Stone-Weierstrass theorem and its forms on different structures

Zipeng Lin

July 11, 2022

In this expository paper, we will discuss the Weierstrass theorem and its derivative, the Stone-Weierstrass theorem. We will also take a closer look at how the Stone-Weierstrass theorem works on other algebraic structures, and what results it brings on those certain structures.

1 Introduction

There are two versions of the Weierstrass theorem, the first being the Weierstrass approximation theorem, stating that every continuous function defined on a closed interval can be approximated as closely as wanted by a polynomial function. The second version, called Stone-Weierstrass theorem, extends the interval to be any compact set. The theorem gives insight on how some sets are dense in some other sets.

The paper will proceed in the following structure: after introducing the history of the theorem after giving some basic notations, we prove the original Weierstrass theorem. After that, we will give a crash course on topology and algebra to prove the Stone-Weierstrass theorem. Then, a few extension and applications will be introduced.

2 Notation

Definition 1 (Fields)**.** Denote $\mathbb{R}$ to be the field of real numbers, $\mathbb{C}$ to be the field of complex numbers.

Definition 2 (Open balls)**.** Denote $B(x, \epsilon)$ to be a open ball centered at x with radius ϵ.

Definition 3 (Function norm)**.** Denote $\|f\|$, where f is a function, to be the supremum of function f with the bound in the context.

Example 4 (Function norm of a linear function)**.** The function norm of the function $y = x$ in the interval $[0, 1]$ is 1.

3 Background of the theorem

Here is the original statement of the Weierstrass theorem, discovered by Karl Weierstrass in 1937 [You06]

Theorem (Original Weierstrass theorem)**.** Suppose f is a continuous complex-valued function defined on the real interval $[a, b]$.

For every $\epsilon > 0$, there exists a polynomial function p over $\mathbb{C}$ such that for all $x \in [a, b]$, we have $|f(x) - p(x)| < \epsilon$

The original theorem, also called Weierstrass approximation theorem, has impact on math analysis [Hip13] since polynomials are the most simplest functions. Notice that it is more direct than Taylor expansion in the sense that the error term (or the difference) could be directly represented by a ϵ.

Marshall H. Stone first generalized the theorem and simplified the proof in 1948 [Hip13]. There are two new changes here: first instead of the real interval $[a, b]$, we can have an arbitrary compact space (rigorously, it is compact Hausdorff space), also the polynomial functions could be replaced by more general algebras.

4 Proof of the approximation theorem

Before introducing the general theorem, we focus on the Weierstrass Approximation theorem first. This is one of the version of the Stone Weierstrass theorem that only focuses on the real space.

Definition 5 (Collection of continuous functions)**.** Denote $C(S, \mathbb{S})$ to be a collection of continuous functions on $\mathbb{S}$ and in the interval S.

Theorem 6 (Weierstrass Approximation Theorem (1885))**.** 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, we want to use a polynomial to approximate the function **as close as possible**, and that means the difference of those two functions should be small enough, which we will figure out later.

In order to approximate a function, the idea of using a power series comes to mind. The following will define a kind of polynomials that are crucial in our proof.

4.1 Bernstein polynomial

Definition 7 (Bernstein polynomial)**.** Take f to be a continuous function on $[0, 1]$, define the nth Bernstein polynomial by

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

We want to prove that this power series could approximate f well.

Lemma 8. For any $x \in \mathbb{R}$ and $n \geq 0$,

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

Proof. We have the left hand side is the expansion of $(x + (1-x))^n$, which is 1. ☐

The reason why we pick Bernstein polynomial is that it includes so many binomial coefficients, making it easy to bound the value of polynomials and calculate the values. Besides, the bigger the function gets in the Bernstein polynomial, the bigger the value becomes. Finally, Bernstein polynomial gives good properties in deal with differences.

Now we try to prove the approximation theorem first. Notice that without loss of generality, we can prove the theorem first on interval $[0,1]$, which is also a **compact** (will define later rigorously) interval. We have function f being continuous implies that it is continuous uniformly.

Definition 9 (Regular continuity)**.** If f is continuous, then for every $x \in dom(f)$ and every $\epsilon > 0$, there exists one $\delta > 0$ such that $|y - x| < \delta \implies |f(y) - f(x)| < \epsilon$

Definition 10 (Uniform continuity)**.** A function f is called uniformly continuous if for every $\epsilon > 0$, there exists δ such that for $x, y \in dom(f)$, $|y - x| < \delta \implies |f(y) - f(x)| < \epsilon$

Remark 11. Notice how regular continuity define different δ for different x, while uniform continuity does not.

Theorem 12. A continuous function is uniformly continuous at a closed interval.

Proof on the Approximation Theorem. So, by the definition, for every $\epsilon > 0$ there exists $\delta > 0$ such that

$$|x - y| \leq \delta \implies |f(x) - f(y)| \leq \frac{\epsilon}{2} \forall x, y \in [0,1]$$

Now, define M to be the "norm" of function, which is the maximum value of f taking under the interval of $[0,1]$. That is to say,

$$M = \|f\|_\infty = \max_{x \in [0,1]} |f(x)|$$

Since the interval $[0,1]$ is closed (in the normal sense, not necessarily the topological one) and bounded, we have M, the norm of the function, exists since the interval is closed. Now we fix a point $e \in [0,1]$. Then, if $|x - e| \leq \delta$, $|f(x) - f(e)| \leq \frac{\epsilon}{2}$ by continuity at e.

Also, if we have $|x-e| \geq \delta$, then

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

which simplifies to

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

According to 8, we have $B_n(x,1)=1$, so

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

Therefore, we have

$$\begin{aligned}
|(B_n(x,f)-f(e))| &= |B_n(x,f-f(e))| \\
&= |B_n(x,|f(x)-f(e)|| \\
&= \left|B_n(x,2M\left(\frac{x-\epsilon}{\delta}\right)\right|^2+\epsilon/2 \\
&\leq B_n(x,2M(\frac{x-\epsilon}{\delta})^2+\frac{\epsilon}{2}) \\
&= \frac{2M}{\delta^2}B_n(x,(x-e)^2)+\frac{\epsilon}{2}
\end{aligned}$$

where in the second step we have $B_n(x,a) \leq B_n(x,b)$ if $a \leq b$.

Since we have

$$\begin{aligned}
B_n(x,(x-e)^2) &= \sum_{k=0}^{n}(\frac{k}{n}-e)^2\binom{n}{k}x^k(1-x)^{n-k} && (1)\\
&= \sum_{k=0}^{n}(\frac{k^2}{n^2}-\frac{2k}{n}e+e^2)\binom{n}{k}x^k(1-x)^{n-k} && (2)\\
&= x^2\sum_{k=1}^{n}\frac{n-1}{n}\binom{n-2}{k-2}x^{k-2}(1-x)^{n-k}+x\sum_{k=1}^{n}(\frac{1}{n}-2e)\binom{n-1}{k-1}x^{k-1}(1-x)^{n-k}+e^2 && (3)\\
&= (1-\frac{1}{n})x^2+\frac{1}{n}x-2ex+e^2 && (4)
\end{aligned}$$

after applying binomial theorem.

Then, we have

$$|B_n(x,f)-f(e)| \le \frac{\epsilon}{2} + \frac{2M}{\delta^2}(x-e)^2 + \frac{2M}{\delta^2}\frac{1}{n}(x-x^2)$$

we then have

$$|B_n(e,f)-f(e)| \le \frac{\epsilon}{2} + \frac{2M}{\delta^2}\frac{1}{n}(e-e^2) \le \frac{\epsilon}{2} + \frac{M}{2\delta^2 n}$$

after realizing the maximum of $e-e^2$ is $\frac{1}{4}$. Therefore, if we take N to be no smaller than $\frac{M}{2\delta^2\epsilon}$ and for $n \ge N$,

$$\|B_n(e,f)-f(e)\|_\infty \le \epsilon$$

which means we have found the polynomial $B_n(e,f)$ that approximates f on $[0,1]$.

Now we can map the interval $[0,1]$ to $[a,b]$ by using function $\phi : x \to (b-a)x - a$ and composite that over the function we have found.

□

5 Extension to the topological space

5.1 Knowledge on topology

Now we will define the notion of metric space and some concepts relevant to it. Intuitively, metric space is a set of elements with distance defined among those elements. The distance has to conform with geometric properties, so the distance will follow a few axioms as below.

To understand the following definitions better, it would help to draw diagrams on a piece of paper.

Definition 13 (Metric and metric space)**.** Let X be a set. A metric is a function $d : X \times X \to \mathbb{R}$ such that:

(a) d is symmetric: $d(x,y) = d(y,x)$ for all x, y in X.

(b) d is positive definite: $d(x,y) \ge 0$ for all x, y in X, and $d(x,y) = 0$ if and only if $x = y$.

(c) The triangle inequality is sufficed: $d(x,y) \le d(x,z) + d(z,y)$ for every $x, y, zz \in X$

and we call a set equipped with metric d as a **metric space**

Example 14 (Metric space in Cartesian coordinate)**.** If we have X are all the points of the form (x, y) in the Cartesian coordinate system, then we have $d(a, b)$ as the distance between two points in the coordinate is a valid metric.

Example 15 (Discrete metric space)**.** Define $d(a, b)$ as 1 if $a = b$, otherwise 0 is a valid metric. The metric space formed by this metric is called the **discrete metric space**.

Now we will define the notion of open ball, which is a set of points whose distance to a fixed point is in a range.

Definition 16 (Open ball)**.** Let X be a set with a metric d and let x be a point in X. An open ball of radius $\epsilon > 0$ centered at x, denoted $B(x, \epsilon)$, is the set $\{y \in X | d(x, y) < \epsilon\}$

Example 17 (An open ball in the metric space for Cartesian coordinate)**.** An open ball with center $(0, 0)$ and radius 1 is the graph of function $(x - 1)^2 + (y - 1)^2 \leq 1$ in the metric space 14

Lemma 18 (Existence of open ball in an arbitrary open ball)**.** Let $B(x, \epsilon)$ be a ball of radius ϵ centered at x. Then, for all y in $B(x, \epsilon)$, there exists an open ball centered at y, $B(y, \delta)$, such that $B(y, \delta) \subset B(x, \epsilon)$

Proof. Define $\delta = \epsilon - d(x, y)$, this is positive because $d(x, y) < \epsilon$, since x in in the ball of radius ϵ centered at y.

Now consider any $z \in B(y, \delta)$. Then $d(y, z) < \epsilon - d(x, y)$, so $d(y, z) + d(x, y) < \epsilon$. By the triangle inequality, $d(x, z) \leq d(y, z) + d(x, y) < \epsilon$, so we have $z \in B(x, \epsilon)$ □

Remark 19 (Intuition)**.** The results of this lemma tells us that there can be an open ball inside an open ball.

Now we will define some points in the set and open/closed set. One can treat open set being a set with no bounds, so each point in it has a circle around it. For a closed set, every circle around every point would intersect the closed set and produce another point.

Definition 20 (Interior point)**.** For a set X and one of its subsets E, $x \in E$ is an interior point of E if there exists a positive real number $r > 0$ such that $B(x, r) \subset E$

Example 21 (Interior point on (0, 1))**.** 0.5 is a interior point of $(0, 1)$ if we consider the metric in a 1d space. The $r = 0.1$ works.

Definition 22 (Open set)**.** A set $E \subset X$ for a set X is open if every point $x \in E$ is an interior point of E. That is to say, for every point $x \in E$ in X, there exists a positive real number $\epsilon > 0$ such that $B(x, \epsilon) \subset E$

Definition 23 (Limit point)**.** Let X be an arbitrary topological space and let A be a subset of X. A point x in X is a **limit point** of A if for all open sets U containing x, U and A have at least one point other than x in common. That is to say, $U \cap A - \{x\} \neq \emptyset$.

The following would be important in forming a closed set by completing the set.

Definition 24 (Closure)**.** Let A be a set in an arbitrary topological space, X. The closure of A, denoted $\hat{A}$, is A together with its limits points.

Definition 25 (Closed set)**.** A set $E \subset X$ of a set X is closed if E contains all its limit points. That is to say, a set is closed if its closure is equivalent to itself.

Example 26 (Examples of open/closed set)**.**

- $(0, 1)$ is open.
- $[0, 1]$ is closed.
- $\emptyset$ is not open nor closed.
- $Q \subset R$ is neither open or closed.

The following definition defines continuity for a function, which means the function is smooth enough, and we can find points within a certain difference in the neighbourhood.

Definition 27 (Continuous at points)**.** A function $f : X \to Y$ is said to be continuous if for every $x \in X$, for every $\epsilon > 0$, there exists a δ for x such that we have for $x_2 \in X$, $|x - x_2| < \delta \to |f(x) - f(x_2)| < \epsilon$

Definition 28 (Uniformly continuous)**.** A function $f : X \to Y$ is said to be continuous if for every $x_1, x_2 \in X$, for every $\epsilon > 0$, there exists a δ for such that we have for $|x_1 - x_2| < \delta \to |f(x_1) - f(x_2)| < \epsilon$

Definition 29 (Continuous in topology definition)**.** A function $f : X \to Y$ is said to be continuous if the image of every open subset of Y is open in X. In other words, if $V \in Y_a$, then its inverse image in $f^{-1}(Y) \in X_a$, where both Y_a and X_a are open sets.

The following sequence is a stronger sequence than a normal sequence that converges, since adjacent terms would also have limited difference.

Definition 30 (Cauchy sequence)**.** Let X be a metric space. A Cauchy sequence is a function $a : \mathbb{N} \to X$ (mapping the natural number indices to X) such that for all $\epsilon > 0$, there exists a natural number N such that if $m, n \geq N$, $d(a(m), a(n)) < \epsilon$.

Definition 31 (Complete)**.** A metric space X is complete if every Cauchy sequence with indices in natural number and elements all in X converges to a point in X.

Lemma 32 (Closer subset also induces a complete metric space)**.** Let X be a complete metric space with metric d. Then, if K is a closed subset of X, the metric space with K and metric d is a complete metric space.

Proof. Suppose (k_n) is a Cauchy sequence in K. Since X is complete, we have (k_n) converges to some x in X.

Since K is closed, K contains all of its limit points. Thus x is in K, which means K is complete. □

Definition 33 (Open cover)**.** An **open cover** for E is a collection of open sets $\{u_\alpha\}_{\alpha\in A}$ such that $E \subset \cup(u_\alpha)_{\alpha\in A}$

Definition 34 (Subcover)**.** A **subcover** of an open cover $\{u_\alpha\}_{\alpha\in A}$ of E is an open cover $\{u_\alpha\}_{\alpha\in B}$ such that $B \subset A$. It is just a sub-collection of the open sets in a open cover.

Definition 35 (Finite open cover)**.** A open cover is **finite** if it contains finitely many sets. In other words, the cardinality of A is finite.

Definition 36 (Compact set)**.** A set $E \subseteq X$ is **compact** if every open cover has finite subcover.

Remark 37 (What compact means in topology)**.** Compact means finite, but in a topology sense. Therefore, it is comparable to treat a compact interval as a finite interval, just as we later can see in the generalized stone theorem.

Lemma 38 (The function of absolute value is a limit point)**.** The absolute value function, $|x|$, on a closed interval $[a, b]$ is a limit point of all polynomials P such that $P(0) = 0$ on $[a, b]$

Proof. By Weierstrass' Approximation theorem 6, there exists a polynomial P_N for P that is $\frac{\epsilon}{2}$ from the absolute value function.

Now, let $P_N^{'} = P_N - P_N(0)$. Then, for all x in $[a, b]$, $|P_N(x) - |x|| \leq \frac{\epsilon}{2}$.

Then, by triangle inequality, we have

$$\epsilon > |P_N(x) - |x|| + ||0| - P_N(0)| \geq |P_N(x) - P_N(0) - |x||$$

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

Therefore, we have $\|P_N - P_N(0) - |x|\|_\infty \leq \epsilon$, which means the function $P_N^{'}$ suffices that it takes value 0 at 0 and is ϵ close to the absolute value function. □

5.2 Knowledge on algebra

Definition 39 (Algebras)**.** A collection of complex functions (functions that work on complex numbers too) on a set X called $\mathbb{A}$, is an algebra if for all functions f and g in $\mathbb{A}$ and all complex constants c, $f + g, fg$ are in A.

Example 40 (Common examples of algebra)**.** All the real or complex polynomials over a set is an algebra, since the polynomials are closed in addition, multiplication and multiplication with a constant.

Theorem 41 (Closure of an algebra is a closed algebra)**.** Let $\overline{A}$ be the closure of an algebra A of bounded functions. Then $\overline{A}$ is a closed algebra.

Proof. Let f and g be in $\overline{A}$, and let c be a complex scalar. First we prove that $f+g$ is in $\overline{A}$ for any two functions f, g in $\overline{A}$.

First we have $+$, as a function, is continuous, so for all open sets U that contains $f+g$, there exists open balls $B(f,\epsilon)$ and $B(g,\delta)$ such that

$$B(f,\epsilon) \times B(g,\delta) \subset +^{-1}(U)$$

here $+^{-1}$ could be treated as the inverse function of $+$ over U, which results in a pair of sets that sum to a subset of U. So the sum of two open balls will be in U.

Then, since f and g are both in $\overline{A}$, a closed set, there exists functions h and j such that (h,j) is in $B(f,\epsilon) \times B(g,\delta)$

Therefore, $h+j$ is in $U \cap A$, so for all open sets U containing $f+g$, there is an element of A in U. Therefore, $f+g$ is a limit point of A by definition, so $f+g$ is in $\overline{A}$.

By considering the inverse function of $*$ and considering the complex constants, the proof for fg and cf are equivalent to what has been written above. □

Example 42 (Examples of algebra)**.** (a) The collection of real-valued functions on $\mathbb{C}$ or $\mathbb{R}$ is an algebra.

(b) The collection of differentiable functions on a open interval is an algebra.

(c) Polynomial functions on $\mathbb{C}$ or $\mathbb{R}$ form an algebra.

(d) Let $K = \{zz \in \mathbb{C} : |z| \leq 1\} = \left\{e^{ix} : x \in [0, 2\pi\right\}$. Then K is a compact set of $\mathbb{C}$ (since we can use a circle to cover it). Then all functions of the form

$$\sum_{n=-k}^{k} c_n e^{inx}$$

with natural number k and $c_n \in \mathbb{C}$ for every n, is an algebra. It might be helpful for the understanding of Stone's theorem on Fourier series.

Lemma 43 (Interpolation)**.** Suppose A is an algebra of complex functions on a set X that separates points and vanishes at no point of X. Suppose that there exist distinct points x_1 and x_2 of X and let c_1 and c_2 be constants. Then there exists a function f in A such that $f(x_1) = c_1$ and $f(x_2) = c_2$

Proof. Because A separates points and vanishes at no point of X, there exists functions g such that $g(x_1) \neq g(x_2)$ because of the definition of separating points, and functions a and b such t hat $a(x_1) \neq 0, b(x_2) \neq 0$

Now consider functions

$$u = gk - g(x_1)k$$

$$v = gh - g(x_2)h$$

Since A is an algebra, both u and v are in A.

Then, we have

$$u(x_1) = (g \cdot k)(x_1) - g(x_1)k(x_1) = g(x_1)k(x_1) - g(x_1)k(x_1) = 0$$

$$v(x_2) = (g \cdot h)(x_2) - g(x_2)h(x_2) = g(x_2)h(x_2) - g(x_2)h(x_2) = 0$$

$$u(x_2) = (g \cdot h)(x_2) = g(x_1)h(x_2) = h(x_2)(g(x_2) - g(x_1)) \neq 0 \text{this is a constant, same as below}$$

$$v(x_1) = (g \cdot h)(x_1) = g(x_2)h(x_1) = h(x_1)(g(x_1) - g(x_2))$$

Therefore, we have $u(x_1) = v(x_2) = 0$, and $u(x_2) \neq 0$, and $v(x_1) \neq 0$.

Now consider a function that has the form

$$f = \frac{c_1 \cdot v}{v(x_1)} - \frac{c_2 \cdot u}{u(x_2)}$$

and we found

$$f(x_1) = c_1 v(x_1)/v(x_1) + c_2 u(x_1)/u(x_2) = c_1$$

$$f(x_2) = c_1 v(x_2)/v(x_1) + c_2 u(x_2)/u(x_2) = c_2$$

So we have found such a function. □

Remark 44 (Intuition and usefulness of the interpolation theorem)**.** The intuition behind creating the function f is that we create two parts of f, one part takes 0 at x_1 and c_2 at x_2, and the other takes 0 at x_2 and c_1 at x_1. To create the value of c_2 and c_1 we can just multiply with the nonzero value we get from the separating property and property of vanishing at no points. To create zero point we just need to create the multiplication like the u and v above.

Theorem 45 (Weierstrass' Theorem on real algebras)**.** Suppose that A is an algebra of real bounded continuous functions over a compact set X that separates points and vanishes at no point of X. Then A is dense in $C_{\mathbb{R}}(X)$. In other words, for each A, there are elements in $C_{\mathbb{R}}(X)$ that is arbitrarily close to A. For each element f in $C_{\mathbb{R}}(X)$ and every $\epsilon > 0$, we have a element in A that differs from f by less than ϵ.

Proof. We prove that the closure of A, $\overline{A}$, is $C_{\mathbb{R}}(X)$.

We will prove two things here:

(a) If g is in $\overline{A}$, then $|g|$ is in $\overline{A}$

(b) If g and h are both in $\overline{A}$, then $\max(g, h)$ and $\min(g, h)$ are in $\overline{A}$.

For (a), suppose g is in $\overline{A}$

we fix $\epsilon > 0$ and we have a open ball $B(|\cdot|, \epsilon) \subset C_{\mathbb{R}}([a, b])$, where $a = -\|g\|_\infty$ and $b = \|g\|_\infty$, because we can pick a function in it and then find a ϵ small enough such that the function that differs ϵ from it would still be in the set.

By the proof of 38, there exists a polynomial P such that $P(0) = 0$, and P is in the ball of $B(|\cdot|, \epsilon)$

We can write the polynomial in $P = \sum_{i=1}^{n} c_i g^i$ for positive integer n and some real number constants c.

By 41, we have $\overline{A}$ is a closed algebra, so for all i, we have the exponential g^i is in $\overline{A}$ (since it is just a product after finite multiplication). Therefore, after multiplying the exponential with constants, we have $c_i g^i$ is in $\overline{A}$. Therefore, all the polynomials are in $\overline{A}$. Since P is in $B(|\cdot|, \epsilon)$, $P(g)$ is in $B(|g|, \epsilon)$. Therefore, $|g|$ is a limit point of $\overline{A}$. Since $\overline{A}$ is closed, we have $|g|$ is in $\overline{A}$.

For the second point, we have

$$\max(g, h) = \frac{g + h}{2} - \frac{|g - h|}{2} \tag{5}$$

which means it is in the algebra of $\overline{A}$, so it is in $\overline{A}$. Similarly, we have $\min(g, h)$ is also in $\overline{A}$.

After proving those two things, fix f in $C_{\mathbb{R}}(X)$ and $\epsilon > 0$. We have a open ball $B(f, \epsilon) \subset C_{\mathbb{R}}(X)$ since f is a limit point.

By theorem 43 we have for every x in X, there exists a function g_x in $\overline{A}$ such that $g_x(x) = f(x)$.

Consider the function $h_x = \max(g_x, f)$ for every x in X. Consider the proof in previous theorem 5, h_x could be expressed into sum of continuous functions and is thus continuous. Therefore, there exists an open set U_x in X such that U_x contains x and $U_x \subset h_x^{-1}(B(f(x)), \epsilon)$ for all x in X by the definition of continuous in topology.

Here $B(f(x), \epsilon)$ denotes an open ball in $\mathbb{R}$.

The collection of all these U_x is an open cover of X. Because X is compact, a finite number of these U_x, U_i must cover X.

By choice of $U_i, h_i(U_i) \subset B(f(x_i), \epsilon)$.

Consider the minimum of h_i, we first have $h_i(x) > f(x) - \epsilon$ for every x in X because of the definition of h_x being the maximum of g_x and f.

Therefore, for all x, $f(x) - \epsilon < \min_{h_i}(x) < f(x) + \epsilon$, and the second inequality is because $h_i(U_i) \subset B(f(x_i), \epsilon)$, the open ball of radius ϵ.

Therefore, we have $\|f - \min(h_i)\|_\infty < \epsilon$, which means $\min_{h_i}$ is in $B(f, \epsilon)$.

By the second point of prove above, we have $\min(h_i)$ is in $\overline{A}$, so it is a limit point of the closure $\overline{A}$, and thus is in $\overline{(A)}$ since $\overline{A}$ is closed.

□

6 Extensions and applications of Stone-Weierstrass Theorem

6.1 Fourier series

Lemma 46. The set of linear combinations of functions $e_n(x) = e^{2\pi i n x}$ with integer n is dense in $C([0,1]/\{0,1\})$ (this is a circle, gluing the endpoint 1 and 0 together).

Proof. Just need to check it does not vanish. Pick two points, and if they have the same value for every n, we can get the two points are the same. □

Remark 47 (Applications after the theorem)**.** e_n are orthonormal on space $L^2([0,1])$ (just calculate the norm and dot product after the definition is given in the following part), and they form a basis since they are dense.

6.2 Certain algebra structures

For specific structures, we also have Stone-Weierstrass theorem on it.

Complex numbers:

Definition 48 (Self-adjoint)**.** A self-adjoint algebra is an algebra A such that for all f in A, $\overline{f}$ is in A

Example 49 (Self-adjoint examples)**.** The algebra of complex number that also includes

Theorem 50. Suppose that A is a self-adjoint algebra of bounded complex functions over a compact set X that separates points and vanishes at no point of X, then A is dense in $C(X)$.

Proof. Let A_R be the set of all real functions on X in A

Suppose $f \in C$ is $f = g + ih$, with g, h being real functions. Then both g and h are in A after expressing them using f and $\overline{f}$.

By interpolation theorem, if $x_1 \neq x_2$, there exists a function f such that $f(x_1) = 1, f(x_2) = 0$. Then A_R separates points since $g(x_1) = 1, g(x_0) = 0$.

Also we have A_R vanishes at no point of x by considering both the real part and imaginary part.

As a result, we have A_R is dense in the set of all real continuous functions on X.

Observing f, we have $f = g + ih$, we have f is in $\overline{A}$, which means A is dense in $C(X)$ □

Definition 51 (Quaternion)**.** A quaternion algebra over F can be described as 4−dimensional F-vector space with basis $\{1, i, j, k\}$ with the following application rules:

$$i^2 = a$$
$$j^2 = b$$
$$ij = k$$
$$ji = -k$$

where a and b are any given nonzero elements of F.

Example 52 (Properties and examples of Quarterions)**.** First we have $k^2 = ijij = -iijj = -ab$.

When we have $F = \mathbb{R}$, $a = b = -1$, we have a quaternion.

Theorem 53 (Stone Weierstrass theorem for quarternion)**.** Suppose X is a compact space and A is a sub algebra of $C(X, H)$, quaternion valued continuous functions which contains a non-zero constant function (since it necessary to show that A contains elements of it multiplied by a constant).

Then A is dense in $C(X, H)$ if and only if it separates points. That is to say, for any two different points x_1, x_2 in the interval, there exists a function f in A such that $f(x_1) \neq f(x_2)$ (so thus separate the points). Then we have there exists a polynomial in $C(X, H)$ that is infinitely close to f.

Proof. Write an arbitrary quaternion as $q = a + bi + jc + kd$ and then we have

$$a = (q - iqi - jqj - kqk)/4$$

$$b = (-qi - iq + jqk - kqj)/4$$

$$c = (-qj - iqk - jq + kqi)/4$$

$$d = (-qk + iqj - jqk - kq)/4$$

Now consider the algebra of A that contains all continuous quaternion-valued functions, it would distinguishes between points, since different points would have different scalars on for either i, j, k or the pure scalar term. Suppose A takes different result on x_1 and x_2, and select one such element from A, f. We can multiply $f(x_1)$ and $f(x_2)$, both are quaternions, by some other quaternion to get the real part. Therefore, we would have the real part of $f(x_1)$ is not equal to the real part of f, as seen above with the example q is equal to $[f - ifi - jfj - kfk]/4$, and this is an element of A since A is an algebra. Notice that the real part of f is a real-valued functions and it also distinguish between points (since if two quaternions are not equal, their real parts are also not equal). Therefore, A contains real valued functions which distinguish between points.

Since A is complete and is an algebra, the set of all real-valued functions in A is also complete and closed under arithmetic operations. (It being complete means every convergent series of functions in A converges to something in A). Therefore, by using Stone-Weierstrass theorem, A contains all continuous real-valued functions on X, and therefore it contains all continuous quaternion-valued functions on X. □

6.3 L2 spaces

Definition 54 (L2 space on [0, 1)**.**] The set of all functions $f : [0,1] \to \mathbb{R}$ whose square are Lebesgue integrable ($\int f^2 d\mu$ is finite). The sum of elements in it and product with constant numbers are also in the set. So $L^2\,[0,1]$ forms a vector space over $\mathbb{R}$.

Definition 55 (L2 norm)**.** For a function $f \in L^2\,[0,1]$, define the norm

$$\|f\|_{L^2} = \left(\int_0^1 f^2 dx\right)^{1/2}$$

We prove that $L^2\,[0,1]$ is separable.

Definition 56 (Countable)**.** A space is separable if it has a dense, countable subset.

We will show the dense, countable subset of $L^2\,[0,1]$ is set of all polynomials, $p : [0,1] \to \mathbb{R}$ we will first realize that the polynomials is dense in $C([0,1]\,,\mathbb{R})$, then $C([0,1]\,,\mathbb{R})$ and use another lemma to prove it.

First realize that polynomials being dense in $C(\mathbb{R}, \mathbb{R})$ is just the Weierstrass approximation theorem [Gad16].

Theorem 57 (Lebesgue Dominated Convergence theorem)**.** Let $\{f_n\}$ be a sequence of measurable functions such that this sequence converges pointwise to some f and $|f_n| \leq g$ for all n and an integrable function g. Then, f is integrable and $\int f d\mu = \lim_{n\to\infty} \int f_n d\mu$

Lemma 58. The set $C([0,1], \mathbb{R})$ is dense in $L^2[0,1]$

Proof. It suffices to show for all $f \in L^2[0,1]$, there exists a sequence of functions $\{g_n\}$, such that

$$\lim_{n\to\infty} \|f - g_n\|_{L^2} = 0$$

Let A be a closed subset of $[0,1]$ and K_A be its indicator function (1 if in the set, 0 otherwise). Now, define $t(x) = \inf_{y\in A}\{|x-y|\}$ and $g_n(x) = \dfrac{1}{1+nt(x)}$. We have $t(x)$ is continuous and thus $|g_n|$ is also continuous. Besides, since $n > 0, t(x) \geq 0$, we have $|g_n(x)| \leq 1$. For all $x \in A$, $t(x) = 0$, so $g_n(x) = 1$.

Also, for $x \in [0,1]\ \ A$, $\lim_{n\to\infty} g_n(x) = 0$ since the denominator will be really big. Therefore,

$$\lim_{n\to\infty} \|g_n(x) - K_A(x)\|_{L^2} = \lim_{n\to\infty} \left(\int_B g_n(x)^2 dx\right)^{1/2} = \left(\int_B \lim_{n\to\infty} g_n(x)^2 dx\right)^{1/2} = 0$$

where the last step is by the Lebesgue dominated convergence theorem. Now we can see the indicator function could be approximated by a sequence of continuous functions. Since simple functions are a finite linear combination of indicator functions (also this is the definition of simple functions), continuity is preserved, so we have any simple functions can be approximated.

Now suppose $f \in L^2[0,1]$ is nonnegative. We have there exists a sequence of nonnegative simple functions $\{s_n\}$ that converges to f. Also, $(f - s_n)^2 \leq f^2$ Therefore, by the Lebesgue dominated convergence theorem, $\|f - s_n\|_{L^2} = 0$.

For every $f \in L^2[0,1]$, we can write f as

$$f^+ - f^-$$

where

$$f^+(x) = \begin{cases} f(x) & f(x) > 0 \\ 0 & \text{otherwise} \end{cases}$$

and

$$f^-(x) = \begin{cases} -f(x) & f(x) < 0 \\ 0 & \text{otherwise} \end{cases}$$

f can be approximated by a sequence of continuous functions because both f^+ and f^- are nonnegative. Therefore, $C([0,1],\mathbb{R})$ is dense in $L^2([0,1])$, as required.

□

The following proof helps to finish the proof on L2 spaces, and it also gives a flavor of what proof in real analysis looks like.

Theorem 59 (Transitiviy of density)**.** Suppose A, B, C are metric spaces with the same metric, d. If A is dense in B , B is dense in C, then A is dense in C.

Proof. Let $\epsilon > 0$ and choose $c \in C$. We can then choose $b \in B, a \in A$ so that $d(b,c) < \epsilon/2, d(a,b) < \epsilon/2$ because of the denseness. Therefore, we have $d(a,c) < \epsilon$ by triangle inequality, so we are done. □

Theorem 60. $L^2\,[0,1]$ is separable.

Proof. By the lemmas above, we have $C([0,1],\mathbb{R})$ is dense in $L^2\,[0,1]$ under the L^2 norm. Also polynomials over $[0,1]$, are dense in $C\,([0,1],\mathbb{R})$ under the supremum norm.

Also,

$$\left(\int_0^1 f^2 dx\right)^{1/2} \leq \left(\int_0^1 \|f\|_u^2\, dx\right)^{1/2} = (\|f\|_u^2)^{1/2} = \|f\|_u$$

So we have polynomials over $[0,1]$ is dense in $C([0,1],\mathbb{R})$ under the L^2 norm. By the transitivity of denseness, polynomials over $[0,1]$ id dense in $L^2\,[0,1]$ under the L^2 norm. Since the polynomials over $[0,1]$ is countable (we can map them to the integers), $L^2\,[0,1]$ is separable. □

7 Further exploration

Though some amount of Stone-Weierstrass theorem's applications are addressed in this paper, there are still a lot more to explore. For instance, further work would explore how the Stone-Weierstrass theorem would help in other fields, like computing in computer science or neural networks.

8 Acknowledgement

The author would like to thank Simon Rubinstein-Salzedo, Sawyer Anthony Dobson, and Tudor for the contribution to the writing of this paper.

References

[Gad16] Philip Gaddy. "THE STONE-WEIERSTRASS THEOREM AND ITS APPLICATIONS TO L2 SPACES". In: *apunts de classe* 1 (2016).

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

[You06] Matt Young. "The stone-weierstrass theorem". In: *MATH 328 Notes*. Queen's University at Kingston, 2006.