

CHAPTER 1

Introduction to Fractals

Unlike many other mathematical notions, the notion of a fractal is not really well-defined by some axiomatic list of properties. The best way to understand what we mean by a fractal is to contrast fractals with classical geometry objects like lines and surfaces. The latter are usually studied using tools from Differential Geometry and, ultimately, Linear Algebra. They are *linear* objects, which are locally well approximated by their tangent planes. This linearity is absent when we deal with fractals.

To formalize this concept of non-linearity, we will introduce the notions of *mini-* and *micro-sets*. We will start with the definition of the Hausdorff distance between subsets of a metric space.

DEFINITION. Let X be a compact metric space and let $\sigma(X)$ be the set of all its closed subsets. If $x \in X$ and $A \in \sigma(X)$, we define the distance from x to A as the minimum of all distances from x to points of A , i.e.,

$$d(x, A) = \min_{y \in A} d(x, y).$$

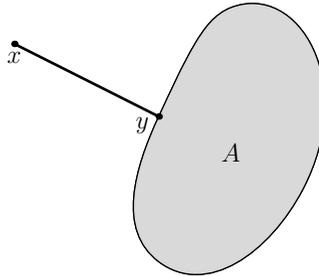


FIGURE 1. The distance from a point x to a set A is the distance from x to a nearest point $y \in A$.

If $A, B \in \sigma(X)$ are two closed subsets of X , we define the distance between them as

$$\mathcal{D}(A, B) = \max\{\max_{x \in A} d(x, B), \max_{y \in B} d(y, A)\}.$$

Thus, the inequality $\mathcal{D}(A, B) < \varepsilon$ means that A is contained in the ε -neighborhood of B and B is contained in the ε -neighborhood of A , i.e., for every $x \in A$, there exists $y \in B$ such that $d(x, y) < \varepsilon$ and vice versa.

Note that the sets close in this metric may be of very different nature from the topological perspective. For instance, the interval $[0, 1]$ is ε -close to the finite set

of points $\{0, \frac{1}{n}, \frac{2}{n}, \dots, 1\}$ if $n > \frac{1}{\varepsilon}$ despite the fact that $[0, 1]$ is a continuum and $\{0, \frac{1}{n}, \frac{2}{n}, \dots, 1\}$ is a discrete set.

Now we are ready to define mini- and micro-sets. Note that this terminology is not universally accepted, so some other books may use different names for the same objects. We will always assume that our starting set A is a compact subset of \mathbb{R}^n contained in the unit cube $Q = [0, 1]^n$.

DEFINITION. *A mini-set of A is any set of the kind $(\lambda A + u) \cap Q$ where $\lambda \geq 1$ and $u \in \mathbb{R}^n$ are such that $Q \subset \lambda Q + u$.*

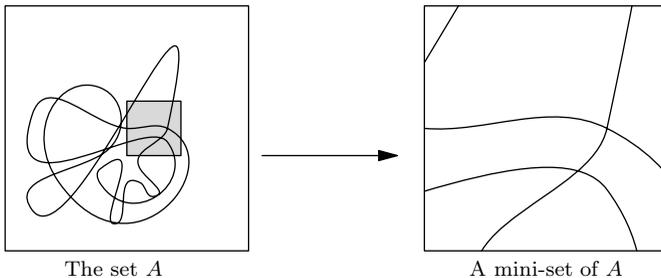


FIGURE 2. A mini-set of A is the scaled intersection of A with a small square window (enlarged on the right).

A micro-set of A is the limit of any sequence A_n of mini-sets of A in the Hausdorff metric.

Informally, a mini-set is what you see if you look at a small portion of a set through a magnifying glass. The condition $Q \subset \lambda Q + u$ is needed to ensure that a mini-set of a mini-set of A is again a mini-set of A . Note also that the scales λ_n and the shifts u_n corresponding to the mini-sets A_n in the definition of a micro-set do not need to be related in any way.¹

A lot of interesting micro-sets can be obtained by zooming at a single point like at the picture below, but we also have an option of moving our window around

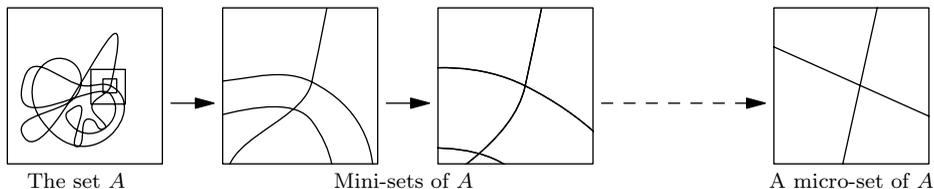


FIGURE 3. A micro-set of a classical set is a union of several points.

simultaneously with the scale reduction.

The most interesting micro-sets are those that correspond to the sequences of mini-sets $A_n = (\lambda_n A + u_n) \cap Q$ with $\lambda_n \rightarrow +\infty$. For all objects studied in classical geometry, they are essentially flat, i.e., either planes or unions of a few planes as

¹Note also that the set of closed subsets of the closed unit cube Q endowed with the Hausdorff metric is compact (Blaschke theorem), so any sequence of mini-sets contains a convergent subsequence.

in Figure 3. What distinguishes fractals from the classical geometry objects is that their micro-sets may be as complex as the original sets.

Consider, for example, the so-called *Sierpinski gasket* S .

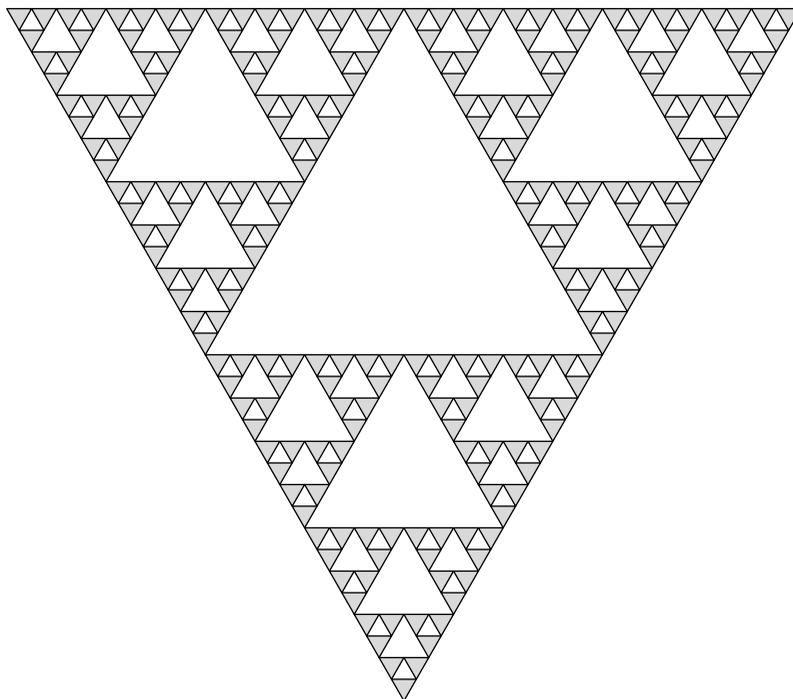


FIGURE 4. Sierpinski gasket.

No matter how much you zoom in on it, it doesn't become flat or gets simpler in any other respect. More precisely, if you start zooming in at some fixed point of S , the corresponding family of mini-sets has periodic behavior: whatever you see at the scale λ , you see again at the scales 2λ , 4λ , 8λ , and so on (provided that the magnifying factor λ is large enough so that you can see only a small piece of S at once). For more complicated fractals the behavior can be more subtle, so that when zooming in one observes a chaotic behavior of shapes. The proper tools to handle this chaos are no longer those of Linear Algebra but those of Ergodic Theory.

Another difference between fractals and classical geometry objects is that fractals usually have non-integer (fractional) *dimension*. The classical notion of dimension is, essentially, a Linear Algebra notion. The dimension of a classical object like a line or a surface in \mathbb{R}^n is nothing but the linear dimension of its tangent plane, so a point has dimension 0, a line has dimension 1, a surface has dimension 2, and so on. However, as a rule, fractals do not have tangent planes and this linear algebra approach becomes meaningless for them. The dimension of fractals has to be defined in a different way and can be any non-negative real number up to the dimension of the ambient Euclidean space.

We will return to this discussion in Chapter 2 and now we will look at how fractals arise in mathematics. The primary sources of fractal objects are infinite

iterations of simple classical processes. Consider, for instance, the standard *middle-third Cantor set* on the line. Its construction starts with a closed interval, say $[0, 1]$:

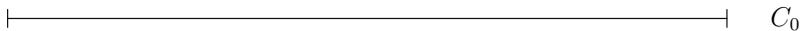


FIGURE 5. The construction of the Cantor set starts with a single interval.

At the first step, the middle third $(\frac{1}{3}, \frac{2}{3})$ is removed and we get the set $C_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$:



FIGURE 6. At the first step the middle third is removed.

Next, the same procedure of removing the middle third is applied to each of the intervals of C_1 resulting in 4 intervals of length $\frac{1}{9}$:



FIGURE 7. This procedure is repeated for each remaining interval.

This process repeats again and again, so C_n consists of 2^n intervals of length 3^{-n} . The Cantor set C is defined as the limit of C_n in the Hausdorff metric, which in this case is the same as $\bigcap_{n=1}^{\infty} C_n$.

Note that the Lebesgue measure of C_n equals $(\frac{2}{3})^n$, which tends to 0 as $n \rightarrow \infty$, so we remove almost all points from the initial interval $[0, 1]$ in the sense of the Lebesgue measure. However, there are still many points left. For instance, all endpoints of the intervals appearing at all stages stay, so C is infinite. It is possible to show that C is of the cardinality of the continuum, so from this point of view, it is as large as the original interval $[0, 1]$.

The Sierpinski gasket S is just a 2-dimensional version of the middle third Cantor set. We start with with a closed equilateral triangle, split it into 4 equal triangles and remove the central one:

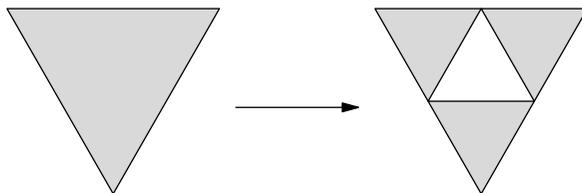


FIGURE 8. At the first step, the central triangle is removed.

Then we repeat this procedure for each of the three remaining triangles: and so on. After n steps, we get 3^n triangles with side-length 2^{-n} . Again, the area of S_n decays as a geometric progression. However, the total length $(\frac{3}{2})^n$ of the sides

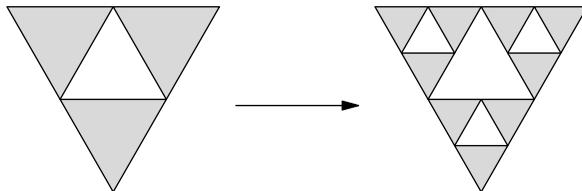


FIGURE 9. The procedure is repeated for each of the remaining triangles.

of the triangles constituting S_n grows exponentially. Thus, S is far too small for a 2-dimensional object and far too large for a 1-dimensional one, so its dimension has to be strictly between 1 and 2.

Another way to construct a fractal set is to start with some simple continuous map $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ and look at the set of points that do not escape to infinity under any number of iterations, i.e., the set of points for which the sequence $x, f(x), f^2(x), \dots$ remains bounded (here and below $f^n(x) = f(f(\dots(f(x))\dots))$ is the n 'th iterate of f).

Let us start with $f(x) = 5x(1 - x)$, $x \in \mathbb{R}$. The graph of f is just a parabola looking down: Note that $f(x) = 5x(1 - x) < 5x$ when $x < 0$, so $f^n(x) < 5^n x$ in this

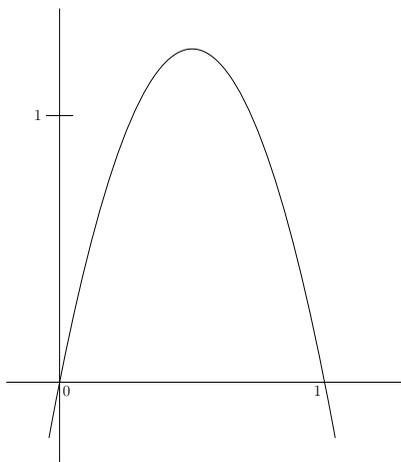
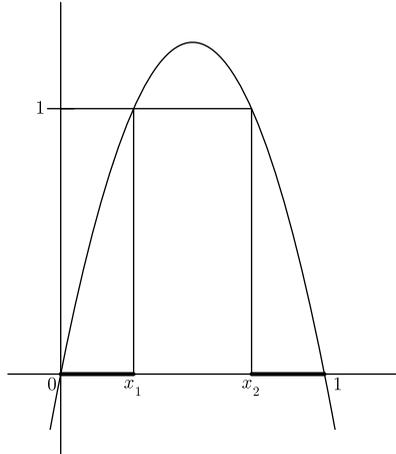
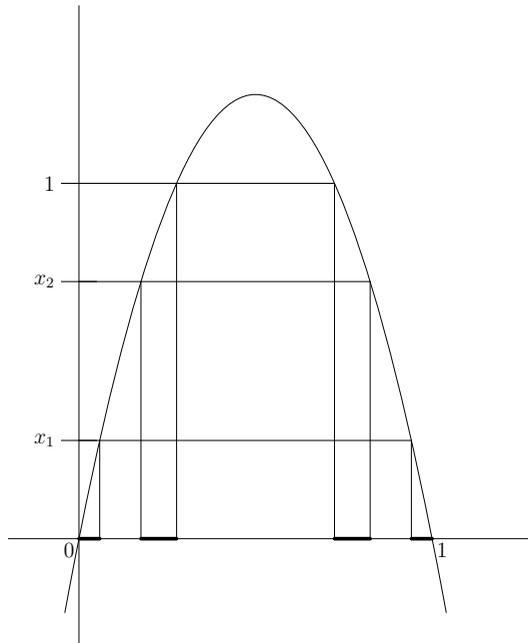


FIGURE 10. The graph of $f(x) = 5x(1 - x)$.

case, and, therefore, $f^n(x) \rightarrow -\infty$. If $x > 1$, then $f(x) < 0$ and, by the previous remark, we still have $f^n(x) \rightarrow -\infty$. Thus, the points x for which $f^n(x)$ does not tend to $-\infty$ must stay in the interval $[0, 1]$ after arbitrarily many iterations. Such points do exist. For example, the fixed point 0 of f satisfies this property, and so do all its pre-images.

To find the set of points with bounded iterates exactly, note that each such point should lie in any iterated pre-image $f^{-n}([0, 1])$. Moreover, since $f^{-1}([0, 1]) \subsetneq [0, 1]$, these pre-images form a nested sequence of closed sets. They can be easily seen on the graph. The set $f^{-1}([0, 1])$ is the union of 2 intervals $[0, x_1]$ and $[x_2, 1]$: The second pre-image $f^{-2}([0, 1])$ is a union of four intervals: and so on. The final

FIGURE 11. The set $f^{-1}([0, 1])$.FIGURE 12. The set $f^{-2}([0, 1])$.

intersection $\bigcap_{n=1}^{\infty} f^{-n}([0, 1])$ has pretty much the same structure as the middle third Cantor set.

The pictures become even more interesting if we consider the quadratic mapping on the complex plane. Let $f_c(z) = z^2 + c$, ($c \in \mathbb{C}$). The set of points $z \in \mathbb{C}$ such that the sequence $z, f_c(z), f_c^2(z), \dots$ does not tend to ∞ is called the *filled Julia set* of the mapping f_c . Several filled Julia sets corresponding to different values of c are shown below.

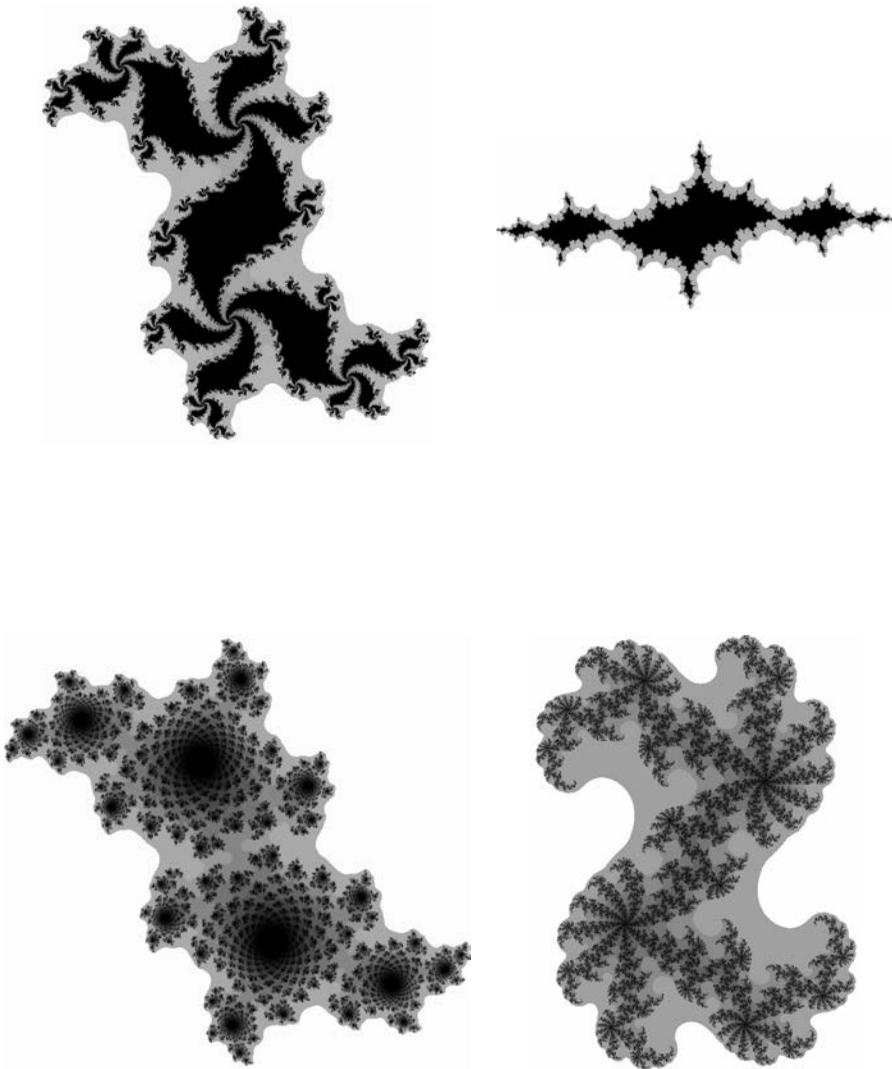


FIGURE 13. Julia sets.

Once we have a parametric set $\{f_c\}_c$ of mappings, we can also consider the set of the values of the parameter c for which f_c has a certain property. The famous Mandelbrot fractal on Figure 14 is defined as the set of all $c \in \mathbb{C}$ for which 0 belongs to the filled Julia set of f_c or, equivalently, for which the Julia set of f_c is connected.

One more example where infinite iteration arises naturally is the Newton method for finding roots of functions.

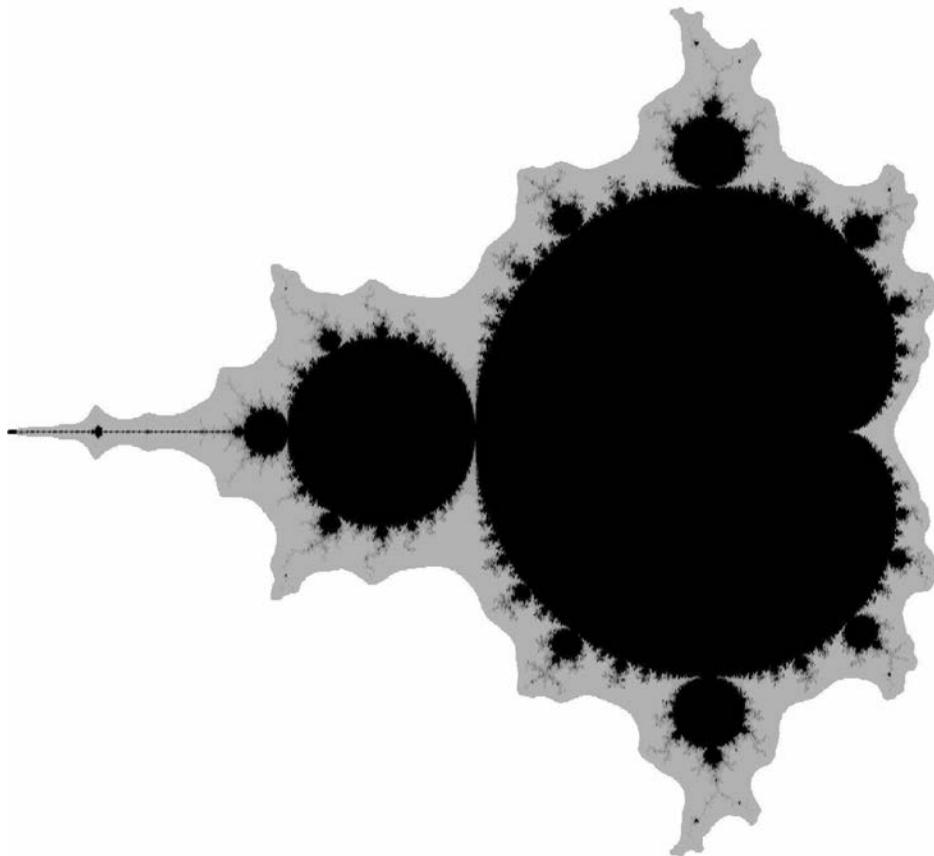


FIGURE 14. The Mandelbrot set

Suppose that we need to solve the equation $f(x) = 0$ and know some initial approximation x_0 to the root. Linearizing f at x_0 , we get the equation

$$f(x_0) + f'(x_0)(x - x_0) = 0,$$

whose solution is $x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$. If the function f is nice enough, and x_0 is close enough to the root we are seeking, x_1 is closer to the root than x_0 : However, that is not always the case:

The classical *Newton method* uses this iteration process to approximate the root. Note that, when the function f has several roots, it is by no means trivial to determine which of them the Newton method will converge to starting with some fixed x_0 (if it converges at all) and the set of initial values leading to a particular root is referred to as the *basin of attraction* of that root. This becomes an even more complicated question when we apply the Newton method to an analytic mapping on the complex plane. The basin of attraction of the root $z = 1$ of the function $f(z) = z^3 - 1$ is depicted in Figure 17. The boundary of this basin is very far from being smooth. It, too, has fractal structure.

For more information regarding fractals in the complex plane and their relation to dynamics, the reader is referred to [B] and [Mi].

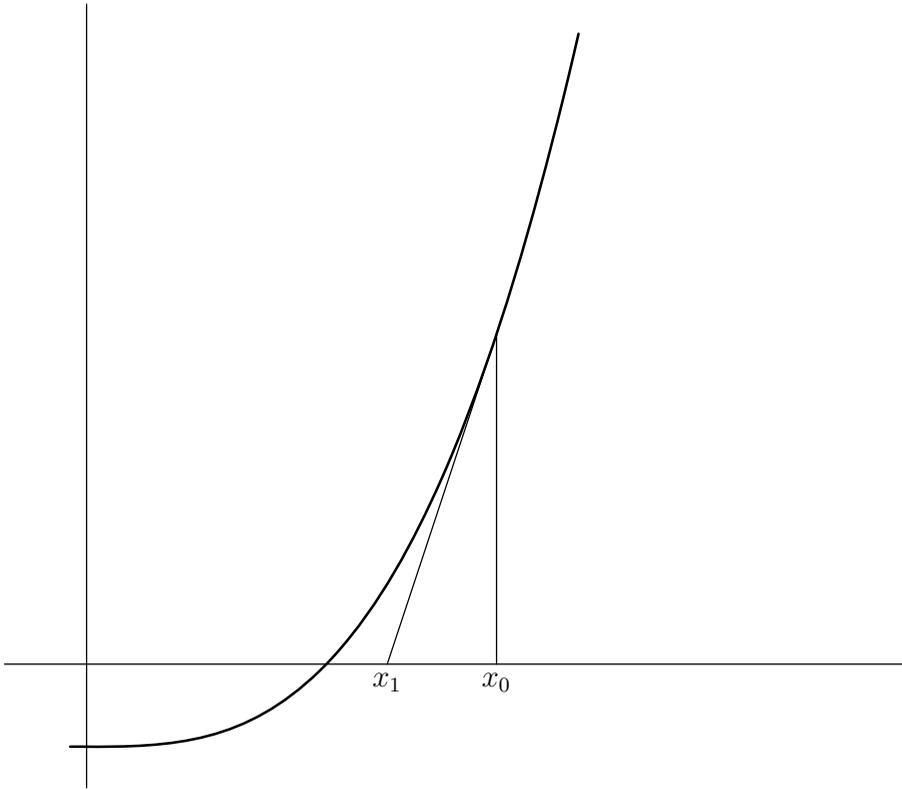


FIGURE 15. In a nice situation x_1 is closer to the root than x_0 .

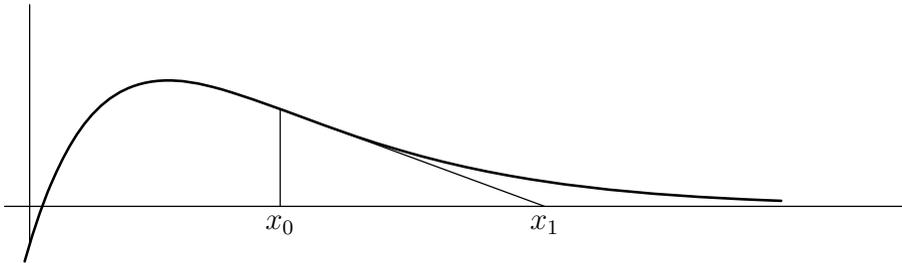


FIGURE 16. Bad situation.

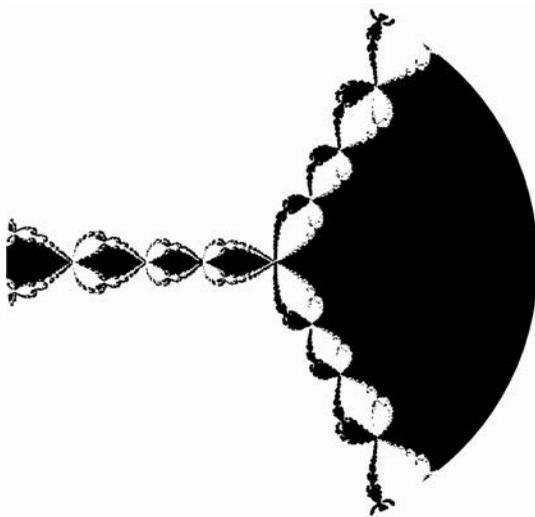


FIGURE 17. The basin of attraction of the root $z = 1$.