A continuous nowhere differentiable function, by Rich Schwartz

These notes give a geometric example of a continuous function that is nowhere differentiable. The example, which is a coordinate function of the standard parametrization of the Koch snowflake, explains one geometric mechanism behind continuous nowhere differentiable functions: *self-similarity*.

Subdivision and Refinement: The *subdivision* of an interval is the collection of two intervals we get by cutting it into equal halves. If we have a collection $\mathcal{I}$ of k intervals, we produce a new collection $\mathcal{J}$ of $2k$ intervals by subdividing each member of $\mathcal{I}$. We write $\mathcal{I} \to \mathcal{J}$. Let $\mathcal{I}_0 = \{I_0\}$, where $I_0 = [0, 1]$. Now we iteratively define collections $\mathcal{I}_0 \to \mathcal{I}_1 \to \mathcal{I}_2 ....$ Here $\mathcal{I}_n$ is the partition of $[0, 1]$ into 2^n intervals of length 2^{-n}. We order these intervals from left to right.

The *refinement* of an obtuse isosceles triangle Δ is the collection of two smaller and similar triangles $\Delta_1, \Delta_2 \subset \Delta$ such that the long side of Δ_j coincides with the (j)th short side of Δ. If we have a collection $\mathcal{T}$ of k obtuse isosceles triangles, we produce a new collection $\mathcal{U}$ of $2k$ obtuse isosceles triangles by refining each member of $\mathcal{T}$. We write $\mathcal{T} \to \mathcal{U}$.

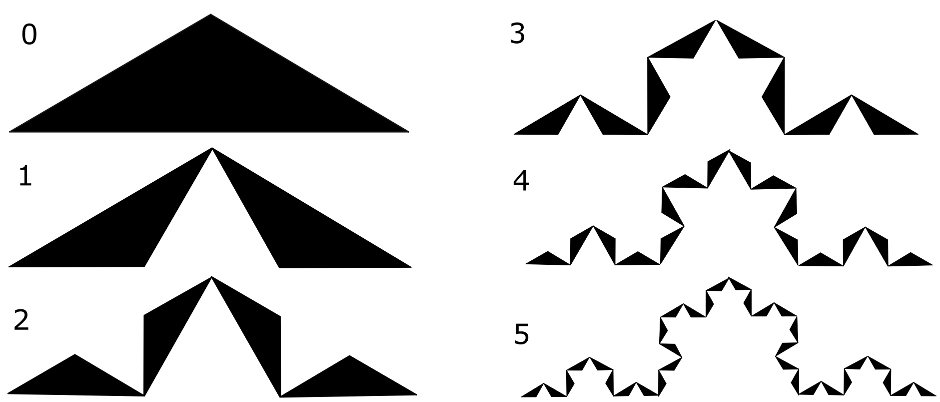

Figure 1: Iterated triangle refinement a

Let $\mathcal{T}_0 = \{T_0\}$, where T_0 is a triangle whose horizontal base has length 1 and whose short sides have length $1/\sqrt{3}$. Now we iteratively define collections $\mathcal{T}_0 \to \mathcal{T}_1 \to \mathcal{T}_2 \to ...$ Here $\mathcal{T}_n$ has 2^n congruent triangles, all having long side length $(\sqrt{3})^{-n}$. Figure 1 shows the triangles of $\mathcal{T}_n$ for $n = 0, 1, 2, 3, 4, 5$. We order the triangles of $\mathcal{T}_n$ so that consecutive ones share a common vertex and the first is leftmost.

A Continuous Map: Figure 2 shows the $\mathcal{T}_n$ superimposed for $n = 0, ..., 9$. We can see a curve emerging. Figure 3 below makes this even more clear.

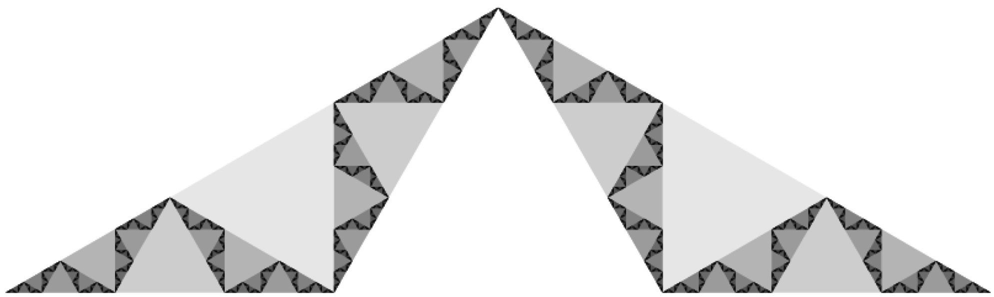

Figure 2: Superimposing the triangles

There is an exact correspondence between the intervals of $\mathcal{I}_n$ and the triangles of $\mathcal{T}_n$. The triangles are nested precisely as the intervals are: Some triangle of $\mathcal{T}_m$ is contained in a triangle of $\mathcal{T}_n$ if and only if the corresponding interval of $\mathcal{I}_m$ is contained in the corresponding interval of $\mathcal{I}_n$.

Every $t \in [0, 1]$ can be written as $t = \bigcap I_n$, where I_n is one of the intervals of $\mathcal{I}_n$. This description is unique unless t is an endpoint of one of the partition intervals. In this case there are two such intersections $t = \bigcap I_n = \bigcap J_n$, and for large n the common endpoint of I_n and J_n is t.

Corresponding to I_n is the triangle T_n. We define $f(t) = \bigcap T_n$. Since we have a nested intersection of compact sets, the intersection is non-empty. Also, the diameter of these triangles decays by a factor of $\sqrt{3}$ each time, so they have a unique intersection point. In case t is the endpoint of a partition interval, the two corresponding nested intersections of triangles intersect at the common vertex, and this common vertex is defined to be $f(t)$. So, f is defined in all cases. Figure 3 below gives another way to visualize f.

Here is why f is continuous: If s and t are very nearby points in $[0, 1]$ they both lie either in the same interval I_n for a large value of n or they lie in adjacent intervals I_n and J_n for some large n. But then the images $f(s)$ and $f(t)$ either lie in a triangle of diameter $(\sqrt{3})^{-n}$ or they lie in adjacent triangles of diameter $(\sqrt{3})^{-n}$. In either case, they are close together in $\boldsymbol{R}^2$. More formally, given $\epsilon > 0$ we choose n so that $2 \times (\sqrt{3})^{-n} < \epsilon$ and then we let $\delta = 2^{-n}$. These choices establish the classic ϵ-δ definition of continuity. (Indeed, these choices directly establish uniform continuity.) Since f is continuous, so are its coordinate functions.

Self-Similarity: A *similarity* of $\boldsymbol{R}^d$ is a map which scales distances by a constant factor and preserves orientation. We only care about the cases $d = 1, 2$. Below, A_1 and A_2 respectively will denote similarities of $\boldsymbol{R}$ and $\boldsymbol{R}^2$.

The map f is *self-similar* in a way we now explain. Let I be an interval of $\mathcal{I}_n$. Let T be the corresponding triangle of $\mathcal{T}_n$. Let $f|_I$ be the restriction of f to I. We have

$$f = A_2 \circ f|_I \circ A_1, \tag{1}$$

where A_1 is the similarity such that $A_1(I_0) = I$ and A_2 is the similarity such that $A_2(T) = T_0$. This works because the subdivision process is compatible with similarities. Note that A_1 shrinks distances by a factor of 2^{-n} and A_2 expands distances by a factor of $(\sqrt{3})^n$.

Here is a useful reformulation of the self-similarity. We note first that A_1 and A_2 are both invertible. Choose some $t \in I$ and let $t^* = A_1^{-1}(t)$. Note that

$$A_2^{-1} \circ f(t^*) = A_2^{-1} \circ f \circ A_1^{-1}(t) = f|_I(t) = f(t). \tag{2}$$

As t ranges over I, the point t^* ranges over all of $[0, 1]$.

Nowhere Differentiability: We write $f(t) = (x(t), y(t))$. Suppose for the sake of contradiction that $x(\cdot)$ is differentiable at some $t \in [0, 1]$. Let $C = \max(1, |x'(t)|)$. There is an interval I of $\mathcal{I}_n$ that contains t. Let $s \in I$ be another point. We have $|s - t| \leq 2^{-n}$. If n is sufficiently large then

$$|x(s) - x(t)| \leq 2C|s - t| \leq \alpha_n = 2C \times 2^{-n}. \tag{3}$$

Geometrically, Equation 3 says that the point $f(s)$ is within α_n of the vertical line L through $f(t)$.

Let $s^* = A_1^{-1}(s)$. By Equation 2, the point $A_2^{-1} \circ f(s^*) = f(s)$ is within α_n of L. Applying A_2 we see that $f(s^*)$ is within

$$\beta_n = (\sqrt{3})^n \alpha_n = 2C \times (\sqrt{3}/2)^n$$

of the line $A_2(L)$. As s ranges over all of I, the point s^* ranges over all of $[0, 1]$. Hence $f([0, 1])$ is contained within β_n of $A_2(L)$. We can make β_n as close as we want to 0 by taking n large enough, and so the above situation is only possible if $f([0, 1])$ lies in a straight line. This contradiction finishes the proof.

Pictures and Discussion: There is another way to think about the map f defined above. We can define f_n to be the map of constant speed $(4/3)^n$ which maps $[0,1]$ to $\boldsymbol{R}^2$ in such a way that $f_n(I)$ is the long side of T, where T is the triangle of $\mathcal{T}_{2n}$ corresponding to the interval I of $\mathcal{I}_{2n}$. Figure 3 shows the images of f_n for $n = 1,2,3,4,5$. We could have defined f as a limit of these maps, but the triangle definition makes the continuity of the limit more clear.

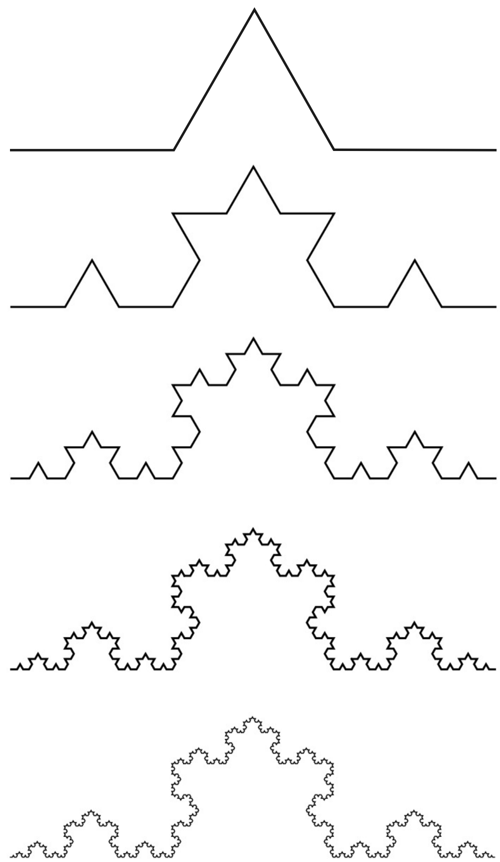

Figure 3: A sequence of maps

Notice that as n increases the map f_n is moving increasingly fast and also is getting increasingly wiggly. We might say, at least informally, that the limit f is moving infinitely fast and wiggling in an infinitely crazy way. This kind of behavior is not compatible with differentiability. Even though the"infinitely fast and wiggly" idea is a bit vague, it gives some intuitive feel for the breakdown of differentiability.

We could have started with a different shape obtuse isosceles triangle. Figure 4 shows the analogue of Figure 2 with respect to a different shape.

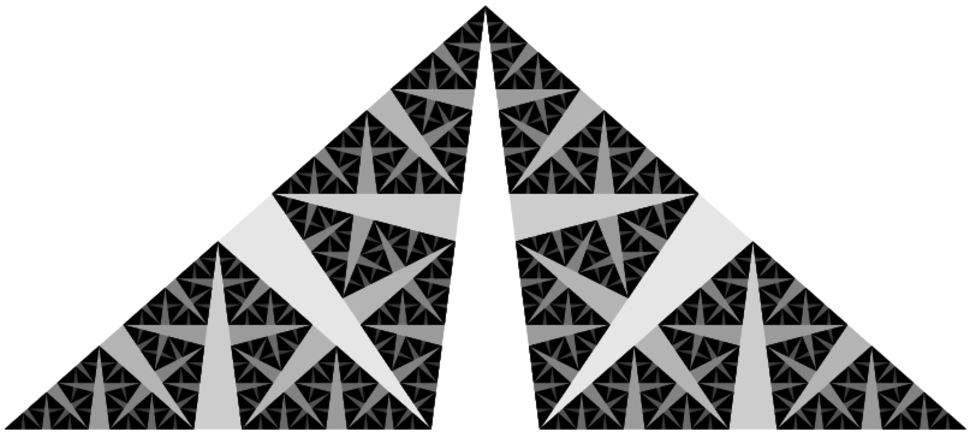

Figure 4: Nested triangles again

Actually, if we took a little bit more care with the definitions, we could base the construction on a triangle that is not isosceles. Figure 5 shows an example.

Figure 5: Nested triangles again

The triangle refinement construction has a limit, for right triangles, and the result is a continuous curve whose image is the entire initial right triangle.

The same proof as above shows, in all these cases, that the corresponding coordinate functions are continuous but nowhere differentiable.