THE WEIERSTRASS APPROXIMATION THEOREM AND ITS GENERALISATIONS

ANIKA CHOPRA

Abstract. In this paper, we present two different proofs of the Weierstrass Approximation theorem, a method of approximating a continuous real-valued function on a closed interval, and explore its generalisation, the Stone-Weierstrass theorem, on real and complex algebras.

1. INTRODUCTION

In 1885, German mathematician Karl Weierstrass proved that a real-valued continuous function on a closed interval could be approximated uniformly by a sequence of polynomials. Until then, it was thought that a function had to be infinitely differentiable in order to be approximated by polynomials.

Weierstrass' result, which came to be known as the Weierstrass approximation theorem, was later simplified and generalised by American mathematician Marshall Harvey Stone in 1948. Stone proved the theorem by replacing the closed real interval $[a, b]$ with a compact set and the set of real-valued continuous functions with a real algebra that separates points and vanishes nowhere.

2. PRELIMINARIES: METRIC SPACE TOPOLOGY

In this section we introduce some background information on metric space topology which will help us define ideas about the Weierstrass approximation theorem.

Definition 2.1. (Metric space) A *metric space* is a pair (X, d) consisting of a set X along with a function d, also called the distance function or metric, which satisfies the following properties:

(1) *Positive definitiveness*: For $x, y \in X, d(x, y) \geq 0$ and $d(x, y) = 0$ only when $x = y$.
(2) *Symmetry*: For $x, y \in X, d(x, y) = d(y, x)$.
(3) *Triangle Inequality*: For $x, y, z \in X$, $d(x, z) \leq d(x, y) + d(y, z)$.

Proposition 2.2. *The set $\mathbb{R}$ of real numbers along with the distance function $d(x, y) = |x-y|$ is a metric space.*

Proof. The absolute value function on $x \in \mathbb{R}$ is defined as:

$$|x| = \begin{cases} -x & x < 0 \\ x & x > 0 \end{cases}$$

Thus, the first property holds. Further, $|x - y| = |y - x|$ because the distance between two points on the real line is independent of the starting point. The triangle inequality on $\mathbb{R}$ with the metric $d(x, y) = |x - y|$ will look like:

Date: July 2023.

$$|x - z| \leq |x - y| + |y - z|.$$

We can consider the numbers

$$a = x - y$$
$$b = y - z.$$

Then, we have

$$a + b = x - y + y - z = x - z$$

and our question reduces to showing that

$$|a + b| \leq |a| + |b|.$$

We have, $a + b \leq |a| + b \leq |a| + |b|$ and $-a - b \leq |a| - b \leq |a| + |b|$. From the definition of the absolute value function, it follows that the third property holds. □

Definition 2.3. (Open ball) The *open ball* of radius r centred at a point $p \in X$ is the set

$$B_r(p) = \{q \in X : d(p, q) < r\}$$

Example. The open ball of radius r centred around a point x in $\mathbb{R}$ equipped with the standard Euclidean metric $d(x, y) = |x - y|$ is the open interval $(x - r, x + r)$.

Proof. We have,

$$B_r(x) = \{y \in X : |x - y| < r\}.$$

Since $|x - y| < r$, we have $x - r < y < x + r$. □

Definition 2.4. (Neighbourhood) The *r-neighbourhood* of a point $p \in X$ is denoted by

$$M_r(p) = \{q \in X : d(p, q) < r\}$$

Remark 2.5. There seems to be no distinction between the definition of a neighbourhood and that of an open ball. However, a neighbourhood of $p \in X$ is a subset of X which contains a open ball containing x.

Definition 2.6. (Open set) A subset S of a metric space X is said to be *open* if for each point $x \in S$ there exist an *open ball* $B_r(x)$, such that $B_r(x) \subset S$.

Proposition 2.7. *Every open interval* (a, b) *in* $\mathbb{R}$ *is an open set.*

Proof. Choose a point $x \in (a, b)$. We consider $\epsilon = \min(x - a, b - x)$. Then $(x - \epsilon, x + \epsilon) \subset (a, b)$. Since x is an arbitrary, (a, b) is an open set.

□

Proposition 2.8. *The infinite union of open sets is open.*

Proof. Take $\mathcal{U}$ to be the union of open sets $U_1, U_2 \cdots, U_n$. Take $x \in \mathcal{U}$. By the definition of union of sets, x must lie in at least one of the U_i, which we know is open. Therefore, for some $r > 0$, we have $B_r(x) \in U_i$. This proves our theorem since x is arbitrary.

□

Definition 2.9. (Closed set) A subset S of a metric space X is said to be *closed* if its complement S^c is open.

Proposition 2.10. *Every closed interval $[a, b]$ in $\mathbb{R}$ is a closed set.*

Proof. $[a, b]^c = (-\infty, a) \cup (b, \infty)$. From the previous proposition we can say that $(-\infty, a)$, (b, ∞) are both open sets and the union of open sets is also open, so $[a, b]^c$ is an open set. Thus $[a, b]$ is a closed set. □

Definition 2.11. (Sequence in metric space) A sequence (a_n) in a metric space X is a function $a_n : \mathbb{N} \to X$.

Definition 2.12. (Convergent sequence) A sequence (a_n) in a metric space X is said to *converge* to a point a if for each $\epsilon > 0$ there exists some $N \in \mathbb{N}$ such that

$$n > N \implies d(a_n, a) < \epsilon.$$

The point a is also called the *limit point* of (a_n) and we write

$$a_n \to a.$$

Remark 2.13. If (a_n) is a convergent sequence converging to a, every neighbourhood of a contains an infinite number of elements of (a_n).

Proof. Let $\epsilon > 0$. Assume that the neighbourhood $M'_\epsilon(a)$ contains a finite number of elements of (a_n), say $a_1, a_2, ..., a_m$.

Let $\delta_1 = d(a, a_1)$, $\delta_2 = d(a, a_2)$,...$\delta_m = d(a, a_m)$.

let $\delta = min\{\delta_1, \delta_2, ..., \delta_m\}$. Then, for $\delta > 0$, $a_i \notin M_\delta(a)$ for $i = 1, 2, 3, ..., m,$. This contradicts the fact that a is a limit point of (a_n). □

Theorem 2.14. *A closed set contains all its limit points.*

Proof. Consider a closed set S and a convergent sequence $x_n \in S$. Assume, by contradiction, that the limit x of x_n lies in S^c, which, by Definition 2.9 is open. Now, by Remark 2.13, each neighbourhood of x contains infinite elements of $(x_n) \in S$, this contradicts the fact that S^c is open. □

Definition 2.15. (Pointwise convergence) A sequence of functions $f_n : [a, b] \to \mathbb{R}$ is said to converge pointwise to a limit function $f : [a, b] \to \mathbb{R}$ if for all $x \in [a, b]$ we have:

$$\lim_{x\to\infty} f_n(x) = f(x)$$

Definition 2.16. (Uniform convergence) A sequence of functions $f_n : [a, b] \to \mathbb{R}$ is said to converge uniformly to a limit function $f : [a, b] \to \mathbb{R}$ if for each $\epsilon > 0$ there exists an $N \in \mathbb{N}$ such that for all $n \geq N$ and for all $x \in [a, b]$,

$$|f_n(x) - f(x)| < \epsilon$$

Remark 2.17. Uniform convergence can be understood by visualising an ϵ-neighbourhood around the graph of the limit function. For large n, the graphs of $f_n(x)$ must lie completely inside the neighbourhood.

Definition 2.18. (Continuous function) The function $f : X \to Y$ is continuous if for each $\epsilon > 0$ and each $x, y \in X$ there exists a $\delta > 0$ such that

$$d_X(y, x) < \delta \implies d_Y(f(y), f(x)) < \epsilon.$$

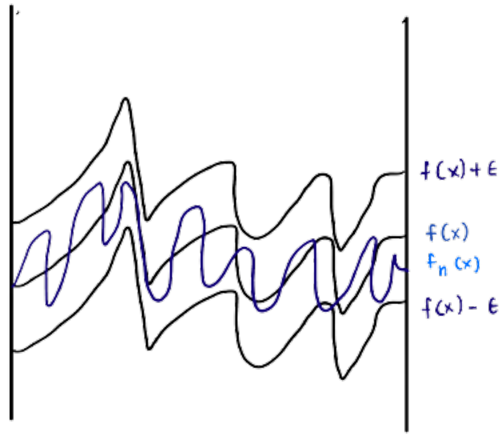

Figure 1. Uniform convergence

Definition 2.19. (Uniformly continuous function) The function $f : X \to Y$ is uniformly continuous if for each $\epsilon > 0$, there exists a $\delta > 0$ such that

$$d_X(y,x) < \delta \implies d_Y(f(y), f(x)) < \epsilon \ \forall x, y \in X$$

Remark 2.20. Notice that in a uniformly continuous function, the choice of δ depends only on ϵ and not on x and y.

Definition 2.21. (Boundedness) A subset S of a metric space (X, d) is said to be bounded if there exists $r > 0$ such that for all $(x, y) \in S$, $d(x, y) < r$.

Theorem 2.22. *Every continuous function on a closed and bounded interval of* $\mathbb{R}$ *is bounded.*

Proof. Let us assume that f is not bounded on $[a, b]$. Then for each $n \in \mathbb{N}$, there exists $x_n \in [a, b]$ such that $|f(x_n)| > n$. As $x_n \in [a, b]$, so $\{x_n\}$ is bounded, by Bolzano-Weierstrass Theorem, there exists a limit point of $\{x_n\}$, so there exists a subsequence of $\{x_{n_k}\}$ such that $x_{n_k} \to x$. Since $a \leq x_{n_k} \leq b$, we also have $a \leq x \leq b$.,i.e. $x \in [a, b]$. Since f is continuous on $[a, b]$, we must have $f(x_{n_k}) \to x$. But this is a contradiction since $|f((x_{n_k})| > n_k \geq k$, for $k \in \mathbb{N}$. □

Definition 2.23. (Covering) Let (X, d) be a metric space. A collection $\mathcal{U}$ of subsets of X *covers* $S \subset X$ if S is contained in the union of the sets belonging to $\mathcal{U}$. If all the sets in the covering of S are open $\mathcal{U}$ is said to be an *open cover* of S.

Definition 2.24. (Subcovering) If $\mathcal{V}$ and $\mathcal{U}$ both cover S and $\mathcal{V} \subset \mathcal{U}$, $\mathcal{V}$ is called a *subcovering* of S.

Definition 2.25. (Covering compact) If every open convering of S reduces to a finite subcovering of S, S is said to be *covering compact.*

Theorem 2.26. *Every continuous function on a compact metric space is uniformly continuous.*

Proof. Let (X, d) and (Y, d') be two metric spaces, where (X, d) is compact and suppose $f : X \to Y$ is a continuous function.

Let $\epsilon > 0$. Since f is continuous on at each point $x \in X$, then there is some $\delta_x > 0$ such that $f(B(x, \delta_x)) \subseteq B(f(x), \frac{\epsilon}{2})$.

Now $\{B(f(x), \frac{\epsilon}{2})\}_{x\in X}$ is an open cover of X, and since X is compact, there exist a finite subcover $\{B(x_i, \frac{\delta_{x_i}}{2})\}$ for $i = 1, 2, ..., n$. Now consider $\delta = \min_i(\frac{\delta_{x_i}}{2})$. Let $d(x, y) < \delta$. Since $x \in B(x_i, \frac{\delta_{x_i}}{2})$ for some i, we get $y \in B(x_i, \delta_{x_i})$. Thus

$$d(y, x_i) \leq d(y, x) + d(x, x_i) < \frac{\delta_{x_i}}{2} + \frac{\delta_{x_i}}{2} = \delta_{x_i}.$$

Finally if $d(x, y) < \delta$, then

$$d'(f(x), f(y)) \leq d'(f(x), f(x_i)) + d'(f(x_i), f(y)) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

□

3. Weierstrass Approximation Theorem: Proof using Bernstein Polynomials

In this section, we provide the statement of the Weierstrass approximation theorem using the Bernstein polynomials [Sur11]

Theorem 3.1. *(Weierstrass) Let $f \in C([a, b], \mathbb{R})$. Then there is a sequence of polynomials $p_n(x)$ that converges uniformly to $f(x)$ on $[a, b]$.*

In other words, $\forall f \in C^0$ and every $\epsilon > 0$, there is a polynomial function $p(x)$ such that $\forall x \in [a, b]$:

$$|f(x) - p(x)| < \epsilon.$$

The proof provided by Bernstein is by far the most well-known of all the proofs of the Approximation theorem since it is constructive (the sequence $p_n(x)$ is explicitly defined) and it relies only on the elementary properties of the Bernstein polynomials, which we shall see below.

Proof.

Theorem 3.2. *(Bernstein) Define the Bernstein polynomial $B_n f(x)$ of a function $f(x)$ by*

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

We have $B_n f \to f$ uniformly on $[0, 1]$.

Remark 3.3. Notice that proving Bernstein's theorem for the interval $[0, 1]$ is sufficient to prove the approximation theorem for any closed interval $[a, b]$. We can use the mapping

$$\phi(x) = (b - a)x + a.$$

Proof. We first consider a few properties of the Bernstein polynomials:

Lemma 3.4.

$$B_n(C) = C$$

where C is a constant polynomial.

Proof. Using binomial expansion, we have

$$B_n(C) = \sum_{k=0}^{n} C\binom{n}{k}x^k(1-x)^{n-k} = C(x+(1-x)^n = C\cdot 1^n = C.$$

□

Lemma 3.5.

$$B_n(f(x) - C) = B_n(f(x)) - C.$$

Proof. Follows directly from Lemma 3.4 and the definition of the Bernstein polynomial (replace $f\left(\frac{k}{n}\right)$ with $f\left(\frac{k}{n} - C\right)$). □

Lemma 3.6. *For $h \geq g$,*

$$B_n(h) \geq B_n(g).$$

Proof. $h \geq g$ implies that $h(\frac{k}{n}) \geq g(\frac{k}{n})$ for all k.

□

Lemma 3.7. *If $f > 0$,*

$$B_n(f) \geq 0$$

Proof. $f \geq 0$ implies that $f(\frac{k}{n}) \geq 0$ for all k.

□

First, we want to find an upper bound for $|f(x) - f(x_0)|$ where x_0 is fixed in $[0,1]$ and x is variable. By Theorem 2.26, $f(x)$ is uniformly continuous. By the definition of uniform continuity, we have for all $\epsilon > 0$.

$$|x - x_0| \leq \delta \implies |f(x) - f(x_0)| \leq \epsilon$$

for some δ which depends on ϵ.

Now, we fix an $\epsilon > 0$. We have 2 possibilities: $|x - x_0| \leq \delta$ and $|x - x_0| > \delta$. he first case is accounted for by the definition of uniform continuity, the bound for $|f(x) - f(x_0)|$ being ϵ. For the second case, we use Theorem 2.22 and take $M = sup\{f\}$. Using the Triangle Inequality on $|f(x) - f(x_0)|$

$$|f(x) - f(x_0)| \leq |f(x)| + |f(x_0)| \leq 2M.$$

Now, since $|x - x_0| > \delta$, $\frac{(x-x_0)^2}{\delta^2} > 1$. Thus,

$$|f(x) - f(x_0)| \leq |f(x)| + |f(x_0)| \leq 2M < 2M\frac{(x-x_0)^2}{\delta^2}.$$

Combining this with the case where $|x - x_0| \leq \delta$,

$$|f(x) - f(x_0)| < 2M\frac{(x-x_0)^2}{\delta^2} + \epsilon.$$

Now, we use our upper bound for $|f(x) - f(x_0)|$ to find an upper bound for $|B_n(f(x) - f(x_0))|$. By Lemma 3.6, we have $|B_n(f(x) - f(x_0))| \leq B_n(2M\frac{(x-x_0)^2}{\delta^2} + \epsilon)$. By Lemma 3.5 we have $B_n(2M\frac{(x-x_0)^2}{\delta^2} + \epsilon) = B_n(2M\frac{(x-x_0)^2}{\delta^2}) + \epsilon$. Now, $\frac{2M}{\delta^2}$ is a constant present in each term of the summation. Thus, we can factor it out so that $B_n(2M\frac{(x-x_0)^2}{\delta^2}) + \epsilon = \frac{2M}{\delta^2}B_n((x-x_0)^2) + \epsilon$. Finally, we have

$$|B_n(f(x)-f(x_0))| \leq \frac{2M}{\delta^2}B_n((x-x_0)^2)+\epsilon.$$

Simplifying the term $B_n((x-x_0)^2)=\sum_{k=0}^{n}\binom{n}{k}x^k(1-x)^{n-k}(\frac{k}{n}-x_0)^2$, we have, using some combinatorial identities and factorisation

$$\sum_{k=0}^{n}\binom{n}{k}x^k(1-x)^{n-k}(\frac{k}{n}-x_0)^2=(x-x_0)^2+\frac{1}{n}(x-x^2).$$

Now, we return to our main goal of finding a bound for $|B_n(f(x_0))-f(x_0)|$, the function and its Bernstein polynomial. By Lemma 3.4, we obtain that $|B_n(f(x))-f(x_0)|=|B_n(f(x)-f(x_0))|$ since $f(x_0)$ is a constant. We substitute $x=x_0$ and use $\max(x_0-x_0^2)=\frac{1}{4}$ in

$$|B_n(f(x))-f(x_0)| \leq \frac{2M}{\delta^2}(x-x_0)^2+\frac{2M}{\delta^2 n}(x-x^2)+\epsilon$$

to obtain

$$|B_n(f(x_0))-f(x_0)| \leq \frac{M}{2\delta^2 n}+\epsilon.$$

Now, we choose n sufficiently large for $\frac{M}{2\delta^2 n}<\epsilon$, so we have

$$|B_n(f(x_0))-f(x_0)|<2\epsilon.$$

Remark 3.8. This proves our theorem since ϵ is arbitrary.

Remark 3.9. Since n in $B_n(x)$ is depends only on ϵ from our definition of uniform continuity and M, the supremum of $f(x)$, and not on x, we have established the condition for uniform convergence.

□

Thus, we found that the Bernstein polynomial of a function f approximates it on a closed interval $[a,b]$.

□

4. Weierstrass's approximation theorem: Weierstrass' proof

Weierstrass published a self-contained proof of the approximation theorem [Hip13].

Theorem 4.1. *For a continuous real-valued bounded function $f:\mathbb{R}\to\mathbb{R}$ and $h>0$, we define the function $S_hf(x)$ by*

$$S_hf(x)=\int_{-\infty}^{\infty}\frac{1}{h\sqrt{\pi}}f(x)e^{-(\frac{u-x}{h})^2}du.$$

The sequence $S_hf(x)$ converges uniformly to $f(x)$ as $h\to\infty$.

Proof. The first part of this proof is similar to the proof of Bernstein's theorem. Based on the uniform continuity of f, we choose an $\epsilon>0$ which guarantees the existence of a $\delta>0$ such that for $|x-y|<\delta$ we have $|f(x)-f(y)|<\frac{\epsilon}{2}$ with $x,y\in\mathbb{R}$. Now, we take M to be the supremum of $f(x)$ so that we have $|f(x)|\leq M$ for all $x\in\mathbb{R}$.

Now, we have

$$\int_{-\infty}^{\infty}e^{-v^2}dv=\sqrt{\pi}.$$

This is known as the *Gaussian integral* and can be proved using polar coordinates or Laplace's method.

Using this, we have

$$\frac{1}{h\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-(\frac{u-x}{h})^2}du = 1.$$

This means we can write $f(x)$ as

$$\frac{1}{h\sqrt{\pi}}\int_{-\infty}^{\infty} f(x)e^{-(\frac{u-x}{h})^2}du.$$

Now, we choose an $h_0 > 0$ which satisfies $h < h_0 < \frac{e\delta\sqrt{\pi}}{2M}$. By the Triangle Inequality, we have

$$|S_hf(x) - f(x)| \leq |S_hf(x)| + |f(x)| = \frac{1}{h\sqrt{\pi}}\int_{-\infty}^{\infty}|f(u) - f(x)|e^{-(\frac{u-x}{h})^2}du.$$

Now, we just try and bound our expressions:

$$\frac{1}{h\sqrt{\pi}}\int_{-\infty}^{\infty}|f(u) - f(x)|e^{-(\frac{u-x}{h})^2}du \leq \ \frac{\epsilon}{2} + \frac{1}{h\sqrt{\pi}}\int_{|x-u|\geq\delta}|f(u) - f(x)|e^{-(\frac{u-x}{h})^2}du$$

By an application of the Triangle Inequality and using $M = sup\{f\}$,

$$\frac{\epsilon}{2} + \frac{1}{h\sqrt{\pi}}\int_{|x-u|\geq\delta}|f(u) - f(x)|e^{-(\frac{u-x}{h})^2}du \leq \ \frac{\epsilon}{2} + \ \frac{1}{h\sqrt{\pi}}\int_{|x-u|\geq\delta}||f(u)| + \ |f(x)||e^{-(\frac{u-x}{h})^2}du$$

$$\leq \ \frac{\epsilon}{2} + \frac{2M}{h\sqrt{\pi}}\int_{|x-u|\geq\delta} e^{-(\frac{u-x}{h})^2}du.$$

We can take $\frac{|x-u|}{h} = v$ in our expression to get

$$\frac{\epsilon}{2} + \frac{2M}{h\sqrt{\pi}}\int_{|x-u|\geq\delta} e^{-(\frac{u-x}{h})^2}du = \frac{\epsilon}{2} + \frac{2M}{\sqrt{\pi}}\int_{|v|\geq\frac{\delta}{h}} e^{-v^2}dv \leq \frac{2Mh}{\delta\sqrt{\pi}}\int_{|v|\geq\frac{\delta}{h}} |v|e^{-v^2}dv.$$

Since $|v|(\frac{h}{\delta}) \geq 1$,

$$\frac{2Mh}{\delta\sqrt{\pi}}\int_{|v|\geq\frac{\delta}{h}} |v|e^{-v^2}dv \leq \frac{\epsilon}{2} + \frac{4Mh}{\delta\sqrt{\pi}}\int_0^{\infty} ve^{-v^2}dv = \frac{\epsilon}{2} + \frac{2Mh}{\delta\sqrt{\pi}}.$$

Since we took $h < h_0 < \frac{e\delta\sqrt{\pi}}{2M}$,

$$\frac{\epsilon}{2} + \frac{2Mh}{\delta\sqrt{\pi}} < \epsilon + \left(\frac{M}{\delta\sqrt{\pi}}\right)\left(\frac{\epsilon\delta\sqrt{\pi}}{4M}\right) = \epsilon$$

This proves our theorem since ϵ is arbitrary.

□

Now we examine Weierstrass' proof of the approximation theorem using Theorem 4.1.

Theorem 4.2. *Let $f \in C([a,b],\mathbb{R})$. Then there is a sequence of polynomials $p_n(x)$ that converges uniformly to $f(x)$ on $[a,b]$.*

Proof. We consider $f \in C([a,b],\mathbb{R})$. We extend f to a bounded uniformly continuous function from $\mathbb{R} \to \mathbb{R}$, which we also call f. We give a piecewise definition of f as follows:

$$f(x) = \begin{cases} f(a)(x-a+1) & x \in [a-1,a) \\ -f(b)(x-b-1) & x \in (b,b+1] \\ 0 & x \notin [a-1,b+1] \end{cases}$$

There must exist a $J > 0$ such that $f(x) = 0$ for all $|x| > J$. We fix an $\epsilon > 0$ and observe that there must exist an M such that $|f(x)| < M$ for all $x \in \mathbb{R}$. By Theorem 4.1 there exists an $h_0 > 0$ such that for all $x \in \mathbb{R}$

$$\left| \frac{1}{h_0\sqrt{\pi}} \int_{-\infty}^{\infty} f(u) e^{-(\frac{u-x}{h_0})^2} du - f(x) \right| < \frac{\epsilon}{2}.$$

For $|u| > J$, $f(u) = 0$. Thus, we have

$$\left| \frac{1}{h_0\sqrt{\pi}} \int_{-J}^{J} f(u) e^{-(\frac{u-x}{h_0})^2} du - f(x) \right| < \frac{\epsilon}{2}.$$

We know that the power series of e^{-v^2} converges uniformly on the interval $[-\frac{2J}{h_0}, \frac{2J}{h_0}]$. Thus, there exists an N such that

$$\left| \frac{1}{h_0\sqrt{\pi}} e^{-(\frac{u-x}{h_0})^2} - \frac{1}{h_0\sqrt{\pi}} \sum_{k=0}^{N} \frac{(-1)^k}{k!} \left(\frac{u-x}{h_0} \right)^{2k} \right| < \frac{\epsilon}{4JM}$$

for all $|x| < J$ and for all $|u| < J$, which would imply $|u-x| < 2J$ (triangle inequality). Thus, we have

$$\left| \frac{1}{h_0\sqrt{\pi}} \int_{-J}^{J} f(u) e^{-(\frac{u-x}{h_0})^2} du - \frac{1}{h_0\sqrt{\pi}} \int_{-J}^{J} f(u) \sum_{k=0}^{N} \frac{(-1)^k}{k!} \left(\frac{u-x}{h_0} \right)^{2k} \right|$$

for all $|x| \leq J$.

We define the function $p(x)$ as:

$$p(x) = \frac{1}{h_0\sqrt{\pi}} \int_{-J}^{J} f(u) \sum_{k=0}^{N} \frac{(-1)^k}{k!} \left(\frac{u-x}{h_0} \right)^{2k}$$

which is a polynomial of degree at most $2N$ such that

$$\left| \frac{1}{h_0\sqrt{\pi}} e^{-(\frac{u-x}{h_0})^2} - p(x) \right| < \frac{\epsilon}{2}.$$

for all $|x| < J$. Thus, we have

$$|f(x) - p(x)| < \epsilon$$

for all $x \in [a,b]$.

□

5. Topology

Definition 5.1. (Topological space) Let X be a set. A topology $\mathcal{T}$ on X is a collection of subsets of X that satisfies the conditions below:

(1) $\mathcal{T}$ contains the set X along with the empty set ϕ: $X, \phi \in \mathcal{T}$
(2) The infinite union of subsets of X is in $\mathcal{T}$: If $U_i \in \mathcal{T}$, $\cup U_i \in \mathcal{T}$.
(3) The finite intersection of subsets of X is in $\mathcal{T}$: If $U_i \in \mathcal{T}$, $\cap U_i \in \mathcal{T}$ for finite i.

The pair $(X, \mathcal{T})$ consisting of X and a collection of its subsets $\mathcal{T}$ satisfying the properties listed above is called a *topological space*.

Definition 5.2. (Continuous) A function $X \to Y$ from a topological space X to a topological space Y is said to be *continuous* if the inverse image of an open set U in Y is open in X.

6. Algebras

Lemma 6.1. *The function $|x|$ on a closed interval $[a, b]$ is a limit point of all polynomials P satisfying $P(0) = 0$ on $[a, b]$.*

Proof. By the Weierstrass Approximation theorem, there exists a polynomial P_N such that

$$P_N(x) - |x| < \epsilon$$

for all x in $[a, b]$.

□

Definition 6.2. (Algebra) A family $\mathcal{A}$ of complex functions defined on a set X is said to be an *algebra* if the following properties are satisfied for all $f, g \in \mathcal{A}$ and all complex constants c:

(1) $f + g \in \mathcal{A}$
(2) $fg \in \mathcal{A}$
(3) $cf \in \mathcal{A}$

The properties above are called closure under addition, closure under multiplication, and closure under scalar multiplication respectively.

Definition 6.3. (Uniformly closed algebra) An algebra $\mathcal{A}$ is said to be *uniformly closed* if it has the property that $f \in \mathcal{A}$ whenever $f_n \in \mathcal{A}$ and $f_n \to f$ uniformly on X.

Definition 6.4. (Uniform closure) The set $\mathcal{B}$ of all functions which are limits of uniformly convergent sequences of members of $\mathcal{A}$ is called the *uniform closure* of $\mathcal{A}$.

Remark 6.5. The Weierstrass Approximation theorem restated using this terminology as follows:

The set of all continuous functions on $[a, b]$ is the uniform closure of the set of all polynomials on $[a, b]$.

Theorem 6.6. *The uniform closure of an algebra of bounded functions is a uniformly closed algebra.*

Proof. Take $\mathcal{B}$ to be the uniform closure of an algebra $\mathcal{A}$ of bounded functions. By the definition of uniform closure, we have that for functions $g, h \in \mathcal{B}$, there exist sequences g_n and h_n in $\mathcal{A}$ of functions such that $h_n \to h$ and $g_n \to g$. By the properties of limits, we have

$$\begin{aligned} h_n + g_n &\to h + g \\ h_n g_n &\to hg \\ ch_n &\to ch. \end{aligned}$$

Now, since $h+g, hg$ and ch all lie in $\mathcal{B}$, it satisfies the condition for a uniformly closed algebra.

□

Definition 6.7. (Separation of points) A family $\mathcal{A}$ of functions on a set X is said to *separate* points on X if there exists a function f for each pair of distinct $x, y \in X$ such that $f(x) \neq f(y)$.

Definition 6.8. (Vanishing) If there exists a function $f \in \mathcal{A}$ for each $x \in X$ such that $f(x) \neq 0$, $\mathcal{A}$ is said to *vanish at no point of* X.

Theorem 6.9. *Let $\mathcal{A}$ be an algebra of functions on a set X which separates points on X and vanishes at no point on X. Let x and y be distinct points of X and let c_1 and c_2 be constants. Then $\mathcal{A}$ contains a function f such that:*

$$f(x) = c_1, f(y) = c_2.$$

Proof. We have assumed that $\mathcal{A}$ separates points on X, so we have

$$g(x) \neq g(y).$$

Also, we presume that $\mathcal{A}$ vanishes at no point on X, so we have

$$h(x) \neq 0, k(y) \neq 0.$$

We introduce u and v such that

$$u = gk - g(x)k, v = gh - g(y)h.$$

Since $\mathcal{A}$ is an algebra, u and v are in $\mathcal{A}$. Further, we have that $u(x) = gk(x) - g(x)k = 0$ and $v(y) = gh(y) - g(y)h = 0$. Also, $u(y) = gk(y) - g(x)k \neq 0$ since $g(x) \neq g(y)$. By a similar argument, we have $v(x) \neq 0$. Using these functions and the constants c_1 and c_2 we construct a function f defined as

$$f = \frac{c_1 v}{v(x)} + \frac{c_2 u}{u(y)}$$

which meets our requirements.

□

Theorem 6.10. *(Stone Weierstrass theorem on real algebras) Let $\mathcal{A}$ be an algebra of real continuous functions on a compact set K. If $\mathcal{A}$ separates points of K and vanishes at no point of K then its uniform closure consists of all real continuous functions on K [Nic82].*

Proof. The proof of this theorem requires a considerable amount of set-up and will use some of the theorems, lemmas and definitions we have already covered. Let $\mathcal{B}$ be the uniform closure of $\mathcal{A}$. First, we must establish the following:

Lemma 6.11. *$g \in \mathcal{B}$ implies that $|g| \in \mathcal{B}$*

Proof. First, we consider $g \in \mathcal{B}$. For a fixed $\epsilon > 0$, we can consider the open ball $B(|\cdot|, \epsilon) \in C([a,b], \mathbb{R})$. We take a and b to be the supremum of $-g$ and g respectively. For this open ball, we can find a function that is ϵ away from a given function if we select a small enough ϵ.
Now, we are guaranteed by Lemma 6.1, the existence of a polynomial P_n such that $P_n(0) = 0$ and $P_n \subset B(|\cdot|, \epsilon)$. We can express P_n as $\sum_{i=1}^{n} c_i g^i$ for $n \in \mathbb{N}$ and $c_i \in \mathbb{R}$. By Theorem 6.6, $\mathcal{B}$ is a closed algebra containing g, so it must also contain g^i for all i since an algebra must be closed under multiplication. Thus, all polynomials of the form $\sum_{i=1}^{n} c_i g^i$ are contained in $\mathcal{B}$.
Now, since $P_n \subset B(|\cdot|, \epsilon)$, $P_n(g) \subset B(|g|, \epsilon)$. We therefore have that $|g|$ is a limit point of $\mathcal{B}$, which is uniformly closed. By Theorem 2.14, $\mathcal{B}$ must contain all its limit points, forcing $|g| \in \mathcal{B}$.

□

Lemma 6.12. *$g, h \in \mathcal{B}$ implies max(g,h), min$(g,h) \in \mathcal{B}$*

Proof. First, we have $\max(g,h) = \frac{1}{2}(g + h + |g - h|)$ and $\min(g,h) = \frac{1}{2}(g + h - |g - h|)$, which can be obtained by some simple algebraic manipulation. By Lemma 6.11 we know that these expressions must be contained in $\mathcal{B}$.

Remark 6.13. This result can be extended by iteration to any finite set of functions $f_1, \cdots, f_n \in \mathcal{B}$.

□

Now, we choose a function $f \in (C, K)$ and fix $\epsilon > 0$. We can consider the open ball $B(f, \epsilon)$, for which f is a limit point. Thus, $B(f, \epsilon) \subset (C, K)$. By Theorem 6.9, we have that for each $x \in K$, there exists a function $g_x \in \mathcal{B}$ such that $g_x(x) = f(x)$.
Now, we introduce a function $h_x = \max(g_x, f(x))$. We know that h_x is continuous since it can be expressed as the sum of continuous functions. By the topological definition of continuity (check definition 5.2), there exists an open set $U_x \in K$ such that $x \in U_x$ and $U_x \subset h_x^{-1}(B(f(x)), \epsilon)$.
Since this is true for all $x \in K$, the collection of all open sets U_x is an open cover of K. Since K is compact, a finite number of these U_x, U_j where $j < x$ must cover K. For these U_j, we have that $U_j \subset B(f(x_i), \epsilon)$.
Now, we take the minimum of all h_j. By the definition of h_x, we have, for all x,

$$h_j(x) > f(x) - \epsilon.$$

Thus, we have $\min(h_i(x)) \in (f(x) - \epsilon, f(x) + \epsilon)$ which gives us that $\min(h_i) \in B(f, \epsilon)$. By Lemma 6.12 we know that $\min(h_i(x)) \in \mathcal{B}$. Therefore, we have that f is a limit point of $\mathcal{B}$. Since $\mathcal{B}$ is closed, $f \in \mathcal{B}$, proving that $\mathcal{B}$ contains all real continuous functions on K. □

Unfortunately, the same argument cannot be extended to complex algebras. Thus, we introduce an additional condition on the complex algebra which will enable us to prove the Stone-Weierstrass theorem on complex algebras.

Definition 6.14. (Self-adjoint algebra) A *self-adjoint algebra* has all the properties of an algebra (closure under addition, multiplication, and scalar multiplication) with an additional property that whenever a function $f \in \mathcal{A}$, its complex conjugate $\bar{f} \in \mathcal{A}$.

Theorem 6.15. *(Stone Weierstrass theorem on complex algebras) Let $\mathcal{A}$ be a self-adjoint algebra of complex continuous functions defined on a compact set K. If $\mathcal{A}$ separates points on K and vanishes at no point of K then its uniform closure consists of all complex continuous functions on K.*

Proof. Let $\mathcal{A}_R$ be the set of all real functions on K belonging to $\mathcal{A}$. Take a function $f = u+iv$ $(i = \sqrt{-1})$ in $\mathcal{A}$ with $u, v \in \mathbb{R}$. Now, we have

$$2u = f + \bar{f}.$$

Since $\mathcal{A}$ is self adjoint, $u \in \mathcal{A}_R$. By a similar argument, we have $v \in \mathcal{A}_R$.

Now, for $x_1, x_2 \in K$ we have the existence of a function $f \in \mathcal{A}$ such that $f(x_1) = 1, f(x_2) = 0$ by Theorem 6.9. Hence, $u(x_2) = 0 \neq u(x_1) = 1$, establishing that $\mathcal{A}_R$ separates points on K.

Since $\mathcal{A}$ vanishes at no point of K, for each $x \in K$ there exists $g \in \mathcal{A}$ such that $g(x) \neq 0$. Considering both the real and imaginary components of f, we must have a function in $\mathcal{A}_R$ such that both these components are nonzero. Thus, we have that $\mathcal{A}_R$ vanishes at no point of K.

Now our algebra A_R fulfills all the conditions of 6.10. It follows that every real continuous function on K lies in the uniform closure of A_R, completing the proof.

□

7. Acknowledgements

The author would like to acknowledge Dr. Simon Rubinstein-Salzedo, Lisa Liu and everyone who provided feedback on early drafts of the paper for their invaluable support.

References

[Hip13] LaRita Barnwell Hipp. *The Weierstrass Approximation Theorem*. PhD thesis, University of South Carolina, 2013.

[Nic82] Shirley Jo Nichols. *Polynomial Approximation: The Weierstrass Approximation Theorem*. University of Hawai'i at Manoa, 1982.

[Sur11] B Sury. Weierstrass's theorem—leaving no 'stone'unturned. *Resonance*, 16:341–355, 2011.