A BRIEF WALK THROUGH THE FRACTAL LANDSCAPE

ZACH STRONG

Date: November 9, 2024.

Contents

1. Introduction

Much of mathematics assumes a certain amount of "smoothness" in the objects it studies. In calculus, the derivative of a function only exists if the function is "smooth enough". In geometry, the formulas we have for the areas and volumes of shapes all assume that the shapes are as perfect as possible (so spheres are perfectly round, lines are impossibly straight, etc.). In Fourier analysis, we concern ourselves with functions that can be represented as a combination of smooth, trigonometric functions. The list goes on.

Of course, there are good reasons why mathematicians do this. Things that are "smooth" are generally easier to understand and manipulate in useful ways. As well, many real-world things can be modelled/approximated pretty well by "smooth" objects (e.g. finding volumes of objects that are *close* to perfect mathematical shapes or calculating physical forces like gravity that can *almost* be modelled by simple polynomial functions). So, in many areas of math, we do away with the minute complications of the real world and restrict ourselves to impossibly perfect curves and shapes.

While much of the world can be adequately modelled by nice, "smooth" things, there are some things that resist such simplifications. Consider the branches of a fern, which consist of rows of densely packed leaves. However, these leaves are themselves rows of densely packed leaves. As are these smaller leaves. And so on. In fact, nature houses some of the most complex shapes we know. Trees have branches that split off of each other exponentially. Stalks of Queen Ann's Lace flower into bunches of petals that are themselves bunches of petals. Romano broccoli is composed of buds that, when one zooms in, appear to be made of the same buds. Though we may not be interested in analysing these specific plants, the point remains that many objects exist in nature that are impossible to reduce down to simple, "smooth" geometric shapes[1].

It may seem plainly obvious that not all things in nature are simple, basic shapes. However, it took a long time before mathematicians fully accepted and embraced this fact. Why was that? How did mathematics expand beyond thinking about "nice" objects? When did it happen? This project, at a high level, aims to explore the rise of "roughness" in mathematical thinking. The main concept we'll focus on is the idea of a *fractal*—objects complex enough to elude succinct definition. We'll investigate some key mathematical constructions and discoveries that forced mathematicians to reconsider their fixation with easily-imaginable, easily-manageable constructs. As

[1]We'll define what constitutes a "smooth" shape later in this project with our discussion on Hausdorff dimension.

well, we'll discuss some of the techniques and ideas that arose out of these discoveries, helping mathematicians make sense of roughness.

FIGURE 1. Fern leaves and Queen Ann's Lace are just two examples of plants that aren't "smooth". Good luck understanding these as simple geometric shapes! Original image of the fern leaf is "Green Leaf Photography" by Arunodhai V from Pexels. Original image of the Queen Ann's Lace is "White-petaled Flower" by Stanislav Kondratiev from Pexels.

2. Grappling with Roughness

To try and find the earliest instance of mathematicians studying "roughness" is a fool's errand. Some of the earliest examples we have of mathematical thinking in humans are of our attempts to grapple with the chaos of the world around us. Early humans etched tally marks representing prime numbers onto pieces of bone, likely in an attempt to understand why these indivisible numbers occurred at random, unpredictable intervals (Jackson [2017]). Egyptians devised formulas for calculating the volumes of perfect geometric solids that approximated the complicated, real-world structures they built (Triola [1973] 357-363). Ancient Greek mathematicians, with their straight edges and compasses, worked harder than any mathematicians before them to explain the world around them through the irrefutable proofs of elementary geometry (Triola [1973] 357-363).

Depending on what we mean by "roughness", we can go back as early as we'd like in the history of math to find examples of mathematicians considering it. Out of a necessity to start *somewhere*, we'll choose 1827 as our starting point. This was the year Robert Brown made a particular observation regarding pollen molecules that sparked a tremendous amount of mathematical interest that persists to this day (albeit in vastly different contexts, such as stock markets!)[2].

[2]Robert Brown wasn't necessarily the first mathematician to notice the behaviour we're about to describe, but he is the mathematician after which the phenomenon is named. See Spencer [1999].

2.1. **Brownian Motion - Part 1.** While looking at pollen molecules that were immersed in water, Robert Brown noticed that the molecules weren't moving in simple lines or curved paths. Rather, the molecules "jiggled" (et. al. [2010]), seeming to choose their direction at random. This was quite the puzzling observation at the time. Up to then, "[a]pplied mathematics had been concentrating for a century on phenomena which were smooth" in nature (Hoffman [2010]), but here was a natural phenomenon that was 100% rough. Because this type of "non-smoothness" was a relatively novel thing to observe, much interest was generated in trying to understand this peculiar particle movement.[3]

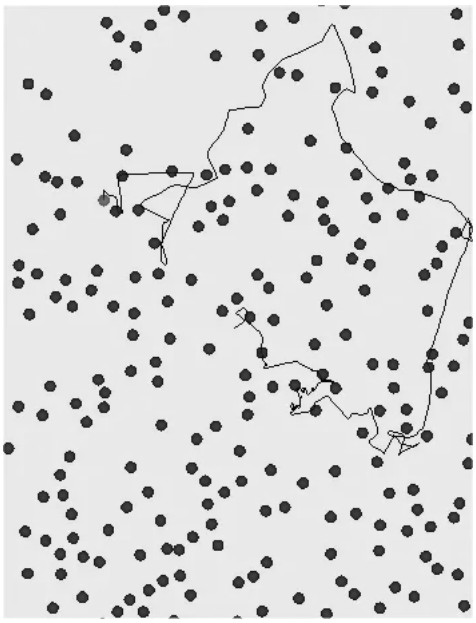


FIGURE 2. An example of the type of path a pollen molecule may take when floating in water. Image taken from Wong [2015].

This type of random, jagged movement is called *Brownian motion*, and it can be used to model much more than just molecules of pollen floating in water (see section 4.1). Eventually, it was understood that Brownian motion, at least in the original context of pollen molecules, was caused by the erratic way in which water molecules are constantly moving around and banging into each other (Mickelin [2017]). Even in a perfectly still tray of water, the molecules that make up that water are always moving. The pollen that Brown was observing happened to be small enough to be affected by this movement, and so it was jostled around as it floated.

While an explanation of Brownian motion is certainly beneficial, it isn't as useful as we'd like for building a model to emulate its behaviour. Using the above insight, the only way we could model Brownian motion is by simulating hundreds

[3]As a fun sidenote, Albert Einstein was among one of the people interested in understanding this motion! See Mickelin [2017].

and hundreds of hyperactive particles! A succinct mathematical model—a set of equations or some set of laws—is ultimately what we'd like. However, as we mentioned earlier, mathematics had, up to this point, been focused mainly on smooth phenomena—things wholly separate from the rough, jagged movement of Brownian motion. Mathematicians didn't yet have the tools for finding such a set of equations or laws. We'll return to Brownian motion later in this project when we've built up some terminology and ideas for working with rougher processes.

2.2. **Weierstrass Function - Part 1.** Fast-forward to 1872. Calculus—perhaps the grandfather of all branches of math dealing with "smooth" things—was in the process of being formalised and made explicit. Many ideas from calculus, such as limits, differentiation, and convergence, had been intuitively understood since the time of Issac Newton and Gottfried Wilhelm Leibniz (and possibly earlier, depending on who one asks), but many of these concepts relied on a sort of faith that functions behaved the way functions ought to behave. Karl Weierstrass, a mathematician of the 19th century, believed that these "definitions" of calculus involved "too much hand waving, and not enough detail" (Kucharski [2014]).

One such definition of interest was differentiability. By the modern definition, a function $f : A \to B$, $A, B \subseteq \mathbb{R}$ is said to be *differentiable* at a point $x \in A$ with a derivative of $c \in \mathbb{R}$ if

$$\forall \epsilon > 0 \;\; \exists \delta > 0 \;\; \forall a \in A \setminus \{x\} \;\; |x - a| < \delta \implies \left| c - \frac{f(x) - f(a)}{x - a} \right| < \epsilon.$$

This definition can be quite difficult to wrap one's head around[4], but the idea behind it is straightforward. Informally, a function is differentiable at x if the function is smooth at x—that is, the function has no sharp points, cusps, jumps, or anything else that doesn't look smooth. Thus, differentiability provides a sort of measure of "smoothness". If a function is differentiable on its domain, it's said to be smooth. Otherwise, it's not. Furthermore, if we create a new function whose output at x is the derivative of f at x, and this new function is *also* differentiable at x, then the function f is *doubly differentiable* at x—even more smooth than merely being "singly differentiable". We can continue this process to define triply differentiable, quadruply differentiable, etc., creating a hierarchy of smoothness for functions.

As was discussed in section 1, mathematicians like things that are smooth since they typically have much nicer properties than things that are not smooth. This is true of differentiability, too: functions which are differentiable behave much nicer than non-differentiable functions. For example, a particularly nice result for functions,

[4]Perhaps looking at the raw definition of differentiability gives the reader some sense as to why mathematicians *informally* worked with derivatives for so long.

the Mean Value Theorem, applies only to functions that are differentiable. Even at a high level, differentiable functions are much nicer to work with simply because they're easier to visualise. Having a visual representation of an object of interest is one of the most valuable things a mathematician can have; it allows observations, conjectures, and explorations led by "gut feeling" to arise more naturally. It's no wonder, then, that mathematicians have a tendency to prefer working with smooth, differentiable functions; they look like nice, smooth curves when plotted, perhaps comparable to the rolling hills of a grassy meadow. Non-differentiable functions, by contrast, could have all sorts of weird behaviour, sharp edges, jumps, asymptotes, and other features that are harder to imagine.

In fact, mathematicians loved working with differentiable functions so much in the 19th century that they made sweeping generalisations about *all* functions in an attempt to make more functions "nice" functions. These generalisations weren't based in measurable definitions, and so there was no way to know for sure whether these generalisations were even valid. However, they *felt* right to most mathematicians, and so they were taken as fact.

One such generalisation was this: "for any continuous curve, [it's] possible to find the [derivative] at all but a finite number of points" (Kucharski [2014]). In simple terms, mathematicians believed that all continuous[5] functions were smooth (expect possibly at a few specific points). This agreed with the types of functions mathematicians were concerned with at the time: simple polynomial equations for gravity, the motion of pendulums, the velocities of objects thrown in the air. These were the only functions on many mathematicians' minds (Kucharski [2014]), and so to them, they were representative of *all* functions. This generalisation was so ingrained in mathematical thinking that there was a supposed "proof" of this fact by the French mathematician André-Marie Ampère (Kucharski [2014]).

This "proof" severely bothered Weierstrass. With the lack of concrete definitions within calculus at the time, it wasn't even clear what was meant by a "continuous curve", so how could anything be said about them? To solve this issue, Weierstrass took it upon himself to create proper definitions for many of the concepts taken for granted by his contemporaries (Kucharski [2014]). With these definitions established, he then went a step further and set out to find a function that completely went against Ampère's "proof". Bernard Riemann had already claimed that such a function existed all the way back in 1861 (Weisstein [2024h]), so it wasn't entirely unbelievable that a function like it was out there somewhere.

[5]Continuity was another property that didn't have a formal definition at the time. They were viewed as functions whose plots were a single connected curve.

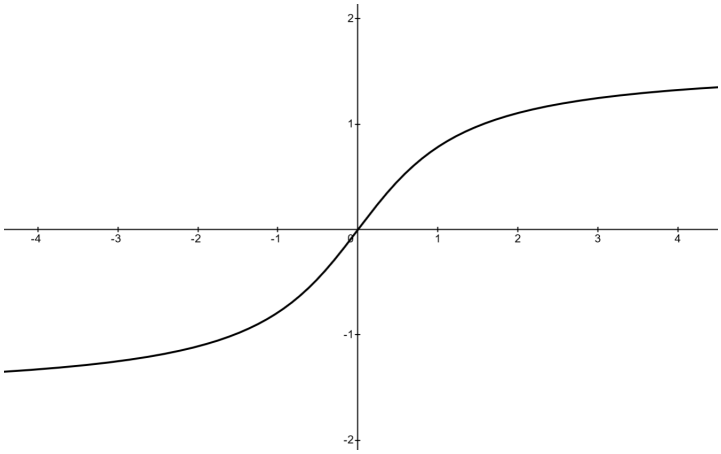


FIGURE 3. A plot of the function $\arctan(x)$ created using Desmos. To a mathematician in the 19th century, all continuous functions were comparable to curves like this.

In 1872, Weierstrass succeeded in finding his function: a function that was continuous but whose derivative couldn't be found at any given point[6]. In other words, the function was continuous but not smooth.

What does this function look like? Well, its formula is given by

$$f(x) = \sum_{k=1}^{\infty} \frac{\cos(3^k \pi x)}{2^k},$$

but this is a rather hard function to visualise given only the expression. A plot of the function would make it much easier to see that f isn't smooth. However, this function was crafted in 1872. Weierstrass couldn't just open Maple and get it to spit out a plot. If he or any other mathematician wanted to plot this function, they'd have to do it by hand. For a function as monstrous as this, that wasn't going to happen.

The fact that the function couldn't be plotted didn't matter to Weierstrass. He had rigorous definitions to play with; his claim that f wasn't differentiable didn't *need* pictures. His argument was a purely abstract one. Whether he could see the function's roughness or not, he knew it was there (Kucharski [2014]).

So, here was another non-smooth object that mathematicians didn't know how to manage. Worse than that, it was an object mathematicians didn't *want* to manage. Henri Poincaré called the function a literal monster[7] (Kucharski [2014]). Like with

[6]More precisely, its derivative can only be found on a set of points with measure zero, meaning if you were to smoosh all the points together, the interval enclosing them would be arbitrarily small (see Weisstein [2024f]).

[7]The author would like to make it known that they believe the term "monster" is quite endearing for Weierstrass' function, and it only makes them like it more.

Brownian motion, we'll return to Weierstrass' function when we've built up some handy tools and ideas for working with roughness.

2.3. **Julia Sets - Part 1.** The year is 1918. Gaston Julia and Pierre Fatou were in the process of studying a particular kind of iterated system along with a corresponding set (Spencer [1999]). At first glance, the systems they were considering appear deceptively simple. We start with some rational function, say $f(z)$. Importantly, this is a function whose domain and codomain are the complex numbers (possibly along with an extra point representing infinity) (Weisstein [2024c]). Then, we take any starting value z_0 and send it through $f(z)$. This gives us a new value z_1. Now, we take z_1 and repeat the process to get z_2. Then, we repeat to get z_3. And so on. More formally, what we're doing is generating a sequence of complex numbers using the recurrence relation

$$z_{n+1} = f(z_n).$$

What can we say about the sequence (z_n)? Does it converge to some fixed value? Does it diverge to infinity? Does it bounce around without ever settling down? It turns out that all three behaviours are possible. To demonstrate this, consider the rational function $f : \mathbb{C} \to \mathbb{C}$, $f(z) = z^2$. If our starting point z_0 is such that $|z_0| < 1$, then repeatedly applying f will result in a sequence of complex numbers whose magnitudes get progressively smaller. If we use polar form to represent this sequence, it's clear that the sequence converges to 0. Conversely, if we choose a starting point z_0 such that $|z_0| > 1$, then the resulting sequence of complex numbers will have magnitudes that get progressively bigger. Again, using polar form, it's clear that this sequence diverges to infinity.

This leaves us with the boundary case. What happens if $|z_0| = 1$? Any complex number with a magnitude of 1 can be represented in the form $e^{i\theta}$, where θ is the angle from the positive real axis to the ray connecting our point to the origin. If we apply f to points of this form, we see that

$$f(e^{i\theta}) = (e^{i\theta})^2 = e^{2i\theta}.$$

After applying f, we get another number whose magnitude is 1. Thus, in this case, our sequence (z_n) will neither converge to zero nor diverge to infinity. Instead, the sequence of points will remain on the unit circle, possibly bouncing around in some way. Other than the special cases where $z_0^{2^n} = 1$ for some $n \in \mathbb{N}$ (meaning z_0 is a 2^n-th root of unity), starting points z_0 where $|z_0| = 1$ will result in sequences that diverge not because successive magnitudes get larger, but because the points simply never settle down to a fixed point.

We see that this third case, the case where our starting points are on the "boundary" of the two simpler cases, results in more interesting sequences. As any curious mathematician would do, we should give out some names to more easily discuss these concepts. Let $\tilde{J}_f$ be the largest invariant subset of $\mathbb{C}$ under f. In other words, $\tilde{J}_f$ is the collection of all starting points that correspond to sequences that *don't* diverge to infinity. We'll say the *Julia set* of f, represented by J_f, is the boundary of $\tilde{J}_f$, i.e. $J_f = \partial \tilde{J}_f$ (Weisstein [2024c])[8].

Because the deciding criteria is so clear for which starting points give each type of sequence behaviour, we can draw a diagram to visually depict the sets $\tilde{J}_f$ and $\mathbb{C} \setminus \tilde{J}_f$ for this particular function f:

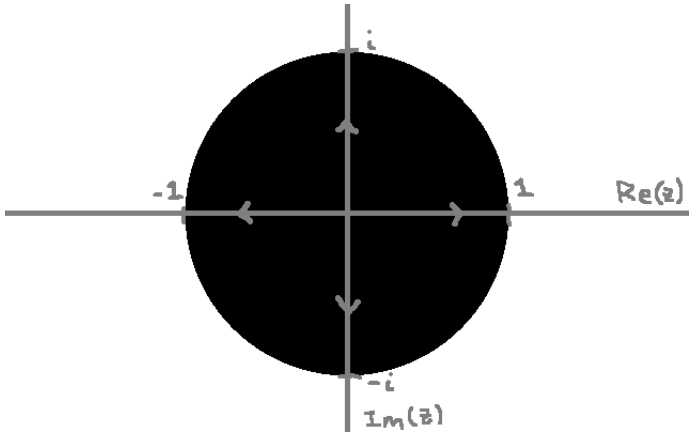


FIGURE 4. A very exciting visual representation of the set $\tilde{J}_f$ for $f(z) = z^2$. The black region represents points in the complex plane whose corresponding sequences don't diverge to infinity. The white region represents starting points with corresponding sequences that do diverge. The boundary of the two regions is the Julia set J_f.

The resulting diagram is about as simple as it can get. The filled-in unit circle is $\tilde{J}_f$, while the white space surrounding it is $\mathbb{C} \setminus \tilde{J}_f$. Admittedly, this isn't very exciting, but it makes the Julia set J_f fairly easy to envision; it's simply the boundary, or "edge", of the circle.

If all Julia sets were as simple as this, then there wouldn't be much to discuss. However, as we may have been able to guess, most Julia sets are far more complex. How much more complex? Well, with the exception of a few particularly nice functions (like $f(z) = z^2 - 2$ and our function $f(z) = z^2$), most Julia sets are complicated enough so as to be practically impossible to draw by hand. Thus, when Julia and

[8]More visually, J_f is the set of all points $z \in \mathbb{C}$ such that, no matter how closely we zoom into z, there will always be a ball around z that contains both points inside $\tilde{J}_f$ and points outside $\tilde{J}_f$. Symbolically, we can say $J_f = \{z \in \mathbb{C} : \forall \epsilon > 0 \; B_\epsilon(z) \not\subseteq \tilde{J}_f \text{ and } B_\epsilon(z) \not\subseteq \mathbb{C} \setminus \tilde{J}_f\}$. We'll discuss what it means to be on the boundary of a set in section 3.4.

Fatou were investigating Julia sets, they found themselves in a situation similar to Weierstrass back in 1872: they wielded a rough, complicated object they couldn't visualise. As such, Julia and Fatou had to develop tools specific to these sets in order to make sense of them. Only with the introduction of the concept of *fractals* did mathematicians have a general set of tools for working with objects as rough as Julia sets.

Once again, we'll return to Julia sets later in this project once we discuss the ideas behind fractal objects.

3. Embracing Roughness

3.1. **Even the Definition is Rough.** Finally, we reach 1975, the year the term *fractal* was coined (Snyder [2006] 1) by Benoit B. Mandelbrot. The word was used to describe objects that possessed some sort of self-similarity, meaning parts of the whole mimicked the whole. Immediately, we see similarities between this idea and the objects of nature we mentioned briefly in section 1. The leaves of a fern mimic the entire fern, and the bundles of petals on a Queen Ann's lace flower look like a miniature version of the entire flower. Then, using this informal definition of a fractal, we could say that much of the roughness observed in nature is "fractally" and employ the techniques of fractal geometry to tackle it. Though, as we learned from Weierstrass back in 1872, we'd be better off having a rigorous definition for fractals so we can make sound, watertight conclusions about them. How, then, do we define a fractal?

This is where things get a little dicey. Mandelbrot, the mathematician responsible for popularising fractals and fractal geometry, could be described as a bit of a wanderer. The man was full of many ideas, but "[i]nstead of rigorously proving his insights in each field, he . . . preferred to 'stimulate the field by making bold and crazy conjectures' — and then move on before his claims had been verified" (Hoffman [2010]). In other words, Mandelbrot often left his ideas only partially fleshed out, leaving other mathematicians to do the "dirty work" of formalising them. As a result, a precise definition for what constitutes a fractal is hard to find. The definition changes depending on the source.

Sandra S. Snyder in their expository paper "Fractals and the Collage Theorem" defines fractals as "mathematically generated pattern[s] that [are] reproducible at any magnification or reduction . . . or at least [have] similar structure" (Snyder [2006] 1). Many fractals indeed possess some level of self-similarity (as we'll see

later in the project), but too much emphasis on this property can obscure the heart of what fractals truly capture: roughness. Brownian motion is often viewed as a type of fractal specifically "because the boundary of the surface [or path] is very rough" (Mickelin [2017]). The Weierstrass function is considered fractally because it constituted an "intricate, beautiful structure" rather than one where the roughness "would disappear when magnified sufficiently" (Kucharski [2014]). Ditto for Julia sets. In all the examples we've considered so far, the key reason why mathematicians apply methods from fractal geometry to study them is because other methods fail; they're too jagged and rough for "traditional" methods to work.

Thus, perhaps a more apt definition for a fractal is any object whose "roughness" prevents it from being measured/manipulated via traditional geometrical or algebraic means. This definition certainly separates fractals from the usual smooth objects of math, yet it remains imprecise. What exactly do we mean by roughness? Intuitively, it's clear what we mean, but mathematicians also had an intuitive understanding of what it meant to be smooth, and this notion led them astray. Or, worse still, even if we're given an object that's sufficiently rough (in whatever sense we eventually decide on), how can we know for sure that "traditional" methods won't work to measure them? What do we even mean by "traditional"? It seems this definition may be more ill-defined than the last.

However, as it surprisingly turns out, having a concrete definition for a fractal really isn't all that important.[9] Many rigorous tools have been developed for understanding the geometry and behaviour of fractals, and intuitively recognising a fractal is enough to decide when to use said tools. This was the approach Mandelbrot used himself when first encountering roughness in his studies. His early work involved studying the behaviour of stock markets[10], but for this project we'll focus on how he tackled a seemingly innocuous question: how long is the coastline of Great Britain?

3.2. **A Complicated Walk Along the Coast.** 1967. Imagine we're walking along some path drawn in chalk on the street. The path is a single continuous curve—just one stroke of chalk. If the path is a straight line, or at least close to one, then following this path will be easy. Now, what happens if the person drawing this path decides to get crafty and scribble knots into it (see figure 5)? Well, if the knots are small enough, then they pose no problem to us. We can simply step over them and continue down the chalk path. We can also ignore any jagged sections of the path, provided the jags are small enough (again, see figure 5). The path forward will be clear enough without having to consider these slight deviations.

[9]In any case, we'll give one possible "rigorous" definition for a fractal in section 3.4.

[10]As an example of some of the economic work Mandelbrot engaged in, see Mandelbrot and Taleb [2006].

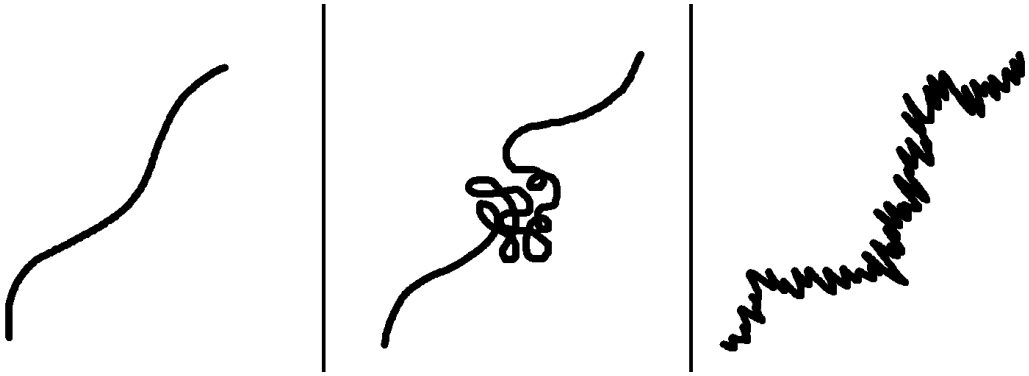


FIGURE 5. On the left, a nice, smooth chalk path. In the middle, a devious chalk path with a knot. On the right, a crinkled chalk path.

Us humans walking down these different sections of the chalk path will notice very little difference in the length of the path. None of the knots we step over force us to walk any further. The length of the path is the same as if the knot wasn't there. The same can be said about any crinkly sections of the chalk path. Whether the path is smooth or jagged, the distance we need to walk remains unchanged.

The same can't be said about an ant following the chalk path. Knots and spikes that are small to us are much bigger for ants (see figure 6). To skip these knots and spikes would be to stray from the path. Thus, for an ant to follow the same knotted, scrunched up chalk path, it must walk a lot further than a human would.

It would be even worse if we were a single-celled blob creature. Such a creature would be so small that the flat surface of the chalk path would turn into a vast stretch of peaks and valleys (again, see figure 6). This, clearly, would add even more distance to our walk.

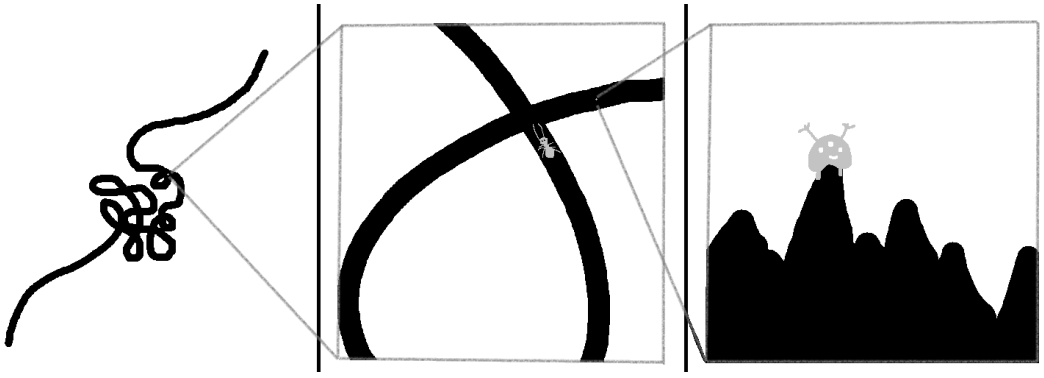


FIGURE 6. If we were ants following the chalk path, knots and crinkles would add a considerable amount of distance to our walk. If we were single-celled blob creatures, the chalk itself would prove to be a mountainous landscape.

Depending on how small a scale we look, the length of the chalk path changes considerably. This begs the question: how long is the chalk path? At what scale should we measure its length? The human scale would be practical for us. It would give us the "practical" length of the path, the distance we'd have to walk to get from one end of the path to the other. The ant scale would be practical from a more mathematical perspective. While knots and squiggles don't add any additional walking distance, they certainly add more length to the path itself. A loop-de-loop path is far longer than a straight path. Thus, the ant scale would provide a "truer" measurement of the path's length. The blob scale, while certainly ridiculous, is technically the most accurate out of the three scales. The mathematical path represented by the chalk is an object purely in two-dimensional space. It doesn't take into account the physical materials required to construct the path. In our three-dimensional world, however, drawing lines with chalk requires that the line exists in three-dimensional space, and so the hills and valleys the blob must climb over are just as much a part of the curve as any other aspect.

This problem may seem pedantic in nature, but the question of determining the proper scale at which to measure length is one that's been encountered in real-world applications (Mandelbrot [1967] 636). Consider the profession of geography. One of the things geographers may want to include on a map or reference is a measure of the length of a certain coastline. Unfortunately, coastlines tend not to be simple straight lines or curves. In fact, they're comparable to our chalk path example: full of nooks, crannies, and jagged edges. Because coastlines are so similar to our chalk paths, they suffer the same problem of their lengths changing depending on the scale at which they're measured.

By comparing the length data of various coastlines (see figure 7), we may notice that the lengths of some coastlines seem to increase faster than other coastlines as the scale decreases. For instance, the length of the west coast of Great Britain seems to increase quite quickly as the scale decreases, while the coast of South Africa barely changes at all. If we compare these two coastlines (see figure 8), one obvious reason for this difference comes to mind: their difference in roughness. The reason our chalk path grew longer at smaller scales was because of how "bunched up" it was in some places. All the knots and crinkled sections added more length for the ant and blob creature to traverse. So, it would make sense for the coastline with more jagged sections (Great Britain) to measure longer at smaller scales than compared to the smoother coastline (South Africa).

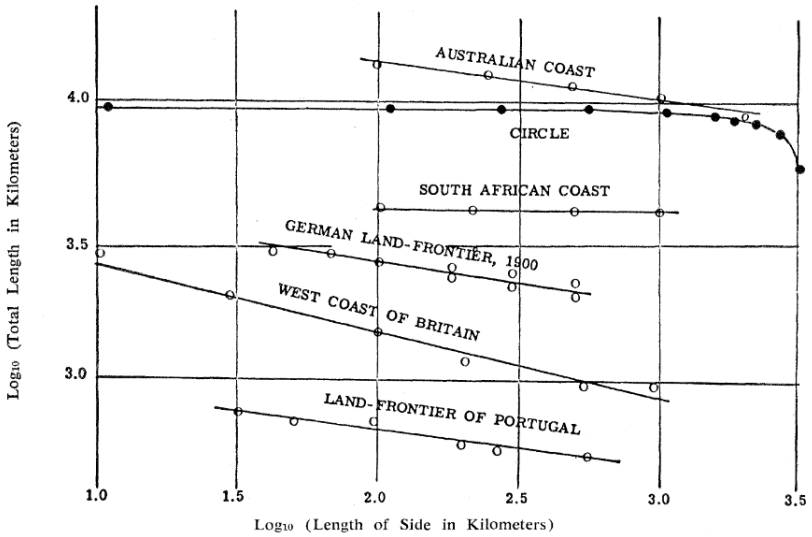


FIGURE 7. Various coastline lengths as functions of the scale at which they're measured. As the scale decreases, the lengths of the coastlines increase. Taken from figure 1 of Mandelbrot [1967].

In fact, this observation was verified empirically by the mathematician L. F. Richardson (Mandelbrot [1967] 636). Using gathered data for the lengths of several coastlines, Richardson devised a formula for relating the measured length of a coastline and the scale at which it's measured. If G is taken to be the smallest length considered in our measurements (so that we ignore any knots or squiggles in the coastline smaller than G), $L(G)$ is the measured length of a coastline as a function of G, and M and D are constants which change per coastline, then Richardson's observation was that

$$L(G) = MG^{1-D}.$$

Importantly, Richardson noted that the constant D had some relation to how rough a coastline appeared. For instance, he found that, for South Africa, $D \approx 1.02$,

FIGURE 8. On the left, the coastline of South Africa (the bottom edge). On the right, the coastline of Great Britain. Images taken from WorldAtlas.com.

whereas for Great Britain, $D \approx 1.25$ (Mandelbrot [1967] 637). In essence, the rougher a coastline, the higher its value of D. For small enough G, a higher value of D means a higher value of $L(G)$ under the formula, which matches our intuition about rougher coastlines having more length at smaller scales. So, according to Richardson's findings, coastlines don't really have an intrinsic length—it depends on the scale at which you measure them.

While there isn't much else to say about the question of coastline length, there's still the question of understanding Richardson's relation. What does the value of D truly represent? Is there some formal interpretation of what it means for one coastline to be "more rough" than another? Mandelbrot believed that there was such a formal interpretation. Rather than merely attributing each coastline a roughness measure D, he declared the *dimension* of the coastline to be D (Mandelbrot [1967] 637). So, rather than leaving dimensions as integer values (e.g. lines are one dimensional, squares are two dimensional, etc.), he allowed dimensions to be fractional. The dimension of South Africa's coastline, then, was 1.02, while the dimension of Great Britain's coastline was 1.25. It wasn't some arbitrary measure of roughness that D was measuring, but rather a coastline's dimension, something which could be precisely determined rather than "eyeballed".

Such an assertion may seem nothing less than insane. How can we possibly have fractional dimensions? If a coastline is 1.25-dimensional, does that mean it exists in a space with an x-axis and 25% of a y-axis? In fact, Mandelbrot's idea of dimension isn't really related to the number of coordinates axes in a space. Rather, it relates to the way in which objects scale in that space.

Consider a line segment of length x, and suppose we stretched it by a factor of s. The new line segment, unsurprisingly, has a length of sx, and so the object grew by a factor of s. As well, the original line segment of length x can be obtained by scaling the new line segment by a factor of $\frac{1}{s}$. This behaviour is how all one-dimensional objects (i.e. all lines) behave, and so there should be a way to link the quantities we just listed to the value of the dimension (i.e. 1). It turns out that there is, via a slightly strange formula:

$$\frac{-\log(s)}{\log(\frac{1}{s})} = \frac{-\log(s)}{\log(s^{-1})} = \frac{-\log(s)}{-\log(s)} = 1.$$

In some sense, this relation is what it means for an object to be one-dimensional. When a one-dimensional object is stretched by some factor s, the resulting object will be s times longer, and the original object should be $\frac{1}{s}$-th the length. This characterisation of dimension may seem overly-complicated, especially for one-dimensional objects, but it ends up generalising to higher dimensions quite nicely.

Now, consider a square with side length x, meaning its area is x^2. Suppose we stretched it by a factor of s. Then, the new side length of the square would be sx, meaning the new area would be s^2x^2. Scaling the square by a factor of s scales the area by s^2. To obtain the original square from our current stretched one, we'd again scale by a factor of $\frac{1}{s}$, just as we did with the line segment. These are properties that all two-dimensional objects share. We can link these quantities to the dimension by the following relation:

$$\frac{-\log(s^2)}{\log(\frac{1}{s})} = \frac{-\log(s^2)}{\log(s^{-1})} = \frac{-2\log(s)}{-\log(s)} = 2.$$

In general, if we have an object that's grown by a factor of N, and requires scaling by $\frac{1}{s}$ to get back to its original size, then its dimension is given by

$$D = \frac{-\log(N)}{-\log(s)}.$$

Mandelbrot treated this expression as the *definition* of dimension, rather than a formula that just so happens to coincide with dimension for the regular shapes of geometry (Mandelbrot [1967] 637). This is the way in which he linked Richardson's value of D for coastlines to the concept of dimension: the two values of D, he

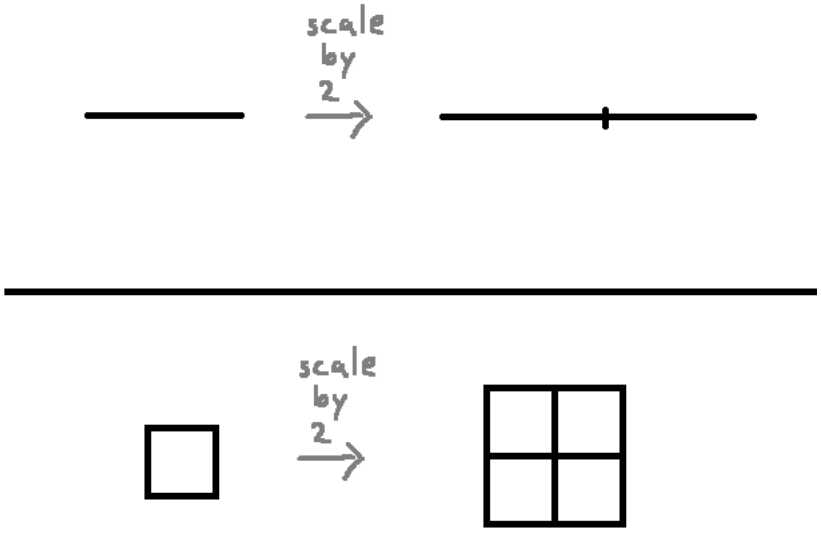


FIGURE 9. An example of scaling a line segment and a square by a factor of 2. In the case of the line, it doubles in length. In the case of the square, it quadruples in area. These differing changes in a shape's "size" can be thought of as an intrinsic property of all objects with the same dimension as these shapes.

asserted, were the same value. To argue this, he made use of a peculiar kind of mathematical object called a "nonrectifiable self-similar curve".

3.3. **Nonrectifiable Self-Similar Curves.** A *rectifiable* curve, according to Mandelbrot, is any smooth curve with a well-defined length (Mandelbrot [1967] 637). A *nonrectifiable* curve, then, is a curve that isn't rectifiable. A perfect example of a nonrectifiable curve is any of the coastlines mentioned in the previous section. At different scales, the coastlines have different lengths, and so their lengths are not well-defined and thus are nonrectifiable.

To establish his link between roughness and dimension, Mandelbrot needed a set of nonrectifiable curves that were easier to work with mathematically. The coastlines he was looking at displayed the exact sort of behaviour he wanted, but since the curves weren't defined in a controlled way, it was hard to argue that his idea was valid. Thus, he turned to an object that had been first described by the mathematician Helge von Koch: the Koch curve (Wolfram [2002] 934).

The Koch curve, much like the Weierstrass function, was an object cooked up specifically to disprove the informal ideas and "proofs" mathematicians had regarding smooth curves in the 1800s (Wolfram [2002] 934). However, unlike the Weierstrass function which was essentially impossible to visualise, the Koch curve can easily be visualised by virtue of how it's constructed.

Start with a line segment of some specified length—usually chosen to be one for simplicity. The Koch curve is created by repeatedly applying the following rule: for every straight line segment, divide the segment into three equal pieces. Then, construct an equilateral triangle whose bottom side is the middle line segment. Erase the bottom line segment. The original straight line segment should now be replaced with four new line segments, each a third the length of the original, that form a sort of triangular spike.

Describing this process in words is much harder than simply showing the process visually. Figure 10 shows the first four iterations of the construction process. We can imagine that, if the process is carried out indefinitely, we'd end up with an incredibly spiky object. The curve would be so spiky, in fact, that the entire curve would be jagged, with no straight line segments remaining. In this way, it's relatively easy to see that, if we interpreted the Koch curve as a function, then it wouldn't be differentiable (i.e. smooth) anywhere. As well, it's also easy to see that the curve is continuous since the entire object is one long jagged line—there aren't any jumps or asymptotes on the curve. Thus, the Koch curve provides another counterexample to Ampère's "proof" that all continuous functions have derivatives

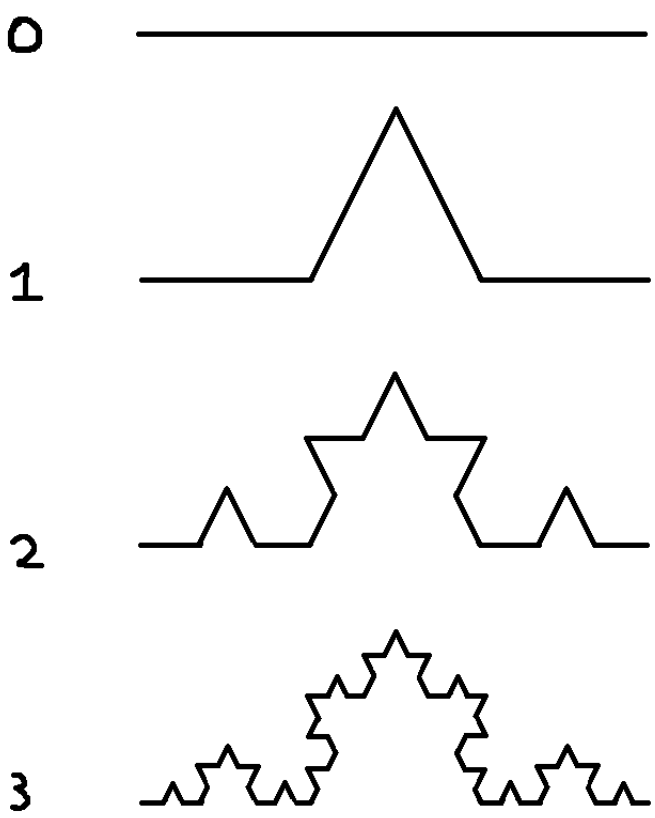


FIGURE 10. The first four iterations of the Koch Curve construction process. Note that the triangles used here are not equilateral triangles, but the resulting curve is essentially the same, just stretched slightly.

at all but a finite number of points (see section 2.2). As an added bonus, unlike the Weierstrass function, the Koch curve can be drawn (given a bit of patience) on a piece of paper.

Of course, we haven't actually *proven* that the Koch curve is continuous and/or nondifferentiable. We'd need to use the definitions for that. In section 3.4, we'll introduce the tools required for showing that something like the Koch curve is continuous and nondfferentiable (as well as the tools for showing that the Koch curve is a well-defined object). For now, we'll concern ourselves with a different property of the Koch curve: its dimension.

If we applied Mandelbrot's definition of dimension to the Koch curve, what would we get? Well, imagine we scale the curve up by a factor of 3. So, if the left tip of the curve originally started at $x = 0$ and the right tip ended at $x = 1$, then the scaled up right tip would end at $x = 3$. If the Koch curve were a normal line, we'd expect the length of the curve to increase by a factor of 3. However, this isn't what happens. The Koch curve has an extra "bump" in its middle which adds more length to the

curve. In fact, we can actually break up the scaled Koch curve into four smaller segments, each one an identical copy of the original curve (see figure 11). Because it has four segments, scaling up the Koch curve by a factor of 3 scales its length by a factor of 4, not 3 (since the "middle third" of the curve is actually composed of two-thirds worth of length due to the bump).

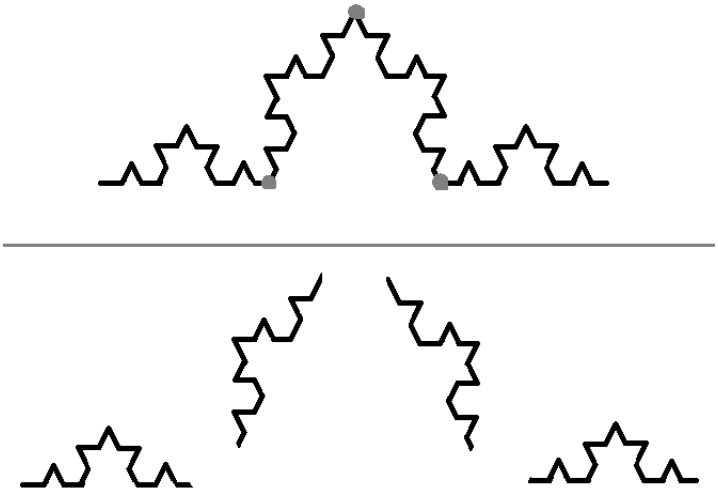


FIGURE 11. If we split the Koch curve at the grey dots, we get four identical smaller curves. If we imagine splitting the "final" Koch curve (i.e. the result of applying the construction process indefinitely) in the same way, the four smaller curves would be identical to the larger curve.

To scale the curve back to its original size, we scale by $\frac{1}{3}$ as we would for any shape. Thus, using Mandelbrot's definition of dimension, the Koch curve has a dimension of

$$D = \frac{-\log(4)}{-\log(3)} \approx 1.2619.$$

However, there's another, equivalent way to obtain this number. Consider again the construction process detailed in figure 10. For simplicity, we'll say that the original line segment at step 0 has length 1. At each step s of our construction, our curve consists of 4^s line segments, each one a length of $\left(\frac{1}{3}\right)^s$. Thus, the total length of our curve at step s is given by $\left(\frac{4}{3}\right)^s$. Rearranging this expression, we get that

$$\left(\frac{4}{3}\right)^s = 4^s \left(\frac{1}{3}\right)^s = \left(3^{s\frac{\log(4)}{\log(3)}}\right)\left(\frac{1}{3}\right)^s = \left(\left(\frac{1}{3}\right)^{-s\frac{\log(4)}{\log(3)}}\right)\left(\frac{1}{3}\right)^s = \left(\frac{1}{3}\right)^{s\left(1-\frac{\log(4)}{\log(3)}\right)}.$$

Notice that $\left(\frac{1}{3}\right)^s$ is the length of the smallest line segment in the Koch curve at step s. Then, using Richardson's notation for measuring the lengths of coastlines, we can say $G = \left(\frac{1}{3}\right)^s$ since the smallest line segment in a step of the Koch curve is the smallest length considered when measuring the length of the curve. Then, our

expression simplifies to

$$\left(\frac{1}{3}\right)^{s\left(1-\frac{\log(4)}{\log(3)}\right)} = G^{1-\frac{\log(4)}{\log(3)}} = G^{1-D}.$$

So, the length of the Koch curve at step s is given by G^{1-D}, where D is the dimension given to the curve and G is the length of the smallest line segment being considered. This is the exact formula Richardson derived empirically for the lengths of coastlines, just without the extra scaling constant. However, in Richardson's formula, the constant D was a sort of measure of roughness for the given coastline. Here, D represents the dimension of the Koch curve. Comparing the two, it seems that dimension and "roughness" play the same role in our expressions for length. In fact, this holds true for many nonrectifiable curves (Mandelbrot [1967] 637), not just the Koch curve.

Seeing this connection between Mandelbrot's definition of dimension and Richardson's measure of roughness, Mandelbrot could then argue that a curve's roughness could be measured in terms of its supposed dimension. However, Mandelbrot wanted a way to make the connection even stronger. He'd shown that the dimension of a curve served the role of Richardson's roughness constants for these peculiar nonrectifiable curves, but was that enough to show that it also applied to natural coastlines? What if the geometry of the world's beaches behave differently from these mathematical models of roughness?

It turns out that they don't, at least not in any way relating to what Mandelbrot wanted to show. Using a probabilistic process for constructing "random" jagged objects like the coastlines Richardson had studied, Mandelbrot argued that the coastlines of the world behaved no differently than these "toy" models he had used (Mandelbrot [1967] 638), and so the roughness of coastlines can also be linked to the notion of dimension.[11]

Therefore, rather than considering the increasing value of a curve's length as the scale is decreased, Mandelbrot showed that mathematicians can quantify the "roughness" of an object by using its "dimension" as it relates to properties of scaling. To differentiate this definition of dimension with other, more conventional definitions, this quantity is often referred to as the *capacity dimension*, *fractal dimension*, or *Hausdorff dimension* (Weisstein [2024a]). This project will use the term fractal dimension.

[11]The specifics of this probabilistic construction for coastline-like objects is beyond the scope of this project, though it can be succinctly described as choosing a series of these "toy" models at random and combining them in such a way as to form "natural" jaggedness as seen in real-world coasts. See Mandelbrot [1967], pages 637-638.

An object's fractal dimension can actually be used as a sort of quantifying value for fractals. So, rather than having to decide for ourselves whether an object "counts" as a fractal or not, we can look at the fractal dimension and get a definite answer. However, to discuss that, we need to get a little more technical.

3.4. **Metric Spaces.** In section 3.3, we constructed the Koch curve and calculated its fractal dimension. This process required that we consider the "final" step of the Koch curve construction process—that is, we had to consider how the resulting object would look/behave if we carried out an infinite number of steps. More formally, we were required to consider the limit of the construction process as the number of steps went to infinity. Limits, much like derivatives, have a very precise mathematical definition to ensure that things are done correctly. Our method of handling the limit of the Koch curve construction process was very much *not* precise. We didn't even mention the word "limit"! How, then, can we be sure that any of our analysis was correct if we made use of a limit we didn't derive rigorously?

In fact, how can we even be sure that such a limit exists? The usual definition of a limit assumes that we're dealing with sequences or functions of numbers. The Koch curve is certainly not a sequence nor a function of numbers, at least not in any obvious sense. Rather, it's a set of lines in a two-dimensional plane. If we did try and rigorously take the limit of our construction process, it isn't clear how the traditional definition of a limit would work.

Thankfully, there exists a type of space more general than the real numbers in which limits can be taken, and this space lends itself quite naturally to *shapes* as opposed to numbers and functions. Once we understand this type of space, we'll have the tools we need to properly handle the limit of the Koch curve creation process.

The type of space in question is called a *metric space*. A *metric space* (X, d) is an arbitrary set X, along with a *metric* $d : X \times X \to \mathbb{R}$. The metric d satisfies the following four properties:

(1) For all $x, y \in X$, $d(x, y) \geq 0$.

(2) For all $x, y \in X$, $d(x, y) = d(y, x)$.

(3) For all $x, y \in X$, $d(x, y) = 0 \implies x = y$.

(4) For all $x, y, z \in X$, $d(x, y) \leq d(x, z) + d(z, y)$.

The definition of a metric space is rather abstract, but its interpretation is quite straightforward: a metric space is a set of objects along with some way to measure "distance" between objects. The value of $d(x, y)$ for points $x, y \in X$ is taken as

the distance between x and y. The four properties the metric d must satisfy ensure that the distances we have between objects actually makes sense as distances (e.g. the distances are nonnegative, they approach zero as the points get closer together, etc.).

Using a metric space's metric, we can redefine the notion of a limit. Given a metric space (X, d) and a sequence of points (x_n) within X, we say that the sequence converges to a point $x \in X$ (i.e. the limit of the sequence is x) if

$$\forall \epsilon > 0 \;\; \exists N \in \mathbb{N} \;\; \forall n \geq N \;\; d(x_n, x) < \epsilon.$$

This definition is practically the same as it is for sequences of real numbers or functions, except we measure the "closeness" of points in the space using a general metric function instead of the absolute value of the difference between numbers. However, this definition is much more powerful as the points in X can be *any* abstract objects we want, so long as we have a way to measure some kind of distance between them. For our Koch curve, we just need to find the appropriate metric space in which the curve lives, and then show that the sequence of curves we get from the construction process converges to the "final" Koch curve.

Because the Koch curve sits in a two-dimensional plane, and because we can draw it on a piece of paper (at least, we can partially draw it using the construction process), an obvious choice for a space to put the curve in would be $\mathbb{R}^2$. For example, at step zero, the Koch curve would be the set of all points in $\mathbb{R}^2$ that sit on the line between the points $(0, 0)$ and $(1, 0)$. At step one, it would be the set of all points that sit between the four lines going from $(0, 0)$ to $\left(\frac{1}{3}, 0\right)$, $\left(\frac{1}{3}, 0\right)$ to $\left(\frac{1}{2}, \frac{\sqrt{3}}{6}\right)$, $\left(\frac{1}{2}, \frac{\sqrt{3}}{6}\right)$ to $\left(\frac{2}{3}, 0\right)$, and $\left(\frac{2}{3}, 0\right)$ to $(1, 0)$.

As well, we know how to measure the distance between two points in $\mathbb{R}^2$ (by subtracting them and calculating the length of the resulting vector), so our metric to use is also obvious. Thus, we can say that the Koch curve is a set of points (given by line segments) within the metric space $(\mathbb{R}^2, d)$, where $d(\vec{x}, \vec{y}) = ||\vec{x} - \vec{y}||$.

Now, to show convergence of the Koch curve creation process, we can't directly use the metric space $(\mathbb{R}^2, d)$. Remember, the objects we're interested in showing convergence for are *sets* of points in $\mathbb{R}^2$, not points in $\mathbb{R}^2$. The definition of convergence for metric space doesn't allow us to directly show convergence for collections/sets of objects within a metric space; we have to work with single points. So, if we want to show convergence of the Koch curve creation process, we actually have to create another metric space, one whose objects of interest are *sets* of points in $\mathbb{R}^2$.

This, however, raises an interesting question: if we want to make a metric space whose objects are sets, how to we measure distance between them? There are many ways to do this, though one way that has proven useful to mathematicians is to use the *Hausdorff metric*. To use the Hausdorff metric, we have to be a bit more specific about what sets of points in $\mathbb{R}^2$ we wish to use for the objects in our metric space. Let's introduce a few more definitions to help us out.

For any given metric space (X, d), we'll say that the *boundary* of a subset $A \subseteq X$, designated as ∂A, is the set of all points in X that, according to the metric d, are a distance zero from both the set A and the complement $X \setminus A$. Thus, the boundary of a subset is literally the boundary between the set itself and everything outside of the set. Not much to explain here.

For the same metric space (X, d), we'll say that a subset $A \subseteq X$ is *closed* if it contains its boundary. Symbolically, A is closed if $\partial A \subseteq A$. So, a set is closed if it has its own boundary "closing it in". In this way, sets are almost like plots of land. To close a piece of land off, we need to fence it in. To close a set, we need its boundary.

Similarly, we'll say a subset $A \subseteq X$ of the metric space (X, d) is *open* if its complement is closed. Open sets, then, can be thought of as the opposites of closed sets.

For the same metric space (X, d), we'll say that a subset $A \subseteq X$ is *bounded* if there exists some point $x \in A$ and some finite $\epsilon > 0$ so that every point in A is at most a distance ϵ from x. Again, this definition isn't too difficult to parse. If every point is at most a distance ϵ from x, then the entire set A can be contained in an enclosing ball of radius ϵ, meaning the set is bounded by that enclosing ball; everything is contained within some finite space.

Unfortunately, the next definition is a little more difficult to understand intuitively. For any given metric space (X, d), we'll say that a subset $A \subseteq X$ is *compact* if any sequence of points in A has a convergent subsequence whose limit is in A.[12] Unlike the previous definitions, there isn't an immediate visual that comes to mind when trying to imagine what a compact subset would look like. Thankfully, the sets we're most concerned about in this section are subsets of $\mathbb{R}^2$, and there's a particularly nice result that makes interpreting compactness much easier for these subsets.

[12]There are actually several different definitions of compactness for metric spaces. The one given here is for "sequential" compactness. No matter the definition used, however, all the same results apply.

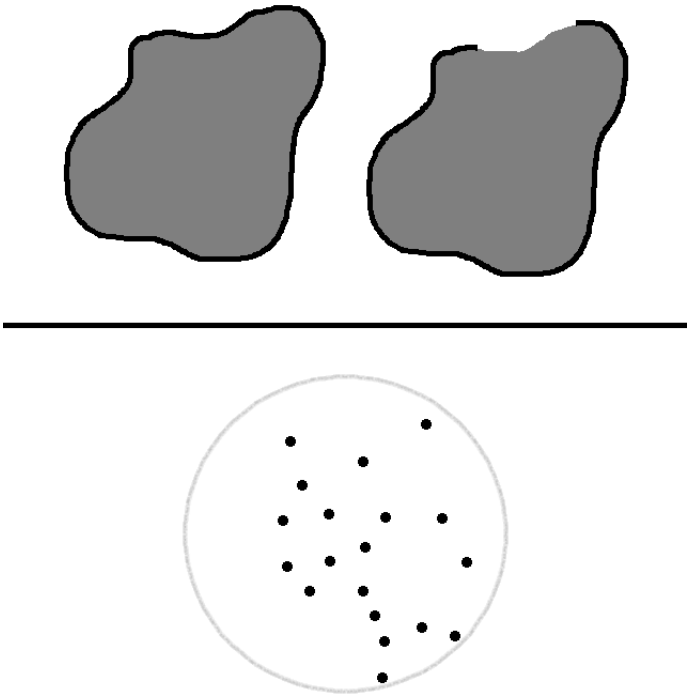


FIGURE 12. Top: two sets, represented by the grey shading and the black boundary. The left set is closed because it contains its own boundary. The similar set on the right is *not* closed because a piece of its boundary is missing. Bottom: a set represented by black points. This set is bounded because a ball of some finite radius can be drawn around it to enclose it.

The Heine-Borel Theorem says that a subset of $\mathbb{R}^2$ (and, more generally, a subset of $\mathbb{R}^n$) is compact if and only if it's closed and bounded. So, for our purposes, being compact simply means you're a closed *and* bounded set.

Now, with these definitions, we can define an appropriate metric space on which the Hausdorff metric can be used. For any metric space (X, d), define the *Hausdorff space*, denoted as $\mathscr{H}(X)$, as the set of all non-empty compact subsets of X. So, for the metric space $(\mathbb{R}^2, d)$ where d is the "usual" distance between vectors, then $\mathscr{H}(\mathbb{R}^2)$ is the set of all non-empty subsets of $\mathbb{R}^2$ which are closed and bounded. One way to think about $\mathscr{H}(\mathbb{R}^2)$ is the set of all non-empty subsets of $\mathbb{R}^2$ which are "small", in some sense. It contains all the finitely-sized subsets of $\mathbb{R}^2$ (except the empty set) which are self-contained enough to contain their own boundaries.

Given a metric space (X, d) with corresponding Hausdorff space $\mathscr{H}(X)$, we define the *Hausdorff* metric, denoted as $d_{\mathscr{H}}$, as the distance between two sets $A, B \in$

$\mathscr{H}(X)$ given by the following expression:

$$d_{\mathscr{H}}(A,B) = \max\left(\sup_{a\in A}\inf_{b\in B} d(a,b),\ \sup_{b\in B}\inf_{a\in A} d(a,b)\right).$$

This is a rather complex expression, but it encodes a simple idea. The Hausdorff metric between two compact sets measures how "isolated" a point from one set can be from the other set.[13] To make the idea more clear, let's imagine ourselves as a point in A. How far away could we be from the set B? Denote this distance as d_a. Now, imagine we're a point in B. How far away could we get from the set A? Denote this distance as d_b. The Hausdorff metric is the maximum of d_a and d_b. It's the furthest possible distance a point from either set could put between it and the other set.

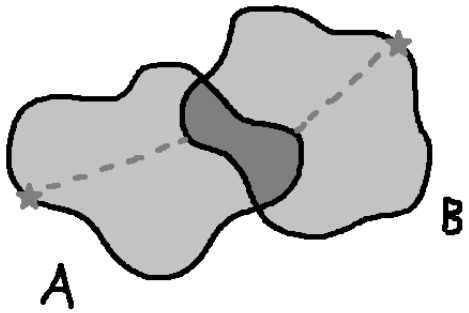


FIGURE 13. Two sets A and B. The starred points represent potential candidates for points in each set that are as isolated as possible from the other set. The Hausdorff distance between A and B would then be the maximum of these two distances.

Using Hausdorff spaces and Hausdorff metrics, we can construct a metric space that's suited for handling the limit of the Koch curve construction process. Let $(\mathbb{R}^2, d)$ be a metric space with d being the "usual" distance between vectors. We've already explained how each step of our Koch curve creation process produces a set of points within this metric space. It's clear to see that these sets of points are bounded (since each step of the construction process can be drawn on a finitely-sized piece of paper). As well, since the Koch curve is composed of only line segments, these sets must contain their own boundaries. Lines have zero thickness, so any points on a line must be zero distance from the points *off* the line. As well, only points on the line can have zero distance from it for the same reason. Thus, every step of the Koch curve construction process produces a compact subset of $\mathbb{R}^2$, meaning each

[13]The reason we restrict our attention to compact sets is so that sets that should be distinct are distinct. If we considered sets that weren't compact, then it would be possible for two different sets to have a distance of zero between them under the Hausdorff metric. Consider, for instance, the intervals $(0,1)$ and $[0,1]$. Under the Hausdorff metric, these sets would be the same.

set is an object in $\mathscr{H}(\mathbb{R}^2)$. So, a fitting metric space to show convergence for the construction process would be $(\mathscr{H}(\mathbb{R}^2), d_{\mathscr{H}})$.

In fact, if we understand what the Hausdorff metric is measuring, then it's rather obvious (intuitively speaking) that the Koch curve construction process really does converge to some final, infinitely-spiky object. At each step of the construction process, the only thing we change about the curve is how spiky it gets—we "push up" more of the flat segments into spikes. Furthermore, the spikes we create get smaller and smaller with each step, protruding less and less from the original line segments from which they're standing. So, if we imagine comparing the curve at one of these steps to the theoretical "final", infinitely-spiky Koch curve object, the maximum distance a point on one curve could put between itself and the other curve gets smaller with each new step in the creation process since the spikes get closer and closer to the "infinitely small" spikes of the final object. Importantly, this maximum distance can never grow bigger as the steps increase. Using the definition of convergence, then, for any $\epsilon > 0$ given, there exists some step N such that, for all steps $n \geq N$, the Koch curve at step n is no more than a distance ϵ from the "final" Koch curve, according to the Hausdorff metric. Thus, the sequence of Koch curve construction steps really does converge to the final, theoretical Koch curve. We can rest easy knowing that the object we've defined is valid.[14]

One last thing worth touching on before we move on to the next section. In section 3.3, we briefly mentioned the idea of using an object's fractal dimension as a way to distinguish fractals from "regular" objects. How exactly do we go about doing that? Well, if an object's fractal dimension is different from the "usual" notion of dimension, then there's a good chance that the object in question is "rougher" than a normal object ought to be (due to the connection between dimension and roughness set forth by Mandelbrot). So, one way to rigorously define a fractal is to declare any object whose fractal dimension is different from its "usual" notion of dimension as a fractal. The "usual" notion of dimension we use is the *Lebesgue covering dimension* (Weisstein [2024a]).

The Lebesgue covering dimension makes use of the notion of an *open cover*. A *cover* is a collection of distinct, non-empty subsets of some set X whose union contains X (Weisstein [2024b]). For instance, given the set $X = \{$☺, 🦇, 👂$\}$, one possible cover could be $\{\{$☺$\}, \{$👂$\}, \{$🦇$\}\}$. Another cover could be $\{\{$🦇, 👂$\}, \{$☺, 👂$\}\}$. An *open cover* is a cover of *open* sets. For any cover of a set X, the *order* of the cover is the

[14]Of course, the explanation given here is a little too loose to be considered a formal proof. However, not much is required to make the argument rigorous. All that's needed is a proper quantification of how close each step in the construction process gets to the final curve, as well as some formal description of this final curve. It's annoying, but not necessarily "difficult".

maximum number of times that any one element of X appears in the elements of the cover. For example, in the first cover we gave for X, the order would be 1, as each element only appears once: one in each element of the cover. In the second cover for X, the order is 2, since 🦻 appears is both elements of the cover. A cover A is said to be a *refinement* of another cover B if, for every element $a \in A$, there exists some $b \in B$ such that $a \subseteq b$. Thus, all the elements of the refinement are subsets of elements of the original cover.

With all this terminology established[15], we say that the *Lebesgue covering dimension* of an object/shape is m if, for every open cover of the object/space, there exists a refinement to that open cover with order at most $m+1$ (Weisstein [2024d]). If no such m exists, then the space is said to have infinite Lebesgue covering dimension.

Once again, this is a fairly abstract definition. However, the Lebesgue covering dimension of a "regular" object/space is the same as the usual dimension. So, lines have a Lebesgue covering dimension of 1, squares have a Lebesgue covering dimension of 2, any space $\mathbb{R}^n$ has a Lebesgue covering dimension of n, and so on. For our purposes, then, the Lebesgue covering dimension is synonymous with the usual notion of dimension: mainly how many coordinates one needs to index all the points in the space.

Using both the Lebesgue covering dimension and the fractal dimension, we can show, under our given "rigorous" definition of fractals, that the Koch curve is indeed a fractal. No matter what step we're at in the Koch curve creation process, the curve is composed of a bunch of straight line segments stitched together end-to-end. Thus, the Koch curve is nothing more than an extremely-crinkled line, meaning we only need one coordinate to index the points on it. This coordinate could tell us, for instance, how far along the line segments we've walked, starting from the leftmost point. So, the Lebesgue covering dimension of the Koch curve is 1. However, from our calculations in section 3.3, we showed that the fractal dimension of the Koch curve is about 1.2619, which is certainly greater than 1. Thus, the Lebesgue covering dimension and the fractal dimension are different, and so the Koch curve is, by our definition, a fractal. The same reasoning can be used to show that the coastlines Richardson was surveying are also fractals.

3.5. **Iterated Function Systems.** Section 3.4 laid down the foundation for metric spaces: spaces that are particularly-well suited for rigorously working with rough,

[15]Unfortunately, there isn't really a more elegant way to introduce these definitions in the context of this project. Topology is a wide, expansive field, and to motivate each concept in a satisfying way would require making this project much longer than it ought to be. The specific ideas introduced here aren't essential to the goal of this project, however, so don't be alarmed if all this new information feels rushed or is difficult to understand. It's provided solely for completeness' sake!

fractal objects. However, as is evident from the complicated definitions and constructions, working with fractals in this way is incredibly cumbersome. Many fractals are constructed using similar creation processes to that of the Koch curve, requiring some simple transformation process to be repeated over and over again *ad nauseam*. Does that mean that, for every fractal we want to create, we have to break out the definition of convergence for metric spaces, find a suitable metric space in which to describe our fractal, and prove that the limit exists rigorously before we can say anything about the fractal itself? If so, that would require a lot of work.

The answer is a welcome *no*. There are a few handy theorems we can make use of to bypass the whole "taking a limit" thing and jump straight into analysing the fractal object itself. It requires describing our fractals as *iterated function systems*, or *IFSs*. Once we translate our fractal construction processes into the language of IFSs, a particular set of theorems will save us from having to take any limits by definition—the theorems will guarantee that our resulting fractal object is well-defined.

IFSs make use of a special type of function called a *contraction*. Given two metric spaces (X, d_X) and (Y, d_Y), a function $f : X \to Y$ is a *contraction* if, for all $a, b \in X$, we have that

$$d_Y(f(a), f(b)) \leq k \cdot d_X(a, b)$$

for some constant $0 \leq k < 1$. The constant k is called the *factor of contraction*. Contractions can be thought of as "squishing" input points together when the function is applied; the points will end up closer to each other after the function is applied.

One very important theorem involving contractions is the Contraction Mapping Theorem, which says that any contraction from a complete metric space to itself has a unique fixed point, and this unique fixed point is approached by repeatedly applying the contraction to any point in the metric space. A *complete* metric space is one where every Cauchy sequence within the metric space is a convergent sequence within the space. A *Cauchy* sequence (x_n) is a sequence such that

$$\forall \epsilon > 0 \;\; \exists N \in \mathbb{N} \;\; \forall n, m \geq N \;\; d(x_n, x_m) < \epsilon,$$

where d is the metric of the given metric space.[16]

[16]Once again, this is another set of definitions that would take a long time to properly motivate. In short, Cauchy sequences are sequences where the terms get closer together as the sequence continues. If we know that Cauchy sequences always converge, then we can show convergence of sequences without having to know where exactly they're converging, providing a potentially easier way to show convergence. This is why spaces where all Cauchy sequences converge have their own special designation; it's a powerful property to have.

Maybe we can see a connection between the Contraction Mapping Theorem and our quest to show the Koch curve construction process produces a well-defined object. If our Koch curves were described inside a complete metric space, and the iterative process we applied to the curve at each step was a contraction, then this theorem would've guaranteed that the "final" Koch curve was indeed well-defined without us having to go through the effort of using the definition of convergence. This theorem is the main reason we'd like to describe our fractals as IFSs. It saves us from having to use the definition of convergence!

So, what exactly is an IFS? We define an IFS on a metric space (X, d) as a finite set of contractions w_1 to w_N, $w_i : X \to X$. The *contractivity constant* of the IFS is the largest factor of contraction of all the IFS' contractions. Under an IFS, completing a step in some fractal's construction process is equivalent to applying all the contractions on some predefined subset of X, then taking the union of the resulting outputs to replace the predefined subset. Symbolically, if we started with some subset $S_0 \subseteq X$, then $S_{n+1} = \bigcup_{i=1}^{N} w_i(S_n)$.

It helps to discuss an example here. Consider the Sierpinski triangle, one of the most well-known fractals. There are numerous ways to construct this fractal, but we'll describe a process similar in spirit to the geometrical construction of the Koch curve explained previously.

Start with an equilateral triangle. At each step of the construction process, remove an inverted triangle from all equilateral triangles in the figure. Repeat *ad nauseam*. The limiting object of this process is a triangle with an infinite number of triangles punched out of it.

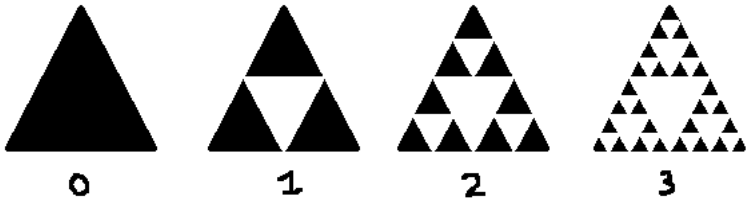


FIGURE 14. The first few steps of the Sierpinski triangle construction process.

Once again, if we'd like to rigorously prove that our construction process results in a well-defined object, we'd have to go through the steps outlined in section 3.4 to show convergence of our object in some appropriate metric space (which would most likely be $(\mathscr{H}(\mathbb{R}^2), d_{\mathscr{H}})$, just like for the Koch curve). However, if we describe our construction process as an IFS, then the tools introduced below can be used to skip all that. By applying a few theorems, we'll know for sure, without having to do any complicated manipulations, that our fractal object is well-behaved in the limit

of the construction process. The question then becomes: how do we translate this construction process into an IFS?

The trick is to encode one of the object's properties within the IFS' mappings. With the "final" Koch curve object, we could split up the curve into four smaller pieces, each one an identical copy of the original curve. For the "final" Sierpinski triangle, we can split up the object into three smaller pieces, each one an identical copy of the original object. To represent this property with an IFS, we can use three mappings, each one mapping the entire object down to a smaller version of itself. If we place these smaller pieces in just the right places, we'll be mimicking the Sierpinski triangle's self similarity, and the resulting IFS should generate the fractal we want.

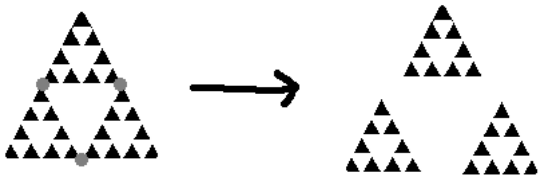


FIGURE 15. Just like with the Koch curve, we can split up the Sierpinski triangle into smaller pieces. Using the "final" Sierpinski triangle, these smaller pieces would be identical to the original triangle.

There are multiple ways we could create a set of three contractions that do this. Here's one such set, which acts on $\mathbb{R}^2$:

$$w_0(\vec{x}) = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \vec{x},$$

$$w_1(\vec{x}) = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \vec{x} + \begin{bmatrix} \frac{1}{2} \\ 0 \end{bmatrix},$$

$$w_2(\vec{x}) = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \vec{x} + \begin{bmatrix} \frac{1}{4} \\ \frac{1}{2} \end{bmatrix}.$$

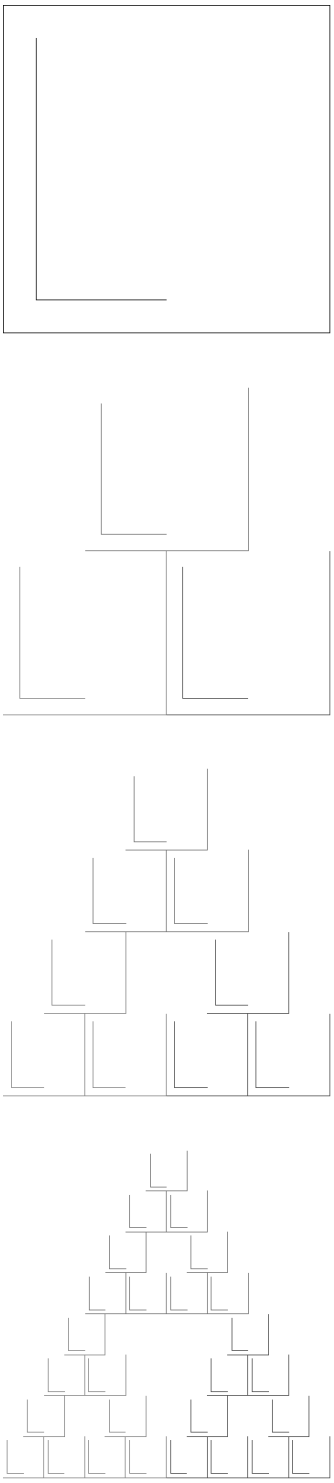
We can visualise what these mappings are doing by taking the unit square

$$\{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq 1, 0 \leq y \leq 1\} \in \mathscr{H}(\mathbb{R}^2)$$

and applying our three mappings to it, then taking the union of the resulting sets as our new starting set. Repeating this process multiple times produces the sets shown in figure 16. We see that as more and more iterations are performed, the resulting set gets closer and closer to the Sierpinski triangle, as we wanted. Despite the process of specifying a fractal this way being more involved, the benefit is that we know 100% that the resulting object is a true, well-defined object. As well,

with the mathematical framework of metric spaces and contractions placed atop our fractal, we can make far stronger observations and conclusions regarding the resulting figure.

First, though, we should explain how exactly the IFS approach lets us skip the "rigorously proving the limit" part. It turns out that, given an IFS $\{w_1, \cdots, w_N\}$



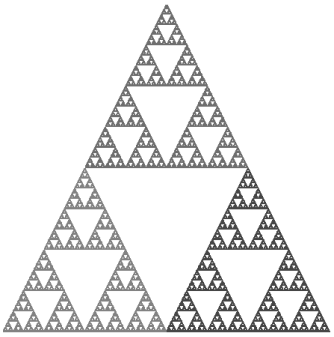


FIGURE 16. The initial square, the first, the second, the third, and the tenth iterations of the Sierpinski triangle IFS. The shading is used to show which of the three mappings is responsible for each part of the shape. Lower left is w_0, lower right is w_1, and top is w_2.

for a metric space (X, d), the mapping $\hat{w} : \mathscr{H}(X) \to \mathscr{H}(X)$ defined by

$$\hat{w}(S) = \bigcup_{i=1}^{N} w_i(S)$$

is a contraction on $\mathscr{H}(X)$ with factor of contraction equal to the IFS' contractivity constant.[17] Notice that $\hat{w}$ is simply the fractal construction process encoded as a single function. Rather than applying each contraction of our IFS individually to some compact subset of some metric space, then taking the union of all the resulting sets, we're applying a single mapping to a single object in the corresponding Hausdorff space of our metric space. Thus, if $\mathscr{H}(X)$ is a complete space, then by the Contraction Mapping Theorem, $\hat{w}$ has a unique fixed point in $\mathscr{H}(X)$ obtained by repeatedly applying $\hat{w}$ to some starting set. This fixed point is special enough to have a name: the *attractor* of the IFS.

Conveniently, if (X, d) is a complete metric space, then $(\mathscr{H}(X), d_{\mathscr{H}})$ is also guaranteed to be a complete metric space.[18]

These results are great news! If we can encode our fractal construction processes as iterations of contractions on some complete starting space, then the Contraction Mapping Theorem guarantees that the limiting object of our process is well-defined!

One question immediately arises from all this: how do we construct an IFS that generates a specific fractal? The Koch curve was created to be connected and infinitely spiky, and so the resulting construction process reflected that desire. However, the process we used was admittedly a bit imprecise (which was why we had to

[17]This result is proven by manipulating some inequalities that hold for the Hausdorff metric. The details get a bit messy, so we'll omit them for this project.

[18]For a more thorough discussion of why this should be true, see Henrikson [1999].

prove the process actually converged to something meaningful!). The language of IFSs doesn't permit such impreciseness: our fractal construction process must be described in terms of contractions from a complete metric space to itself. If we have some fractal shape we'd like to translate into an IFS, how should we go about doing that?

The Collage Theorem is one tool to do this. Say we're given a complete metric space (X, d) and an IFS of contractions w_1 to w_N from the metric space to itself, where the contractivity constant of the IFS is $c < 1$. By the Contraction Mapping Theorem, there exists some fixed point $\bar{x} \in \mathscr{H}(X)$ for this IFS. Then, for any $y \in \mathscr{H}(X)$, we have that

$$d_{\mathscr{H}}(y, \bar{x}) \leq \frac{1}{1-c} d_{\mathscr{H}}(y, \hat{w}(y)),$$

where $\hat{w}$ is given by the union of all the IFS' contractions (H. E. Kunze [2006] 3). What is this theorem saying? In a nutshell, the Collage Theorem says that "[i]f the [fractal] you want to get is called L, then you need to find functions $[w_i]$ such that $[\hat{w}(L) = L]$. Then no matter what initial image you start with, if you iterate $[\hat{w}]$, you'll 'eventually' get L" (Snyder [2006] 11). An added stipulation is that "[i]f $[\hat{w}(S) = S]$ and S contains L, then $S = L$. That is, no larger set is 'fixed' by $[\hat{w}]$" (Snyder [2006] 11). Without this added condition, trivial functions like the identity function will technically satisfy our condition, but they clearly won't generate any sort of interesting fractal under repeated iteration of the map.

We already implicitly used the Collage Theorem when we created an IFS for the Sierpinski triangle. The Sierpinski triangle is composed of three smaller versions of itself, each arranged to form a bigger version of the same self-similar object. The mappings we specified for our IFS above encoded this exact self-similarity. Then, by the Collage Theorem, the fixed point guaranteed by the Contraction Mapping Theorem ended up being the Sierpinski triangle. Using this same principle, if we have some fractal image we'd like to represent as an IFS, we simply need to find a set of mappings such that the fractal is unchanged under those mappings. Then, no matter what we use as a starting set, repeated iteration of the mappings will eventually transform the set into the fractal we want!

Now, it may seem like restricting our attention to IFSs is limiting. What if there are fractals we wish to analyse that can't be represented as an IFS? It's certainly true that not every rough phenomena we encounter can be translated into a set of simple, deterministic contractions. Brownian motion, for instance, relies on a random process almost by definition, so we can't use some predetermined set of

contractions to describe *all* Brownian motion.[19] And, indeed, any sort of fractal that doesn't possess the "perfect" self-similarity of objects like the Koch curve are more difficult to describe using IFSs, as they rely on some amount of randomness in their construction.

However, these "perfect" self-similar fractals are often the starting point for understanding more "random" fractals, as we saw in section 3.3 when we used the Koch curve to understand Richardson's measurement of roughness for coastlines. And, by the Collage Theorem, so long as we recognise the symmetries of a perfectly self-similar fractal, there's guaranteed to be a set of contractions whose corresponding attractor is the fractal object of interest. So, although IFSs can't perfectly describe *every* fractal object out there, they serve as a good starting point for grappling with roughness and understanding the properties that other, more "random" fractals may have.

3.6. **de Rham Curves.** To close off our introduction to metric spaces and IFSs, let's again consider the Koch curve. Although we originally described it as a purely geometrical construction, the Collage Theorem says we can represent it as the attractor of an IFS. In fact, the Koch curve turns out to be a single element of a continuum of fractal curves called *de Rham curves*, originally described by Georges de Rham in 1957 (Vepstas [2006] 1). Much like the Koch curve, all de Rham curves are continuous yet are infinitely spiky, meaning they aren't differentiable on all but a finite number of points (Vepstas [2006] 1).

The idea behind de Rham curves is straightforward. Assume we have two contractions $w_0, w_1 : \mathbb{R}^2 \to \mathbb{R}^2$ where the fixed points x_0 and x_1, respectfully, lie in the basin of attraction of the other map. A *basin of attraction* is simply the set of points around a fixed point such that, under repeated iteration of the contraction, the points will approach the fixed point (Israel [2010]). For our contractions, the basin of attraction will be the entire space $\mathbb{R}^2$ since a contraction, by definition, contracts the entirely of its domain. On top of this, we also assume that each contraction applied to the other map's fixed point will result in the same point. Symbolically,

$$w_0(x_1) = w_1(x_0). \tag{1}$$

Under these assumptions, we then construct a curve between x_0 and x_1 by considering the fixed points of all contractions of the form

$$c_t(x) = w_{t_1} \circ w_{t_2} \circ \cdots (x),$$

[19] In section 4.1, we'll discuss some ways that mathematicians and physicists model Brownian motion.

where t is a real number on the interval $[0, 1]$, and t_n denotes the n-th digit after the decimal place of the binary representation of t. For example, if $t = \frac{1}{4}$, then the binary representation of t would be 0.01000..., and so $t_1 = 0$, $t_2 = 1$, and all t_i with $i > 2$ would be 0. The collection of all the fixed points for each c_t, with t being a real-valued parameter from 0 to 1, forms the de Rham curve corresponding to the maps w_0 and w_1. Georges de Rham proved that, as a function of t, the curve of all fixed points for each c_t forms a continuous curve that is nowhere differentiable, just like the Koch curve (Vepstas [2006] 1).

The condition imposed by equation 1 ensures that differing representations of the same binary number yield the same value under the mapping c_t. For example, if $t = \frac{1}{2}$, then $t = 0.1000...$ and $t = 0.0111...$ are both correct representations in binary. However, due to the condition of equation 1, the resulting maps c_t would behave identically. Without this condition, our curve of fixed points wouldn't be continuous.

Due to how de Rham curves are constructed, we can actually show the self-similarity of the curves in a fairly easy way. For any point $t \in [0, 1]$, we have that

$$p(t) = w_0(p(2t)) \text{ for } t \in \left[0, \frac{1}{2}\right], \quad \text{and}$$
$$p(t) = w_1(p(2t-1)) \text{ for } t \in \left[\frac{1}{2}, 1\right],$$

where $p(t)$ is a point on the resulting de Rham curve parameterized by t (de Rham [1993]). These relations are due to the relationship between a number's binary expansion and its corresponding sequence of contractions w_0 and w_1 that define its place on the curve. Thus, we have a pair of explicit formulas that indicate that the curves we obtain should possess some sort of self-similarity.

And, indeed, figures 17, 18, and 20 show that de Rham curves do indeed possess self-similarity. For completeness, we also list the Koch curve itself in figure 19. In all these examples, we certainly end up with a curve which isn't obviously differentiable, yet the curves are (at least visually) continuous, constituting a single curve.

So, via our newly-understood metric spaces and IFSs systems that formalise fractal objects, we're able to extend Helge von Koch's results on the Koch curve to an infinite number of related curves, showing that non-differentiable curves like it and the Weierstrass function aren't actually the "monsters" mathematicians initially believed them to be, but are commonplace denizens in the world of continuous curves. This mirrors nature: as we noted in section 1, much of the natural world

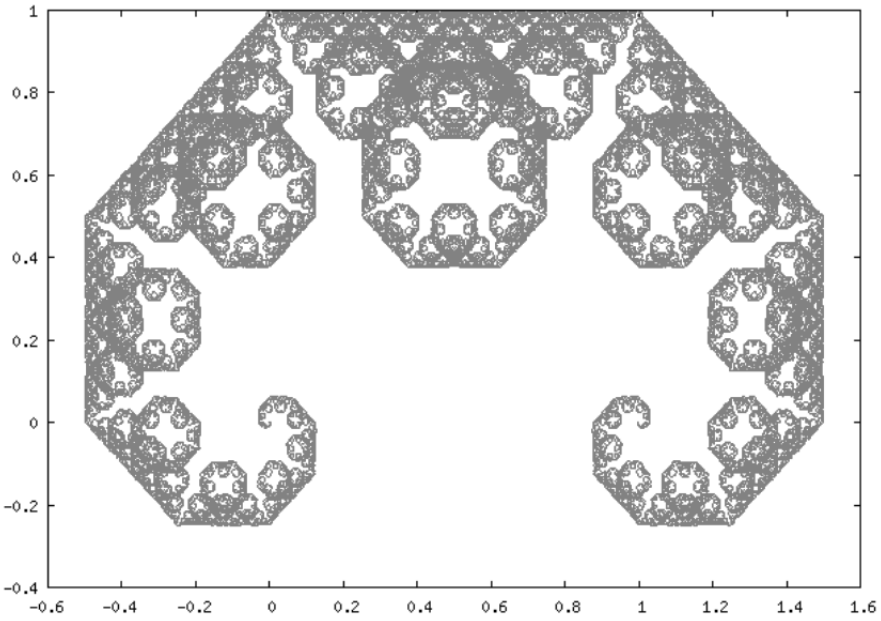


FIGURE 17. The de Rham curve obtained using the mappings $w_0(\vec{x}) = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \vec{x}$ and $w_1(\vec{x}) = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \vec{x} + \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$. This curve is known as the Cesàro curve or Lévy C-curve. Figure obtained from Vepstas [2006].

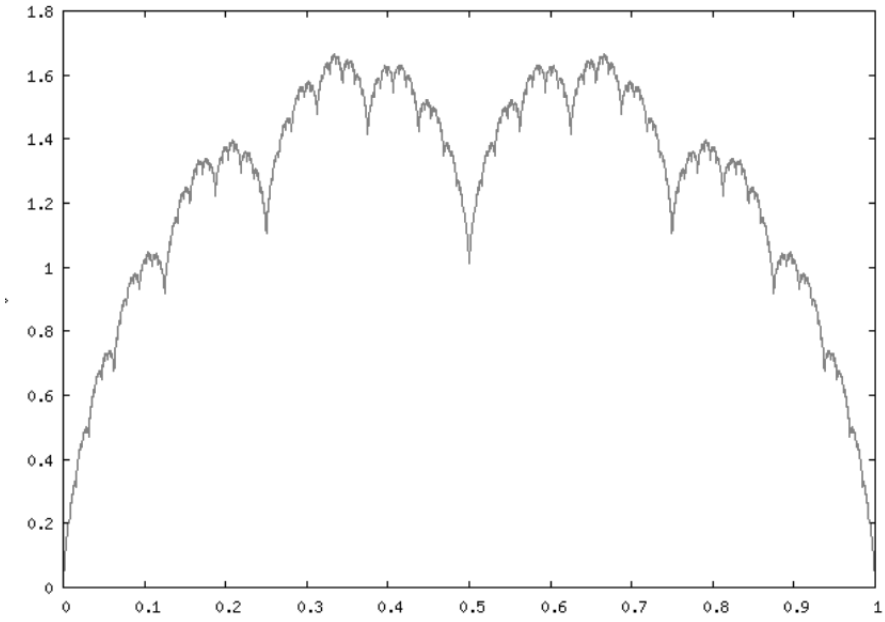


FIGURE 18. The de Rham curve obtained using the mappings $w_0(\vec{x}) = \begin{bmatrix} \frac{1}{2} & 0 \\ 1 & \frac{3}{5} \end{bmatrix} \vec{x}$ and $w_1(\vec{x}) = \begin{bmatrix} \frac{1}{2} & 0 \\ -1 & \frac{3}{5} \end{bmatrix} \vec{x} + \begin{bmatrix} \frac{1}{2} \\ 1 \end{bmatrix}$. This curve is known as the Takagi curve or blancmange curve. Figure obtained from Vepstas [2006].

consists not of the nice, smooth constructions of abstract mathematics, but of rough, jagged objects that require great care both to model and to understand[20].

[20]...and to appreciate!

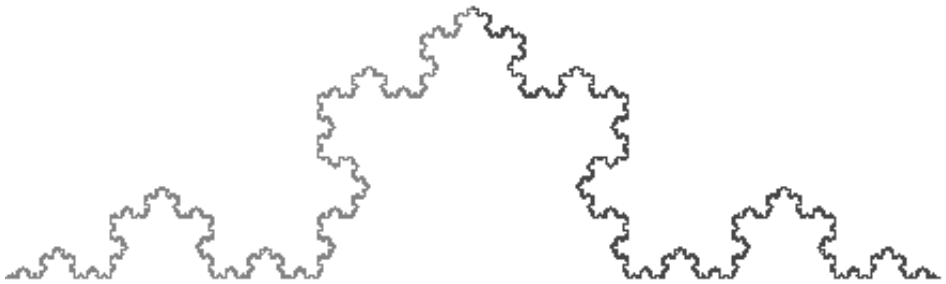


FIGURE 19. The de Rham curve obtained using the mappings $w_0(\vec{x}) = \begin{bmatrix} \frac{1}{2} & \frac{\sqrt{3}}{6} \\ \frac{\sqrt{3}}{6} & -\frac{1}{2} \end{bmatrix}\vec{x}$ and $w_1(\vec{x}) = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{6} \\ -\frac{\sqrt{3}}{6} & -\frac{1}{2} \end{bmatrix}\vec{x} + \begin{bmatrix} \frac{1}{2} \\ \frac{\sqrt{3}}{6} \end{bmatrix}$. This curve is our familiar Koch curve!

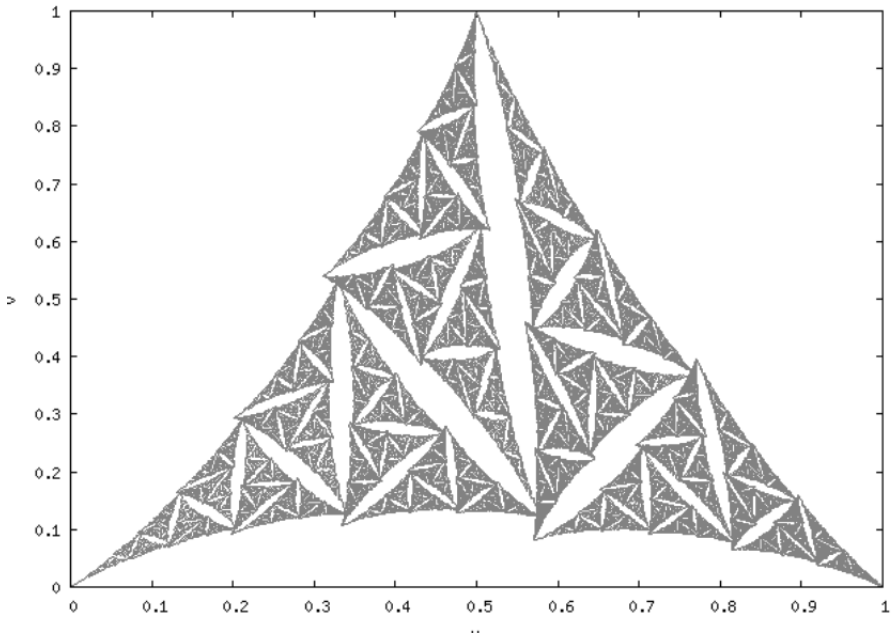


FIGURE 20. An example of an arbitrary de Rham curve. Figure obtained from Vepstas [2006].

4. Conquering Roughness

4.1. **Brownian Motion - Part 2.** Let's return to the topic of Brownian motion. Our problem from before was devising a way to model the motion in a useful way. We couldn't simply use some simple, straight line approximations for Brownian motion since the thing that makes Brownian motion unique is its jaggedness, its *resistance* to being modelled via simple functions. However, once mathematicians (and, in this particular case, physicists) began to understand more about roughness and fractal-like objects, they realised that the roughness of Brownian motion had to be embraced in order to make any sort of sense of it. In fact, in 1904, it was Einstein who popularised the usage of non-smooth functions to model Brownian motion. His models "used functions that were infinitely jagged [like the Weierstrass function]. It

set a longstanding precedent: Physicists have used non-smooth functions as a proxy for Brownian motion ever since" (Kucharski [2014]).

Researchers didn't stop there. Another characteristic of Brownian motion is its seemingly random nature (Mickelin [2017]). In the original context of pollen molecules dancing around in water, Brownian motion is caused by the unpredictable movements of water molecules bumping into the pollen, nudging it in equally-unpredictable directions. If we want to model Brownian motion more faithfully, we should somehow incorporate randomness into our functions and equations. This is exactly what the mathematician Kiyoshi Itō did in the 1940s. To add randomness into the models for Brownian motion, Itō "introduced a method for handling a mathematical function that depends on a non-smooth quantity—like Brownian motion—rather than a more traditional variable, like distance. Using his new methods, he derived 'Itō's Lemma' to calculate how such a function changes over time" (Kucharski [2014]). Itō's work eventually led to the creation of stochastic calculus in the 1970s, which "is used to study all sorts of phenomena, from neurons firing in a brain to diseases spreading through a population. It is also at the heart of financial mathematics, where it helps banks estimate option prices. It can account for the bumpy behavior of a stock price, and hence reveal how the value of an option changes over time" (Kucharski [2014]). Learning to model Brownian motion opened up the doors to modelling all sorts of rough phenomena.

One example of such a rough phenomenon is lightning. Imagine air particles as a bunch of small, hyperactive orbs that are jittering around and colliding with one another. For lightning to strike Earth, there needs to be a conductive bridge between the clouds in the sky and the ground. One such bridge can be created by air particles bumping into each other close enough to transfer electricity (Mickelin [2017]). This setup can be modelled using Brownian motion: move the air particles around the air in a jagged, random way. If they bump into each other, we'll "lock" the particles in place, forming a connection. Over time, more and more air particles will collide, creating longer and longer connections through which lightning can travel. In reality, air particles don't lock together like this, but they're densely-packed enough (and move fast enough) so that connections like ours will be formed at particular instances in time, so our model isn't totally inaccurate. If we tried to model our particles as densely as they are in reality, we'd likely suffer precision errors (and/or the code would be too slow), so this is a good alternative.

Using such a simple model, we may expect simple lightning shapes to emerge: straight lines, maybe a few zigzags, but nothing more. This is not what happens. In fact, we get the same branching, capillary structures of lightning. Importantly,

we notice that the resulting pathways are very jagged and winding (see figure 21). Thus, our simulated lightning ends up exhibiting fractal structure. This is perhaps unsurprising since, to generate the pathways, we used Brownian motion. However, because we end up with something fractal-like, we can use the tools and techniques developed for fractals to analyse and model these pathways. In other words, we don't need to run a simulation of hundreds of particles to simulate lightning. Instead, we can use tools like the ones developed by Itō and Mandelbrot!

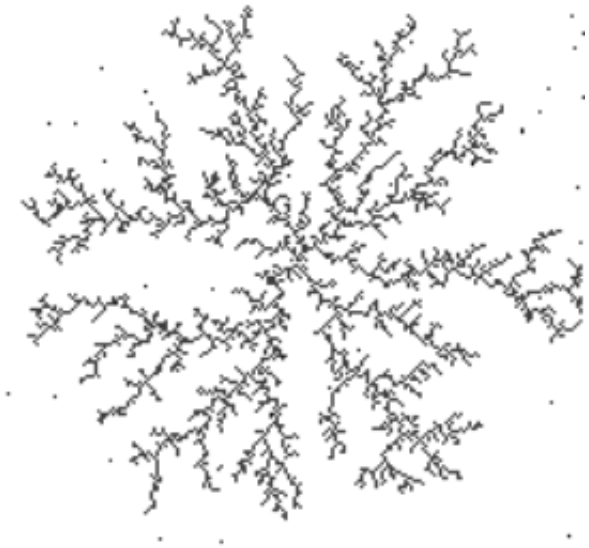


FIGURE 21. A simulation of diffusion limited aggregation provided at Mickelin [2017], where particles can only stick to those that have stuck themselves to the central structure.

The model we described for simulating lightning strikes is known as *diffusion limited aggregation*, and it happens to describe many other real-world phenomena, like how blood vessels wind through the human body, or how bacteria form colonies (Mickelin [2017]). Like coastlines, the fractal dimension of diffusion limited aggregation can be calculated to give some measure of roughness to the resulting patterns. The fractal dimension comes out to be around 1.71 (Mickelin [2017]), making it far spikier than the Koch curve, whose fractal dimension was just over 1.26 (since the higher the number, the more jagged the object is, according to Richardson's observations).

Because diffusion limited aggregation can explain so many real-world phenomena, knowing its fractal dimension immediately tells us how rough and jagged we should expect all these phenomena to be! As well, because Mandelbrot's fractal dimension is intrinsically related to length and scaling, knowing the fractal dimension of an object/design tells us something about how large, in some sense, we expect that

FIGURE 22. Much like tree branches, lightning features many branching pathways and jagged limbs. Original image is "Scenic View of Thunderstorm" by Amol Mande from Pexels.

object/shape to be. For example, comparing a bacteria colony modelled using diffusion limited aggregation to the Koch curve, we expect the pathways of the bacteria to be longer and more winding than the Koch curve, giving us a sense as to the density of the colony.

Diffusion limited aggregation is a useful model, but it isn't the only one of its type. A more complex model often used to model surface growth (like snow piling atop the ground) is called the *Kardar-Parisi-Zhang (KPZ) equation*, and is given by the following stochastic (random process) differential equation:

$$\frac{\partial h}{\mathrm{d}t} = \nu \frac{\partial^2 h}{\mathrm{d}x^2} + \frac{\lambda}{2}\left(\frac{\partial h}{\mathrm{d}x}\right)^2 + \frac{\mathrm{d}B}{\mathrm{d}t},$$

where $B(t)$ is a function used to describe a particle's path under Brownian motion over time, ν and λ are parameters of the system we adjust depending on the context, and $h(x,t)$ is the function describing whatever random, jagged process we wish to model (like lightning strikes) (Mickelin [2017]). An entire project could be made just on this stochastic differential equation alone. Many other interesting models describing random, jagged processes are derived from the KPZ equation, and the KPZ equation itself is still being researched by mathematicians and physicists alike (Mickelin [2017]).

Like the tools we built up in sections 3.4 and 3.5, the KPZ equation is yet another way mathematicians/physicists can make sense of roughness. By studying its

properties, we can deduce more about how fractal-like objects behave and construct themselves, whether in nature or in the toy models we create using IFSs.

4.2. **Weierstrass Function - Part 2.** In section 2.2, we mentioned how it was practically impossible to plot the Weierstrass function at the time it was initially being discussed. However, with the advent of computers and programs like Maple, we can plot the function almost instantly with two lines of code:

```
w := (x, n) -> sum(cos((3^k)*Pi*x)/(2^k), k=1..n);
plot(w(x, 100), x=-0.5..0.5, y=-1.5..1.5)
```

This gives us the following plot:

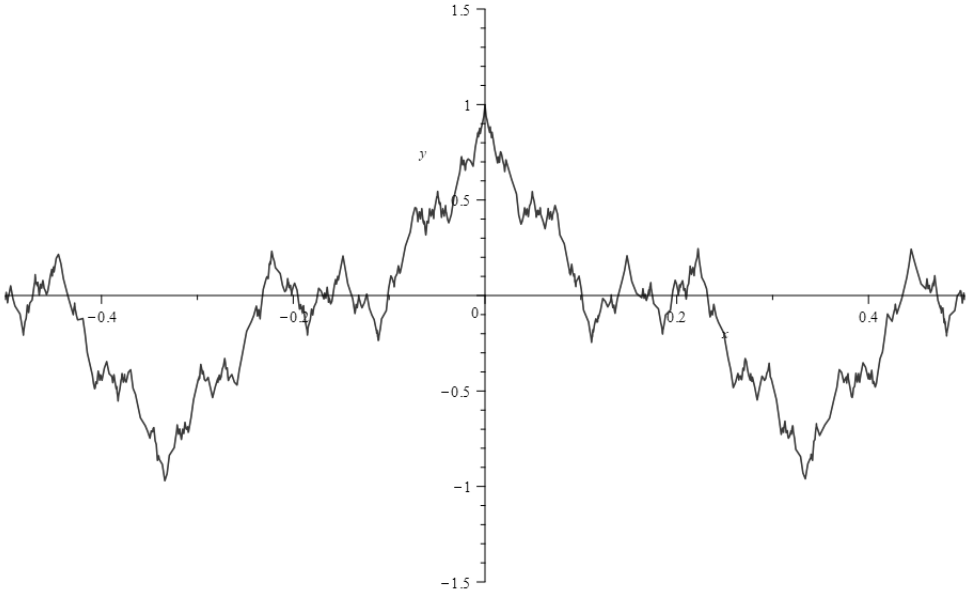


FIGURE 23. A plot of the Weierstrass function from $x = -0.5$ to $x = 0.5$.

Immediately, we see that the Weierstrass function is very much like the rough objects we've been considering throughout this project. It's clear from the plot that the function is indeed continuous (one contingent line), and is not differentiable at most points (due to its spikiness). As well, with a little zooming (see figure 24), we note that the function is self-similar.

Thus, to understand the Weierstrass function, we'd need to approach it as a fractal, something that mathematicians at the time it was being discussed weren't equipped to do.

However, once the roughness of the Weierstrass function was accepted, it found usefulness in applications like Brownian motion[21] where the jagged, almost random

[21]What a coincidence! We were just talking about that!

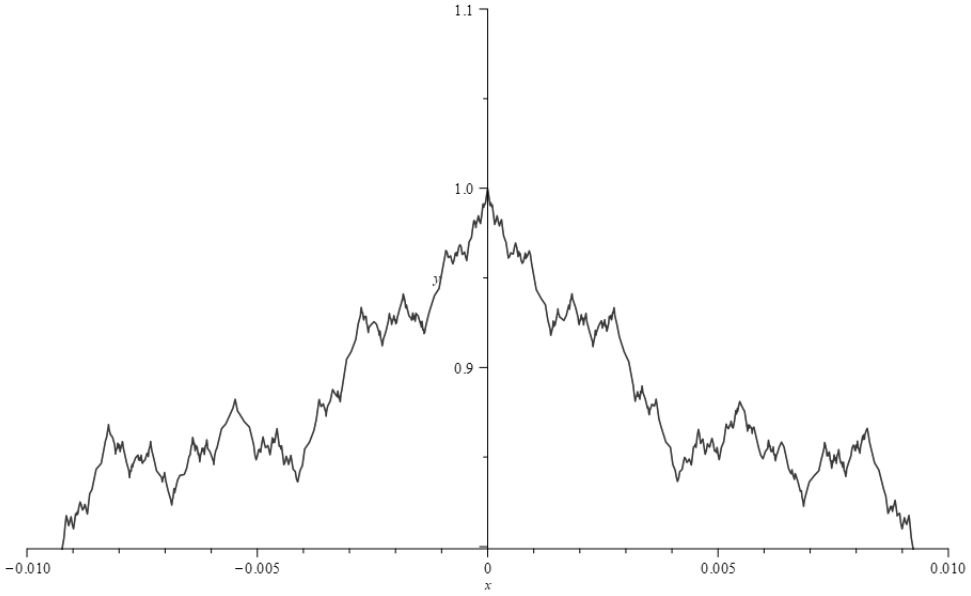


FIGURE 24. A plot of the Weierstrass function from $x = -0.01$ to $x = 0.01$, zoomed into the "tip" of the mountain at $x = 0$.

paths of particles needed functions to model them (Kucharski [2014]). The unpredictability of the Weierstrass function happened to fit that bill nicely.

4.3. **Julia Sets - Part 2.** When computers hit the scene, mathematicians were finally able to properly see what Julia sets looked like. Before that, they could only get a rough idea as to the general shapes of the sets. In particular cases (such as the one in section 2.3), they could draw exact Julia sets, but these cases were rare. So, imagine their shock when, after only being able to visualise overly-simple Julia sets, they were greeted to pictures like the following:

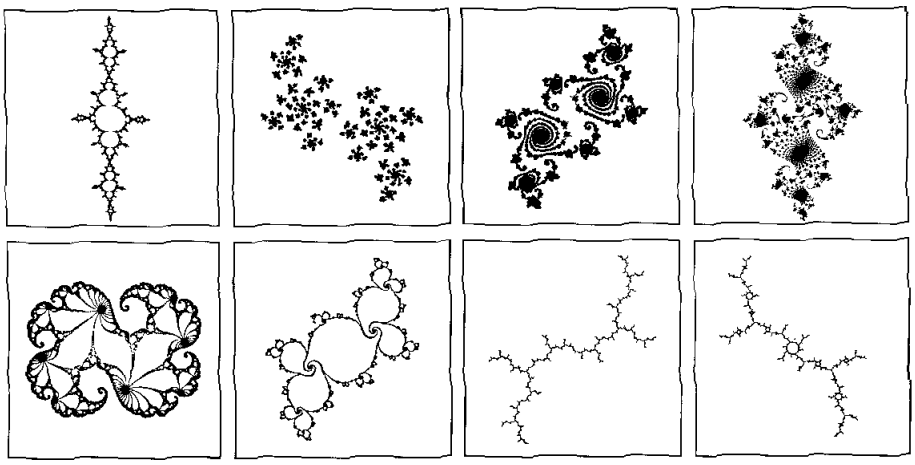


FIGURE 25. Plots of various Julia sets under functions of the form $f(z) = z^2 + c$ for complex z and c. Image taken from Pyke [2003].

The black region of the images in figure 25 represent the Julia sets for each respective system. These are a far cry from the simple circle boundary of figure 4! In fact, the vast majority of Julia sets are complex. With the exception of a small number of values (like $c = 0$ and $c = 2$), all Julia sets corresponding to the quadratic map $f(z) = z^2 + c$ generate wildly-complex shapes, too complex to adequately draw without a computer (Weisstein [2024c]).

The shapes of Julia sets end up being complex enough to warrant their own classification system. For the second Julia set on the top row of figure 25, we notice that it seems to be more "disconnected" than the others—it's a collection of scattered points rather than some spiky, enclosed area. Such disconnected Julia sets are called *Fatou dust* (Weisstein [2024c]). For the more "connected-looking" sets, mathematicians have given some of the more distinctly-shaped sets names. Figures 26 and 27 showcase some of the more well-known types of Julia sets.

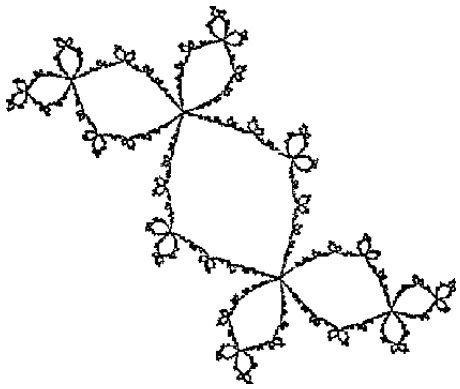


FIGURE 26. A particular Julia set known as the "Douady's Rabbit Fractal", corresponding to the function $f(z) = z^2 - \frac{1}{8} + \frac{3}{4}i$ (Weisstein [2024c], PrimeFan [2013]). The fractal was named after French mathematician Adrien Douady, who worked on dynamical systems (Mravinci [2013]). Image taken from Moulay and Baguelin [2007].

From figure 26, and less obviously from figure 25, we see that Julia sets are both intricately detailed and possess self-similarity. It turns out that the iterative process for creating Julia sets produces fractals! However, for all of the complexity within these sets, they're still the result of an iterated process, just like the attractors of our IFSs. So, we can treat them just as we did for all the previous fractals we considered (such as the Koch curve). In fact, for any particular Julia set, we can create an IFS whose attractor has a boundary equal to a given Julia set.

Say we're given a Julia set whose defining function is given by $f(z) = z^2 + c$ for some $c \in \mathbb{C}$. So, some $a \in \mathbb{C}$ is in $\tilde{J}_f$ if, under repeated iteration of f, the sequence $(f^i(a))$, $i \in \mathbb{N}$, remains bounded. The Julia set J_f is then the boundary of $\tilde{J}_f$, just like before. Earlier, the only way we had to compute J_f was to iterate over

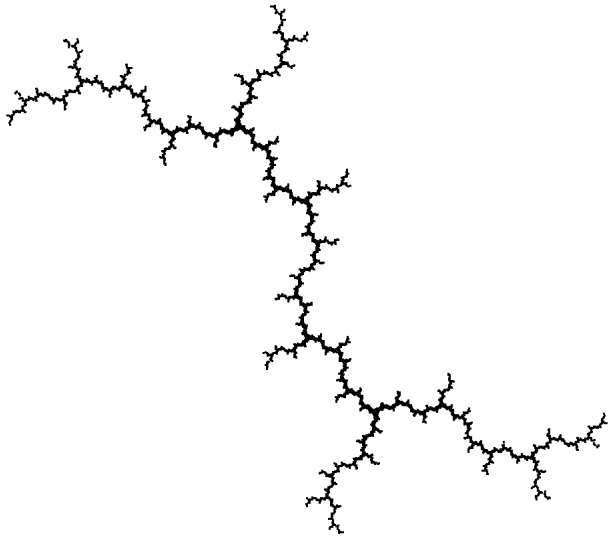


FIGURE 27. A particular type of Julia set known as a "dendrite" fractal. Several functions generate dendrite Julia sets, one such example being $f(z) = z^2 + i$ (Weisstein [2024c]). Dendrite fractals are a common type of fractal that appear in various places (like the diffusion limited aggregation models discussed in section 4.1), not just in Julia sets. Image taken from HandWiki [2023].

all complex numbers, determine the shape of $\tilde{J}_f$, and then take its boundary. Even with the tools of metric spaces we developed in section 3.4, finding the boundary of an arbitrary set is rather difficult, especially if we can't visualise the specific set. As figure 25 shows, Julia sets fall into the category of "difficult to visualise".

However, using IFSs, we have an easier way to find the boundary. Take the two inverse maps of f: $w_+(z) = \sqrt{z-c}$ and $w_-(z) = -\sqrt{z-c}$. These two mappings aren't quite contractions since, for all $z \in \mathbb{C}$ such that $|z - c| < \frac{1}{4}$, both w_+ and w_- will take input points farther away from each other rather than squishing them together. However, for all points outside this region (which is a circle of radius $\frac{1}{4}$ centred at c), both mappings are contractions. Then, the mapping $\hat{w} : \mathscr{H}(\mathbb{C}) \to \mathscr{H}(\mathbb{C})$, $\hat{w}(S) = w_+(S) \cup w_-(S)$ isn't strictly a contraction, and so the Contraction Mapping Theorem doesn't guarantee us an attractor for the IFS.

However, it turns out that $\hat{w}$ is "close enough" to a contraction so that the IFS *does* have an attractor! As an example, take $c = 0.274 - 0.008i$, so that the defining function for the Julia set is $f(z) = z^2 + 0.274 - 0.008i$. Starting with a circle of radius 2 centred at the origin, repeatedly applying the map $\hat{w}$ results in the following sets:

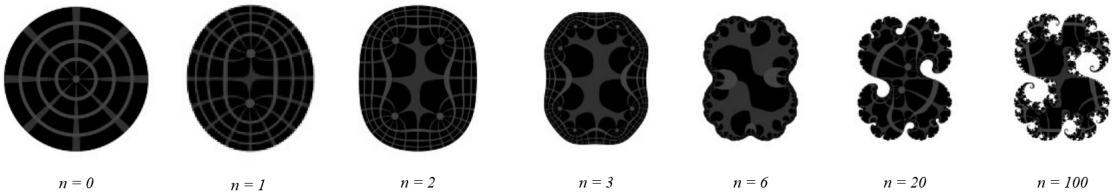


FIGURE 28. Starting with a circle of radius 2 centred at the origin, repeatedly applying the map $\hat{w}$ results in a fractal shape emerging. Images taken from Sims.

The resulting shape represents all those complex numbers who, under repeated iteration of f, remain bounded. Thus, the boundary of the resulting shape gives us the Julia set J_f. Because this process is formalised as an IFS, mathematicians can make use of all the tools of metric spaces and IFS available to them to analyse the Julia set! This allows for far more sophisticated results to be shown for these sets. For example, we can show that J_f is the smallest closed set containing all the repelling periodic points of f (SCHWICK [1997]). As well, the above IFS suggests another way to characterise Julia sets: for all but at most two points in $\mathbb{C}$, J_f is the set of limit points of the full backwards orbit $\bigcup_n f^{-n}(z)$, $z \in \mathbb{C}$. The two exception points would be the fixed points of the mappings w_+ and w_-.

With an IFS description of Julia sets, it's also possible to find an approximate fractal dimension for Julia sets with $c \ll 1$:

$$D \approx 1 + \frac{|c|^2}{4\ln(2)} \quad \text{(Weisstein [2024c])}.$$

We see that, once again, having a formal set of tools and terminology for working with fractals allows us to uncover properties of fractals that would otherwise go unnoticed. As well, the use of computers to visualise said fractals allows us to get a sense for how jagged these objects really are, which is important for building the right sort of intuition. We'd rather not fall into the same traps mathematicians did during the time of Weierstrass! By both working with fractals carefully (i.e. using properly-defined definitions) and using computer-generated plots to make conjectures and guide our reasoning, we can be sure that our conclusions are sound and correct.

Sometimes, even with the rigorous tools we've built up throughout this project, there are some questions which prove terribly difficult to answer. Computer visualisation, then, proves to be an invaluable tool. For example, consider the following question: given the defining function $f(z) = z^2 + c$, for what values of $c \in \mathbb{C}$ is the resulting Julia set J_f connected? Here, a *connected* set is one that cannot be represented as the union of two nonempty, disjoint, open subsets of the original set. This definition

matches with our usual understanding of what it means for something to be connected: if the set can be sectioned off into separate chunks, then it's disconnected. Otherwise, it must be connected.

From figure 25, we can see that some Julia sets are sets of "scattered" points, resembling dust more than a single connected path. Other Julia sets like the Douady rabbits seem more like connected paths, although it isn't entirely clear if they actually are connected. If we try various different values for c, we'll find that values of c close to zero produce Julia sets that look connected, while values of c farther away appear more likely to be disconnected. Is there some underlying shape to the values of c that produce connected sets? In other words, is there some shape that, if we pick c inside of this shape, we end up with a connected Julia set?

In 1981, Robert Brooks and Peter Matelski were the first to attempt to understand this shape (Brooks and Matelski [1981] 68). While a specific description of the shape isn't given in their paper on the subject (which is really more about the properties of a particular mathematical group), they provide a rough image of the shape:

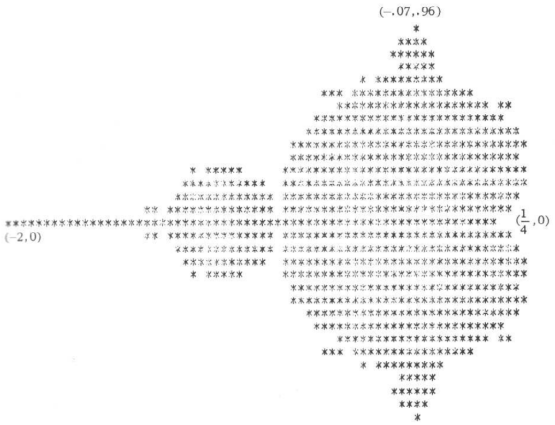


FIGURE 29. An asterisk is plotted on a complex number c if the Julia set defined by the function $f(z) = z^2 + c$ is connected. Image taken from (Brooks and Matelski [1981] 71).

It suffices to say that the shape produced in figure 29 is unexpected. Rather than being some simple geometric shape or a completely chaotic mess, we get something in between: a strange cardioid shape with what appear to be circles or bulbs protruding from it. Though, the image is quite low resolution, so it's difficult to know for sure whether that's actually what we're seeing.

A year or so later, Mandelbrot independently rediscovered this shape and published his findings in the book *The Fractal Geometry of Nature* (Munafo [2023]). This

rediscovery stirred some attention in the mathematical community, and soon, the shape was rendered in a much better resolution:

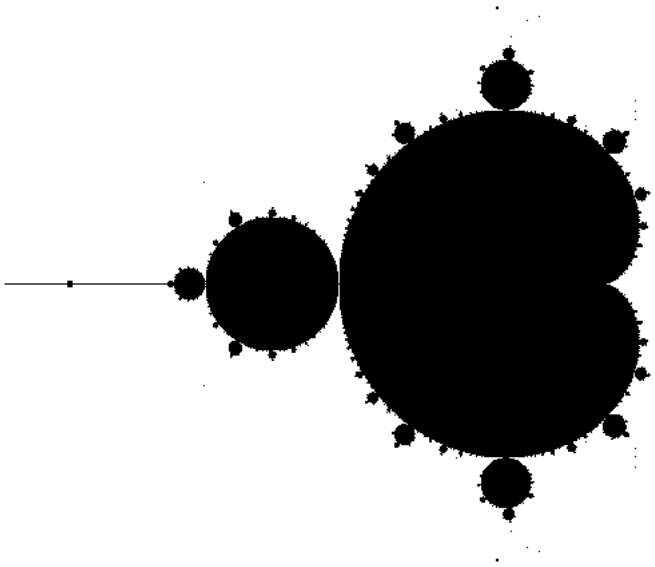


FIGURE 30. The same shape as plotted in figure 29, but at a much higher resolution. Image taken from (Fonseca [2024]).

The resulting shape is even more bizarre at a higher resolution. More than that, the shape appears to be a fractal in its own right! If we look closely, we'll see that figure 30 has these small black dots floating just around the main shape. If we were to zoom in, we'd find that these dots are miniature copies of the larger shape, showing that this shape indeed possesses some amount of self-similarity. Even if we didn't know that, however, the seemingly-infinite number of circles tangent to the main cardioid would suggest that this shape is infinitely rough, just like Richardson's coastlines. No matter how much we zoom into the figure, it'll never be perfectly smooth; there will always be more circles popping out of it. Ditto for the circles themselves.

This shape ended up being so interesting that it was given its own name: the Mandelbrot set. Just as much interest, if not *more* interest, has gone into understanding this fractal (Weisstein [2024e]). We know that the boundary of the set has fractal dimension 2, and we know that the area of the set is approximately 1.506 (Weisstein [2024e]). However, many open questions still exist. For instance, it's unknown whether the Mandelbrot set is *pathwise-connected*, a special type of connected set A where, given any two points $x, y \in A$, there exists a continuous function $f : [0, 1] \to A$ with $f(0) = x$ and $f(1) = y$. This is particularly funny since

the Mandelbrot set was intended as a sort of visual aid for determining whether certain Julia sets were connected or not! One problem about connected sets leads to another.

What *is* known is that, for any given Julia set, it's either pathwise-connected or totally disconnected, meaning for all distinct points x and y in the set, there doesn't exist a continuous function f on $[0, 1]$ such that $f(0) = x$ and $f(1) = y$ (Belk). This is known as the *Fundamental Dichotomy Theorem*, and it was actually proven by Gaston Julia and Pierre Fatou before computer images of Julia sets were created (Belk)! This is yet another example of how having a formal space in which to work with fractals (e.g. metric spaces) can drastically aid in proving results for said fractals. Even without images, Julia and Fatou were able to prove a remarkably-complex property of Julia sets, all without the full knowledge that they were working with fractals!

As is common with topics about fractals, the Mandelbrot set is an extensive enough object to fill its own project. The takeaway from all this is that, given rigorous tools and computer visualizations for fractals, mathematicians are able to go much deeper into their investigations, sometimes uncovering completely new fractal objects in the process! Without these things, we'd likely be stuck toying around with the ideas of Julia sets without making much headway (as was the case in section 2.3).

5. Conclusion

As is probably evident from reading this project, the history of fractals is long, winding, and not at all straightforward—almost like fractals themselves. We attempted to touch on some of the major ideas and developments in the field: understanding Brownian motion, the creation of the Weierstrass function, Julia sets, Richardson and Mandelbrot's exploration into the roughness of coastlines, and finally the formal tools and definitions used in the analysis of fractal objects. Through these topics, a few key ideas emerged. To understand fractals, we must accept the roughness that comes along with them, rather than trying to brush it away through generalisations and oversimplifications; ignoring the details leads to problems, as Weierstrass demonstrated with his function. And, as Weierstrass also made clear, we need properly-formulated tools and definitions in order to make rigorous conclusions about fractal objects (or any sufficiently complex object in math, for that matter). Observations and intuition are essential in guiding mathematics, too, but they can easily obscure us from the finer details.

Many topics were covered in this project, but we only scratched the surface of the vast world of fractals and their history. Even individual fractals like the Sierpinski triangle could fill a project of their own. The thing appears everywhere: as the emergent shape of the Wolfram rule 90 cellular automaton (Weisstein [2024g]), as the resulting shape of colouring all the odd numbers on Pascal's Triangle (Weisstein [2024g]), as the figure produced by a strange plotting process known as the Chaos Game, and so on. Not only do fractals appear everywhere, but *specific* fractals also occur in vastly different contexts!

Fractals also extend far beyond the self-similar patterns we see in nature. They can also be used "to explain how galaxies cluster, how wheat prices change over time and how mammalian brains fold as they grow" (Hoffman [2010]). Perhaps the only thing surprising about fractals at this point is that it took humans so long to give a proper name to them! For a type of object that constantly surrounds us, it's a wonder we weren't working with them sooner!

In any case, the development of fractal geometry and all its related disciplines caused a massive shift in mathematical thinking. No longer could the world be described with simple, smooth abstractions. Mathematicians realised that the roughness of the world couldn't always be ignored in their thinking. Through the developments explored in this project (as well as innumerable others), we not only created the tools necessary for working with these rough objects, but we discovered entirely new objects that emerged from old ones (such as the Mandelbrot Set), a result of our broadening understanding. As with fractals themselves, the study of fractal geometry ended up being an endeavour whose complexity and beauty grows the closer one looks.

Perhaps Mandelbrot himself put it best in his introduction to *The Fractal Geometry of Nature*: "Clouds are not spheres, mountains are not cones, coastlines are not circles, and bark is not smooth, nor does lightning travel in a straight line." To see what's truly in front of us may be one of the hardest tasks mathematics asks of us.

References

Jim Belk. Notes – julia sets and the mandelbrot set. URL `https://e.math.cornell.edu/people/belk/dynamicalsystems/NotesJuliaMandelbrot.pdf`. [Online; accessed 02-November-2024].

Robert Brooks and Peter Matelski. *Riemann Surfaces Related Topics (AM-97): Proceedings of the 1978 Stony Brook Conference. (AM-97)*. Princeton University

Press, 05 1981. ISBN 9780691082677. URL http://www.jstor.org/stable/j.ctt1bd6kzd.

Georges de Rham. *Classics on Fractals*. Addison-Wesley, 1993.

P. Pearle et. al. What brown saw and you can too. Physics at Hamilton, 2010. URL https://physerver.hamilton.edu/Research/Brownian/index.html. [Online; accessed 29-September-2024].

Renato Fonseca. The mandelbrot set, 03 2024. URL https://renatofonseca.net/mandelbrotset. Online; accessed 27-October-2024].

E. R. Vrscay H. E. Kunze, D. La Torre. Random fixed point equations and inverse problems by collage theorem. *UNIMI - Research Papers in Economics, Business, and Statistics*, 06 2006. URL https://links.uwaterloo.ca/papers/waterloo/kulavr06jmaa_ranfp.pdf.

HandWiki. Dendrite (mathematics) — handwiki{,}, 2023. URL https://handwiki.org/wiki/index.php?title=Dendrite_(mathematics)&oldid=2624157. [Online; accessed 02-November-2024].

Jeffrey T. Henrikson. Completeness and total boundedness of the hausdorff metric. 1999. URL https://api.semanticscholar.org/CorpusID:2937913.

Jascha Hoffman. Benoît mandelbrot, novel mathematician, dies at 85. The New York Times, 2010. URL https://www.nytimes.com/2010/10/17/us/17mandelbrot.html. [Online; accessed 29-September-2024].

Robert Israel. Lesson 9: Basins of attraction., 2010. URL https://personal.math.ubc.ca/~israel/m210/lesson9.pdf. Online; accessed 28-October-2024].

T. Jackson. *Mathematics: An Illustrated History of Numbers*. 100 Ponderables. Shelter Harbor Press, 2017. ISBN 9781627950954. URL https://books.google.ca/books?id=_n-ytAEACAAJ.

Adam Kucharski. Math's beautiful monsters. Nautilus, 2014. URL https://nautil.us/maths-beautiful-monsters-234859/. [Online; accessed 02-October-2024].

Benoit Mandelbrot. How long is the coast of britain? statistical self-similarity and fractional dimension. *Science*, 156(3775):636–638, 05 1967.

Benoit Mandelbrot and Nassim Taleb. A focus on the exceptions that prove the rule. Financial Times, 2006. URL https://www.ft.com/content/

5372968a-ba82-11da-980d-0000779e2340. [Wayback Machine; accessed 06-October-2024].

Oscar Mickelin. Random surface growth. Massachusetts Institute of Technology, 2017. URL https://web.mit.edu/8.334/www/grades/projects/projects17/OscarMickelin/. [Online; accessed 29-September-2024].

Emmanuel Moulay and Marc Baguelin. Meta-dynamical adaptive systems and their applications to a fractal algorithm and a biological model. *Physica D: Nonlinear Phenomena*, 207:79–90, 04 2007. doi: 10.1016/j.physd.2005.05.013.

Mravinci. Adrien douady, 2013. URL https://planetmath.org/adriendouady. [Online; accessed 02-November-2024].

Robert P. Munafo. History, 03 2023. URL http://www.mrob.com/pub/muency/history.html. Online; accessed 27-October-2024].

PrimeFan. Douady rabbit, 2013. URL https://planetmath.org/douadyrabbit. [Online; accessed 09-November-2024].

R. Pyke. Some julia sets, 2003. URL http://www.sfu.ca/~rpyke/335/juliasets.html. Online; accessed 02-November-2024].

WILHELM SCHWICK. Repelling periodic points in the julia set. *Bulletin of the London Mathematical Society*, 29(3):314–316, 1997. doi: 10.1112/S0024609396007035.

Karl Sims. Understanding julia and mandelbrot sets. URL http://www.karlsims.com/julia.html. Online; accessed 26-October-2024].

Sandra S. Snyder. Fractals and the collage theorem. *MAT Exam Expository Papers*, 07 2006.

Philip Spencer. Fractals and their history. University of Toronto Mathematics Network, 1999. URL https://www.math.utoronto.ca/mathnet/plain/questionCorner/fracthist.html. [Online; accessed 29-September-2024].

Mario F. Triola. *Mathematics and the Modern World*. Cummings Publishing Company, 2727 Sand Hill Road, Menlo Park, California. 94025., 1973.

Linas Vepstas. A gallery of de rham curves. 08 2006. URL https://www.linas.org/math/de_Rham.pdf.

Eric W. Weisstein. Capacity dimension. Wolfram MathWorld, 2024a. URL https://mathworld.wolfram.com/CapacityDimension.html. [Online; accessed 15-October-2024].

Eric W. Weisstein. Cover. Wolfram MathWorld, 2024b. URL https://mathworld.wolfram.com/Cover.html. [Online; accessed 18-October-2024].

Eric W. Weisstein. Julia set. Wolfram MathWorld, 2024c. URL https://mathworld.wolfram.com/JuliaSet.html. [Online; accessed 04-October-2024].

Eric W. Weisstein. Lebesgue covering dimension. Wolfram MathWorld, 2024d. URL https://mathworld.wolfram.com/LebesgueCoveringDimension.html. [Online; accessed 18-October-2024].

Eric W. Weisstein. Mandelbrot set. Wolfram MathWorld, 2024e. URL https://mathworld.wolfram.com/MandelbrotSet.html. [Online; accessed 27-October-2024].

Eric W. Weisstein. Measure zero. Wolfram MathWorld, 2024f. URL https://mathworld.wolfram.com/MeasureZero.html. [Online; accessed 03-October-2024].

Eric W. Weisstein. Rule 90. Wolfram MathWorld, 2024g. URL https://mathworld.wolfram.com/Rule90.html. [Online; accessed 03-November-2024].

Eric W. Weisstein. Weierstrass function. Wolfram MathWorld, 2024h. URL https://mathworld.wolfram.com/WeierstrassFunction.html. [Online; accessed 29-September-2024].

Stephen Wolfram. *A New Kind of Science*. Wolfram Media Inc., Champaign, IL, 2002.

Edwin Wong. Brownian motion, tragedy, comedy, and history. Doing Melpomene's Work, 2015. URL https://melpomeneswork.com/brownian-motion-tragedy-comedy-history/. [Online; accessed 04-November-2024].