CONTINUOUS, NOWHERE DIFFERENTIABLE FUNCTIONS

SARAH VESNESKE

ABSTRACT. The main objective of this paper is to build a context in which it can be argued that most continuous functions are nowhere differentiable. We use properties of complete metric spaces, Baire sets of first category, and the Weierstrass Approximation Theorem to reach this objective. We also look at several examples of such functions and methods to prove their lack of differentiability at any point.

Contents

List of Figures

Date: May 10, 2019.
1991 *Mathematics Subject Classification.* 26A27.

Introduction

Mathematical intuition is often what guides our pursuit of further knowledge through the development of rigorous definitions and proofs. To illustrate this idea, let us consider the conception of a continuous function. Initially, we have an idea that a function should be continuous if it can be "drawn" without lifting one's pen. It is a continuous movement, with no jumping around.

Of course, general intuitive concepts cannot take us very far, and thus we develop the $\epsilon - \delta$ definition of continuity with which we are familiar. This definition was carefully constructed in ways that lean into our intuition. We hope to ensure that the functions we perceive to be continuous stay that way within our formal definition, and that those which seem to have obvious discontinuities are also preserved.

However, once formalizing these intuitive ideas, we are often faced with mathematical truths that are more difficult to accept. When we are first introduced to the concept of differentiability, or rather of non-differentiability, we are shown functions with a handful of problematic points on an otherwise differentiable function. It seems at first glance that for any function to remain continuous it must have a finite - or at most a countable - number of these cusps. It is here that we see the mathematics push back against our intuition.

In 1872, Karl Weierstrass became the first to publish an example of a continuous, nowhere differentiable function [5]. It is now known that several mathematicians, including Bernard Bolzano, constructed such functions before this time. However, the publishing of the Weierstrass function (originally defined as a Fourier series) was the first instance in which the idea that a continuous function must be differentiable almost everywhere was seriously challenged.

Differentiability, what intuitively seems the default for continuous functions, is in fact a rarity. As it turns out, chaos is omnipresent, and the order required for differentiability is in no way guaranteed under the weak restrictions of continuity.

In the first section of this paper we provide an overview of key results from several areas of mathematics. First, we present a proof of the Weierstrass Approximation Theorem and develop an obvious corollary. Then, we review properties of complete metric spaces and apply them to the space of continuous functions. We then prove the Baire Category Theorem and develop definitions for Baire first and second category sets.

In the second half of the paper we are formally introduced to everywhere continuous, nowhere differentiable functions. We begin with a proof that Weierstrass' famous nowhere differentiable function is in fact everywhere continuous and nowhere differentiable. We then address the main motivation for this paper by showing that the set of continuous functions differentiable at any point is of first category (and so is relatively small). We conclude with a final example of a nowhere differentiable function that is "simpler" than Weierstrass' example.

The standard notation in this paper will follow that used in Russell A. Gordon's *Real Analysis: A First Course* [1]. Any unusual uses of notation or terms will be defined as they are used.

1. Review and Preliminaries

1.1. **Weierstrass Approximation Theorem.** To begin this section, we introduce Bernstein polynomials and prove several facts about them. We use the construction of these polynomials in our proof of the Weierstrass Approximation Theorem.

Definition 1.1. *If f is a continuous function on the interval $[0,1]$, then the **nth Bernstein polynomial of f** is defined by*

$$B_n(x,f) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k}x^k(1-x)^{n-k}.$$

Note that the degree of B_n is less than or equal to n.

Remark 1.2. *If f is a constant function, namely $f(x)=c$ for all x, then*

$$B_n(x,f) = \sum_{k=0}^{n} c\binom{n}{k}x^k(1-x)^{n-k} = c$$

given that the binomial expansion of $\sum\limits_{k=0}^{n}\binom{n}{k}x^k(1-x)^{n-k} = (x+(1-x))^n = 1$.

Let $p,q>0$. Then $(p+q)^n = \sum\limits_{k=0}^{n}\binom{n}{k}p^kq^{n-k}$. Taking the derivative with respect to p and then multiplying by p on both sides of the resulting equation leads us to

$$np(p+q)^{n-1} = \sum_{k=0}^{n}\binom{n}{k}kp^kq^{n-k},$$

and repeating this process gives us

$$n(n-1)p^2(p+q)^{n-2} + np(p+q)^{n-1} = \sum_{k=0}^{n}\binom{n}{k}k^2p^kq^{n-k}.$$

We note that in the special case when $p+q=1$, we find that

$$np = \sum_{k=0}^{n} \binom{n}{k} kp^k q^{n-k}, \text{ and so } \; n^2p^2 - np^2 + np = \sum_{k=0}^{n} \binom{n}{k} k^2 p^k (1-p)^{n-k}.$$

Theorem 1.3. *If f is continuous on the interval $[0,1]$, then $\{B_n(x,f)\}$ converges uniformly to f on $[0,1]$.*

Proof. We first note that the continuity of f on the closed interval $[0,1]$ implies that f is uniformly continuous and bounded on $[0,1]$. Let M be a bound for f on $[0,1]$.
Let $\epsilon > 0$ be given. Then there exists $\delta > 0$ guaranteed by the definition of uniform continuity such that $|f(y) - f(x)| < \epsilon$ for all $x, y \in [0,1]$ which satisfy $|y - x| < \delta$. Choose $N \in \mathbb{N}$ such that $N \geq \frac{M}{2\epsilon\delta^2}$, and fix $n \geq N$. Then, for any $x \in [0,1]$ we have

$$\begin{aligned}
|f(x) - B_n(x,f)| &= \left| \sum_{k=0}^{n} \left(f(x) - f\left(\frac{k}{n}\right) \right) \binom{n}{k} x^k (1-x)^{n-k} \right| \\
&\leq \sum_{k=0}^{n} \left| f(x) - f\left(\frac{k}{n}\right) \right| \binom{n}{k} x^k (1-x)^{n-k}
\end{aligned}$$

At this point, we can divide the sum into two parts. When $|\frac{k}{n} - x| < \delta$, we use the uniform continuity of f, and when $|\frac{k}{n} - x| \geq \delta$, we use the bound M. Let T be the set of all k such that $|\frac{k}{n} - x| < \delta$ and let S be the set of all k such that $|\frac{k}{n} - x| \geq \delta$. Thus, the previous expression is equivalent to

$$\begin{aligned}
&\sum_{k\in T} \left| f(x) - f\left(\frac{k}{n}\right) \right| \binom{n}{k} x^k (1-x)^{n-k} + \sum_{k\in S} \left| f(x) - f\left(\frac{k}{n}\right) \right| \binom{n}{k} x^k (1-x)^{n-k} \\
&\leq \sum_{k=0}^{n} \epsilon \binom{n}{k} x^k (1-x)^{n-k} + \frac{1}{\delta^2} \sum_{k\in S} \delta^2 2M \binom{n}{k} x^k (1-x)^{n-k} \\
&\leq \epsilon + \frac{2M}{\delta^2} \sum_{k=0}^{n} \left(\frac{k}{n} - x \right)^2 \binom{n}{k} x^k (1-x)^{n-k} \\
&= \epsilon + \frac{2M}{n^2\delta^2} \sum_{k=0}^{n} \left(k^2 - n2kx + n^2x^2 \right) \binom{n}{k} x^k (1-x)^{n-k} \\
&= \epsilon + \frac{2M}{\delta^2} \left[\sum_{k=0}^{n} k^2 \binom{n}{k} x^k (1-x)^{n-k} - 2nx \sum_{k=0}^{n} k \binom{n}{k} x^k (1-x)^{n-k} + n^2x^2 \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} \right] \\
&= \epsilon + \frac{2M}{n^2\delta^2} \left((n^2x^2 - nx^2 + nx) - 2nx(nx) + n^2x^2 \right)
\end{aligned}$$

$$= \epsilon + \frac{2M}{n\delta^2}(x(1-x))$$
$$\leq \epsilon + \frac{2M}{N\delta^2} \cdot \frac{1}{4}$$
$$\leq 2\epsilon.$$

Given that x was arbitrary in the interval $[0, 1]$, we have shown that for any $\epsilon > 0$ there exists an $N \in \mathbb{N}$ such that $|f(x) - B_n(x, f)| \leq 2\epsilon$ for all $x \in [0, 1]$ and for all $n \geq N$. Thus, the sequence $\{B_n(x, f)\}$ converges uniformly to f on the interval $[0, 1]$. □

We now use the uniform convergence of $\{B_n(x, f)\}$ to prove the following theorem.

Theorem 1.4. *(Weierstrass Approximation) If f is continuous on $[a, b]$, then there exists a sequence $\{p_n\}$ of polynomials such that $\{p_n\}$ converges uniformly to f on $[a, b]$.*

Proof. Let $g(t) = f(a + t(b - a))$. Then g is continuous on $[0, 1]$. By Theorem 1.3, the sequence $\{B_n(t, g)\}$ converges to g uniformly on $[0, 1]$. That is, for each $\epsilon > 0$ there exists some $N \in \mathbb{N}$ such that $|B_n(t, g) - g(t)| < \epsilon$ for all $t \in [0, 1]$ and for all $n \geq N$.

Now, consider the sequence $\{p_n\}$ defined by $p_n(x) = B_n\left(\frac{x-a}{b-a}, g\right)$ for each n. We can see that $p_n(x)$ is a polynomial for every n. Now, for any $x \in [a, b]$, the number $\frac{x-a}{b-a} \in [0, 1]$, and thus,

$$|p_n(x) - f(x)| = \left| B_n\left(\frac{x-a}{b-a}, g\right) - g\left(\frac{x-a}{b-a}\right) \right| < \epsilon$$

for all $n \geq N$. Therefore, the sequence defined by $\{p_n\}$ converges uniformly on $[a, b]$. Since f was an arbitrary continuous function, this holds for all continuous functions on the interval $[a, b]$. □

Corollary 1.5. *If f is continuous on an interval $[a, b]$, then for each $\epsilon > 0$ there exists a polynomial p such that $|p(x) - f(x)| < \epsilon$ for all $x \in [a, b]$.*

This corollary follows easily from Theorem 1.4, as we can choose our p from the tail end of $\{p_n\}$. This greatly simplifies several of the arguments that follow later in the paper.

1.2. **Metric Spaces.** We now review several properties of metric spaces and introduce the metric space of continuous functions that the remainder of the paper utilizes.

Definition 1.6. *A* ***metric space*** (X, d) *consists of a set* X *and a function* $d\colon X \times X \to \mathbb{R}$ *that satisfies the following four properties:*

1. $d(x, y) \geq 0$ *for all* $x, y \in X$.
2. $d(x, y) = 0$ *if and only if* $x = y$.
3. $d(x, y) = d(y, x)$ *for all* $x, y \in X$.
4. $d(x, y) \leq d(x, z) + d(z, y)$ *for all* $x, y, z \in X$.

The function d *which gives the distance between two points in* X *is known as a metric.*

We are most familiar with the metric space $\mathbb{R}$ with $d(x, y) = |x - y|$, where the distance between two real numbers is denoted by the absolute value of their difference. In n dimensions, we are also familiar with the Euclidean distance in $\mathbb{R}^n$ with $d(x, y) = \sqrt{\sum\limits_{i=1}^{n}(x_i - y_i)^2}$.

Here, we consider the metric space consisting of the set of all continuous functions on the closed interval $[a, b]$, with the metric $d_\infty(x, y) = \sup\{|x(t) - y(t)| : t \in [a, b]\}$. We refer to this metric space as $C([a, b])$.

Definition 1.7. *A metric space* (X, d) *is* ***complete*** *if every Cauchy sequence in* (X, d) *converges to a point in* X.

Definition 1.8. *The open set* $N_\epsilon(x)$ *denotes the* ***neighborhood*** *of* x *with radius* $\epsilon > 0$. *That is, if* X *is a metric space with* $x \in X$, *then* $N_\epsilon(x) = \{y \in X : d(x, y) < \epsilon\}$.

Definition 1.9. *Let* (X, d) *be a metric space and let* E *be a subset of* X. *The set* E *is* ***dense*** *in* X *if and only if* $E \cap U \neq \varnothing$ *for every nonempty open set,* U, *in* X. *The set* E *is* ***nowhere dense*** *in* X *if and only if the interior of the closure of* E *is empty. This means that for all* $x \in E$ *and for all* $\epsilon > 0$ *there exists a point* $y \in N_\epsilon(x)$ *such that* $y \notin E$.

Definition 1.10. *A set* E *is* ***separable*** *if it contains a countable, dense subset.*

Theorem 1.11. *The set of polynomials is dense in* $C([a, b])$.

Proof. Let P be the set of polynomials on the interval $[a, b]$. Consider an arbitrary open set $U \subseteq C([a, b])$. Since U is nonempty, there must exist at least one function, say f, in $U \subseteq C([a, b])$. Then, given that U is open, for some $\epsilon > 0$ there exists a neighborhood of f, $N_\epsilon(f)$, that is completely contained in U.

Now, by Corollary 1.5, there exists a polynomial p in $C([a, b])$ such that $|p(x) - f(x)| < \epsilon$ for all $x \in [a, b]$, i.e., that $p \in N_\epsilon(f) \subseteq U$. This means that $P \cap U$ is nonempty. Since U was arbitrary, this holds for every nonempty open set in $C([a,b])$. By definition, the set P is dense in $C([a,b])$. □

Lemma 1.12. *$C([a,b])$ is separable.*

Proof. Let P_Q be the set of polynomials with rational coefficients on the interval $[a,b]$. As in the proof of Theorem 1.11, let $U \subseteq C([a,b])$ be an arbitrary nonempty open set, containing at least one continuous function, say f. We showed by Corollary 1.5, there exists a polynomial p in $C([a,b])$ such that $|p(x) - f(x)| < \frac{\epsilon}{2}$ for all $x \in [a,b]$.

Now, we also know that the set of rational numbers, $\mathbb{Q}$, is dense in $\mathbb{R}$. This means that for any $r \in \mathbb{R}$ and for any $\epsilon > 0$, there exists some rational number $q \in \mathbb{Q}$ such that $q \in N_\epsilon(r)$, i.e., $|r - q| < \epsilon$.

Let $p(x) = r_0 + r_1 x + r_2 x^2 + \cdots + r_n x^n$ where $r_i \in \mathbb{R}$. For any $\epsilon > 0$, let $\epsilon_i = \frac{\epsilon}{2nb^i}$. We can choose $q_i \in \mathbb{Q}$ such that $|q_i - r_i| < \epsilon_i$ for all i. Using these q_i, let $g(x) = q_0 + q_1 x + q_2 x^2 + \cdots + q_n x^n$ We can see that

$$\begin{aligned}
|g(x) - f(x)| &\leq |g(x) - p(x)| + |p(x) - f(x)| \\
&< \sum_{k=1}^{n} \epsilon_k x^k + \frac{\epsilon}{2} \\
&= \epsilon_0 + \epsilon_1 x + \epsilon_2 x^2 + \cdots + \epsilon_n x^n + \frac{\epsilon}{2} \\
&= \frac{\epsilon}{2n} + \frac{\epsilon}{2nb} x + \frac{\epsilon}{2nb^2} x^2 + \cdots + \frac{\epsilon}{2nb^n} x^n + \frac{\epsilon}{2} \\
&\leq \frac{\epsilon}{2n} + \frac{\epsilon}{2nx} x + \frac{\epsilon}{2nx^2} x^2 + \cdots + \frac{\epsilon}{2nx^n} x^n + \frac{\epsilon}{2} \\
&= n\left(\frac{\epsilon}{2n}\right) + \frac{\epsilon}{2} = \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.
\end{aligned}$$

This means that $g \in N_\epsilon(f) \subseteq U$, which implies $P \cap U$ is nonempty. Since U was arbitrary, this holds for every nonempty open set in $C([a,b])$. By definition, the set P_Q is dense in $C([a,b])$.

Further, P_Q is a countable union of countable sets (as the number of possible terms are the integers and the possible coefficients are the rationals). Thus, the metric space $C([a,b])$ contains a countable, dense subset, and is separable. □

Remark 1.13. *The set of polynomials is a subspace of $C([a,b])$, but it is not a closed subset of $C([a,b])$.*

1.3. **Baire Category Sets and the Baire Category Theorem.**

Theorem 1.14. *(Baire Category Theorem) Let (X, d) be a complete metric space. If $\{O_n\}$ is a sequence of open, dense sets in X, then $\bigcap_{n=1}^{\infty} O_n$ is dense in X. In particular, $\bigcap_{n=1}^{\infty} O_n \neq \varnothing$.*

Proof. Let U be a nonempty, open set. We need to show that there exists an x such that for all positive integers n, the set $U \cap O_n$ contains x. Choose $x_0 \in U$ and $\epsilon > 0$ so that $N_\epsilon(x_0)$ is a subset of U.
Since $O_1 \cap N_{\frac{\epsilon}{4}}(x_0) \neq \varnothing$, there exists some $x_1 \in O_1$ such that $d(x_0, x_1) < \frac{\epsilon}{4}$. Choose $\epsilon_1 < \frac{\epsilon}{2}$ so that $S_1 = N_{\epsilon_1}(x_1) \subseteq O_1$. Then, $\bar{S_1} \subseteq N_\epsilon(x_0) \subseteq U$.

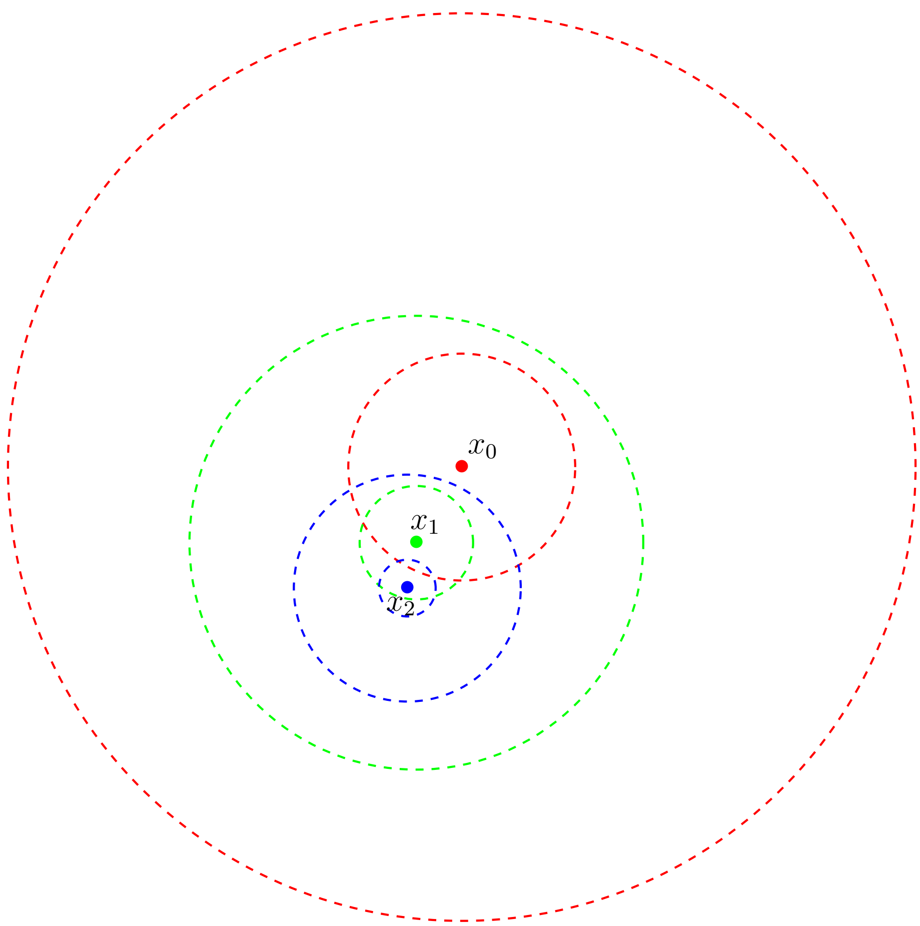

FIGURE 1. A two-dimensional illustration of the nested sets $\{S_n\}$

Since $O_2 \cap N_{\epsilon_1/4}(x_1) \neq \varnothing$, there exists $x_2 \in O_2$ such that $d(x_1, x_2) < \frac{\epsilon_1}{4}$. Choose $\epsilon_2 < \frac{\epsilon_1}{2}$ so that $S_2 = N_{\epsilon_2}(x_2) \subseteq O_2$. Then, $\bar{S_2} \subseteq N_{\epsilon_1}(x_1) \subseteq S_1$.

Figure 1 illustrates how the x_i might be chosen and how the S_i are strictly contained in one another. Each outer circle corresponding to its center x_i has radius $\epsilon_i < \frac{\epsilon_{i-1}}{2}$. The smaller circle has radius $\frac{\epsilon_i}{4}$, illustrating the open disk in which the next x_{i+1} can be chosen. This preserves the condition that $d(x_i, x_{i+1}) < \frac{\epsilon_i}{4}$.
We can see that each large circle is contained within the previous. More rigorously, we will show why $\overline{S_{i+1}} \subseteq S_i$ for all i.

Let $y \in \overline{S_{i+1}}$. Then, given the way each x_i and ϵ_i are chosen, and using the triangle inequality, we see that

$$\begin{aligned} d(y, x_i) &\leq d(y, x_{i+1}) + d(x_{i+1}, x_i) \\ &\leq \frac{\epsilon_i}{2} + d(x_{i+1}, x_i) \\ &\leq \frac{\epsilon_i}{2} + \frac{\epsilon_i}{4} < \epsilon_i. \end{aligned}$$

Thus, the distance from the center of S_i, given by x_i, to any point in the set S_{i+1}, will always be less than the radius given by ϵ_i. This means that $S_{i+1} \subseteq N_{\epsilon_i}(x_i)$.

This process is repeated so that we obtain a sequence $\{S_n\}$ of nonempty sets such that

i) $S_n \subseteq O_n$,
ii) $\overline{S_{n+1}} \subseteq S_n$,
iii) $S_1 \subseteq U$, and
iv) the diameter of S_n, or the maximum distance between any two elements of S_n, is given by $\frac{2\epsilon_n}{4^n}$ which approaches 0 as n approaches infinity.

Let $\{x'_n\}$ be a sequence such that $x'_n \in S_n$ for each n. Given that $S_{n+1} \subseteq S_n$, we can see that $x'_n \in S_n \subseteq S_N$ for all $n \geq N$. So, for any $n, m \geq N$, it follows that $x_n, x_m \in S_n$. From item (iv), we know that $d(x'_n, x'_m) < \frac{2\epsilon_N}{4^N}$. Then $\{x'_n\}$ is a Cauchy sequence and converges, say to the point x. Given that (X, d) is a complete metric space, x must be an element of (X, d). Thus,

$$x \in \bigcap_{n=2}^{\infty} \bar{S_n} \subseteq \bigcap_{n=1}^{\infty} S_n \subseteq U \cap \bigcap_{n=1}^{\infty} S_n \subseteq U \cap \bigcap_{n=1}^{\infty} O_n.$$

Given that x is contained in this intersection of all O_n with any nonempty, open set U, by definition $\bigcap_{n=1}^{\infty} O_n$ must be dense in X. It follows that $\bigcap_{n=1}^{\infty} O_n$ is also nonempty. $\square$

Definition 1.15. *Let (X, d) be a metric space and let A be a subset of X. The set A is of* ***first category*** *if there exists a countable collection $\{E_n\}$ of nowhere dense sets such that $A = \bigcup_{n=1}^{\infty} E_n$. The set A is of* ***second category*** *if it is not of first category.*

Sets of first category are also referred to as **meager** sets, implying a sense of relative smallness when compared to the larger set X.
As an example, the set of rational numbers is of first category in $\mathbb{R}$. We know that the rationals are countable and the reals are not, and so in this way the rationals are a smaller set. Further, the irrationals are of second category in $\mathbb{R}$, which agrees with our conception that both are uncountable.

2. Continuous, Nowhere Differentiable Functions

2.1. **Weierstrass' nowhere differentiable function.** At this point, we consider a particular function that was first presented by Weierstrass in 1872. We prove that this function is continuous on the real numbers but is not differentiable at any point.

Theorem 2.1. [4] *Let b be a real number such that $0 < b < 1$ and let a be a positive odd integer. If $ab > 1$ and $\frac{2}{3} > \frac{\pi}{ab-1}$, then*
$$W(x) = \sum_{n=0}^{\infty} b^n \cos(a^n x\pi)$$
is continuous on $\mathbb{R}$ and is not differentiable at any point in $\mathbb{R}$.

Proof. Throughout the proof of this theorem, we fix a and b, assuming that the values satisfy the given conditions. Within figures 2 and 3, we let $a = 13$ and $b = \frac{1}{2}$.

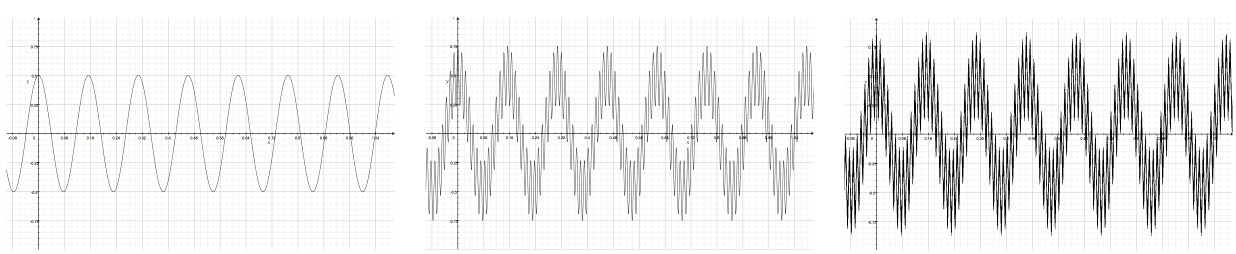

FIGURE 2. Plots of the partial sums of the Weierstrass function for $n = 1, 2, 3$

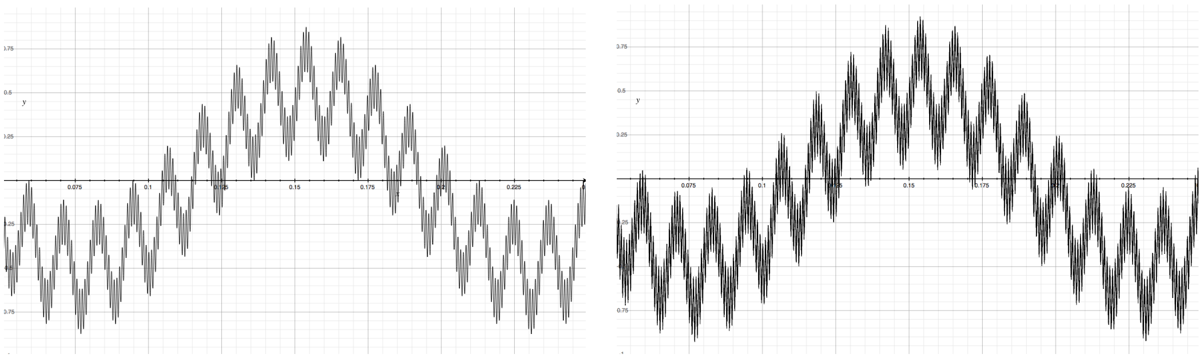

FIGURE 3. A closer look at the development of $W(x)$ for $n = 3$ and $n = 4$.

First, we consider the continuity of the function. Note that $|b^n \cos(a^n x\pi)| \leq b^n$ for all real values of x. Given that our conditions require that $b < 1$, we know that $\sum_{n=0}^{\infty} b^n$ is a convergent geometric series. Thus, by the Weierstrass M-test, the function W is continuous on $\mathbb{R}$.
From this, we can see that the claim of continuity for our function W is not the difficult part of our proof.

Now, we focus on the proof that W is not differentiable at any point in $\mathbb{R}$.

Remark 2.2. *If $W_N(x) := \sum_{n=0}^{N} b^n \cos(a^n x\pi)$, then $\{W_N\}$ converges uniformly to W. Each W_N can be differentiated an infinite number of times, but we claim that W does not have a derivative at any point.*

Fix a value x_0. For each $m \in \mathbb{N}$, let β_m be the integer which satisfies the conditions $\frac{1}{2} \leq \beta_m - a^m x_0 < \frac{3}{2}$. We can restate this in two inequalities, so that $x_0 \leq \frac{\beta_m}{a^m} - \frac{1}{2a^m}$ and $\frac{\beta_m}{a^m} - \frac{3}{2a^m} < x_0$. Let us call $\frac{\beta_m}{a^m} = \alpha_m$ Given that $a > 1$, we can see that $\lim_{m\to\infty} \alpha_m = x_0$.
Now, if we were to suppose that W was differentiable at the point x_0, then

$$\lim_{m\to\infty} \frac{W(\alpha_m) - W(x_0)}{\alpha_m - x_0} = W'(x_0).$$

To form a contradiction, will show that in fact

$$\lim_{m\to\infty} (-1)^{\beta_m} \frac{W(\alpha_m) - W(x_0)}{\alpha_m - x_0} = \infty,$$

and therefore that $W'(x_0)$ does not exist.

To begin, express

$$\begin{aligned}
(-1)^{\beta_m}\frac{W(\alpha_m)-W(x_0)}{\alpha_m-x_0} &= (-1)^{\beta_m}\frac{\sum_{n=0}^{\infty} b^n\cos(a^n(\alpha_m)\pi)-\sum_{n=0}^{\infty} b^n\cos(a^n(x_0)\pi)}{\alpha_m-x_0}\\
&= \sum_{n=0}^{\infty}(-1)^{\beta_m}b^n\frac{\cos(a^n(\alpha_m)\pi)-\cos(a^n(x_0)\pi)}{\alpha_m-x_0}\\
&= \sum_{n=0}^{m-1}(-1)^{\beta_m}b^n\frac{\cos(a^n\alpha_m\pi)-\cos(a^nx_0\pi)}{\alpha_m-x_0}\\
&\qquad+\sum_{n=m}^{\infty}(-1)^{\beta_m}b^n\frac{\cos(a^n\alpha_m\pi)-\cos(a^nx_0\pi)}{\alpha_m-x_0}\\
&= A_m+B_m.
\end{aligned}$$

We show below that:

$$\text{(i) } |A_m|\le (ab)^m\frac{\pi}{ab-1} \qquad \text{and} \qquad \text{(ii) } B_m\ge\frac{2}{3}(ab)^m.$$

From this, it follows that

$$\begin{aligned}
(-1)^{\beta_m}\frac{W(\alpha_m)-W(x_0)}{\alpha_m-x_0} &\ge B_m-|A_m|\\
&\ge (ab)^m\frac{2}{3}-(ab)^m\frac{\pi}{ab-1}\\
&= (ab)^m\left(\frac{2}{3}-\frac{\pi}{ab-1}\right).
\end{aligned}$$

Given that $ab>1$ and $\frac{\pi}{ab-1}<\frac{2}{3}$, we can see that

$$\lim_{m\to\infty}(ab)^m\left(\frac{2}{3}-\frac{\pi}{ab-1}\right)=\infty$$

and thus

$$\lim_{m\to\infty}(-1)^{\beta_m}\frac{W(\alpha_m)-W(x_0)}{\alpha_m-x_0}=\infty$$

Therefore, the proof will be complete upon verification of (i) and (ii).

i. To begin, we note that

$$A_m=A_m\cdot\frac{a^n\pi}{a^n\pi}=\sum_{n=0}^{m-1}(-1)^{\beta_m}a^nb^n\pi\frac{\cos(a^n\alpha_m\pi)-\cos(a^nx_0\pi)}{a^n(\alpha_m-x_0)\pi}$$

$$= \sum_{n=0}^{m-1} (-1)^{\beta_m} a^n b^n \pi (-\sin(c_{n,m}))$$

for some $c_{n,m} \in [a, b]$ guaranteed by the Mean Value Theorem. Taking the absolute value, we find

$$\begin{aligned}
|A_m| &= \left| \sum_{n=0}^{m-1} a^n b^n \pi \sin(c_{n,m}) \right| \\
&\leq \sum_{n=0}^{m-1} a^n b^n \pi \\
&= \pi \frac{(ab)^m - 1}{ab - 1} \\
&< (ab)^m \frac{\pi}{ab - 1}.
\end{aligned}$$

Thus, we can see that item (i) holds.

ii. For all terms in B_m, by definition $n \geq m$. For these terms,

$$\begin{aligned}
(-1)^{\beta_m} \cos(a^n \alpha_m \pi) &= (-1)^{\beta_m} \cos\left(a^n \frac{\beta_m}{a^m} \pi\right) \\
&= (-1)^{\beta_m} \cos(a^{n-m} \beta_m \pi).
\end{aligned}$$

Given that a is an odd integer, $\cos(a^{n-m}\beta_m \pi) = -1$ when β_m is odd, and $\cos(a^{n-m}\beta_m\pi) = 1$ when β_m is even. Thus,

$$(-1)^{\beta_m} \cos(a^{n-m} \beta_m \pi) = (-1)^{\beta_m} (-1)^{\beta_m} = 1.$$

It then follows that

$$\begin{aligned}
(-1)^{\beta_m} \cos(a^n x_0 \pi) &= (-1)^{\beta_m} \cos(a^{n-m} a^m x_0 \pi) \\
&= (-1)^{\beta_m} \cos(a^{n-m} \beta_m \pi - a^{n-m} \beta_m \pi + a^{n-m} a^m x_0 \pi) \\
&= (-1)^{\beta_m} \cos(a^{n-m} \beta_m \pi + a^{n-m} (a^m x_0 - \beta_m) \pi) \\
&= (-1)^{\beta_m} \big[\cos(a^{n-m} \beta_m \pi) \cos(a^{n-m} (a^m x_0 - \beta_m) \pi) \\
&\qquad - \sin(a^{n-m} \beta_m \pi) \sin(a^{n-m} (a^m x_0 - \beta_m) \pi) \big] \\
&= (-1)^{\beta_m} \cos(a^{n-m} \beta_m \pi) \cos(a^{n-m} (a^m x_0 - \beta_m) \pi) \\
&= \cos(a^{n-m} (a^m x_0 - \beta_m) \pi).
\end{aligned}$$

Using these two identities, we can establish that

$$\begin{aligned} B_m &= \sum_{n=m}^{\infty} b^n \frac{1-\cos(a^{n-m}(a^m x_0 - \beta_m)\pi)}{\frac{\beta_m}{a^m} - x_0} \\ &= b^m \frac{1-\cos((a^m x_0 - \beta_m)\pi)}{\frac{\beta_m}{a^m} - x_0} + \sum_{n=m+1}^{\infty} b^n \frac{1-\cos(a^{n-m}(a^m x_0 - \beta_m)\pi)}{\frac{\beta_m}{a^m} - x_0}. \end{aligned}$$

Recall that β_m was chosen to satisfy $\frac{1}{2} \leq \beta_m - a^m x_0 < \frac{3}{2}$. This means that $\cos((a^m x_0 - \beta_m)\pi) \leq 0$ and $\frac{\beta_m}{a^m} - x_0 > 0$. Thus,

$$\begin{aligned} B_m &\geq b^m \frac{1-0}{\frac{\beta_m}{a^m} - x_0} + \sum_{n=m+1}^{\infty} b^n \frac{0}{\frac{\beta_m}{a^m} - x_0} \\ &= b^m \frac{1}{\frac{\beta_m}{a^m} - x_0} \\ &= a^m b^m \frac{1}{\beta_m - a^m x_0} \\ &\geq (ab)^m \frac{1}{\beta_m - a^m x_0} > (ab)^m \frac{1}{\frac{3}{2}} = (ab)^m \frac{2}{3}. \end{aligned}$$

This verifies part (ii).

Now, given that we have verified parts (i) and (ii), we have completed the proof and have shown that

$$\lim_{m\to\infty} (-1)^{\beta_m} \frac{W(\alpha_m) - W(x_0)}{\alpha_m - x_0} = \infty.$$

This means that $W'(x_0)$ does not exist, and thus $W(x)$ is not differentiable at any point on $\mathbb{R}$. Additionally, we have shown that $W(x)$ has unbounded difference quotients on $\mathbb{R}$. $\square$

2.2. **Somewhere differentiable functions.** We now consider the set of continuous functions which have at least one point on $[0,1]$ where they are differentiable.

Theorem 2.3. *If $A = \{f \in C([0,1]) :$ there exists an $x_0 \in [0,1]$ such that $f'(x_0)$ exists$\}$, then A is a set of first category.*

Proof. For each $N \in \mathbb{N}$ we define $F_N = \{f \in C([0,1]) :$ there exists an $x_0 \in [0,1]$ such that $|f(x) - f(x_0)| \leq N|x - x_0|$ for all $x \in [0,1]\}$. If we prove the following three statements:

i) F_N is closed for all N,
ii) F_N is nowhere dense for all N, and
iii) $A \subset \bigcup_{N=1}^{\infty} F_N$,

then we have shown that A is a set of first category.

i. Let f be a limit point of F_N. Then there exists a sequence $\{f_n\}$ in F_N such that $\{f_n\}$ converges to f. For each n, choose $x_n \in [0,1]$ to be the point defined as x_0 "corresponding" to each f_n within the definition of F_n. That is, for each n, the inequality $|f_n(x) - f_n(x_n)| \leq n|x - x_n|$ holds for all $x \in [0,1]$. We can see that the sequence $\{x_n\}$ is bounded given that each $x_n \in [0,1]$.

By the Bolzano-Weierstrass Theorem, the sequence $\{x_n\}$ has a convergent subsequence $\{x_{n_k}\}$ that converges to $x_0 \in [0,1]$. For all k, let $y_k = x_{n_k}$ and let $g_k = f_{n_k}$. Since convergence in $C([0,1])$ is uniform convergence, and since $\{f_{n_k}\}$ converges to f and $\{y_k\}$ converges to x_0, for each $\epsilon > 0$ there exists a k such that

$$\begin{aligned} |g_k(y_k) - f(x_0)| &= |f_{n_k}(y_k) - f(x_0)| \\ &\leq |f_{n_k}(y_k) - f(y_k)| + |f(y_k) - f(x_0)| \\ &\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \end{aligned}$$

Therefore, $\{g_k\}$ converges uniformly to f.
Now, suppose $x \in [0,1]$ such that $x \neq x_0$. Then there exists some $K \in \mathbb{N}$ such that $y_k \neq x$ for all $k \geq K$. Thus, for all $x \in [0,1]/\{x_0\}$, we find that

$$\left| \frac{g_k(y_k) - g_k(x)}{y_k - x} \right| \leq N$$

implies

$$\left| \frac{f(x_0) - f(x)}{x_0 - x} \right| \leq N.$$

Hence, for all $x \in [0,1]$, we see that $|f(x) - f(x_0)| \leq N|x - x_0|$, so $f \in F_N$. Since f was an arbitrary limit point in F_N, the set F_N must contain all of its limit points and thus is closed.

ii. To show that F_N is nowhere dense, we must show that for any arbitrary $g \in C([0,1])$ and any $\epsilon > 0$, there exists some $\phi \in C([0,1])$ such that $d_\infty(\phi, g) < \epsilon$ and $\phi \notin F_N$. That is, we show that there is a continuous function $\phi \in N_\epsilon(g)$ that is not in F_N.

By Corollary 1.5, there exists a polynomial p which converges uniformly to g on $[0,1]$. This means that for each $\epsilon > 0$ there is some p such that $|p(x) - g(x)| < \frac{\epsilon}{2}$ for all $x \in [0,1]$.

Recall that in our proof that the Weierstrass function is not differentiable at any point, we also showed that W has unbounded difference quotients for all x. Thus, we can multiply W by any constant c and be left with a function cW which also has unbounded difference quotients for every x. Let $M = \sup\{W(x) : x \in [0,1]\}$ and let $\phi = p(x) + \frac{\epsilon}{2M}W(x)$. Given that p and W are both elements of $C([0,1])$, it is clear that ϕ must be as well.

Now,

$$\begin{aligned} d_\infty(\phi, g) &= \sup\{|\phi(x) - g(x)| : 0 \le x \le 1\} \\ &= \sup\left\{|p(x) + \frac{\epsilon}{2M}W(x) - g(x)| : 0 \le x \le 1\right\} \\ &\le \sup\left\{|p(x) - g(x)| + \left|\frac{\epsilon}{2M}W(x)\right| : 0 \le x \le 1\right\} \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2M}M = \epsilon. \end{aligned}$$

We have shown that for any arbitrary continuous function g and for each $\epsilon > 0$ we can find another continuous function $\phi \in N_\epsilon(g)$. Further, F_N is defined by functions which have bounded difference quotients, and we have established that. We previously established that ϕ has only unbounded difference quotients on $[0,1]$, and so is not an element of F_N. Thus, we have shown that F_N is nowhere dense.

iii. Let $f \in A$ and choose x_0 such that $f'(x_0)$ exists. Then there is a $\delta > 0$ such that $0 < |x - x_0| < \delta$ implies

$$\left|\frac{f(x) - f(x_0)}{x - x_0}\right| < |f'(x_0)| + 1.$$

There is a bound for f, say M, such that $|f(x)| \le M$ for all $x \in [0,1]$ which satisfy the inequality

$$\left|\frac{f(x) - f(x_0)}{x - x_0}\right| \le \frac{2M}{\delta}$$

when $|x - x_0| \geq \delta$.

If $N \geq \max\left\{\frac{2M}{\delta}, |f'(x_0)| + 1\right\}$, then for all $x \in [0,1]$, we find that $|f(x) - f(x_0)| \leq N|x - x_0)|$. Thus $f \in F_N$. Since f was chosen to be an arbitrary function in A, we conclude that $A \subseteq \bigcup_{N=1}^{\infty} F_N$.

Therefore, given that we have shown F_N is both closed and nowhere dense, and that $A \subseteq \bigcup_{N=1}^{\infty} F_N$, we have shown that A is a set of the first category. □

Remark 2.4. *Since $C([0,1])$ is a set of second category, the set A is a proper subset of $C([0,1])$.*

Within this framework, we can claim that functions which are differentiable at even a single point are, in a sense, rare when considering all possible continuous functions. From a different angle, we say that most continuous functions are not differentiable anywhere.

While initially startling, after contemplating for a while we can be convinced that it makes sense for most arbitrary continuous functions to be chaotic in this way. Our definition of continuity is much weaker than that of differentiability. While it seems that the continuous functions we can visualize must be differentiable at most points, we can see that this is not due to the natures of continuity and differentiability, but only to the limitations of our visualizations.

2.3. **An algebraic nowhere differentiable function.** Now that we have shown that most continuous functions are pathological in this way, we might hope to find a "simpler" function than Weierstrass' famous example. While the Weierstrass function is an interesting example in that all of its partial sums are differentiable, it does require the use of a transcendental function to express. We will give another example that uses only algebraic functions in its construction. In fact, the function is defined using a piecewise linear function. This example is found in Gordon's text [1] in the proof of his Theorem 7.31.

Theorem 2.5. *Define $g : \mathbb{R} \to \mathbb{R}$ by letting $g(x) = |x|$ for $-1 \leq x < 1$ and $g(x+2) = g(x)$ for all other x. The function f defined by*

$$f(x) = \sum_{k=0}^{\infty} \left(\frac{3}{4}\right)^k g(4^k x)$$

is continuous on $\mathbb{R}$ but is not differentiable at any point on $\mathbb{R}$.

Proof. By the Weierstrass M-test, we see that this series converges uniformly. Further, each term of the series is continuous. It is easy then to see that f is continuous on $\mathbb{R}$. Again, the more interesting fact is that this function is nowhere differentiable.

For each x, we will find a sequence $\{\delta_n\}$ of nonzero real numbers such that $\{\delta_n\}$ converges to zero, but the sequence

$$\left\{\frac{f(x+\delta_n)-f(x)}{\delta_n}\right\}$$

is unbounded and thus does not converge. This will show that f is not differentiable at x.
Fix $x \in \mathbb{R}$, and let n be an integer. We choose $\delta_n = \pm\frac{1}{2\cdot 4^n}$ such that there are no integers between $4^n x$ and $f^n(x+\delta_n)$. We note here that the sequence $\{\delta_n\}$ converges to 0.
We now consider the value of $|g(4^k x + 4^k\delta_n) - g(4^k x)|$.
For $k > n$, we note that $4^k\delta_n$ is a multiple of 2 and that g is a periodic function with a period of 2. This means $|g(4^k x + 4^k\delta_n) - g(4^k x)| = 0$.
When $k = n$, we know $|4^k\delta_n| = \frac{1}{2}$. Further, the function g is linear with a slope of either 1 on the interval $[4^k x + 4^k\delta_n, 4^k x]$ or -1 on the interval $[4^k x, 4^k x + 4^k\delta_n]$, depending on the choice of δ_n. Thus in this case, $|g(4^k x + 4^k\delta_n) - g(4^k x)| = \frac{1}{2}$.
Consider when $k < n$. Given that $|g(y)-g(x)| \leq |y-x|$, we find that $|g(4^k x + 4^k\delta_n) - g(4^k x)| \leq |4^k\delta_n|$.
Using these results, we conclude that

$$\begin{aligned}
\left|\frac{f(x+\delta_n)-f(x)}{\delta_n}\right| &= \left|\sum_{k=0}^{n}\left(\frac{3}{4}\right)^k \frac{g(4^k x + 4^k\delta_n) - g(4^k x)}{\delta_n}\right| \\
&\geq \left(\frac{3}{4}\right)^n 4^n - \sum_{k=0}^{n-1}\left|\left(\frac{3}{4}\right)^k \frac{g(4^k x + 4^k\delta_n) - g(4^k x)}{\delta_n}\right| \\
&\geq 3^n - \sum_{k=0}^{n-1} 3^k \\
&= 3^n - \frac{3^n - 1}{2} > \frac{3^n}{2}.
\end{aligned}$$

We have shown that the sequence

$$\left\{\frac{f(x+\delta_n)-f(x)}{\delta_n}\right\}$$

is unbounded and thus does not converge. Given $\{\delta_n\}$ converges to 0, it follows that f is not differentiable at x. □

Conclusions

Within this paper, we used the prevalence of nowhere differentiable functions to show how mathematics often forces us to rethink our intuition. Proving the continuity and nowhere differentiablity of two very different example functions, we convince ourselves that these functions are not as impossible as we may have initially thought. Within the context of Baire first and second category sets, we make the claim that most continuous functions are nowhere differentiable.

Acknowledgements

I would like to thank my supervisor, Russ Gordon, for his help with this paper and for four years of mentorship, kindness, and support. I would also like to thank Sarah Fix and Pat Keef for their thoughtful comments and edits on countless drafts.

References

[1] Russell A. Gordon. *Real Analysis: A First Course*. Addison-Wesley Higher Mathematics, 2nd edition, 2002.

[2] Madeleine Hanson-Colvin. Everywhere continuous nowhere differentiable functions. 2014.

[3] Michael E. McLaughlin. An introduction to everywhere continuous, nowhere differentiable functions. 2013.

[4] Steen Pedersen. *From Calculus to Analysis*, chapter 11, pages 236–240. Springer International Publishing, 2015.

[5] Wikipedia. Weierstrass function.

Department of Mathematics, Whitman College, Walla Walla, WA 99362

E-mail address: vesneskesm@gmail.com