1/2) \) must lie in the chaotic attractor. There is a small interval interval on one side of 1/2 that also lies in the attractor. A small-enough interval must transform smoothly under a

---

\[ \dagger \] It is not a group in the strict mathematical sense. In a real group every group operation has an inverse. In this renormalization group, we only have a prescription to go from \( k \) to \( k+1 \), but not the reverse. Indeed, we saw earlier in the course that each \( x_i \) has two possible predecessors \( x_{i-1} \).

\[ \dagger \dagger \] For the logistic map and its relatives, a single stable attractor attracts all but a vanishing fraction of initial \( x \)'s. So there is only one chaotic attractor for a given \( f \). We have not proven this.

\[ \dagger \dagger \dagger \] Conversely if \( f(1/2) \) were to lie in a finite cycle, then some \( f^{(n)}(1/2) \) is a superstable fixed point. When this occurs, chaos ceases, as noted below.
given number of iterations \( n \). But since \( f'(1/2) = 0 \), this interval gets strongly contracted with each iteration. Thus points of the attractor get concentrated near \( f(1/2) \), \( f^{(2)}(1/2) \), etc. These concentrated regions are readily visible as “scars” on a the bifurcation diagram of iterates of \( f \) vs. \( r \). The first image of \( 1/2 \), i.e., \( r/4 \) is a straight line on the diagram. Since there is no \( f(x) \) larger than \( f(1/2) \), this line is an upper bound of the attractor for all \( r \). The second image, \( f^{(2)}(1/2) \), is a cubic function of \( r \), namely \( r(r/4)(1-r/4) \). It is a lower bound of the attractor.

Figure 5.4 Bifurcation diagram for the logistic map, \( f(x) = rx(1-x) \), for \( 3.5 < r < 4 \). The images \( f^{(n)}(1/2) \) for \( n = 1, 2, \ldots, 8 \) are plotted in green. Dashed red line indicates \( x = 1/2 \). Blue ticks along bottom edge indicate intervals of .01 in \( r \). Region indicated by heavy line at bottom is expanded in Figure 5.5.

Figure (5.5) Expanded view of the typical interval \( 3.726 < r < 3.79 \) of Figure 5.4 showing the 5-cycle where \( f^{(5)}(1/2) \simeq 1/2 \) and the 7-cycle where \( f^{(7)}(1/2) \simeq 1/2 \).

Higher images of \( 1/2 \) are shown in Figure 5.4, 5.5. We see that as \( r \) increases from \( r_\infty \), these images spread apart and cross. When they cross, the sub-bands of the attractor merge. The last two images to cross are \( f^{(3)}(1/2) \) and \( f^{(4)}(1/2) \). At the crossing point \( f^{(4)}(1/2) = f^{(3)}(1/2) \) all the bands have merged into a single chaotic region. Taking \( f \) of both sides of this equation, we see that \( f^{(5)}(1/2) = f^{(4)}(1/2) \), which in turn equals \( f^{(3)}(1/2) \). Repeating this reasoning, we see that all the higher images of \( 1/2 \) must be equal at this point. Thus all the scar lines must cross here.

These images of \( 1/2 \) explain another feature of the diagram. Many of the images beyond the second cross \( x = 1/2 \) at some \( r \). When \( f^{(n)}(1/2) = 1/2 \), the subsequent images can no longer all be distinct. Instead,
they form a superstable \(n\)-cycle. Since there is only one attractor, it must be this \(n\)-cycle: chaos must cease. On the diagram we see this happening: bands of complexity abruptly give way to intervals or windows of \(r\) with only a finite cycle in the attractor. The transition occurs as soon as an \(n\)-cycle becomes stable, for slightly larger \(r\) within the stable region the cycle becomes superstable. This \(r\) within the stable window is the point where \(f^{(n)}(1/2) = 1/2\). Another sign of such a superstable point is clear if one takes \(f\) of both sides: \(f^{(n+1)}(1/2) = f(1/2)\). Thus one of the images must graze the upper bound, viz. \(f(1/2)\). These stable \(n\)-cycles undergo period doubling and chaos returns within narrow bands as \(r\) continues to increase. Again the chaotic bands begin to merge. But now, before the merging process is complete, the attractor abruptly jumps to encompass the whole band. Such behavior is called a crisis according to Hilborn. A crisis occurs when a chaotic band crosses an unstable cycle. Thus, e.g., when the 5-cycle in Figure 5.5 became stable, \(f^{(5)}' = 1\): \(f^{(5)}\) was tangent to the fixed-point line. As \(r\) increased, \(f^{(5)}(x)\) passed through the fixed-point line, giving rise to a stable point and an unstable twin. When this unstable fixed point touches a chaotic band, the crisis occurs, as Hilborn explains in Chapter 7. See the External Links of the Chalk site for an applet that shows these features on a fine scale.

C. Universality in other guises

We now return to the phenomenon of universality demonstrated above for iterative maps. We make a big deal out of this universality, because this same phenomenon occurs in a much broader range of situations. The phenomenon is this: we impose two qualitative symmetry conditions at the same time, and find a quantitative determination of these symmetries. The symmetry of crystals provides a prosaic example of this phenomenon. A two dimensional lattice is a set of points with a certain translational invariance. Such a lattice can also have rotational symmetries. If we require both symmetries at the same time, we find that the only possible rotational symmetries are 60, 90, 120, or 180 degrees. i.e. 6, 4, 3, or 2-fold rotations. The same type of mechanism constrains the rotational symmetries of three-dimensional crystals.

The iterative map example of Section A above is more intriguing, because the symmetry in question is dilation symmetry, relating an object to a subset of itself. The object studied was a sequence of points and the symmetry determined was the dilation symmetry of a geometric sequence. A number of extended physical systems also show universal dilation symmetry. The subunits of these systems are equivalent; thus the dilation symmetry is that of a fractal.

Fractals are observed many places in nature, in systems as diverse as bacterial colonies, mountain ranges, and clouds. One of the great open questions in physics is why fractals are observed in so many different systems. Here we discuss a few other situations in nature in which fractals appear, and point out cases of universality.

\[C1: \text{The random walk}\]

A random walk is very simple: start at a point (in a 2 dimensional plane, say) and take a step of size \(L\) in a randomly chosen direction. Randomly choose a new direction and take another step of size \(L\). Continue this process to produce a random walk of arbitrary length. Random walks (in 1, 2, and 3 dimensions) are a very useful model for many physical processes, such as Brownian motion or the diffusion of a particle.

![Figure 5.6. Two random walks.](image)

\[\text{Problem 5.1}\] The random walk. Write a program which implements the two-dimensional random walk procedure described above and which calculates how the average (mean-square) distance from
the starting point varies with i) the step size $L$, and ii) the number of steps $N$. Since the process involves random numbers (which you can generate with the \texttt{Math.random} method), you should do a fairly large number of random walks for each $L$ and $N$, and average the results. First, fix $L$ and obtain average distances for, say, ten values of $N$ (none smaller than about 1000). You should take data “on the fly” in the following manner. Let your particle walk 1000 steps, calculate and store the distance from the origin, let it walk another 1000 steps, calculate and store the distance from the origin, etc. The average distance, $\langle d \rangle$, should vary as $\langle d \rangle \sim N^\nu$ for some exponent $\nu$. Now fix $N$, and try varying $L$. Again, $\langle d \rangle$ should vary as a power of $L$, $\langle d \rangle \sim L^\alpha$. Can you guess, a priori, what $\nu$ and $\alpha$ should be? Does this computer experiment verify your intuition? A random walk does not look exactly the same when it is magnified (since, after all, a different random number gets picked at every step). Thus the dilation symmetry applies not to individual walks but rather to the statistical ensemble of walks of length $N$.

**C2: the self-repelling polymer** Some molecules take the form of a long, flexible chain of subunits connected by freely-rotating bonds. Such a molecule dispersed in a liquid is like the track of random walk, but with an additional constraint. Atoms from different parts of the molecule cannot intersect. (There is no such constraint on a randomly walking particle; it is free to visit places it has visited before.) It’s more realistic to think of these polymers as feeling some repulsion between the subunits. This self-repelling polymer is one of the simplest examples of universal fractal symmetry. As in the iterative maps, universality arises from two qualitative symmetries. First, such a walk is a fractal: small pieces sampled intersect. (There is no such constraint on a randomly walking particle; it is free to visit places it has visited before.) It’s more realistic to think of these polymers as feeling some repulsion between the subunits. This self-repelling polymer is one of the simplest examples of universal fractal symmetry. As in the iterative maps, universality arises from two qualitative symmetries. First, such a walk is a fractal: small pieces sampled at random are indistinguishable statistically from subsets of those pieces. Second, the polymer is \textit{locally constructed}. That is, we can decide whether a set of subunits is a legal self-avoiding polymer by looking in a small region around each subunit. The region remains the same size regardless of how large the length $N$ becomes. When both the dilation symmetry and the locality are taken into account, the value of the fractal dimension $D$ is determined \cite{3}. This $D$ depends on the dimension of space. In two dimensions, it was shown in the 1980’s that $D = 4/3$, using methods of renormalized field theory and conformal symmetry \cite{4}. In three dimensions, we have no precise value for $D$; we have only approximate methods as with the Feigenbaum $\beta$ and $\gamma$ discussed above. These methods suggest that $D \approx 1.70$. Many experimental and numerical tests have demonstrated the universality of these exponents by showing that the polymer’s local makeup can be altered greatly without changing the asymptotic $D$ value for large length $N$.

**Menu Project: Diffusion Limited Aggregation (DLA)** This project explores the use of random walks to model an interesting growth process of dendritic structures made up of particles that stick together when they come in contact. The particles are assumed to diffuse slowly through space until they come in contact with another particle, at which time they stick. The computer implementation is simple in concept: Put a particle in the middle of an applet window at $\langle i_0, j_0 \rangle$. Then start another particle from some point $(i, j)$ far away, and allow it to execute a random walk with integer steps. After a time, it will land on one of the points adjacent to the original particle; call it $(i_1, j_1)$. At this point the random walk terminates and the new particle at $(i_1, j_1)$ becomes part of the aggregate, along with $(i_0, j_0)$. Now one introduces a second random walker like the first one. It stops when it reaches any point adjacent to the existing aggregate and becomes a part of it. One continues in this way until a large aggregate has been made.

The structure that develops has a tree-like, fractal shape. Unfortunately, it grows very slowly. You can speed up the growth by starting from a point $(i, j)$ that is closer to the existing cluster. It can be shown that the structure remains unaffected if one chooses a starting point $(i, j)$ at random on a circle around $(i_0, j_0)$ large enough to contain the cluster which has grown thus far. This is a program that you will want to run as long as possible. Compute the fractal dimension of your clusters, defined by the relation $M \propto R^D$, where $M$ is the number of particles in the cluster, $R$ is the “radius” of the cluster (typically the root-mean-square radius distance of a particle from the starting point), and $D$ is the fractal dimension. Grow several large aggregates to get an idea of the accuracy of your estimate for $D$. Test whether this $D$ is universal by varying the rules. Instead of stopping a particle when it moves adjacent to the cluster, allow it to continue with some fixed probability $p$. This this can lead to cluster points which are multiply occupied. This is OK. Choose a value of $p$ large enough to make a clearly noticeable difference in the structure, and determine $D$. 

Physics 251/CS 279/Math 292 Chapter 5 typeset October 26, 2012 page 13
Figure 5.7. Topography of river drainage basin. From Robert Barnes, Armidale, NSW, Australia: Area shown is northeast of Armindale and 70 km east of Glen Innes in Australia, which includes the drainage basins of the Mann, Boyd, Nymboida, and Clarence rivers. The image is a false-color map of the digital elevation data of the area.


again. Is it consistent with the previous value with $p = 0$?

C3: river networks Figure 5.7 is an image of a river network in Australia. Analysis of pictures like Figure 5.6 or Figure 5.7 suggests that these systems are statistically dilation-symmetric (i.e., fractal). Why river basins should be fractal is a matter of some debate. Several plausible models have been proposed that lead to dilation-symmetric drainage networks (some of which are in the book by Ignacio Rodriguez-Iturbe and Andrea Rinaldo, Fractal River Basins: Chance and Self-Organization, Cambridge University Press (1997)). It is an open question whether there is a fundamental unifying mechanism that causes fractals to emerge from many different models. One interesting view on this subject is presented by Per Bak in his book How Nature Works: The Science of Self-Organized Criticality, (Springer-Verlag, 1996).

Menu Project. A river network model. Chapter 19 of Gould and Tobochnik (2nd edition) discusses one model of river networks (from R.L. Leheny, Phys. Rev. E 52, 5610 (1995)), in which a rectangular lattice of points describes an eroding terrain with the height of the land, $h(x,y)$, specified at each point. The simulation begins with the landscape as a featureless incline: $h(x,y) = ly$. Then the following rules are implemented:

1) A raindrop lands at a random site on the lattice.

2) The raindrop flows from this site to one of the four nearest neighbors with a probability proportional to $e^{E\Delta h}$, where $\Delta h$ is the height difference between the site and the neighbor, and $E$ is a parameter of the model. If $\Delta h < 0$, this probability is set equal to zero.

3) Step 2) is repeated until the water reaches the bottom of the lattice, $y = 0$.

4) Each point that has been visited by the flowing water has its height reduced by a constant amount $b$. This process represents erosion.

5) Any site at which the height difference $\Delta h$ with a neighbor exceeds a threshold $M$ is reduced in height by an amount $\Delta h/S$, where $S$ is another parameter in the model.

Write a program that implements this model. You can get an idea of suitable parameters to use from Leheny’s paper. The resulting river network is defined as follows: every lattice point receives one unit of precipitation which traces a path of steepest descent, without eroding the terrain, until it reaches the lattice edge, $y = 0$; the river network is defined as all points through which at least $R$ units flow. Analyze the network that is generated at different times. Does the river network appear to be fractal? How does evolving the model for longer times affect the network’s properties?
REFERENCES

2. Steven Finch, MathSoft Inc.. 101 Main Street. Cambridge, MA, USA 02142, <sfinch@mathsoft.com> http://www.mathsoft.com/asolve/constant/fgnbaum/fncteqn.html