

# The Stone-Weierstrass Theorem

Matt Young  
MATH 328 Notes  
Queen's University at Kingston  
Winter Term, 2006

The Weierstrass Approximation Theorem shows that the continuous real-valued functions on a compact interval can be uniformly approximated by polynomials. In other words, the polynomials are uniformly dense in  $C([a, b], \mathbb{R})$  with respect to the sup-norm. The original proof was given in [1] in 1885. There are now several different proofs that use vastly different approaches. One well-known proof was given by the Russian Sergei Bernstein in 1911. His proof uses only elementary methods and gives an explicit algorithm for approximating a function by the use of a class of polynomials now bearing his name. It will be seen that the Weierstrass Approximation Theorem is in fact a special case of the more general Stone-Weierstrass Theorem, proved by Stone in 1937, who realized that very few of the properties of the polynomials were essential to the theorem. Although this proof is not constructive and relies on more machinery than that of Bernstein, it is much more efficient and has the added power of generality.

First, Bernstein's proof of the Weierstrass Approximation Theorem, which is taken from [4], is examined. The Bernstein polynomials, which play a central role in Bernstein's proof, are introduced below.

**Definition 14.1 (Bernstein Polynomials).** For each  $n \in \mathbb{N}$ , the  $n^{\text{th}}$  Bernstein Polynomial of a function  $f \in C([0, 1], \mathbb{R})$  is defined as

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

**Theorem 14.1 (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  $[a, b]$ .

**Proof.** Consider first  $f \in C([0, 1], \mathbb{R})$ . Once the theorem is proved for this case, the general theorem will follow by a change of variables. Since  $[0, 1]$

is compact, the continuity of  $f$  implies uniform continuity. So, given  $\epsilon > 0$ , there exists  $\delta > 0$  such that

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

Now, let  $M := \|f\|_\infty$ . Note that  $M$  exists since  $f$  is a continuous function on a compact set. Now, fix  $\xi \in [0, 1]$ . Then, if  $|x - \xi| \leq \delta$ , then  $|f(x) - f(\xi)| \leq \frac{\epsilon}{2}$  by continuity. Alternatively, if  $|x - \xi| \geq \delta$ , then

$$|f(x) - f(\xi)| \leq 2M \leq 2M \left( \frac{|x - \xi|}{\delta} \right)^2 + \frac{\epsilon}{2}.$$

Combining the above two inequalities, we see that

$$|f(x) - f(\xi)| \leq 2M \left( \frac{|x - \xi|}{\delta} \right)^2 + \frac{\epsilon}{2} \quad \forall x \in [0, 1].$$

The Bernstein Polynomials can be used to approximate  $f$  on  $[0, 1]$ . First, note that

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

And,

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

where the Binomial Theorem was used in the second equality. Thus,

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

where in the second step the fact that  $0 \leq B_n(x, f)$  for  $0 \leq f$  and  $B_n(x, g) \leq B_n(x, f)$  if  $g \leq f$  were used. Both can be proven directly from the definition of  $B_n(x, f)$ . It can also be shown that

$$B_n(x, (x - \xi)^2) = x^2 + \frac{1}{n}(x - x^2) - 2\xi x + \xi^2.$$

So

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

In particular,

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

A simple calculation shows that on  $[0, 1]$ , the maximum of  $z - z^2$  is  $\frac{1}{4}$ . Thus,

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

So, take  $N \geq \frac{M}{2\delta^2 \epsilon}$ . Then, for  $n \geq N$ ,

$$\|(B_n(\xi, f) - f(\xi))\|_\infty \leq \epsilon.$$

This proves the theorem for continuous functions on  $[0, 1]$ . Now, let  $g \in C([a, b], \mathbb{R})$ . Consider the function  $\phi : [0, 1] \rightarrow [a, b]$  defined by  $\phi : x \mapsto (b - a)x + a$ .  $\phi$  is clearly a homeomorphism. Thus, the composite function  $f := g \circ \phi$  is a continuous function on  $[0, 1]$ . By an application of the theorem for functions on  $[0, 1]$ , the case for an arbitrary interval  $[a, b]$  follows.  $\square$

Although tedious, the above proof relied on only elementary methods. The introduction of some new concepts and results is required before the proof of the Stone-Weierstrass Theorem can be approached.

**Definition 14.2 (Unital Sub-algebra, Separating Points).** Let  $K$  be a compact metric space. Consider the Banach algebra

$$C(K, \mathbb{R}) := \{f : K \rightarrow \mathbb{R} \mid f \text{ is continuous}\}$$

equipped with the sup-norm,

$$\|f\|_\infty := \sup_{t \in K} |f(t)|.$$

Then,

- 1)  $A \subset C(K, \mathbb{R})$  is a *unital sub-algebra* if  $1 \in A$  and if  $f, g \in A$ ,  $\alpha, \beta \in \mathbb{R}$  implies that  $\alpha f + \beta g \in A$  and  $fg \in A$ .
- 2)  $A \subset C(K, \mathbb{R})$  *separates points* of  $K$  if for all  $s, t \in K$  with  $s \neq t$ , there exists  $f \in A$  such that  $f(s) \neq f(t)$ .

The above definition can also be generalized by replacing the requirement that  $K$  be a compact metric space by requiring that  $K$  be a compact topological space.

**Remarks.** (a) It follows from the above definition that if  $A$  is a unital sub-algebra, then all constant functions are elements of  $A$ .

(b) Let  $\mathcal{P}([a, b], \mathbb{R})$  be the space of polynomials from  $[a, b]$  to  $\mathbb{R}$ . It is easily checked that  $\mathcal{P}([a, b], \mathbb{R})$  is a unital sub-algebra and separates points.

Before stating and proving the Stone-Weierstrass Theorem a useful lemma about closed sub-algebras will be proven. The following definition must first be made.

**Definition 14.3 (Lattice).** A subset  $S \in C(K, \mathbb{R})$  is a *lattice* if, for all  $f, g \in S$ ,  $f \vee g \in S$  and  $f \wedge g \in S$ , where

$$(f \vee g)(x) := \max\{f(x), g(x)\} \text{ and } (f \wedge g)(x) := \min\{f(x), g(x)\}.$$

**Remark.** In the following lemma, the Taylor series of the function  $\sqrt{1 - g(t)}$ , where  $0 \leq g(t) \leq 1$ , is investigated. It is useful to first study the Taylor series of  $\sqrt{1 - t}$ . Formally,

$$\sqrt{1 - t} = 1 - \sum_{n=1}^{\infty} a_n t^n$$

where, for  $n \in \mathbb{N}$ ,

$$a_n = (-1)^{n-1} \binom{\frac{1}{2}}{n} = \frac{(-1)^{n-1}}{n!} \prod_{k=0}^{n-1} \left( \frac{1}{2} - k \right) = 2^{1-2n} \frac{(2n-2)!}{n!(n-1)!}.$$

Note that for all  $n \in \mathbb{N}$ ,  $a_n \geq 0$ . Observe that

$$\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = \lim_{n \rightarrow \infty} \frac{2n-1}{2(n+1)} = 1.$$

Hence, by the ratio test, the series converges pointwise for  $t \in (-1, 1)$ . We will show that the series does indeed converge to  $\sqrt{1 - t}$ . Define  $\psi(t) := 1 - \sum_{n=1}^{\infty} a_n t^n$ . So,  $\psi(t)$  converges for  $t \in (-1, 1)$ . Similarly, we obtain  $\frac{d\psi(t)}{dt} = -\sum_{n=1}^{\infty} n a_n t^{n-1}$ . Multiplying, we see that

$$\psi(t) = -2(1 - t) \frac{d\psi(t)}{dt}.$$

Then,

$$-\frac{1}{2} \int \frac{dt}{1-t} = \int \frac{d\psi}{\psi} \implies \psi(t) = c\sqrt{1-t}$$

for some  $c \in \mathbb{R}$ . Evaluating both sides at  $t = 0$  gives  $c = 1$ . Hence,  $\psi(t) = \sqrt{1-t}$  pointwise for  $t \in [0, 1)$ . Now, we must show that  $\psi(1) = 0$ . Stirling's inequality states

$$e^{\frac{7}{8}-n} n^{n+\frac{1}{2}} < n! < e^{1-n} n^{n+\frac{1}{2}}.$$

Hence, for  $n \geq 2$ ,

$$a_n < 2^{1-2n} \frac{1}{e^{\frac{7}{4}}} \frac{(2n-2)^{2n-\frac{3}{2}}}{n^{n+\frac{1}{2}}(n-1)^{n-\frac{1}{2}}} < \frac{1}{\sqrt{2}e^{\frac{7}{4}}} \frac{(n-1)^{2n-\frac{3}{2}}}{(n-1)^{2n}} = \frac{1}{\sqrt{2}e^{\frac{7}{4}}} \frac{1}{(n-1)^{\frac{3}{2}}}.$$

So,

$$\sum_{n=1}^{\infty} a_n < \frac{1}{2} + \frac{1}{\sqrt{2}e^{\frac{7}{4}}} \sum_{n=2}^{\infty} \frac{1}{(n-1)^{\frac{3}{2}}} < \infty$$

where the comparison test was used in the first inequality and p-series test in the second. Since  $\left(\sum_{n=1}^k a_n\right)_{k \geq 1}$  is monotonically increasing and bounded above,  $\psi(1)$  exists. Then, by Abel's Theorem<sup>1</sup>,

$$\psi(1) = 1 - \sum_{n=1}^{\infty} a_n = 1 - \lim_{t \rightarrow 1^-} \sum_{n=1}^{\infty} a_n t^n = \lim_{t \rightarrow 1^-} \sqrt{1-t} = 0.$$

Hence, again using Abel's Theorem,  $\psi(t) = \sqrt{1-t}$  uniformly for  $t \in [0, 1]$ .

**Lemma 14.1** *Let  $A \subset C(K, \mathbb{R})$  be a closed unital sub-algebra. Then*

- i) *if  $f \in A$  and  $f \geq 0$ , then  $\sqrt{f} \in A$ ;*
- ii) *if  $f \in A$ , then  $|f| \in A$ ;*
- iii)  *$A$  is a lattice.*

---

<sup>1</sup>**Theorem (Abel).** *Let  $f(x) := \sum_{n=0}^{\infty} c_n(x-x_0)^n$ , and assume that the series converges at  $x = x_0 + R$ , for some  $R \in (0, \infty)$ . Then the series is uniformly convergent on  $[x_0, x_0 + R]$  and*

$$\lim_{x \rightarrow (x_0+R)^-} f(x) = f(x) = f(x_0 + R) = f(x) = \sum_{n=0}^{\infty} c_n R^n.$$

**Proof.** To see i), consider without restriction

$$0 \leq f \leq 1.$$

Then, we can write

$$f = 1 - g$$

with

$$0 \leq g \leq 1.$$

Using a Taylor series expansion, we can write, formally,

$$\sqrt{f(t)} = \sqrt{1 - g(t)} = 1 - \sum_{n=1}^{\infty} a_n g^n(t)$$

where the coefficients are as in the above remark. The Taylor series approximates  $\sqrt{f}$  uniformly in  $\|\cdot\|_{\infty}$ . Indeed,

$$\begin{aligned} \|\sqrt{f} - (1 - \sum_{n=1}^N a_n g^n)\|_{\infty} &\leq \sup_{t \in K} |\sqrt{f(t)} - (1 - \sum_{n=1}^N a_n g^n(t))| \\ &= \sup_{\zeta \in [0,1]} |\sqrt{\zeta} - (1 - \sum_{n=1}^N a_n \zeta^n)|. \end{aligned}$$

So, given  $\epsilon > 0$ , by the uniform convergence of the Taylor series of  $\sqrt{1-t}$  on  $[0, 1]$ , there exists  $\bar{N} \in \mathbb{N}$  such that

$$\|\sqrt{f} - (1 - \sum_{n=1}^N a_n g^n)\|_{\infty} < \epsilon \quad \forall N \geq \bar{N}.$$

That is,

$$\lim_{N \rightarrow \infty} \|\sqrt{f} - (1 - \sum_{n=1}^N a_n g^n)\|_{\infty} = 0.$$

Since for all  $n \in \mathbb{N} \cup \{0\}$   $g^n \in A$ , and  $A$  is a sub-algebra,  $1 - \sum_{n=1}^N a_n g^n \in A$ . And, since  $A$  is closed by hypothesis,  $\sqrt{f} \in A$ . This completes the proof of i). To prove ii), note that

$$|f| = \sqrt{f^2}.$$

So, for  $f \in A$ ,  $f \cdot f = f^2 \in A$ , since  $A$  is an algebra. Applying i) to  $f^2$  implies  $|f| \in A$ . To prove iii), note that

$$f \wedge g = \frac{1}{2}(f + g - |f - g|) \text{ and } f \vee g = \frac{1}{2}(f + g + |f - g|).$$

Applying ii) to the above identities gives the desired result.  $\square$

We are now in a position to state and prove the Stone-Weierstrass theorem. The theorem was first proved by Stone in 1937 in [2]. However, he greatly simplified his proof in 1948 into the one that is commonly used today. See [3].

**Theorem 14.2 (Stone-Weierstrass Theorem (1937)).** *Let  $K$  be a compact metric space and  $A \subset C(K, \mathbb{R})$  a unital sub-algebra which separates points of  $K$ . Then  $A$  is dense in  $C(K, \mathbb{R})$ .*

**Remark.** An equivalent statement is that if  $A$  is a closed unital sub-algebra that separates points of a compact set  $K$ , and  $A \subset C(K, \mathbb{R})$ , then  $A = C(K, \mathbb{R})$ . We will proceed using this formulation.

**Proof.** Let  $A \subset C(K, \mathbb{R})$  be a closed unital sub-algebra that separates points of  $K$ . Let  $\epsilon > 0$  be given. For any  $f \in C(K, \mathbb{R})$  we will show that there exists  $g \in A$  such that

$$\|f - g\|_\infty < \epsilon.$$

Consider points  $s, t \in K$ . Since  $A$  separates points, there exists  $h \in A$  such that  $h(s) \neq h(t)$ . For some  $\lambda, \mu \in \mathbb{R}$ , define  $\tilde{h} : K \rightarrow \mathbb{R}$  by

$$\tilde{h}(v) := \mu + (\lambda - \mu) \frac{h(v) - h(t)}{h(s) - h(t)} \quad \forall v \in K.$$

Note that  $\tilde{h} \in A$  and  $\tilde{h}(s) = \lambda, \tilde{h}(t) = \mu$ . Thus, for  $s \neq t$ , there exists  $f_{s,t} \in A$  such that

$$\begin{aligned} f_{s,t}(s) &= f(s) \\ \text{and } f_{s,t}(t) &= f(t). \end{aligned}$$

So,  $f_{s,t}$  approximates  $f$  in neighbourhoods around  $s$  and  $t$ . Now, fix  $s$ , and let  $t$  vary. Put

$$U_t := \{v \in K | f_{s,t}(v) < f(v) + \epsilon\}.$$

$U_t$  is open because it is the pre-image of an open set. Also,  $t \in U_t$ . So,  $\bigcup_{t \in K} U_t$  is clearly an open cover of  $K$ . By the compactness of  $K$ , there exists finitely many  $t_1, \dots, t_n \in K$  such that

$$K \subset \bigcup_{i=1}^n U_{t_i}.$$

Put

$$h_s := \min_{1 \leq i \leq n} f_{s, t_i}.$$

Then,

$$\begin{aligned} h_s &\in A \\ \text{and } h_s(s) &= f(s) \\ \text{and } h_s &< f + \epsilon. \end{aligned}$$

Now, define

$$V_s := \{v \in K \mid h_s(v) > f(v) - \epsilon\}.$$

Note that  $V_s$  is open and  $K \subset \bigcup_{s \in K} V_s$ . By compactness, there exists finitely many  $s_1, \dots, s_m \in K$  such that

$$K \subset \bigcup_{j=1}^m V_{s_j}.$$

Put  $g = \max_{1 \leq j \leq m} h_{s_j}$ . Then,  $g \in A$  and

$$f - \epsilon < g < f + \epsilon.$$

That is,

$$\|f - g\|_\infty < \epsilon.$$

So,  $A$  is dense in  $C(K, \mathbb{R})$ . And, since  $A$  is closed,  $A = C(K, \mathbb{R})$ .  $\square$

**Corollary 14.1** *Let  $X$  be a compact subset of  $\mathbb{R}^n$  for some  $n \in \mathbb{N}$ . Then the algebra of all polynomials  $\mathcal{P}(X, \mathbb{R})$  in the coordinates  $x_1, \dots, x_n$  is dense in  $C(K, \mathbb{R})$ .*

**Remarks.** (a) The case in which  $n = 1$  in the above corollary is the Weierstrass Approximation Theorem.

(b) A result similar to the Weierstrass Approximation Theorem occurs in the theory of Fourier series, and was also first proved by Weierstrass. It states that a continuous  $2\pi$  periodic function can be uniformly approximated on  $\mathbb{R}$  by the trigonometric polynomials. The result is stated formally and proved below.



**Definition 14.4 (Trigonometric Polynomials).** The space of *real-valued trigonometric polynomials*  $\mathcal{TP}(\mathbb{R}, \mathbb{R})$  are functions  $f : \mathbb{R} \rightarrow \mathbb{R}$  which are finite sums of the form

$$f(x) = a_0 + \sum_{n=1}^N (a_n \cos(nx) + b_n \sin(nx)).$$

**Corollary 14.2** *The set of all trigonometric polynomials are uniformly dense in  $C_{per}([0, 2\pi], \mathbb{R})$ .*

**Proof.** First note that  $\mathcal{TP}(\mathbb{R}, \mathbb{R})$  is an algebra. The only non-trivial requirement to check is closure under multiplication. Indeed, for  $m, n \in \mathbb{R}$ , we have

$$\sin mt \cos nt = \frac{1}{2} (\cos(m-n)t - \cos(m+n)t) \in \mathcal{TP}(\mathbb{R}, \mathbb{R}).$$

Similar identities hold for other products. Also, note that 1 is the trigonometric polynomial with  $a_0 = 1$ ,  $a_n = b_n = 0$  for all  $n \in \mathbb{N}$ . To use the Stone-Weierstrass Theorem, we identify  $C_{per}([0, 2\pi], \mathbb{R})$  with  $C(\mathbb{T}, \mathbb{R})$ , where

$$\mathbb{T} = \{(x_1, x_2) \in \mathbb{R}^2 | x_1^2 + x_2^2 = 1\}$$

is the unit circle. Let  $t \mapsto (\cos t, \sin t)$ . Then, the trigonometric polynomials are polynomials in  $x_1, x_2$  on  $\mathbb{T}$ , which are dense in  $C(\mathbb{T}, \mathbb{R})$  by the previous corollary with  $n = 2$ .  $\square$

**Remark.** Note that the above corollary can be easily generalized for  $C_{per}([a, b], \mathbb{R})$ .

We can also use the Stone-Weierstrass Theorem to prove the complex Stone-Weierstrass Theorem. The complex version, however, requires additional assumptions. The following proof is taken from [5].

**Theorem 14.3 (Complex Stone-Weierstrass Theorem)** *Let  $A$  be a (complex) unital sub-algebra of  $C(K, \mathbb{C})$ , such that if  $f \in A$ , then  $\bar{f} \in A$ , and  $A$  separates points of  $K$ . Then,  $A$  is dense in  $C(K, \mathbb{C})$ .*

**Proof.** Let  $f \in A$ . Then,  $Re(f) = \frac{1}{2}(f + \bar{f}) \in A$  and  $Im(f) = \frac{1}{2i}(f - \bar{f}) \in A$ . Denote by  $A_{\mathbb{R}}$  the unital sub-algebra of  $A$  containing all real-valued functions. Note that  $A$  separates points in  $K$  because  $A_{\mathbb{R}}$  does. By the Stone-Weierstrass Theorem for real valued functions,  $A_{\mathbb{R}}$  is dense in  $C(K, \mathbb{R})$ . Since  $A = \{f + ig | f, g \in A_{\mathbb{R}}\}$ , we see that  $A$  is indeed dense in  $C(K, \mathbb{C}) = C(K, \mathbb{R}) + iC(K, \mathbb{R})$ .  $\square$

## References

- [1] K. Weierstrass (1885). Über die analytische Darstellbarkeit sogenannter willkrlicher Functionen einer reellen Veränderlichen. Sitzungsberichte der Kniglich Preuischen Akademie der Wissenschaften zu Berlin, 1885 (11).
- [2] M. Stone (1937). Applications of the Theory of Boolean Rings to General Topology. Translations of the American Mathematical Society **41** (3), 375-481.
- [3] M. Stone (1948). The Generalized Weierstrass Approximation Theorem. Mathematics Magazine **21** (21), 167-184 and **21** (5), 237-254.
- [4] K. Davidson and A. Donsig. *Real Analysis with Real Applications*. Prentice Hall, Upper Saddle River, N.J., 2002.
- [5] H. Sohrab. *Basic Real Analysis*. Birkhäuser, New York, N.Y., 2003.