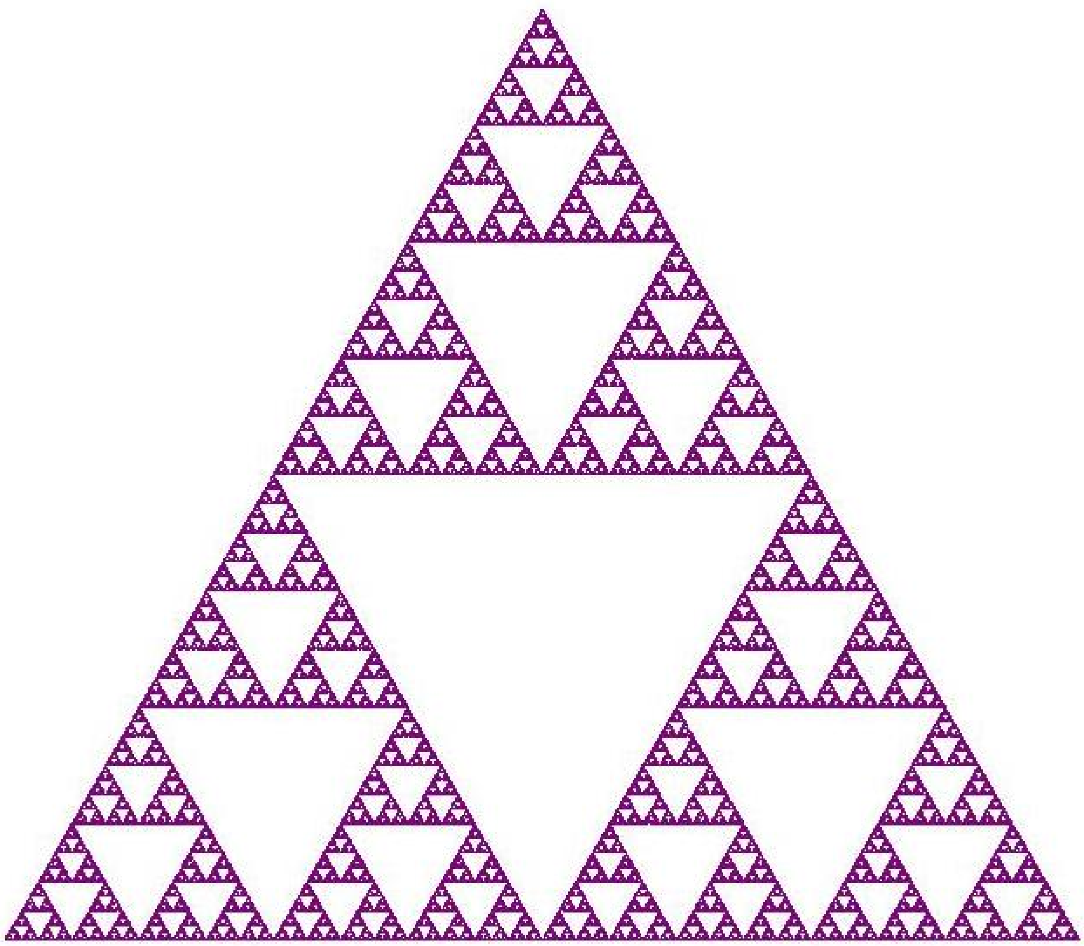
The Sierpinski Triangle: An Aesthetically Pleasing Limit Point



Jim Bell

Table of Contents

Section 0: Many Methods With One Result

Over the years people have developed many methods to create the Sierpinski triangle (or gasket), some of which are more visual and intuitive and some of which are more technical and demanding. We will start with some of the more intuitive strategies and build into a rigorous definition according to equations and their limits.

To begin, we will play a game called Sir Pinski's Game, cleverly named in Manfred Schroeder's *Fractals, Chaos, Power Laws*. This game requires two people. Consider an equilateral triangle. Each person picks a starting point in the triangle and the goal is to keep your point in the triangle longer than your opponent can. The rules are simple: on each turn the player must move his or her point to twice the distance from the nearest corner along the line that the point and the corner make. It must be moved along the line connecting the corner and original point so that the new point is on the same side of the corner as the original point. The player who keeps the dot in the triangle for the most moves is the victor. Some interesting phenomena occur in this game, the result of which is that there are actually infinitely many starting points such that if the player chooses them the player's dot will remain in the triangle for all of eternity. Ironically and unfortunately, if the player chooses the dot at random, the player has a probability of zero of choosing one of these points.

As can be seen from a few examples in Figure 1, any point chosen within the large, centered, upside down triangle will move out of the triangle in one move. This is therefore a poor choice for a player and it is recommended that a player start from one of the three surrounding right-side up triangles. It also turns out that if a centered equilateral upside-down triangle is placed inside of each of the three smaller triangles, any dot inside the smaller upside down triangles will in one move be transported to the large upside down triangle in the center. From there the triangle will be moved outside of the main triangle and the player's game is over.

This pattern can be continued indefinitely so a player can find the number of moves he or she has left simply by finding the number of iterations needed to place the original dot in an upside down triangle. On the other hand, if a player strategically places his or her original point precisely in the shape that remains after an infinite number of these iterations, the player can make an unlimited number of moves and the resulting point will always be within the

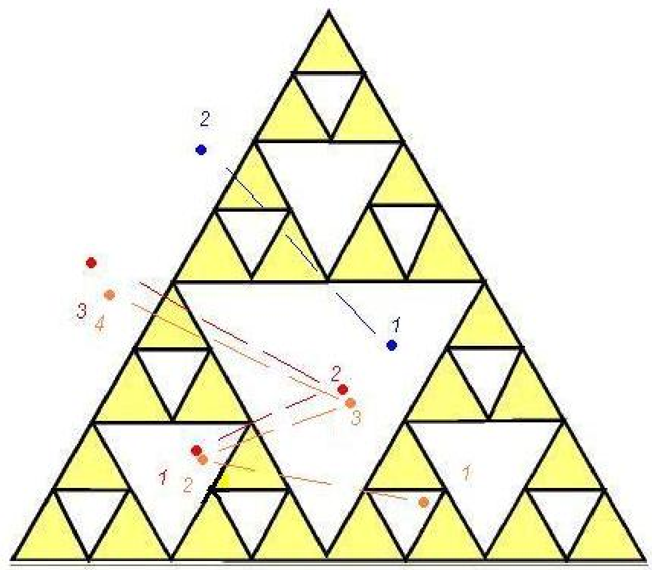


Figure 1: Jumping out

bounds of the original triangle. The shape where the original point should be strategically placed is the Sierpinski triangle and can be seen on the cover of this paper (only a finite number of its lines can be seen of course). It is obvious that with every step of removing triangles there is still a positive area for the remaining shape. Interestingly enough, the resulting shape from an infinite amount of iterations has an area of 0.

Yet another method for creating the Sierpinski triangle is via the Chaos Game, described by Michael F. Barnsley in *Fractals Everywhere*. Create an equilateral triangle with three vertices, V_0, V_1, and V_2. Now put a point anywhere in the plane in which the triangle exists. A method which selects one of the three vertices at random is needed. A die works well, with one and two corresponding to V_0, three and four corresponding to V_1, and five and six corresponding to V_2. Essentially, we have created a 3-sided die. Each time the die is rolled, the point is moved to one half of its original distance to the vertex corresponding to the number on the die. The new point must remain on the line connecting the vertex and the original point. Suprisingly enough, if this process is repeated enough times and the first few (few meaning up to the hundreds if necessary) points created from this iterated process are removed the Sierpinski Triangle is the result! A few steps of the process can be seen in Figure 2. The triangle at the left of the figure has the least amount of iterated points and each triangle to its right has more points created from the Chaos Game.

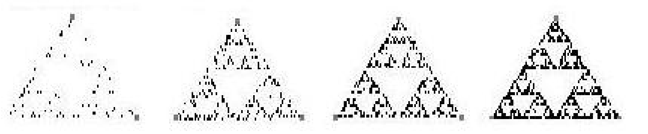


Figure 2: Chaos Game Iterations

Interestingly enough, not only does it not matter where the original point is, it does not even matter what order the vertices are in as long as they are chosen randomly and with equal probability. This method of creating the Sierpinski triangle is more similar to the method using equations and limits which is the focus of this paper. The sequence of points created in this method is called the orbit of x_0 with x_0 being the first point in the sequence. Note that no matter what the starting point, x_0, is, if the "function" we are using is iterated enough times we find ourselves with the same result.

Section 1: The Hausdorff Metric Space

So what is the Sierpinski triangle? How can we define it? In how many dimensions does it exist? These are all questions which should be answered but to do so we must first delve into the world of fractals and iterated function systems. A very general and loose definition of a fractal is an image that contains copies of itself at infinitely many different scales. This means it is "self-similar" which is a term we will visit later. The Sierpinski triangle is a fractal; if any of the smaller right-side-up triangles within the Sierpinski triangle were to be blown up to the size of the Sierpinski triangle itself, the same image would be viewed. It should be noted now that in many of the examples and much of the discussion to follow, unless otherwise noted we will be working in $(\mathbb{R}^2, d)$.

To more rigorously define fractals it is necessary that we define something called the Hausdorff metric denoted $\mathbb{H}(X)$ for some metric space X.

Definition 1 *If (X, d) is a complete metric space then, excluding the null set, $\mathbb{H}(X)$ is the space whose points are the compact subsets of X.*

This means, for example, that if we are working in $\mathbb{H}(\mathbb{R})$, then [0, 1] and [-1, 5] are each different points in $\mathbb{H}(\mathbb{R})$. We must now define distance from a point to a set so that we can later properly define distance in the Hausdorff space.

Definition 2 *If (X, d) is a complete metric space, $x \in X$, and $B \in \mathbb{H}(X)$ then the* ***distance*** *d* ***from*** *x* ***to*** *B is $d(x, B) = \min\{d(x, b) : b \in B\}$.*

It is not obvious that such a point $b \in B$ would necessarily exist. The proof that follows is similar to the one found in Barnsley's *Fractals Everywhere*.

Proof:
Let (X, d) be a complete metric space, $x \in X$, and $B \in \mathbb{H}(X)$. Remember that $\forall\ y \in B,\ d(x, y) \geq 0$. This means that $d(x, y)$ is bounded below. Let $D = \inf\{d(x, y) : y \in B\}$. From the definition of infimum, we know that for all $n \in \mathbb{N}$, $\exists\ p \in B$ such that $d(x, p) - D \leq \frac{1}{n}$. We can now create an infinite sequence of points, $(b_n)_{n\geq 1}$ in B such that $d(x, b_n) - D \leq \frac{1}{n}$. Because B is compact we know (b_n) has a subsequence $(b_{n_i})_{i\geq 1}$ that converges to a point in B. We will call it b. Because $(\frac{1}{n_i})$ converges to 0 and $0 \leq d(x, b_{n_i}) - D \leq \frac{1}{n_i}$, we know $(d(x, b_{n_i}) - D)$ converges to 0 so $\lim_{i\to\infty}(d(x, b_{n_i})) = D$ and $d(x, \lim_{i\to\infty}(b_{n_i})) = D$. Because $\lim_{i\to\infty}(b_{n_i}) = b$ it must be true that $d(x, b) = D$. Therefore, there exists $b \in B$ such that $d(x, b) = \min\{d(x, y) : y \in B\}$. ∎

Now we must define distance in the Hausdorff metric.

Definition 3 *If (X, d) is a complete metric space and $A, B \in \mathbb{H}(X)$ then the* ***distance*** *d* ***from*** *A* ***to*** *B is defined as $d(A, B) = \max\{d(a, B) : a \in A\}$.*

The proof that such an a exists is similar to the proof above.

Proof:
Let (X, d) be a complete metric space and let $A, B \in \mathbb{H}(X)$. Because A is compact we know A is bounded. Then there exists $D = \sup\{d(a, B) : a \in A\}$. By the definition of supremum we can create an infinite sequence $(a_n)_{n\geq 1}$ such that $a_n \in A$ and $D - d(a_n, B) \leq \frac{1}{n}\ \forall\ n \in \mathbb{N}$. Because A is compact, (a_n) has

a subsequence $(a_{n_i})_{i\geq 1}$ such that the limit point of (a_{n_i}) is in A. We will call this limit point a. Because $(\frac{1}{n_i})$ converges to 0 and $0 \leq D - d(a_{n_i}, B) \leq \frac{1}{n_i}$, we know $(D - d(a_{n_i}, B))$ converges to 0. Then $(d(a_{n_i}, B))$ converges to D so for the same reasons as in the previous proof, $d(a, B) = D$. Then there exists $a \in A$ such that $d(a, B) = \max\{d(x, B) : x \in A\}$. ▮

Intuitively, the Hausdorff distance h is the same as putting an "epsilon net" around one set so you can capture the other. This means if there are two sets, A and B, and you are finding the distance from A to B then you must find the point b in B closest to A. Then, as shown in Figure 3, find the smallest radius possible that, when one end is fixed at b, will inscribe A. This

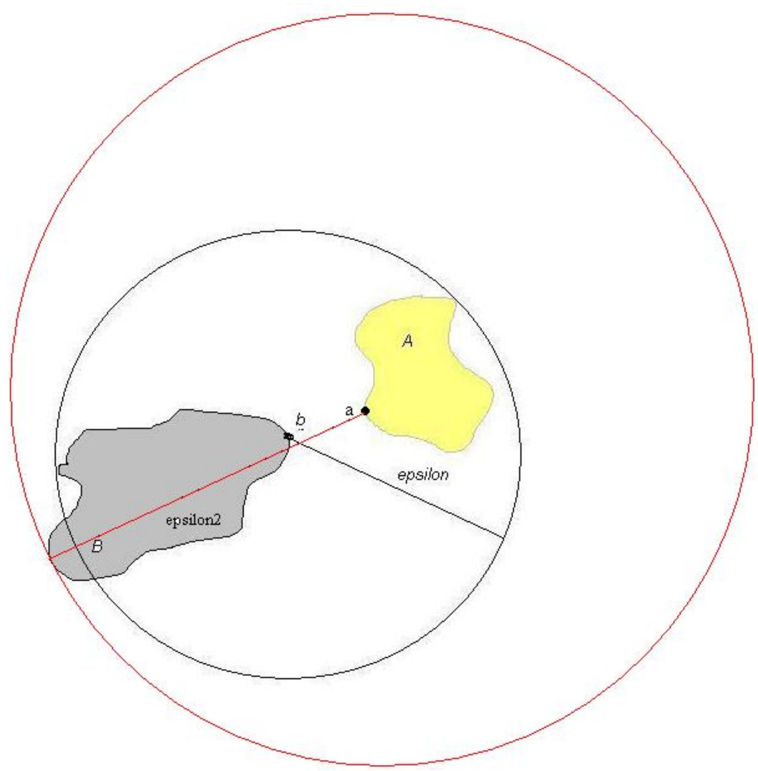


Figure 3: Epsilon Net from A to B where $d(A, B)$ = Epsilon and Epsilon2 Net from B to A where $d(B, A)$ = Epsilon2

radius is the distance from A to B and is sometimes called an epsilon net.

Note that this distance is not commutative; $d(A,B)$ is not necessarily the same as $d(B,A)$. We thus come to the definition of the Hausdorff distance:

Definition 4 *If (X,d) is a complete metric space and $A,B \in \mathbb{H}(X)$ then the **Hausdorff distance,** h, is defined as follows; $h(A,B) = \max\{d(A,B), d(B,A)\}$.*

This just means you make an epsilon net from A to B and from B to A and the largest of the two is $h(A,B)$. We must now show that h is a metric for $\mathbb{H}(X)$.

Theorem 5 *If X is a complete metric space then h is a metric for $\mathbb{H}(X)$.*

Proof:
Assume X is a complete metric space. We must show that h is a metric for $\mathbb{H}(X)$. To do this we have to prove four statements are true.

This first is that for all $A,B \in \mathbb{H}(X), h(A,B) \geq 0$. We know the following three statements are true:

$$h(A,B) = \max\{d(A,B),\ d(B,A)\}$$

$$d(A,B) = \max\{d(a,B) : a \in A\}$$
$$d(B,A) = \max\{d(b,A) : b \in B\}.$$

Delving once more into a more basic definition, $d(a,B) = \min\{d(a,b) : b \in B\}$ and $d(b,A) = \min\{d(b,a) : a \in A\}$. We know because d is a metric on X that $\min\{d(a,b) : b \in B\} \geq 0$ and that $\min\{d(b,a) : a \in A\} \geq 0$. This means that $d(a,B) \geq 0$ and that $d(b,A) \geq 0$ which implies that $\max\{d(a,B) : a \in A\} \geq 0$ and that $\max\{d(b,A) : b \in B\} \geq 0$. Finally, we know that $d(A,B) \geq 0$ and that $d(B,A) \geq 0$ so we find that $h(A,B) = \max\{d(B,A), d(A,B)\} \geq 0$.

The second item on our list to prove is that if $A,B \in \mathbb{H}(X)$, then $h(A,B) = 0$ iff $A = B$. To begin, assume $A = B$. Then

$$d(A,B) = \max\{d(a,B) : a \in A\}.$$

Because $A = B$, if $a \in A$ then $a \in B$ as well. Then because

$$d(a,B) = \min\{d(a,b) : b \in B\}$$

and for all $a \in A$, a is in B as well, $d(a, B) = 0$. This implies that $d(A, B) = 0$. The same argument can be made for the fact that $d(B, A) = 0$. Then it must be true that $\max\{d(A, B), d(B, A)\} = 0$ so $h(A, B) = 0$. Now assume $h(A, B) = 0$. Then it must be true that $\max\{d(A, B), d(B, A)\} = 0$. Consider the fact that $d(A, B) = 0$. Then $\max\{d(a, B) : a \in A\} = 0$. This means if we choose some point $a \in A$ that $\min\{d(a, B) : b \in B\} = 0$. Thus, for any $a \in A$, there is some $b \in B$ such that $a = b$. Therefore, for all $a \in A, a \in B$ as well. The same argument can be made for the fact that for all $b \in B$, $b \in A$ as well. Therefore, $A = B$.

Thirdly, we must show that if $A, B \in \mathbb{H}(X)$ then $h(A, B) = h(B, A)$. This is relatively simple. Because both the statements

$$h(A, B) = \max\{d(A, B), d(B, A)\}$$

$$h(B, A) = \max\{d(B, A), d(A, B)\}$$

are true, it must be true that either $h(A, B) = d(A, B)$ and $h(B, A) = d(A, B)$ or $h(A, B) = d(B, A)$ and $h(B, A) = d(B, A)$. Then it must be true that $h(A, B) = h(B, A)$.

Finally we must show that if $A, B, C \in \mathbb{H}(X)$ then

$$h(A, C) \leq h(A, B) + h(B, C).$$

Assume (X, d) is a metric space. Let $A, B, C \in \mathbb{H}(X)$. We know

$$h(B, C) = \max\{d(B, C), d(C, B)\}.$$

Assume

$$h(B, C) = d(B, C) = \max\{d(b, C) : b \in B\} = d(b_1, C)$$

for some $b_1 \in B$. We know

$$d(b_1, C) = \min\{d(b_1, c) : c \in C\} = d(b_1, c_1)$$

for some $c_1 \in C$. Also, $h(A, B) = \max\{d(A, B), d(B, A)\}$. Assume

$$h(A, B) = d(A, B) = \max\{d(a, B) : a \in A\} = d(a_1, B)$$

for some $a_1 \in A$. We know

$$d(a_1, B) = \min\{d(a_1, b) : b \in B\} = d(a_1, b_2)$$

for some $b_2 \in B$. Finally, we know $h(A, C) = \max\{d(A, C), d(C, A)\}$. Assume

$$h(A, C) = d(A, C) = \max\{d(a, C) : a \in A\} = d(a_2, C)$$

for some $a_2 \in A$. We know

$$d(a_2, C) = \min\{d(a_2, c) : c \in C\} = d(a_2, c_2)$$

for some $c_2 \in C$.

From the above we know that there exists $b_3 \in B$ such that $d(a_2, b_3) \leq d(a_1, b_2)$ because $d(a_1, b_2) = \max\{d(a, B) : a \in A\}$. Also, there exists $c_3 \in C$ such that $d(b_3, c_3) \leq d(b_1, c_1)$ because $d(b_1, c_1) = \max\{d(b, C) : b \in B\}$. Finally, we know that $d(a_2, c_2) \leq d(a_2, c_3)$ because $d(a_2, c_2) = \min\{d(a_2, c) : c \in C\}$.

From all this we can deduce that

$$\begin{aligned} h(A, C) &= d(a_2, c_2) \\ &\leq d(a_2, c_3) \\ &\leq d(a_2, b_3) + d(b_3, c_3) \\ &\leq d(a_1, b_2) + d(b_3, c_3) \\ &\leq d(a_1, b_2) + d(b_1, c_1) \\ &= h(A, B) + h(B, C). \end{aligned}$$

Then it is true that $h(A, C) \leq h(A, B) + h(B, C)$. The cases when $h(A, B) = d(B, A)$, $h(B, C) = d(C, B)$, and when $h(A, C) = d(C, A)$ are all similar. From this and the previous three results we know that h is a metric on $\mathbb{H}(X)$. ∎

Now that we know that h is a metric on $\mathbb{H}(X)$ we know that $(\mathbb{H}(X), h)$ is a metric space. Fortunately for us, this is a metric space where it will be possible for us to define distances between fractals and convergence of shapes in $\mathbb{R}^2$. Hence, the Hausdorff metric space, $(\mathbb{H}(X), h)$, is a metric space where we can truly define and understand fractals. Many fractals, such as the Sierpinski triangle, are the limit points of iterated functions. It does not make sense, however, to speak of an entire shape being a limit point in a metric space such as $(\mathbb{R}^2, d)$. How would we decide whether an iterated point was within a certain distance of the limit point? The Hausdorff metric space is a space where we can treat entire compact sets as points which simplifies definitions and equations immensely.

Section 2: Fractal Dimensions

According to Robert L. Devany in his book *A First Course in Chaotic Dynamical Systems*, "A fractal is a subset of $\mathbb{R}^n$ which is self-similar and whose fractal dimension exceeds its topological dimension" (178). What in the world does that mean? To understand this definition we must learn about self similarity and magnification. It must be noted at this point that this section is a bit of a tangent from the ultimate goal of this paper. It seemed necessary, however, to at least give fractals a more rigorous definition and touch on their complex dimensions. This is a very large field so what follows is only a scratch on the surface.

The Sierpinski triangle is self-similar. One can look at the figure on the title page of this paper and see that any "right side up" equilateral triangle, if blown up to the right size, will yield the original triangle. To be more explicit, one of the ways to create the Sierpinski triangle is to take the midpoint of each segment of an equilateral triangle, connect them to make a triangle, and remove that triangle from the set. This leaves three identical triangles, each with sides that are half the length of the sides of the original triangle. For the rest of this paper I will use the fact that the sides become half as long to mean that the triangle becomes half as big, in spite of the fact that the newly created triangle actually has only a quarter of the original triangle's area. This process can be repeated with the three new triangles and one would have nine new triangles, each one fourth the size of the original and one half the size of the second generation of triangles. If this process is repeated an infinite number of times, the Sierpinski triangle is the result.

This has certain interesting implications. One can therefore take the Sierpinski triangle and find 3^n identically-sized triangles within it and zoom any one of them up to 2^n times its original size to yield the original Sierpinski triangle! From this we can define the term affine self-similar.

Definition 6 *A set is considered* ***affine self-similar*** *if it can be divided into some integer m matching subsets. The matching subsets must be able to be magnified by some positive constant n to yield the original set.*

To understand Devany's definition of a fractal, we must still define topological dimension.

Definition 7 *A set has a* ***topological dimension of 0*** *if all its points have arbitrarily small boundary ovals such that the boundary ovals do not intersect the set.*

In this situation, "boundary ovals" just means small circles or ellipses circumscribing points. For instance, in Figure 4, the set depicted by the points has a topological dimension of zero because around each point can be placed an ellipse that does not intersect any of the other points.

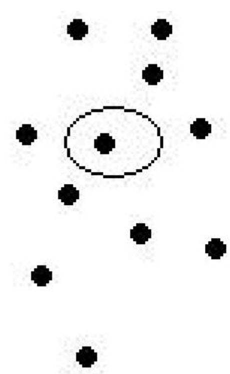


Figure 4: 0 Topological Dimensions

Now we can define topological dimension in general.

Definition 8 *A set has* ***topological dimension*** *m if at each point an arbitrarily small boundary oval can be created which intersects the original set in a topological dimension of $m - 1$ and m must be the smallest nonnegative integer for which this is true.*

This basically means that if the boundary ovals around the points in some set K intersect K so that the intersection has a topological dimension of 0, the topological dimension of K is 1 because the intersection had a topological dimension of 0 = 1 - 1. This is illustrated in the top figure in Figure 5. For another object, say a filled in circle, any boundary oval around any point in that circle will yield an intersection with a topological dimension greater than one. If we take a boundary oval around a point in the intersection we get a new set with a topological dimension of 0. This means that the intersection of the circle with the original boundary oval had a topological dimension of 1 which implies the filled in circle had a topological dimension of 2. This is illustrated in the bottom figure in Figure 5.

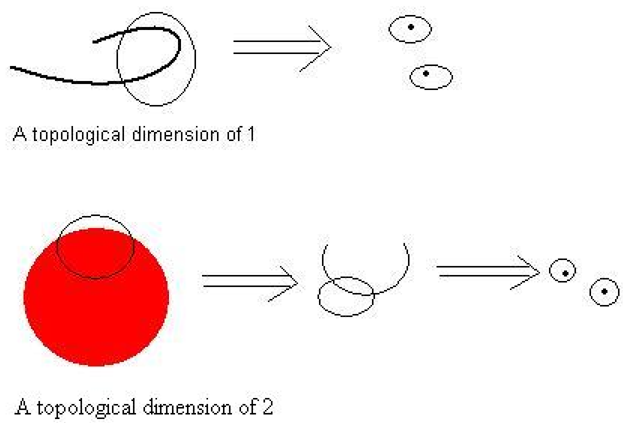


Figure 5: Topological Dimensions

This basically indicates that finding the topological dimension of a set can be an iterative process. The Sierpinski triangle therefore has a topological dimension of 1 because no matter where a boundary oval is created the boundary oval will intersect the Sierpinski triangle in non-connected points. This intuitively makes sense because the Sierpinski triangle is composed entirely of straight lines while all boundary ovals are some form of an ellipse. Therefore, no two consecutive points of the ellipse will intersect any lines of the Sierpinski triangle because of the curvature of the ellipse so the intersection of the Sierpinski triangle with a boundary oval will yield non-connected points. Keep in mind that this is not a rigorous proof that the Sierpinski triangle has a topological dimension of 1; it hopefully is just a helpful illustration of why it has a topological dimension of 1.

Finally, we must define fractal dimension.

Definition 9 *If a set K can be broken into some integer m identical parts and those parts can each be magnified by some factor R to yield K then K has a **fractal dimension** D of $D = \frac{\log(m)}{\log(R)}$.*

Let us apply this new knowledge to the Sierpinski triangle. We know that for any $n \in \mathbb{N}$, the Sierpinski triangle can be broken into 3^n identical pieces, each of which may be magnified by a factor of 2^n to yield the Sierpinski triangle. This means it has a fractal dimension of $D = \frac{\log(3^n)}{\log(2^n)} = \frac{\log(3)}{\log(2)} \approx 1.584962501$. We may come to the conclusion with our result that the Sierpinski triangle is, according to Devany, a fractal because it is a subset of $\mathbb{R}^2$, it is affine self-similar, and it has a fractal dimension which is greater than its topological dimension.

So what does a fractal dimension of 1.584962501 mean when normally we deal with dimensions that are integers? We know a line is one dimensional. Intuitively if we were placed on the line, we could either move forward or backward but that is it. We would only have choices in one dimension. A filled in circle is two dimensional with the same idea. We could move up, down, left, right, or some combination which involves two dimensions. With the Sierpinski Triangle, however, at each point there are more choices than just back and forth, however, not as many choices as if we were in an infinite plane. This is why its dimension is stuck between 1 and 2.

Section 3: The Contraction Mapping Theorem

The Sierpinski triangle is actually a limit point of an iterated map of a function. Odd as it is to think of it this way, it is true. How can an entire shape be considered a point? Remember that we are working in the Hausdorff metric space so compact sets are actually considered points. Not only is it a limit point of the iterated map of this function, but it is the limit point no matter what shape you start off with or where in $\mathbb{R}^2$ this shape is located. How can we show the Sierpinski triangle is the limit point of any initial point's orbit under this special function? We can do this by the contraction mapping theorem.

Definition 10 *Let X be a metric space and let $0 < k < 1$. Then if $f : X \to X$ is a function such that $d(f(a), f(b)) \leq k(d(a, b))$ for all $a,\ b \in X$ then f is a k-**contraction** of X.*

Theorem 11 (The Contraction Mapping Theorem) *If X is a complete metric space and $f : X \to X$ is a k-contraction of X then for all $x \in X$, the orbit of x under f converges to some unique fixed point x_0.*

This is a very powerful theorem to have. To use it, however, we must prove it. The following proof was outlined in Carol Schumacher's *Closer and Closer*.
Proof:
Let X be a complete metric space and $f : X \to X$ be a k-contraction of X. The first step in proving this theorem is to show that f cannot have more than one fixed point. This is relatively simple and we shall prove this by contradiction. Assume f has more than one fixed point. Let two of them be x and y. Then $f(x) = x$ and $f(y) = y$. Then we know $d(f(x), f(y)) = d(x, y)$ so f is not a contraction. This is a contradiction so f can have at most one fixed point.

The second step is to show that for all $n \in \mathbb{N}$ and for some $x \in X$,

$$d(f^n(x), f^{n+1}(x)) \leq k^n d(x, f(x)).$$

This will be a proof by induction. We know because f is a k-contraction that

$$d(f(x), f^2(x)) \leq kd(x, f(x)).$$

Now assume that

$$d(f^n(x), f^{n+1}(x)) \leq k^n d(x, f(x)).$$

Then because f is a k-contraction, it must be true that

$$\begin{aligned} d(f^{n+1}(x), f^{n+2}(x)) &= d(f(f^n(x)), f(f^{n+1}(x))) \\ &\leq k \cdot k^n d(x, f(x)) \\ &= k^{n+1} d(x, f(x)). \end{aligned}$$

Induction holds and this completes this part of the proof.

Finally, we must prove that if X is complete then for any $x \in X$, the orbit of x under f converges to some fixed point x_0. Because X is complete, every Cauchy sequence in X must converge to some point in X. We know that because $0 < k < 1$ that $\lim_{n\to\infty}(k^n) = 0$. Set $\varepsilon \geq 0$ and let $x \in X$. We know from the properties of geometric series that for all $N \in \mathbb{N}$ that

$$\begin{aligned} \sum_{i=N}^{\infty} k^i d(x, f(x)) &= d(x, f(x)) \sum_{i=N}^{\infty} k^i \\ &= d(x, f(x)) k^N \sum_{i=0}^{\infty} k^i \\ &= d(x, f(x)) k^N (\frac{1}{1-k}). \end{aligned}$$

Because $d(x, f(x))$ and $\frac{1}{1-k}$ are constant and (k^n) converges to 0 we know we can fix $N \in \mathbb{N}$ such that $d(x, f(x))k^N(\frac{1}{1-k}) < \varepsilon$. Fix N such that this is true. Then $\sum_{i=N}^{\infty} k^i d(x, f(x)) < \varepsilon$. Now choose $m, n > N$ such that $m < n$ and $m, n \in \mathbb{N}$. Then from the triangle inequality we know

$$\begin{aligned} d(f^m(x), f^n(x)) &\leq d(f^m(x), f^{m+1}(x)), +d(f^{m+1}(x), f^{m+2}(x)) + ... + d(f^{n-1}(x), f^n(x)) \\ &\leq \sum_{i=m}^{n-1} k^i d(x, f(x)) \\ &< \sum_{i=N}^{\infty} k^i d(x, f(x)) \\ &< \varepsilon. \end{aligned}$$

This means $(f^n(x))$ is Cauchy so it converges to some point x_0 in X. x_0 must be a fixed point because $(f^n(x))$ converges to it. This completes our proof of the contraction mapping theorem. ∎

Now we should use this powerful new theorem to show that we may iterate a certain function an infinite number of times on any starting point in $\mathbb{H}(X)$ and the result is the Sierpinski triangle. To apply this theorem properly though, we must show that $\mathbb{H}(X)$ is complete.

This proof is not simple. It is similar to the one proven in Barnsley's *Fractals Everywhere*. To prove it, however, we need to prove a few lemmas first. The first is called the Dilation Lemma so we should define dilation.

Definition 12 *Let X be a metric space and let $A \in X$ and let $\delta \in \mathbb{R}$. Then*

$$A + \delta = \{x \in X : d(x, a) \leq \delta \; \forall \; a \in A\}.$$

*This is the **dilation of** A **by a ball of radius** δ.*

Now we are ready to work with the Dilation Lemma.

Lemma 13 (The Dilation Lemma) *If (X, d) is a complete metric space, and $A, B \in \mathbb{H}(X)$, and $\varepsilon > 0$ then $h(A, B) \leq \varepsilon$ iff $A \subseteq B + \varepsilon$ and $B \subseteq A + \varepsilon$.*

Proof:
Let (X, d) be a complete metric space, and $A, B \in \mathbb{H}(X)$, and $\varepsilon > 0$. To prove this lemma we will begin with the fact that we know that $h(A, B) = \max\{d(A, B), d(B, A)\}$. Assume $h(A, B) = d(A, B)$. Assume $d(A, B) \leq \varepsilon$. Then $\max\{d(a, B) : a \in A\} \leq \varepsilon$ so $d(a, B) \leq \varepsilon \; \forall \; a \in A$ and therefore $\forall \; a \in A, \; a \in B + \varepsilon$ and finally $A \subseteq B + \varepsilon$. We know $d(B, A) \leq d(A, B)$ so $d(B, A) \leq \varepsilon$ so $\max\{d(b, A) : b \in B\} \leq \varepsilon$ and $d(b, A) \leq \varepsilon \; \forall \; b \in B$. Hence, $\forall \; b \in B, \; b \in A + \varepsilon$ so $B \subseteq A + \varepsilon$. A similar argument can be made if $h(A, B) = d(B, A)$. Now suppose $A \subseteq B + \varepsilon$ and $B \subseteq A + \varepsilon$. Because $A \subseteq B + \varepsilon$ we know $\forall \; a \in A, \; \exists \; b \in B$ such that $d(a, b) \leq \varepsilon$. Let $a \in A$. Then $d(a, B) \leq \varepsilon$. Because we know this is true $\forall \; a \in A$, it must be true that $d(A, B) \leq \varepsilon$. A similar argument can be made for the fact that $d(B, A) \leq \varepsilon$. Then because $h(A, B) = \max\{d(A, B), d(B, A)\}$ we know $h(A, B) \leq \varepsilon$. Thus $h(A, B) \leq \varepsilon$ iff $A \subseteq B + \varepsilon$ and $B \subseteq A + \varepsilon$. ∎

Another lemma we must prove before trying to tackle the fact that $(\mathbb{H}(X), h)$ is complete is the Extension Lemma which is as follows.

Lemma 14 (The Extension Lemma) *Let (X, d) be a metric space. Let (A_n) be a Cauchy sequence of points in $(\mathbb{H}(X), h)$. Let (n_i) be a sequence of integers such that $0 < n_1 < n_2 ...$. Assume there exists a Cauchy sequence of points $\{a_{n_i} \in A_{n_i} : i = 1, 2, 3...\}$ in (X, d). Then there exists a Cauchy sequence of points $\{a^*_n \in A_n : n = 1, 2, 3, ...\}$ such that $a^*_{n_i} = a_{n_i} \; \forall \; i = 1, 2, 3....$*

Proof:
Let (X, d) be a metric space. Let (A_n) be a Cauchy sequence of points in $(\mathbb{H}(X), h)$. Let (n_i) be a sequence of integers such that $0 < n_1 < n_2 ....$

Assume there exists a Cauchy sequence of points $\{a_{n_i} \in A_{n_i} : i = 1, 2, 3...\}$ in (X, d). To prove this lemma, we must first show such a sequence (a_n^*) exists. To do this we will create it. For all $n \in \{1, 2, ..., n_1\}$, pick $a_n^* \in \{a \in A_n : d(a, a_{n_1}) = d(a_{n_1}, A_n)\}$. We know a_n^* exists because A_n is compact. Parallel proofs to this are the proofs corresponding to Definitions 2 and 3. Now for all $i \in \{2, 3, 4...\}$ and all $n \in \{n_{i-1} + 1, n_{i-1} + 2, ..., n_i\}$ let $a_n^* \in \{a \in A_n : d(a, a_{n_i}) = d(a_{n_i}, A_n)\}$. Intuitively, each a_n^* is just the closest point in A_n to its respective a_{n_i}. This means that for $a_{n_i}^*$, that the closest point in A_{n_i} to a_{n_i} is of course a_{n_i} itself so $a_{n_i}^* = a_{n_i}$.

Now we must show that (a_n^*) is Cauchy. Let $\varepsilon > 0$. Because (a_{n_i}) is Cauchy, there exists $N_1 \in \mathbb{N}$ such that if $n_r, n_s > N_1$ then $d(a_{n_r}, a_{n_s}) < \frac{\varepsilon}{3}$. Also, because (A_n) is Cauchy, there exists $N_2 \in \mathbb{N}$ such that if $e, f > N_2$ then $h(A_e, A_f) < \frac{\varepsilon}{3}$. Now let $N = \max\{N_1, N_2\}$. Let $m, n > N$ such that $m \in \{n_{r-1} + 1, n_{r-1} + 2, ..., n_r\}$ and $n \in \{n_{s-1} + 1, n_{s-1} + 2, ..., n_s\}$. Because $h(A_m, A_{n_r}) < \frac{\varepsilon}{3}$ we know there exists $a_m^* \in A_m$ such that $d(a_m^*, a_{n_r}) < \frac{\varepsilon}{3}$. A similar argument can be made for $d(a_{n_s}, a_n^*) < \frac{\varepsilon}{3}$. Now we know that $d(a_m^*, a_n^*) \le d(a_m^*, a_{n_r}) + d(a_{n_r}, a_{n_s}) + d(a_{n_s}, a_n^*) < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon$. Hence, (a_n^*) is Cauchy. ∎

Lemma 15 *Let X be a metric space. Let A be a compact subset of X. Then if $\delta > 0$ it follows that $A + \delta$ is closed.*

Proof:
Let X be a metric space and let $A \subseteq X$ be compact. Set $\delta > 0$. Let (x_n) be a sequence in $A + \delta$ that converges to some x. Then, from the definition of $A + \delta$, there exists a sequence (a_n) in A such that $d(a_n, x_n) \le \delta$ for all $n \in \mathbb{N}$. Because A is compact, we know there exists a subsequence (a_{n_i}) of (a_n) such that $\lim_{n\to\infty}(a_{n_i}) = a$ and $a \in A$. We can now create (x_{n_i}) such that (x_{n_i}) is a subsequence of (x_n) and for all $i \in \mathbb{N}, d(a_{n_i}, x_{n_i}) \le \delta$. Set $\varepsilon > 0$. We can find $I_1 \in \mathbb{N}$ such that for all $i > I_1$, $d(a_{n_i}, a) < \frac{\varepsilon}{2}$. We can find $I_2 \in \mathbb{N}$ such that for all $i > I_2$, $d(x_{n_i}, x) < \frac{\varepsilon}{2}$. Set $I = \max\{I_1, I_2\}$. Now we know for all $i > I$ that

$$d(a, x) \le d(a, a_{n_i}) + d(a_{n_i}, x_{n_i}) + d(x_{n_i}, x) < \frac{\varepsilon}{2} + \delta + \frac{\varepsilon}{2} = \varepsilon + \delta.$$

Since ε was arbitrarily chosen we know $d(a, x) \le \varepsilon$. Because $a \in A$, we know $x \in A + \delta$. Therefore $A + \delta$ contains all its limit points and is closed. ∎

We now need a definition and one more lemma.

Definition 16 *If X is a metric space and A is a subset of X then A is* ***totally bounded*** *if for any $\varepsilon > 0$, A can be covered by a finite number of closed balls with a radius of ε.*

Lemma 17 *Let X be a complete metric space and let A be complete and a totally bounded subset of X. Then A is compact.*

This lemma is essentially proven in Carol Schumacher's standard proof of the Heini-Borel Theorem in *Closer and Closer*. She helpfully points out that any totally bounded subset of a compact metric space that is also closed can be shown to be compact as well. We will not go into the details of this proof in this paper, but again, the intricacies of the argument can be found in *Closer and Closer*.

Now we should have the information we need to prove that if (X, d) is a complete metric space then $(\mathbb{H}(X), h)$ is a complete metric space.

Theorem 18 *Let (X, d) be a complete metric space. Then $(\mathbb{H}(X), h)$ is a complete metric space. Thus, if $(A_n)_{n\geq 1}$ is a Cauchy sequence in $\mathbb{H}(X)$ then $A = \lim_{n\to\infty}(A_n) \in \mathbb{H}(X)$ can be rewritten as $A = \{x \in X :$ there is a sequence, (x_n) with $x_n \in A_n$, that converges to $x\}$.*

Proof:
Let (X, d) be a complete metric space and let (A_n) be a Cauchy sequence in $\mathbb{H}(X)$. Let $A = \{x \in X :$ there is a sequence (x_n) such that $x_n \in A_n$ that converges to $x\}$. We must prove the following:
(1) $A \neq \emptyset$.
(2) A is closed which implies it is complete since X is complete.
(3) For all $\varepsilon > 0$, $\exists\ N \in \mathbb{N}$ such that for all $n \geq N, A \subseteq A_n + \varepsilon$.
(4) A is totally bounded which, combined with the results of (2) and Lemma 17 implies A is compact.
(5) $\lim_{n\to\infty}(A_n) = A$.

Proof of (1): $A \neq \emptyset$.

We must show there exists a Cauchy sequence (a_n) such that $a_n \in A_n$. We know there exists a sequence $\{n_1, n_2, n_3, ...\}$ such that if we pick some n_i from this sequence then $h(A_m, A_n) < \frac{1}{2^i}\ \forall\ m, n \geq n_i$. Pick $a_{n_1} \in A_{n_1}$. Because $h(A_{n_1}, A_{n_2}) < \frac{1}{2}$ we know from the Dilation Lemma that there exists $a_{n_2} \in A_{n_2}$ such that $d(a_{n_1}, a_{n_2}) < \frac{1}{2}$. Because of the way we created the sequence $\{n_1, n_2, n_3, ...\}$ we know that we can pick any n_j and similarly, $h(A_{n_j}, A_{n_{j+1}}) < \frac{1}{2^j}$. Then, again from the Dilation Lemma, we know that there exists $a_{n_j} \in A_{n_j}$ and $a_{n_{j+1}} \in A_{n_{j+1}}$ such that $d(a_{n_j}, a_{n_{j+1}}) < \frac{1}{2^j}$. In this fashion we can create an entire sequence which we will call (a_{n_i}). Now we have a sequence (a_{n_i}) such that $a_{n_i} \in A_{n_i}$ and $d(a_{n_i}, a_{n_{i+1}}) < \frac{1}{2^i}$. We must now show (a_{n_i}) is Cauchy in X. Let $\varepsilon > 0$. Pick N_ε such that $\sum_{i=N_\varepsilon}^{\infty} \frac{1}{2^i} < \varepsilon$. Then for all $t > s \geq N_\varepsilon$ we know

$$\begin{aligned} d(a_{n_s}, a_{n_t}) &\leq d(a_{n_s}, a_{n_{s+1}}) + d(a_{n_{s+1}}, a_{n_{s+2}}) + ... + d(a_{n_{t-1}}, a_{n_t}) \\ &< \sum_{i=s}^{t} \frac{1}{2^i} \\ &< \sum_{i=N_\varepsilon}^{\infty} \frac{1}{2^i} \\ &< \varepsilon. \end{aligned}$$

From the Extension Lemma we know there exists a Cauchy sequence (a_n^*) such that $a_n^* \in A_n$ and $a_{n_i}^* = a_{n_i}$. Because (a_n^*) is in X and X is complete, (a_n^*) converges to some a in X. By the definition of A, $a \in A$ and therefore $A \neq \emptyset$.

Proof of (2): A is closed which implies it is complete since X is complete.

Let (x_i) be a sequence in A that converges to some point x. Because of the way we defined A, we know for each x_i, there exists a sequence (a_{ij}) such that $a_{ij} \in A_j$ and (a_{ij}) converges to x_i. For each $n \in \mathbb{N}$, we can pick x_{i_n} such that $d(x_{i_n}, x) < \frac{1}{2n}$. We can then pick a_{ij_n} such that $d(a_{ij_n}, x_{i_n}) < \frac{1}{2n}$. Then $d(a_{ij_n}, x) \leq d(a_{ij_n}, x_n) + d(x_n, x) < \frac{1}{2n} + \frac{1}{2n} = \frac{1}{n}$. Thus, for any $\varepsilon > 0$, we can find $N \in \mathbb{N}$ such that for all $n > N$, $d(a_{ij_n}, x) < \varepsilon$. Then (a_{ij_n}) converges to x. From the Extension Lemma, we know there exists a sequence (a_{in}^*) such that (a_{in}^*) converges to x and $a_{in}^* \in A_n$ and $a_{in}^* = a_{ij_n}$. Then from the way we created A, $x \in A$ and A is closed. Because A is a closed subset of X and X is complete, A is complete as well.

Proof of (3): For all $\varepsilon > 0$, $\exists\ N \in \mathbb{N}$ such that for all $n \geq N, A \subseteq A_n + \varepsilon$.

Let $\varepsilon > 0$. Let $a \in A$ and let (a_i) be a sequence such that $a_i \in A_i$ and (a_i) converges to a. Pick $N \in \mathbb{N}$ such that for all m, n, and $j > N$, $h(A_m, A_n) \leq \varepsilon$

and $d(a_j, a) < \varepsilon$. Set $n > N$. If $m > n$ then $A_m \subseteq A_n + \varepsilon$ and $a_m \in A_n + \varepsilon$. Because A_n is compact, A_n is closed. We can then deduce from Lemma 15 that $A_n + \varepsilon$ is closed as well. This fact combined with the fact that for all $j > m,\ a_j \in A_n + \varepsilon$ shows that (a_i) must converge to a point in $A_n + \varepsilon$. Then $a \in A_n + \varepsilon$. Because a was arbitrarily chosen, it must be true that $A \subseteq A_n + \varepsilon$. Hence, for all $\varepsilon > 0,\ \exists\ N \in \mathbb{N}$ such that for all $n > N,\ A \subseteq A_n + \varepsilon$.

Proof of (4): A is totally bounded which, combined with the results of (2) and Lemma 17 implies A is compact.

This is a proof by contradiction. Let A not be totally bounded and set $\varepsilon > 0$. Then we can find a sequence (a_i) in A such that for all i and j, $d(a_i, a_j) \geq \varepsilon$ for $i \neq j$. From part (3) we know that there exists n such that $A \subseteq A_n + \frac{\varepsilon}{3}$. Then for all a_i, there exists an $x_i \in A_n$ such that $d(a_i, x_i) < \frac{\varepsilon}{3}$. We know A_n is compact so a subsequence (x_{i_n}) of (x_i) converges in A_n. Then we can find two points x_{i_n} and x_{i_m} such that $d(x_{i_n}, x_{i_m}) < \frac{\varepsilon}{3}$. For these points x_{i_n} and x_{i_m}, there exists $a_{i_n}, a_{i_m} \in A$ such that $d(a_{i_n}, x_{i_n}) < \frac{\varepsilon}{3}$ and $d(a_{i_m}, x_{i_m}) < \frac{\varepsilon}{3}$. Thus

$$d(a_{i_n}, a_{i_m}) \leq d(a_{i_n}, x_{i_n}) + d(x_{i_n}, x_{i_m}) + d(x_{i_m}, a_{i_m}) < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

This is a contradiction so A is totally bounded. Therefore, because A is also closed, A is compact so $A \in \mathbb{H}(X)$.

Proof of (5): $\lim_{n\to\infty}(A_n) = A$.

From the Dilation Lemma, we know we must show that for all $\varepsilon > 0,\ \exists\ N \in \mathbb{N}$ such that for all $n > N,\ A_n \subseteq A + \varepsilon$. Let $\varepsilon > 0$ and let N be such that for all $m, n > N,\ h(A_m, A_n) < \frac{\varepsilon}{2}$. Then for all $m, n > N,\ A_m \subseteq A_n + \frac{\varepsilon}{2}$. Set $n > N$. Let $y \in A_n$. Then because (A_n) is Cauchy and because of the Dilation Lemma we know there exists an increasing sequence (N_i) in $\mathbb{N}$ such that $n < N_1 < N_2 < ... < N_k < ...$ and for all $r, s > N_j$ for some $j,\ A_r \subseteq A_s + \frac{\varepsilon}{2^{j+1}}$. Then it must be true that $A_n \subseteq A_{N_1} + \frac{\varepsilon}{2}$. Because $y \in A_n,\ \exists\ a_{N_1} \in A_{N_1}$ such that $d(y, a_{N_1}) < \frac{\varepsilon}{2}$. We also know from the construction of (A_{N_i}) that $A_{N_1} \subseteq A_{N_2} + \frac{\varepsilon}{2^2}$. Because $a_{N_1} \in A_{N_1},\ \exists\ a_{N_2} \in A_{N_2}$ such that $d(a_{N_1}, a_{N_2}) < \frac{\varepsilon}{2^2}$. Through induction we can make a sequence $a_{N_1}, a_{N_2}, ...$ such that $a_{N_j} \in A_{N_j}$ and $d(a_{N_j}, a_{N_{j+1}}) < \frac{\varepsilon}{2^{j+1}}$. Then

$$d(y, a_{N_j}) \leq d(y, a_{N_1}) + d(a_{N_1}, a_{N_2}) + ... + d(a_{N_{j-1}}, a_{N_j}) < \sum_{i=1}^{j} \frac{\varepsilon}{2^j} < \varepsilon.$$

It is true that (a_{N_i}) is Cauchy so because X is complete, (a_{N_i}) converges to some a. In addition, because of the way we chose n and the way (N_i) was constructed, $A_{N_i} \subseteq A_n + \varepsilon$. Then (a_{N_i}) is in $A_n + \varepsilon$. We know from Lemma 15 $A_n + \varepsilon$ is closed so it must be true that $a \in A_n + \varepsilon$. We know $d(y, a_{N_i}) < \varepsilon \; \forall \; N_i$ and (a_{N_i}) converges to a so it must be true that $d(y, a) \leq \varepsilon$. We can prove this by contradiction. Assume $d(y, a) > \varepsilon$. If $d(y, a) > \varepsilon$ then we can rewrite $d(y, a)$ as $\varepsilon + b$ where $b > 0$. Then it is true that for all N_i, $d(y, a) \leq d(y, a_{N_i}) + d(a_{N_i}, a)$ so $d(y, a) - d(y, a_{N_i}) \leq d(a_{N_i}, a)$. There exists $q > 0$ such that $q < \varepsilon + b - d(y, a_{N_i})$ because $d(y, a_{N_i}) < \varepsilon$. Hence, for all N_i, $q \leq d(a_{N_i}, a)$ and this is a contradiction because (a_{N_i}) converges to a. We know that, according to the Extension Lemma, there is some sequence (a_n^*) that converges to a where $a \in A$ and $a_{n_i}^* = a_{n_i}$ and $a_n^* \in A_n$. Hence, $A \subseteq A_n + \varepsilon \; \forall \; n > N$. This shows $\lim_{n \to \infty}(A_n) = A$ and because A_n was an arbitrary Cauchy sequence in $(\mathbb{H}(X), h)$, it is true that $(\mathbb{H}(X), h)$ is complete.
▮

Section 4: The Hutchinson Operator

We have now successfully shown that the Hausdorff Metric Space is complete so we can apply the contraction mapping theorem to k-contractions within $\mathbb{H}(X)$. This means it is time for us to become more familiar with the Sierpinski triangle itself, and not just general formulas. As has been mentioned before, there are many ways to create the Sierpinski triangle. There are Sir Pinksi's Game, the Chaos Game, and others. There is also the method which iterates functions on points in the Hausdorff Metric Space. A specific function exists which, when iterated an infinite number of times, yields the Sierpinski triangle. It can be described as a Hutchinson operator.

Definition 19 *Let X be a metric space and let A be a subset of X. Let $f_0, \; f_1, \; ... \; , \; f_n$ be a finite number of affine transformations. The* ***Hutchinson Operator*** *of $f_0, \; f_1, \; ... \; , \; f_n$ on A is $f_0(A) \bigcup f_1(A) \bigcup ... \bigcup f_n(A)$.*

This definition basically means that if there is a group of functions where each function changes an image in $\mathbb{R}^2$, the Hutchinson operator of these functions creates their resulting images simultaneously. There are many choices that are possible for the individual affine transformations and lead to different interesting fractals. In the case of the Sierpinski triangle, however, we will be working with the three affine transformations $A_0 : \mathbb{R}^2 \to \mathbb{R}^2, A_1 : \mathbb{R}^2 \to \mathbb{R}^2$, and $A_2 : \mathbb{R}^2 \to \mathbb{R}^2$ which are defined as follows:

$$A_0(\vec{z}) = \tfrac{1}{2}\begin{pmatrix} x \\ y \end{pmatrix}$$

$$A_1(\vec{z}) = \tfrac{1}{2}\begin{pmatrix} x-1 \\ y \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$A_2(\vec{z}) = \tfrac{1}{2}\begin{pmatrix} x-\frac{1}{2} \\ y-\frac{\sqrt{3}}{2} \end{pmatrix} + \begin{pmatrix} \frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{pmatrix}$$

where $\vec{z} = \begin{pmatrix} x \\ y \end{pmatrix}$.

Now we can extend each of these functions so they act on compact sets. This is relatively simple. They will each perform their respective operations on every point in each set, the only difference is that once they are extended they can act on more than one point at once. Because we are only concerned with compact sets we can define these functions in the Hausdorff Metric Space. We, for clarification purposes, will denote each of these affine transformations when acting on compact sets differently than A_0, A_1, and A_2 which only act on points in $\mathbb{R}^2$. We will refer to them as $A_0' : \mathbb{H}(\mathbb{R}^2) \to \mathbb{H}(\mathbb{R}^2)$, $A_1' : \mathbb{H}(\mathbb{R}^2) \to \mathbb{H}(\mathbb{R}^2)$, and $A_2' : \mathbb{H}(\mathbb{R}^2) \to \mathbb{H}(\mathbb{R}^2)$.

From A_0', A_1', and A_2' we can deduce that the Hutchinson operator for the Sierpinski triangle when acting on $Z \in \mathbb{H}(\mathbb{R}^2)$ is

$$H(Z) = A_0'(Z) \bigcup A_1'(Z) \bigcup A_2'(Z)$$

where $H : \mathbb{H}(\mathbb{R}^2) \to \mathbb{H}(\mathbb{R}^2)$.

Let us take a minute and try to understand intuitively what each of these functions actually does to sets in $\mathbb{R}^2$ or to points in $\mathbb{H}(\mathbb{R}^2)$. We will start with A_0'. A_0' basically just takes sets in $\mathbb{R}^2$, contracts them to half their original size, and shifts them linearly toward (0, 0). A_1' takes sets in $\mathbb{R}^2$, contracts them to half their original size, and shifts them linearly toward (1, 0). A_2' takes sets in $\mathbb{R}^2$, contracts them to half their size, and shifts them linearly toward $(\frac{1}{2}, \frac{\sqrt{3}}{2})$. What does this mean when the three functions are combined and how does this create the Sierpinski triangle? Each function's output is an image half the size of the original. That means the Hutchinson operator takes an element in $\mathbb{H}(\mathbb{R}^2)$ and returns the union of three sets, each the same shape as the original but only half its size. Figure 6 illustrates this nicely. The figure is a simple case because, in reality, the original set can be anywhere in $\mathbb{R}^2$ but for the sake of making it easier to understand the picture, the original set is in the center of the future Sierpinski triangle.

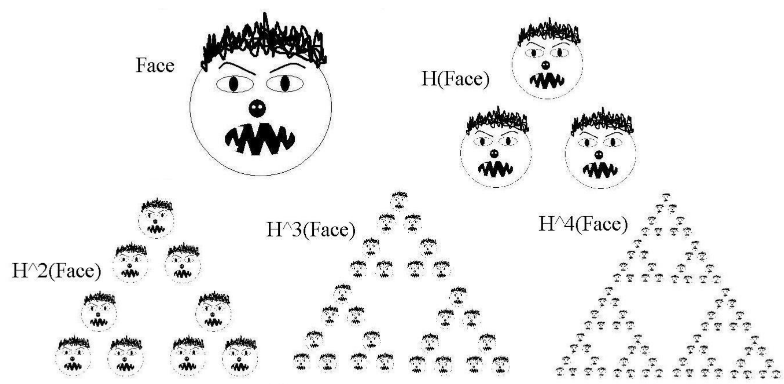


Figure 6: Hutchinson Operator Iterations

Section 5: An Aesthetically Pleasing Limit Point

Figure 6 helps us understand how if the Sierpinski triangle can be created from an angry face with fangs then the Sierpinski triangle can be created from any compact subset of $\mathbb{R}^2$. It is an iterated process which occurs in $\mathbb{H}(\mathbb{R}^2)$. With each iteration of the specific Hutchinson operator, which will from now on be referred to as H, the set in $\mathbb{R}^2$, or the point in $\mathbb{H}(\mathbb{R}^2)$ becomes more like the Sierpinski triangle. A more mathematical way of phrasing this is that $\lim_{n\to\infty}(H^n(S)) = \mathbb{S}$ where $S \in \mathbb{H}(\mathbb{R}^2)$ and $\mathbb{S}$ denotes the Sierpinski triangle. Once again, even though the idea of an entire geometrical shape being a limit point is fairly abstract, it must be remembered that we are working in $\mathbb{H}(\mathbb{R}^2)$ and shapes are in fact points.

We of course cannot be satisfied with the idea that figures converge to $\mathbb{S}$ with just some equations and a figure illustrating what the equations do to a poorly drawn face. We must prove that no matter what the initial point

$S_0 \in \mathbb{H}(\mathbb{R}^2)$ is, the limit of $(H^n(S_0))$ is $\mathbb{S}$. This will be the focus of the rest of this paper, but how can we accomplish such a task? We can use the contraction mapping theorem. We have already shown that because $\mathbb{R}^2$ is complete, $\mathbb{H}(\mathbb{R}^2)$ is complete as well. Now we must show H is a k-contraction and we must show that $\mathbb{S}$ is a fixed point of H.

We will begin showing that H is a k-contraction on $\mathbb{H}(\mathbb{R}^2)$ by showing that each of the "point versions" of its component functions is a k-contraction on $\mathbb{R}^2$. To do this we will need to prove the following lemma.

Lemma 20 *Let* $\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}, \begin{pmatrix} x_3 \\ y_3 \end{pmatrix}, \begin{pmatrix} x_4 \\ y_4 \end{pmatrix} \in \mathbb{R}^2$. *If*

$$d\left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}\right) \leq d\left(\begin{pmatrix} x_3 \\ y_3 \end{pmatrix}, \begin{pmatrix} x_4 \\ y_4 \end{pmatrix}\right)$$

then it must be true that

$$d\left(A_0\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, A_0\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}\right) \leq d\left(A_0\begin{pmatrix} x_3 \\ y_3 \end{pmatrix}, A_0\begin{pmatrix} x_4 \\ y_4 \end{pmatrix}\right).$$

Proof:
Let $\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}, \begin{pmatrix} x_3 \\ y_3 \end{pmatrix}, \begin{pmatrix} x_4 \\ y_4 \end{pmatrix} \in \mathbb{R}^2$. We know that for any $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$, $A_0\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{x}{2} \\ \frac{y}{2} \end{pmatrix}$ so we know that

$$\begin{aligned}
d\left(A_0\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, A_0\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}\right) &= \sqrt{\left(\frac{x_1}{2} - \frac{x_2}{2}\right)^2 + \left(\frac{y_1}{2} - \frac{y_2}{2}\right)^2} \\
&= \sqrt{\left(\frac{1}{2}(x_1 - x_2)\right)^2 + \left(\frac{1}{2}(y_1 - y_2)\right)^2} \\
&= \sqrt{\frac{1}{4}((x_1 - x_2)^2 + (y_1 - y_2)^2)} \\
&= \frac{1}{2}\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \\
&= \frac{1}{2}d\left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}\right).
\end{aligned}$$

A parallel argument can be made to show that

$$d\left(A_0\begin{pmatrix} x_3 \\ y_3 \end{pmatrix}, A_0\begin{pmatrix} x_4 \\ y_4 \end{pmatrix}\right) = \frac{1}{2}d\left(\begin{pmatrix} x_3 \\ y_3 \end{pmatrix}, \begin{pmatrix} x_4 \\ y_4 \end{pmatrix}\right).$$

This means that if

$$d\left(\left(\begin{array}{c} x_1 \\ y_1 \end{array}\right),\ \left(\begin{array}{c} x_2 \\ y_2 \end{array}\right)\right) \leq d\left(\left(\begin{array}{c} x_3 \\ y_3 \end{array}\right),\ \left(\begin{array}{c} x_4 \\ y_4 \end{array}\right)\right)$$

then it must be true that

$$d\left(A_0\left(\begin{array}{c} x_1 \\ y_1 \end{array}\right),\ A_0\left(\begin{array}{c} x_2 \\ y_2 \end{array}\right)\right) \leq d\left(A_0\left(\begin{array}{c} x_3 \\ y_3 \end{array}\right),\ A_0\left(\begin{array}{c} x_4 \\ y_4 \end{array}\right)\right).$$

$\blacksquare$

Overall this implies that if in a set there is a pair of points such that the distance between them is the maximum distance between any pair of points in the set, the distance between the output of A_0 of those two points will be the maximum of the distance between the output of A_0 of any pair of points in the set. The same conclusion can be drawn from similar arguments for A_1, and A_2.

With this result we can show that A_0' is a k-contraction in $(\mathbb{H}(\mathbb{R}^2), h)$.

Lemma 21 *A_0' is a k-contraction in $(\mathbb{H}(\mathbb{R}^2), h)$.*

Proof:
Let $A,\ B \in (\mathbb{H}(\mathbb{R}^2), h)$. For this proof assume that $h(A, B) = d(A, B)$. A parallel argument exists for when $h(A, B) = d(B, A)$. We know from our proof stemming from Definition 3 that there exist some $\left(\begin{array}{c} a_x \\ a_y \end{array}\right) \in A$ and some $\left(\begin{array}{c} b_x \\ b_y \end{array}\right) \in B$ such that $h(A, B) = d\left(\left(\begin{array}{c} a_x \\ a_y \end{array}\right), \left(\begin{array}{c} b_x \\ b_y \end{array}\right)\right)$. Now we can apply A_0 to these points and we know that

$$\begin{aligned}
d\left(A_0\left(\begin{array}{c} a_x \\ a_y \end{array}\right), A_0\left(\begin{array}{c} b_x \\ b_y \end{array}\right)\right) &= \sqrt{\left(\frac{a_x}{2} - \frac{b_x}{2}\right)^2 + \left(\frac{a_y}{2} - \frac{b_y}{2}\right)^2} \\
&= \sqrt{\left(\frac{1}{2}(a_x - b_x)\right)^2 + \left(\frac{1}{2}(a_y - b_y)\right)^2} \\
&= \sqrt{\frac{1}{4}((a_x - b_x)^2 + (a_y - b_y)^2)} \\
&= \frac{1}{2}\sqrt{(a_x - b_x)^2 + (a_y - b_y)^2} \\
&= \frac{1}{2} d\left(\left(\begin{array}{c} a_x \\ a_y \end{array}\right),\ \left(\begin{array}{c} b_x \\ b_y \end{array}\right)\right).
\end{aligned}$$

We have already proven Lemma 20 which says that since

$$d\left(\begin{pmatrix} a_x \\ a_y \end{pmatrix}, \begin{pmatrix} b_x \\ b_y \end{pmatrix}\right) = \max\{d\left(\begin{pmatrix} x \\ y \end{pmatrix}, B\right) : \begin{pmatrix} x \\ y \end{pmatrix} \in A\}$$

it must be true that

$$d\left(A_0\begin{pmatrix} a_x \\ a_y \end{pmatrix}, A_0\begin{pmatrix} b_x \\ b_y \end{pmatrix}\right) = \max\{d\left(A_0\begin{pmatrix} x \\ y \end{pmatrix}, A_0'(B)\right) : \begin{pmatrix} x \\ y \end{pmatrix} \in A\}$$

by showing that if A_0 is applied to every point in two sets, the distance relationships between all pairs of points between the two sets will not change. We can apply what we just showed to show that

$$\begin{aligned} h(A_0'(A), A_0'(B)) &= d\left(A_0\begin{pmatrix} a_x \\ a_y \end{pmatrix}, A_0\begin{pmatrix} b_x \\ b_y \end{pmatrix}\right) \\ &= \frac{1}{2}d\left(\begin{pmatrix} a_x \\ a_y \end{pmatrix}, \begin{pmatrix} b_x \\ b_y \end{pmatrix}\right) \\ &= \frac{1}{2}h(A, B). \end{aligned}$$

We therefore know that A_0' is a k-contraction in $\mathbb{H}(\mathbb{R}^2)$ with a contraction factor of $\frac{1}{2}$. ▮

As a side note not relevant to these proofs, we can use the contraction mapping theorem on A_0' alone. We know that $A_0\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ so $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ is a fixed point of A_0 and because A_0' acts in the same way on individual points as A_0, $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ is a fixed point of A_0' as well. With that fact and Lemma 21, we know that if $S \in (\mathbb{H}(\mathbb{R}^2), h)$ then $(A_0'^n(S))$ converges to $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

Now back to the focus of the paper. The method used to show A_0' is a k-contraction in $\mathbb{H}(\mathbb{R}^2)$ can be applied to A_1' and A_2' to show that they are also k-contractions in $\mathbb{H}(\mathbb{R}^2)$ with contraction factors of $\frac{1}{2}$. We will now use these new results to show that H is a k-contraction as well.

Theorem 22 *H is a k-contraction in $\mathbb{H}(\mathbb{R}^2, h)$.*

Proof:
Let $A,\ B \in (\mathbb{H}(\mathbb{R}^2), h)$. For this proof assume that $h(A,B) = d(A,B)$. A parallel argument exists for when $h(A,B) = d(B,A)$. We know that there exist some $\begin{pmatrix} a_x \\ a_y \end{pmatrix} \in A$ and some $\begin{pmatrix} b_x \\ b_y \end{pmatrix} \in B$ such that $h(A,B) = d\left(\begin{pmatrix} a_x \\ a_y \end{pmatrix}, \begin{pmatrix} b_x \\ b_y \end{pmatrix}\right)$. Now we can apply A_0' to these sets and we know that

$$h(A_0'(A), A_0'(B)) = d\left(A_0\begin{pmatrix} a_x \\ a_y \end{pmatrix}, A_0\begin{pmatrix} b_x \\ b_y \end{pmatrix}\right) = \frac{1}{2}h(A,B).$$

This means that

$$d\left(A_0\begin{pmatrix} a_x \\ a_y \end{pmatrix}, A_0\begin{pmatrix} b_x \\ b_y \end{pmatrix}\right) = \min\{d\left(A_0\begin{pmatrix} a_x \\ a_y \end{pmatrix}, A_0\begin{pmatrix} x \\ y \end{pmatrix}\right) : \begin{pmatrix} x \\ y \end{pmatrix} \in B\}.$$

Then either

$$\min\{d\left(A_0\begin{pmatrix} a_x \\ a_y \end{pmatrix}, H\begin{pmatrix} x \\ y \end{pmatrix}\right) : \begin{pmatrix} x \\ y \end{pmatrix} \in B\} = d\left(A_0\begin{pmatrix} a_x \\ a_y \end{pmatrix}, A_0\begin{pmatrix} b_x \\ b_y \end{pmatrix}\right)$$

or there exists $\begin{pmatrix} x \\ y \end{pmatrix} \in B$ such that

$$\begin{aligned}
\min\{d\left(A_0\begin{pmatrix} a_x \\ a_y \end{pmatrix}, H\begin{pmatrix} x \\ y \end{pmatrix}\right) : \begin{pmatrix} x \\ y \end{pmatrix} \in B\} &= d\left(A_0\begin{pmatrix} a_x \\ a_y \end{pmatrix}, A_i\begin{pmatrix} x \\ y \end{pmatrix}\right) \\
&\leq \frac{1}{2}d\left(\begin{pmatrix} a_x \\ a_y \end{pmatrix}, \begin{pmatrix} b_x \\ b_y \end{pmatrix}\right) \\
&= \frac{1}{2}h(A,B)
\end{aligned}$$

such that $i = 1, 2$. Now we know that $d\left(A_0\begin{pmatrix} a_x \\ a_y \end{pmatrix}, H(B)\right) \leq \frac{1}{2}h(A,B)$ and we know that

$$d\left(A_0\begin{pmatrix} a_x \\ a_y \end{pmatrix}, H(B)\right) = \max\{d\left(A_0\begin{pmatrix} x \\ y \end{pmatrix}, H(B)\right) : \begin{pmatrix} x \\ y \end{pmatrix} \in A\}.$$

This is true because if any other point in $A_0'(A)$ were chosen there would be some point in $A_0'(B)$ such that the distance between the two points would be smaller than the distance from $A'_0(A)$ to $A'_0(B)$. A parallel argument can be used for $A_1\begin{pmatrix} a_x \\ a_y \end{pmatrix}$ and $A_2\begin{pmatrix} a_x \\ a_y \end{pmatrix}$. In any of the three cases,

$$\max\{d\left(A_n\begin{pmatrix} x \\ y \end{pmatrix}, H(B)\right) : \begin{pmatrix} x \\ y \end{pmatrix} \in A\} \leq \frac{1}{2}h(A,B).$$

Therefore $d(H(A), H(B)) \leq \frac{1}{2}h(A,B)$. The same argument can be used to show $d(H(B), H(A)) \leq \frac{1}{2}h(B,A)$. Then

$$h(H(A), H(B)) = \max\{d(H(A), H(B)), d(H(B), H(A))\} \leq \frac{1}{2}h(A,B),$$

so H is a k-contraction with contraction factor $\frac{1}{2}$ in the metric space $(\mathbb{H}(\mathbb{R}^2), h)$. ∎

We have only one proof left until we can use the contraction mapping theorem. We must prove $\mathbb{S}$ is a fixed point in $(\mathbb{H}(\mathbb{R}^2), h)$. Before we can begin with the proof, however, we must devise a way of giving each point in $\mathbb{S}$ an "address." Obviously the Sierpinski triangle can be divided into triangles within triangles. At each level, there are three new triangles, each one-half the size of the triangles at the level before. Let us define these "sub-triangles" as R for the bottom right triangle, L for the bottom left triangle, and T for the top triangle. Each point in $\mathbb{S}$ can be given an address consisting of an infinite series of these letters such as $L,\ R,\ R,\ R,\ T,\ L,\ T,\ R\ldots$. The L at the beginning denotes that the point is located within the bottom left triangle. The R that follows denotes the point is within the bottom right triangle of the bottom left triangle, and so on. Figure 7 helps illustrate this point.

To prove $\mathbb{S}$ is a fixed point of H we will need to show that if $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{S}$ then $H\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{S}$ as well and we will need to show that if $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{S}$ then there is some $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \in \mathbb{S}$ such that $A_i\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$ where $i = 0, 1, 2$.

Lemma 23 *Let $\vec{s} \in \mathbb{S}$. Then $H(\vec{s}) \in \mathbb{S}$.*

Proof:
Consider some point $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{S}$. We know it has some address as described

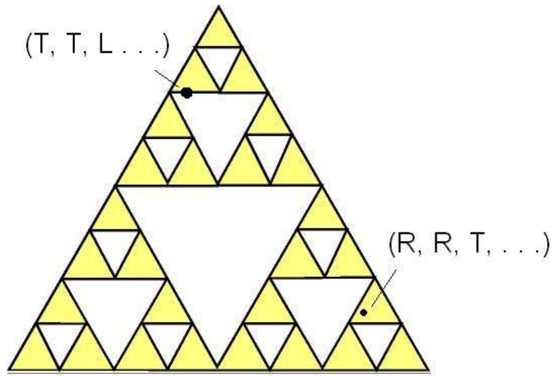


Figure 7: Sierpinski Triangle Addresses

above. For the sake of argument, let us choose the first address letter to be R and let us apply A_0 to $\begin{pmatrix} x \\ y \end{pmatrix}$. There are parallel arguments for if the first address letter is L or T. We know that because the first address letter of $\begin{pmatrix} x \\ y \end{pmatrix}$ is R that $\frac{1}{2} \leq x \leq 1$ and $0 \leq y \leq \frac{\sqrt{3}}{4}$. We can therefore rewrite $\begin{pmatrix} x \\ y \end{pmatrix}$ as $\begin{pmatrix} \frac{1}{2} + \varepsilon_1 \\ \frac{\sqrt{3}}{4} - \varepsilon_2 \end{pmatrix}$ where $0 \leq \varepsilon_1 \leq \frac{1}{2}$ and $0 \leq \varepsilon_2 \leq \frac{\sqrt{3}}{4}$. If we apply A_0 to our point written with its newly written coordinates we get $\begin{pmatrix} \frac{1}{4} + \frac{\varepsilon_1}{2} \\ \frac{\sqrt{3}}{8} - \frac{\varepsilon_2}{2} \end{pmatrix}$.

The strategy for the rest of this proof is to show that if we apply A_0 to our point then we will get a new point located in the same "geometrical" position of another triangle in $\mathbb{S}$ as our original point was located in the triangle with its first address label as R. Because $\mathbb{S}$ is self-similar this means the rest of the new point's address labels will be the same as the original point's address labels so it will be an element of $\mathbb{S}$. This concept is illustrated in Figure 8 where point a was shifted to point b. Point a has the same geometrically relevant position within the bottom right triangle as point b does within the bottom right triangle within the bottom left triangle. Because of the self similarity of $\mathbb{S}$, if each of these triangles were scaled to be the same size they would be identical. This means that other than the first couple of address labels for each point, the two points have identical addresses within $\mathbb{S}$.

Let us examine our new coordinates for $A_0 \begin{pmatrix} x \\ y \end{pmatrix}$ more closely. We know

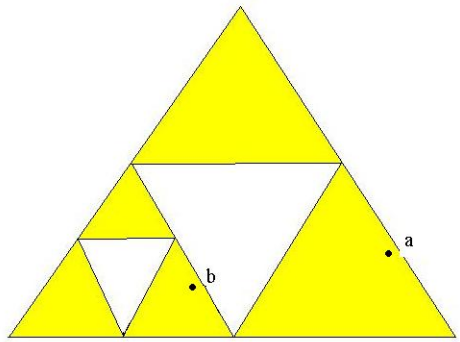


Figure 8: Proportional Spatial Geometry

our new point will have an address starting with L. This is because the x coordinate is $\frac{1}{4}+\frac{\varepsilon_1}{2}$. Because $0 \leq \varepsilon_1 \leq \frac{1}{2}$, it must be true that $0 \leq \frac{\varepsilon_1}{2} \leq \frac{1}{4}$. Therefore, $\frac{1}{4} \leq \frac{1}{4}+\frac{\varepsilon_1}{2} \leq \frac{1}{2}$, which means either the new point has a first address letter of either L or T because all of the points in R have an x coordinate greater than $\frac{1}{2}$.

Now we will turn to the y coordinate. Our new y coordinate is $\frac{\sqrt{3}}{8}-\frac{\varepsilon_2}{2}$. We know that $0 \leq \varepsilon_2 \leq \frac{\sqrt{3}}{4}$ so it must be true that $0 \leq \frac{\varepsilon_2}{2} \leq \frac{\sqrt{3}}{8}$ which implies that $0 \leq \frac{\sqrt{3}}{8}-\frac{\varepsilon_2}{2} \leq \frac{\sqrt{3}}{8}$. All of the points with addresses that start with T have y coordinates greater than $\frac{\sqrt{3}}{4}$ so our point must have an address that starts with L.

We can now examine our point's new address even more microscopically. The triangle with the first two address labels as $L,\ R$ has x coordinates that range from $\frac{1}{4}$ to $\frac{1}{2}$. It has y coordinates that range from 0 to $\frac{\sqrt{3}}{8}$. Therefore, because our new point's coordinates are $\begin{pmatrix} \frac{1}{4}+\frac{\varepsilon_1}{2} \\ \frac{\sqrt{3}}{8}-\frac{\varepsilon_2}{2} \end{pmatrix}$ with $0 \leq \frac{\varepsilon_1}{2} \leq \frac{1}{4}$ and $0 \leq \frac{\varepsilon_2}{2} \leq \frac{\sqrt{3}}{8}$ we know that if our new point is an element of $\mathbb{S}$ then its first two address labels will be L and R. The triangle $(L,\ R)$ is one half the size of the triangle (R). In the triangle (R) our original point was shifted a distance of ε_1 to the right of the leftmost point. In the square containing the triangle $(L,\ R)$ the point resulting from A_0 will be shifted by a factor of $\frac{\varepsilon}{2}$ to the right of the leftmost point. In addition, in the triangle (R) our original point was shifted a distance of ε down from the topmost point. In the square containing the triangle $(L,\ R)$ our newly created point is shifted a value of $\frac{\varepsilon}{2}$ down from the topmost point. This means, because the triangle $(L,\ R)$ is half the size of the triangle (R), that not only is $\begin{pmatrix} \frac{1}{4}+\frac{\varepsilon_1}{2} \\ \frac{\sqrt{3}}{8}-\frac{\varepsilon_2}{2} \end{pmatrix}$ in the triangle $(L,\ R)$, but it is in the same exact geometrically relative position in that triangle as $\begin{pmatrix} x \\ y \end{pmatrix}$ was in the triangle (R). This means that if the first address point of $\begin{pmatrix} x \\ y \end{pmatrix}$ is R, then $A_0 \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{S}$ and its first two address points are $L,\ R$ followed by the same sequence of address points $\begin{pmatrix} x \\ y \end{pmatrix}$ had. This same argument can be used for the functions A_1 and A_2 as well. They will, however, shift points in different directions, which in turn, will add different address labels than will A_0.

This means that if any one of these functions is applied to any point in $\mathbb{S}$, its output will be an element of $\mathbb{S}$ as well. Therefore if H is applied to any point in $\mathbb{S}$ the set of all three of its affine transformations' outputs will be an element of $\mathbb{S}$. ∎

Now we must show that if $\vec{s} \in \mathbb{S}$ then there exists $\vec{s_1} \in \mathbb{S}$ such that $A_n(\vec{s_1}) = \vec{s}$.

Lemma 24 *Let $\vec{s} \in \mathbb{S}$. Then there exists $\vec{s_1} \in \mathbb{S}$ such that $A_n(\vec{s_1}) = \vec{s}$ where $n = 0, 1,$ or 2.*

Proof:
This proof will be similar to the previous one. We will use an address that starts with the points L and R but any pair of the three possible address points will suffice as a starting point. Also, any of the three functions will work but we will prove this using A_0. Let $\vec{s} = \begin{pmatrix} x \\ y \end{pmatrix}$ and let its first two address points be L and R. This will allow us to write $\begin{pmatrix} x \\ y \end{pmatrix}$ as $\begin{pmatrix} \frac{1}{4}+\frac{\varepsilon_1}{2} \\ \frac{\sqrt{3}}{8}-\frac{\varepsilon_2}{2} \end{pmatrix}$ with $0 \leq \varepsilon_1 \leq \frac{1}{2}$ and $0 \leq \varepsilon_2 \leq \frac{\sqrt{3}}{4}$ for the reasons explained in the proof of Lemma 23.

The strategy for the rest of this proof is to show that there is some other point $\vec{s_1}$ such that $H(\vec{s_1})$ will yield $\vec{s}$. To do this we will find a point $\vec{s_1}$ such that $A_0(\vec{s_1}) = \vec{s}$. We will then need to show $\vec{s_1} \in \mathbb{S}$. We can do this through similar methods as the ones used in the proof to Lemma 23. We can show that $\vec{s_1}$ is in the same geometrically relevant position in one of the triangles of $\mathbb{S}$ as $\vec{s}$ was in another triangle of $\mathbb{S}$ and because the triangles are similar and $\vec{s} \in \mathbb{S}$, $\vec{s_1} \in \mathbb{S}$ as well.

Now consider the point $\begin{pmatrix} \frac{1}{2} + \varepsilon_1 \\ \frac{\sqrt{3}}{4} - \varepsilon_2 \end{pmatrix}$. We know that $A_0\left(\begin{pmatrix} \frac{1}{2} + \varepsilon_1 \\ \frac{\sqrt{3}}{4} - \varepsilon_2 \end{pmatrix}\right) = \begin{pmatrix} \frac{1}{4} + \frac{\varepsilon_1}{2} \\ \frac{\sqrt{3}}{8} - \frac{\varepsilon_2}{2} \end{pmatrix}$. Now we must show that $\begin{pmatrix} \frac{1}{2} + \varepsilon_1 \\ \frac{\sqrt{3}}{4} - \varepsilon_2 \end{pmatrix} \in \mathbb{S}$. We know that the triangle (R) is twice the size of the triangle $(L,\ R)$. Let us denote $\begin{pmatrix} \frac{1}{2} + \varepsilon_1 \\ \frac{\sqrt{3}}{4} - \varepsilon_2 \end{pmatrix}$ as s_1. The x coordinate of $\vec{s_1}$ is twice the distance to the right of the leftmost point of the the triangle (R) as the x coordinate of $\vec{s}$ was in $(L,\ R)$. The y coordinate of $\vec{s_1}$ is twice the distance down from the topmost point of the triangle (R) as the y coordinate of $\vec{s}$ was in $(L,\ R)$. This means that $\vec{s_1}$ is in the same geometrically relevant position in (R) as $\vec{s}$ was in $(L,\ R)$. Because all the triangles in $\mathbb{S}$ are similar and $\vec{s} \in \mathbb{S}$ it must be true that $\vec{s_1} \in \mathbb{S}$ as well. Thus, for every point $\vec{s} \in \mathbb{S}$, there exists some other point, $\vec{s_1} \in \mathbb{S}$ such that $H(\vec{s_1})$ will yield $\vec{s}$ (it will also give two other points, but we have already proven that they will be in $\mathbb{S}$ also). ▮

Theorem 25 $\mathbb{S}$ *is a fixed point of* H *in* $(\mathbb{H}(\mathbb{R}^2), h)$.

Proof:
From Lemmas 23 and 24 we know that if $\vec{s} \in \mathbb{S}$ then $H(\vec{s}) \in \mathbb{S}$ and there is some $\vec{s_0} \in \mathbb{S}$ such that $H(\vec{s_0})$ will yield $\vec{s}$. Therefore, $H(\mathbb{S}) = \mathbb{S}$ so $\mathbb{S}$ is a fixed point of H in $(\mathbb{H}(\mathbb{R}^2), h)$. ▮

We can therefore conclude that because $(\mathbb{H}(\mathbb{R})^2, h)$ is complete, H is a k-contraction with a contraction factor of $\frac{1}{2}$, and $\mathbb{S}$ is a fixed point of H, if $S \in (\mathbb{H}(\mathbb{R})^2, h)$ then $H^n(S)$ converges to $\mathbb{S}$. This fact is fascinating in and of itself. More generally, any compact shape in $\mathbb{R}^2$ can be taken and molded into a very precise and ordered shape with the many interesting properties encountered throughout this paper. There are many more interesting fractals similar to the Sierpinski triangle such as the Koch Snowflake and the Cantor Set. Each can be written as the limit of an iterated map of a Hutchinson operator on a point in the Hausdorff metric space. There are also many more complicated fractals which would require a more in-depth study of fractal dimension, complex functions, and other areas to understand. Each of these fractals has its own special qualities but all of them share qualities such as self-similarity. From what we have learned here, however, we know that the Sierpinski triangle is justifiably described as an aesthetically pleasing limit point.

Bibliography

Howard Anton, Chris Rorres. *Elementary Linear Algebra*. New York: John Wiley and Sons Inc., 2000.

Barnsley, Michael F. *Fractals Everywhere*. San Diego: Morgan Kaufmann, 1993.

Devany, Robert L. *A First Course In Chaotic Dynamical Systems*. Reading, Massachusetts: Addison-Wesley Publishing Company, Inc., 1992.

Devany, Robert L. *An Introduction to Chaotic Dynamical Systems*. Reading, Massachusetts: Addison-Wesley Publishing Company, Inc., 1989.

Hart, John C. "Iterated Function Systems and Recurrent Iterated Function Systems." May 14, 1996.

Hartmut Jurgens, Heinz-Otto Peitgen, Dietmar Saupe. *Chaos and Fractals New Frontiers of Science*. New York: Springer-Verlag, 1992.

Holmgren, Richard A. *A First Course in Discrete Dynamical Systems*. New York: Springer, 1991.

McCauley, J. L. *Chaos, Dynamics and Fractals*. Cambridge: Cambridge University Press, 1995.

Pickover, Clifford A. *Fractal Horizons*. New York: St. Martin's Press, 1996.

Schroeder, Manfred. *Fractals, Chaos, Power Laws*. New York: W. H. Freeman and Company, 1991.

Schumacher, Carol. *Closer and Closer An Introduction to Real Analysis*. Jones
and Bartlett Publishers (not yet in print).

The cover art was taken from the website with the address http://www.spsu.edu/math/tile/im/sierp.gif.

The triangles in Figure 2 were taken from the website with the address
http://mathforum.org/advanced/robertd/chaos_game.html

Special thanks in alphabetical order to Professors Judy Holdener, Robert Milnikel, and Carol Schumacher who all helped me with this paper.