1-1-2013

The Weierstrass Approximation Theorem

LaRita Barnwell Hipp
University of South Carolina

The Weierstrass Approximation Theorem

by

LaRita Barnwell Hipp

Bachelor of Arts
Furman University, 1996

Master of Mathematics
University of South Carolina 2011

Submitted in Partial Fulfillment of the Requirements

For the Degree of Master of Science in

Mathematics

College of Arts and Sciences

University of South Carolina

2013

Accepted by:

Maria Girardi, Major Professor

Ralph Howard, Committee Member

Lacy Ford, Vice Provost and Dean of Graduate Studies

Dedication

To Christian, my amazing husband and best friend; to Ronny and Tommy and Ansley, my precious children. Without them, this task would have been impossible. Their enduring patience, unconditional love and unwavering faith in me has been the core for my perseverance. I am truly blessed!

Acknowledgments

This work owes so much to many. Family, friends, fellow graduate students, colleagues, and professors have supported and mentored me throughout this intense process. First, my thesis supervisor and mentor Dr. Maria Girardi has given so much of her time, research, guidance and council. She has played an irreplaceable part in this work. I am grateful to her for helping me develop personally and professionally. I also benefitted, directly and indirectly, from comments, conversations, and suggestions from Dr. Ralph Howard as well. Expressing gratitude seems inadequate to communicate the weighty significance I feel as I reflect upon your investment in me. Last, but certainly not least, I want to say thank you to my family and friends. Their love, support and prayers have been invaluable as I have journeyed through this season of my life. Many words of encouragement came to the forefront of my mind when the most difficult challenges emerged. Their confidence in my success spurred me on to the end. With a grateful heart, I give my thanks.

Abstract

In this thesis we will consider the work began by Weierstrass in 1855 and several generalization of his approximation theorem since. Weierstrass began by proving the density of algebraic polynomials in the space of continuous real-valued functions on a finite interval in the uniform norm. His theorem has been generalized to an arbitrary compact Hausdorff space and the approximation with elements from more general algebras of continuous real-valued functions. We will consider proofs that use brute force and proofs based on convolutions and approximate identities, trudge through probability and the use of the Bernstein polynomials, and become intimately close to what is meant by an algebra. We will also see what inspired Weierstrass and Stone throughout their life as we take a sneak peak into their Biographies.

Table of Contents

Chapter 1

Definitions and Notations

As we begin, let us consider some important definitions and clarify our notation. Throughout this work, we will consider the field of the real numbers and the field of the complex numbers.

Definition 1.1. We denote by $\mathbb{K}$ the field of real numbers, $\mathbb{R}$, or the field of complex numbers, $\mathbb{C}$.

The approximation theory is concentrated on the consideration of elements in a topological space.

Definition 1.2. Throughout this work K will denote a compact Hausdorff topological space. We will also assume that K contains at least two points. We will consider functions defined from K into $\mathbb{K}$.

We consider continuous functions. So the following definitions are important for our discussion.

Definition 1.3. We define $\mathcal{C}(K, \mathbb{K})$ to be the collection of all continuous functions from K to $\mathbb{K}$.

Definition 1.4. We endow $\mathcal{C}(K, \mathbb{K})$ with the supremum norm given by

$$\|f\|_{\mathcal{C}(K,\mathbb{K})} := \sup_{x \in K} |f(x)|_{\mathbb{K}} \tag{1.1}$$

for $f \in \mathcal{C}(K, \mathbb{K})$.

Note that since K is compact, $\|\cdot\|_{\mathcal{C}(K,\mathbb{K})} : \mathcal{C}(K,\mathbb{K}) \to [0,\infty)$. Also, we see $\|\cdot\|_{\mathcal{C}(K,\mathbb{K})}$ is a norm; indeed, this follows easily from the fact $|\cdot|_{\mathbb{K}} : \mathbb{K} \to [0,\infty)$ is a norm. Thus $\left(\mathcal{C}(K,\mathbb{K}), \|\cdot\|_{\mathcal{C}(K,\mathbb{K})}\right)$ is a normed linear space.

Since $\mathcal{C}(K,\mathbb{K})$ is a linear space, we can use the norm in equation (1.1) to turn $\mathcal{C}(K,\mathbb{K})$ into a metric space, $(\mathcal{C}(K,\mathbb{K}), \rho)$, in the usual manner.

Definition 1.5. We define $\rho : \mathcal{C}(K,\mathbb{K}) \times \mathcal{C}(K,\mathbb{K}) \to [0,\infty)$ by

$$\rho(f,g) := \sup_{x\in K} |f(x) - g(x)|, \tag{1.2}$$

where $f, g \in \mathcal{C}(K,\mathbb{K})$.

For such a metric space, $(\mathcal{C}(K,\mathbb{K}), \rho)$, for $\epsilon > 0$ and $f \in \mathcal{C}(K,\mathbb{K})$ we denote by $N_\epsilon(f)$ the set

$$N_\epsilon(f) := \{g \in \mathcal{C}(K,\mathbb{K}) : \rho(f,g) < \epsilon\}.$$

Now such a ρ induces a topology on $\mathcal{C}(K,\mathbb{K})$.

Definition 1.6. Let ρ be as given in Definition 1.5. The topological space,

$$(\mathcal{C}(K,\mathbb{K}), \tau_\rho)$$

is defined as follows. A subset F of $\mathcal{C}(K,\mathbb{K})$ is open in τ_ρ if and only if for all $f \in F$ there exists an $\epsilon > 0$ such that $N_\epsilon(f) \subseteq F$.

Now we introduce two familiar functions, the maximum and the minimum functions, which will be important components in our discussion of the Weierstrass Approximation Theorem.

Definition 1.7. For real-valued functions f and g on a domain D into $\mathbb{R}$ define the maximum function $f \vee g$ and the minimum function $f \wedge g$ from D into $\mathbb{R}$ by

$$(f \vee g)(x) := \max\{f(x), g(x)\}$$

$$(f \wedge g)(x) := \min\{f(x), g(x)\}$$

where $x \in D$. Clearly

$$f \vee g = \frac{f + g + |f - g|}{2}$$
$$f \wedge g = \frac{f + g - |f - g|}{2}.$$

Given the above definition of the two functions, maximum and minimum, we introduce the term 'lattice'.

Definition 1.8. A collection $\mathcal{L}$ of real-valued functions is called a *lattice* provided

$$f \vee g \in \mathcal{L}$$

$$f \wedge g \in \mathcal{L}$$

whenever $f, g \in \mathcal{L}$.

Definition 1.9. Let $f, g \in K$ be real-valued functions. We will define *lattice operations* to be $f \vee g$ and $f \wedge g$.

Remark 1.10. If $f, g \in \mathcal{C}(\Omega, \mathbb{K})$ then $f \vee g, f \wedge g \in \mathcal{C}(\Omega, \mathbb{K})$ for any topological space Ω.

Definition 1.11. Let f be any function. Then we can define the real part of f, $\Re f$, and the imaginary part of f, $\Im f$, by

$$\Re f = \frac{f + \overline{f}}{2}$$
$$\Im f = \frac{f - \overline{f}}{2i}.$$

Then

$$f = \Re f + i \Im f.$$

Also $\Re f$ and $\Im f$ are real-valued.

The following are two similar definitions. We define a Dirac sequence and an approximate identity.

Definition 1.12. We shall define a *Dirac sequence* as a sequence of real-valued functions, $\{K_n\}$, defined on all of $\mathbb{R}$, which satisfies the following properties.

1. We have $K_n(x) \geq 0$ for all $n \in \mathbb{N}$ and all $x \in \mathbb{R}$.

2. Each K_n is continuous, and $\int_{\mathbb{R}} K_n(t)\, dt = 1$.

3. For every $\epsilon > 0$ and $\delta > 0$, there exists N such that if $n \geq N$ then

$$\int_{|x|>\delta} K_n(x)\, dx < \epsilon.$$

Note that $K_n \in L_1(\mathbb{R}, \mathbb{R})$, which are used to form integrals like the convolution, are sometimes called *kernel functions*. Often the functions K_n are even, that is $K_n(-x) = K_n(x)$ for all $x \in \mathbb{R}$.

Definition 1.13. Let $\varphi \in L_1(\mathbb{R}, \mathbb{K})$. Then φ is called an *approximate identity* if

$$\int_{\mathbb{R}} \varphi(t)\, dt = 1.$$

Remark 1.14. For an approximate identity, we can define the family of functions $\{\varphi_\epsilon\}_{\epsilon>0}$ in $L_1(\mathbb{R}, \mathbb{K})$ by

$$\varphi_\epsilon(t) = \frac{1}{\epsilon}\varphi\left(\frac{t}{\epsilon}\right).$$

Note that $\{\varphi_\epsilon\}_{\epsilon>0}$ satisfies the following properties with a simple change of variables; namely $s = \frac{t}{\epsilon}$.

(AI1) $\int_{\mathbb{R}} \varphi_\epsilon(t)\, dt = 1$.

We can see this by the following justification and the fact that φ is an approximate identity.

$$\int_{\mathbb{R}} \varphi_\epsilon(t)\, dt = \int_{\mathbb{R}} \frac{1}{\epsilon}\varphi\left(\frac{t}{\epsilon}\right) dt = \int_{\mathbb{R}} \varphi(s)\, ds = 1.$$

(AI2) $\|\varphi_\epsilon\|_{L_1} = \|\varphi\|_{L_1}$.

We use the definition of the L_1 norm in the following justification.

$$\|\varphi_\epsilon\|_{L_1} = \int_{\mathbb{R}} |\varphi_\epsilon(t)|\, dt = \int_{\mathbb{R}} \frac{1}{\epsilon}\left|\varphi\left(\frac{t}{\epsilon}\right)\right| dt = \int_{\mathbb{R}} |\varphi(s)|\, ds = \|\varphi\|_{L_1}.$$

(AI3) For $\delta > 0, \lim_{\epsilon\to 0^+} \int_{|t|\geq\delta} |\varphi_\epsilon(t)|\, dt = 0$.

To show this property is true we fix $\delta > 0$. Then

$$\lim_{\epsilon\to 0^+} \int_{|t|\geq\delta} |\varphi_\epsilon(t)|\, dt = \lim_{\epsilon\to 0^+} \int_{|t|\geq\delta} \left|\frac{1}{\epsilon}\varphi\left(\frac{t}{\epsilon}\right)\right| dt = \lim_{\epsilon\to 0^+} \int_{|s|\geq\frac{\delta}{\epsilon}} |\varphi(s)|\, ds = 0$$

by the Lebesgue's Differentiation Convergence Theorem.

Note that if φ is a nonnegative continuous approximate identity, then $\left\{\varphi_{\frac{1}{n}}\right\}_{n\in\mathbb{N}}$ is a Dirac sequence.

Definition 1.15. Let $f \in L_\infty(\mathbb{R},\mathbb{K})$ and $g \in L_1(\mathbb{R},\mathbb{K})$. Define the *convolution* $f * g : \mathbb{R} \to \mathbb{K}$ by

$$(f * g)(x) = \int_{\mathbb{R}} f(t)\, g(x-t)\, dt.$$

Note that by a simple change of variables

$$(f * g)(x) = \int_{\mathbb{R}} f(x-t)\, g(t)\, dt.$$

Remark 1.16. In the setting of Definition 1.15, $f * g$ exists for all $x \in \mathbb{R}$. Futhermore $f * g$ is bounded and uniformly continuous and

$$\|f * g\|_\infty \leq \|f\|_\infty \|g\|_1 .$$

A proof of the existence of $f * g$ and the estimate for $\|f * g\|_\infty$ can be found in [cf. [6], Proposition 8.8]. The key idea of the proof is Hölder's inequality.

Now we define the *Bernstein polynomials.* In 1911, Sergei Bernstein first used these polynomials to prove the Stone-Weierstrass Theorem over the interval $[0, 1]$.

Definition 1.17. Let $f : [0, 1] \to \mathbb{R}$ be a function. We define the n^{th} *Bernstein polynomial* associated with f to be

$$B_n(f)(x) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right) \binom{n}{k} x^k (1-x)^{n-k} \tag{1.3}$$

for each $n \in \mathbb{N}$, where $\binom{n}{k}$ is the binomial coefficient.

Chapter 2

An Algebra: What is it? what does it do? How do we use it?

In this chapter, we will introduce the concept of an algebra and define related terms, discuss basic facts and give examples of algebras. The following definitions and examples assume K is a non-empty set.

Definition 2.1. An *algebra of functions* $\mathcal{A}$ over $\mathbb{K}$ is a nonempty set of functions from K into $\mathbb{K}$ such that if $f, g \in \mathcal{A}$ and $c \in \mathbb{K}$ then

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

As we will consider the limits of functions, we need to define what is meant by a closure.

Definition 2.2. Let $(\mathcal{C}(K, \mathbb{K}), \tau_\rho)$ be as in Definition 1.6. Let $F \subseteq \mathcal{C}(K, \mathbb{K})$. We define the closure, $\overline{F}$, of F to be the intersection of all closed sets containing F. Thus

$$\overline{F} := \cap \{C : F \subset C, C \text{ is closed}\}.$$

Note that $\overline{F}$ is the smallest closed set containing F.

One may ask the question, is the closure of an algebra again an algebra? We will answer that question with the proof of the following lemma.

Lemma 2.3. *If $\mathcal{A}$ be an algebra of functions, then $\overline{\mathcal{A}}$ is also an algebra.*

Proof. Let $\mathcal{A}$ be an algebra of functions. Fix $f, g \in \overline{\mathcal{A}}$. Then there exists $f_n, g_n \in \mathcal{A}$ such that

$$f_n \to f$$

$$g_n \to g$$

uniformly. Let $c \in \mathbb{K}$. Then the following uniform convergences are true.

$$f_n + g_n \to f + g.$$

$$f_n g_n \to fg.$$

$$cf_n \to cf.$$

Thus $f + g, fg, cf \in \overline{\mathcal{A}}$. Therefore by Definition 2.1, $\overline{\mathcal{A}}$ is an algebra. □

Below we have compiled a collection of algebras which is, by far, not an exhaustive list.

Example 2.4. Examples of algebras, $\mathcal{A}$, of functions from K to $\mathbb{K}$.

1. The collection, $\mathcal{A}$, of all real-valued functions on K is an algebra.

2. Let K be an open interval of $\mathbb{R}$. Then the collection, $\mathcal{A}$, of all differentiable functions on K is an algebra.

3. Let K be the interval $[0, 1]$. Then the collection, $\mathcal{A}$, of polynomial functions on K form an algebra.

4. Let K be a subset of $\mathbb{R}^n$. Again, the collection, $\mathcal{A}$, of polynomial functions on K form an algebra.

5. Let φ be a function on K. Then the collection, $\mathcal{A}$, of all functions which can be written in the form

$$f(x) = a_0 + a_1 \varphi(x) + ... + a_n [\varphi(x)]^n$$

where $a_j \in \mathbb{K}$ is an algebra. This algebra, $\mathcal{A}$, is called the *algebra of polynomials* in φ.

6. Let K be the interval $[0, 2\pi]$. Then the collection, $\mathcal{A}$, of all functions on K into $\mathbb{R}$ which can be written in the form

$$a_0 + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx)$$

with $a_k, b_k \in \mathbb{R}$ is an algebra. This algebra is called the *algebra of trigonometric polynomials*.

7. Let $K = \{z \in \mathbb{C} \colon |z| \leq 1\} = \{\mathrm{e}^{ix} \colon x \in [0, 2\pi)\}$. Then K is a compact subset of $\mathbb{C}$. Then the collection, $\mathcal{A}$, of all functions of the form

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

where $k \in \mathbb{N}$ and $c_n \in \mathbb{C}$, is an algebra. This algebra is called the *algebra of complex trigonometric polynomials*.

Next we will define what is meant by 'separation of points'. It is a key characteristic of an algebra that meets the qualifications given in the Stone-Weierstrass Theorem.

Definition 2.5. We say that an algebra $\mathcal{A}$ of functions, from K into $\mathbb{K}$, *separates points* of K provided if points $x, y \in K$ and $x \neq y$, there exists a function $f \in \mathcal{A}$ such that $f(x) \neq f(y)$.

Definition 2.6. We say that an algebra $\mathcal{A}$ of functions, from K into $\mathbb{K}$, *strongly separates points* of K if for all distinct $x, y \in K$, and for all $\alpha, \beta \in \mathbb{K}$ there exists $f \in \mathcal{A}$ such that

$$f(x) = \alpha \quad \text{and} \quad f(y) = \beta \ .$$

Observe that for each $x \in K$ there exists $f \in \mathcal{A}$ such that

$$f(x) = 1 \ .$$

Remark 2.7. Clearly *strongly separates points* implies *separates points*.

Definition 2.8. Let $\mathcal{A}$ be an algebra of complex valued functions on the set K. If $f \in \mathcal{A}$, its complex conjugate $\overline{f}$ is defined by $\overline{f}(x) = \overline{f(x)}$. For instance, if $f(x) = \mathrm{e}^{ix}$ then $\overline{f}(x) = \mathrm{e}^{-ix}$.

Definition 2.9. If $\mathcal{A}$ is an algebra over $\mathbb{C}$ of complex valued functions, we say that $\mathcal{A}$ is self conjugate if whenever $f \in \mathcal{A}$ the conjugate functions $\overline{f}$ is also in $\mathcal{A}$.

Chapter 3

History: From Weierstrass to Stone

The Weierstrass approximation theorem was originally discovered by Karl Weierstrass in 1855. The original statement as proved by Weierstrass is as follows. Suppose f is a continuous complex-valued function defined on the real interval $[a, b]$. For every $\epsilon > 0$, there exist a polynomial function p over $\mathbb{C}$ such that for all $x \in [a, b]$, we have $|f(x) - p(x)| < \epsilon$, or equivalently, the supremum norm $\|f - p\|_{\mathcal{C}([a,b],\mathbb{K})} < \epsilon$. If f is real-valued, the polynomial function can be taken over $\mathbb{R}$. In other words, The Weierstrass approximation theorem states that every continuous function defined on a closed interval $[a, b]$ can be uniformly approximated arbitrarily close by a polynomial function. This theorem has had a big impact in the mathematical analysis community because polynomials are among the simplest functions.

As a consequence of the Weierstrass approximation theorem, one can show that the space $\mathcal{C}([a, b], \mathbb{R})$ is separable: the polynomial functions are dense, and each polynomial function can be uniformly approximated by one with rational coefficients.

In 1937, Marshall H. Stone considerably generalized the theorem and then simplified the proof in 1948. His result is known as the Stone-Weierstrass Theorem. This theorem generalizes the Weierstrass approximation theorem in two ways: instead of the real interval $[a, b]$, an arbitrary compact Hausdorff space K is considered, and instead of the algebra of polynomial functions, Stone investigated the approximation with elements from more general algebras of $\mathcal{C}(K, \mathbb{K})$.

Stone started with an arbitrary compact Hausdorff space K and considered the algebra $\mathcal{C}(K, \mathbb{R})$ of real-valued continuous functions on K, with the topology of uniform convergence. He wanted to find algebras of $\mathcal{C}(K, \mathbb{R})$ which are dense. It turns out that the necessary property that an algebra must satisfy is that it separates points. The Stone-Weierstrass Theorem states the following. Let K be a compact Hausdorff space and $\mathcal{A}$ an algebra of $\mathcal{C}(K, \mathbb{R})$ which contains a non-zero constant function. Then $\mathcal{A}$ is dense in $\mathcal{C}(K, \mathbb{R})$ if and only if it separates points. This implies Weierstrass' original statement since the polynomials on $[a, b]$ form an algebra of $\mathcal{C}([a, b], \mathbb{R})$ which contains the constants and separates points.

The Stone-Weierstrass Theorem can be used to prove the following two statements which go beyond Weierstrass' results.

1. If f is a continuous real-valued function defined on the set $[a, b] \times [c, d]$ and $\epsilon > 0$, then there exists a polynomial function p in two variables such that $|f(x, y) - p(x, y)| < \epsilon$ for all $x \in [a, b]$ and $y \in [c, d]$.

2. If X and Y are two compact Hausdorff spaces and $f : X \times Y \to \mathbb{R}$ is a continuous function, then for every $\epsilon > 0$ there exist $n > 0$ and continuous functions $f_1, f_2, ..., f_n$ on X and continuous functions $g_1, g_2, ..., g_n$ on Y such that $\|f - \sum f_i g_i\|_{\mathcal{C}([a,b]\times[c,d],\mathbb{R})} < \epsilon$.

There is also a complex version of the Stone-Weierstrass Theorem. It is a slightly more general version where we consider the algebra $\mathcal{C}(K, \mathbb{C})$ of complex-valued continuous functions on the compact space K, again with the topology of uniform convergence. Let K be a compact Hausdorff space and let $\mathcal{A}$ be an algebra of $\mathcal{C}(K, \mathbb{C})$ containing a non-zero constant function such that $\mathcal{A}$ separates points and if $f \in \mathcal{A}$, then $\overline{f} \in \mathcal{A}$. Then $\mathcal{A}$ is dense in $\mathcal{C}(K, \mathbb{C})$.

Chapter 4

A collection of Approximation Theorems

There are many versions of the well-known Stone-Weierstrass Theorem. We will compile several versions here and then present many proofs in following chapter.

In 1885, Karl Weierstrass proved that continuous real-valued polynomials on a compact interval are dense in the set of all continuous functions, that is to say that any continuous real-valued function defined on a closed bounded interval of the reals can by uniformly approximated by polynomials. The first two theorems in this sections are examples of this significant result.

Theorem 4.1. *The Weierstrass Approximation Theorem (Case: $\mathcal{C}([0,1],\mathbb{R})$)*
Let $f : [0,1] \to \mathbb{R}$ be a continuous function. Then for each $\epsilon > 0$, there exists a polynomial function P such that for all $x \in [0,1]$, $|f(x) - P(x)| < \epsilon$. Equivalently, for any such f, there exists a sequence $\{P_n\}_{n=1}^{\infty}$ of polynomials such that P_n converges to f uniformly on $[0,1]$.

We will prove in Lemma 5.16 that the following theorem is an extension of the previous theorem by rescaling.

Theorem 4.2. *The Weierstrass Approximation Theorem (Case: $\mathcal{C}([a,b],\mathbb{R})$)*
Let $[a,b]$ be a closed interval, and let f be a continuous function on $[a,b]$. Then f can be uniformly approximated by polynomials on $[a,b]$. That is given any $\epsilon > 0$, there is a polynomial P_ϵ such that $|f(x) - P_\epsilon(x)| < \epsilon$ for each $x \in [a,b]$.

Polynomials form an algebra with nice properties but in the 1940's, M. H. Stone isolated the necessary properties to generalize the Weierstrass Theorem to algebras.

He also expanded the domain of the real interval $[a, b]$ to an arbitrary compact Hausdorff space. Hence the Weierstrass Approximation Theorem can be generalized to any algebra, $A \subset \mathcal{C}(K, \mathbb{K})$, with sufficiently nice properties as stated in the next theorems now known as the Stone-Weierstrass Theorem.

Theorem 4.3. *Stone-Weierstrass Theorem (Case: $\mathcal{C}(K, \mathbb{R})$)*
Let K be a compact Hausdorff space. Suppose that $\mathcal{A} \subset \mathcal{C}(K, \mathbb{R})$ satisfies the following conditions.

1. *$\mathcal{A}$ is an algebra.*

2. *$\mathcal{A}$ separates points of K.*

3. *$1_K \in \mathcal{A}$.*

Then $\overline{\mathcal{A}} = \mathcal{C}(K, \mathbb{R})$.

The next theorem combines the above conditions 2 and 3 of Theorem 4.3 into one condition. Recall, 'strongly separartes' was defined in Definition 2.6.

Theorem 4.4. *Stone-Weierstrass Theorem (Case : $\mathcal{C}(K, \mathbb{R})$)*
Let $\mathcal{A}$ be a collection of real-valued functions on K that forms an algebra and strongly separates points. Then any continuous real-valued function on K can be uniformly approximated by functions in $\mathcal{A}$.

In the previous theorems, we have considered only $\mathbb{K} = \mathbb{R}$. Now we look at when $\mathbb{K} = \mathbb{C}$, that is $A \subset \mathcal{C}(K, \mathbb{C})$.

Theorem 4.5. *Stone-Weierstrass Theorem (Case : $\mathcal{C}(K, \mathbb{C})$)*
Let K be a compact Hausdorff space. Suppose that $\mathcal{A} \subset \mathcal{C}(K, \mathbb{C})$ satisfies the following conditions.

1. *$\mathcal{A}$ is an algebra.*

2. $\mathcal{A}$ separates points of K.

3. $1_K \in \mathcal{A}$.

4. If $f \in \mathcal{A}$ then $\overline{f} \in \mathcal{A}$. In words, $\mathcal{A}$ is self-conjugate.

Then $\overline{\mathcal{A}} = \mathcal{C}(K, \mathbb{C})$.

We close this chapter with some remarks and examples.

Remark 4.6. If we remove the assumption that $\mathcal{A}$ separates points of K in the statement of Theorem 4.3, the resulting statement is false. Indeed, if $\mathcal{A}$ does not separate points, then there exists distinct points $x_0, y_0 \in K$ such that

$$f(x_0) = f(y_0)$$

for all $f \in \mathcal{A}$. Then

$$\overline{\mathcal{A}} \subseteq \{f \in \mathcal{C}(K, \mathbb{R}) : f(x_0) = f(y_0)\} \neq \mathcal{C}(K, \mathbb{R}).$$

Thus $\mathcal{A}$ is not dense in the set of continuous real-valued functions.

Remark 4.7. In Theorem 4.5, we claim that condition 4 of the assumptions is necessary. Let K be defined by

$$K := \{z \in \mathbb{C} : |z| \leq 1\} \quad \text{and so} \quad K^0 = \{z \in \mathbb{C} : |z| < 1\}.$$

Now let $\mathcal{A}$ be defined as follows

$$\mathcal{A} := \left\{ f \colon K \to \mathbb{C} \mid f(z) = \sum_{n=0}^{m} \alpha_n z^n, \alpha_n \in \mathbb{C}, m \in \mathbb{N} \cup \{0\} \right\}.$$

Note that

$$\mathcal{A} \subseteq \Big\{ f : K \to \mathbb{C} | f \text{ analytic on } K^0 \text{ and continuous on } K \Big\} := H.$$

Then $\mathcal{A}$ satisfies conditions 1, 2, and 3, but not 4 of Theorem 4.5 but

$$\overline{\mathcal{A}} \subseteq \overline{H} = H \subsetneq \mathcal{C}(K, \mathbb{C}).$$

Therefore $\overline{\mathcal{A}} \neq \mathcal{C}(K, \mathbb{C})$.

In 1885 Weierstrass proved the density of algebraic polynomials in the collection of continuous real-valued functions on a compact interval, and the density of trigonometric polynomials in the collection of 2π-periodic continuous real-valued functions. His approximation results were surprising given his construction, in 1861, of a continuous nowhere differentiable function, which has become known as Weierstrass' function.

The following example is also surprising for it gives that Weierstrass' continuous nowhere differentiable function can be uniformly approximated by a differentiable function whose derivative is continuous.

Example 4.8. Let our collection of real-valued functions be defined by

$$\mathcal{C}^1([a,b],\mathbb{R}) = \{f\colon\ [a,b] \to \mathbb{R} \mid f' \text{ exists and is continuous}\}.$$

We can see that $\mathcal{C}^1([a,b],\mathbb{R})$ is an algebra, it separates points of $[a,b]$ and it contains the constant functions. Thus $\overline{\mathcal{C}^1([a,b],\mathbb{R})} = \mathcal{C}([a,b],\mathbb{R})$.

Example 4.9. Let $\mathcal{A}_{2\pi}$ be the set of real-valued trigonometic polynomials in $\mathcal{C}([0,2\pi],\mathbb{R})$ as defined in Example 2.4 part 6. If $f \in \mathcal{A}_{2\pi}$, then

$$f(0) = f(2\pi).$$

Thus $\mathcal{A}_{2\pi}$ does not separate points, so by Remark 4.6,

$$\overline{\mathcal{A}_{2\pi}} \neq \mathcal{C}([0,2\pi],\mathbb{R}).$$

However, let us consider the algebra $\mathcal{A}_\pi$ of all real-valued trigonometric polynomials on $[0,\pi]$. Then $\mathcal{A}_\pi$ satisfies the conditions of Theorem 4.3, thus

$$\overline{\mathcal{A}_\pi} = \mathcal{C}([0,\pi],\mathbb{R}).$$

Example 4.10. The algebra of complex trigonometric polynomials, as defined in Example 2.4 part 7, is an algebra satisfying the assumptions of Theorem 4.5. Thus $\mathcal{A}$ is dense in $\mathcal{C}(K,\mathbb{C})$.

Chapter 5

A collection of useful lemmas and corollaries

Remark 5.1. We now collect some basic definitions and facts from undergraduate real analysis.

1. A topological space, K, is called a Hausdorf space provided for all distinct points x and y in K there exists disjoint open sets U_x and U_y with $x \in U_x$ and $y \in U_y$.

2. *Principle of Uniform Continuity*

   A function that is continuous in a closed, bounded interval is uniformly continuous in that interval.

3. A point x is a *limit point* of $\{x_n\}$ if and only if every open set that contains x contains a term from $\{x_n\}$ distinct from x.

4. *Weierstrass' Principle*

   Every bounded sequence in $\mathbb{R}$ has at least one limit point; thus, the sequence has at least one convergent subsequence.

5. A bounded non-decreasing or non-increasing sequence $\{x_n\}$ converges.

6. *Least Upper Bound Principle*

   A bounded set of real numbers has a greatest lower bound and a least upper bound.

7. *Extremum Principle for a Continuous Function*

   A continuous function on a closed interval is bounded and attains its maximum and minimum values at some points in the interval.

8. *Rolle's Theorem*

   Suppose g is continuous on an interval $[a, b]$, differentiable on (a, b) and $g(a) = g(b)$. Then there is a point c in (a, b) such that $g'(c) = 0$.

9. *Mean Value Theorem*

   Suppose $f : [a, b] \to \mathbb{R}$ is continuous on an interval $[a, b]$ and differentiable on (a, b). Then there is a point c in (a, b) such that $f'(c) = \frac{f(b)-f(a)}{b-a}$.

10. *Dini's Theorem Version 1*

    Let K be a compact Hausdorff space. Suppose that $F \subset \mathcal{C}(K, \mathbb{R})$ has the following two properties.

    (a) If $f, g \in F$, then there is an $h \in F$ such that $h \leq f \wedge g$.

    (b) The function f_0 defined by $f_0(x) := \inf\{f(x) : f \in F\}$ is real-valued and continuous.

    Then for each $\epsilon > 0$ there exists an $f \in F$ such that $\|f - f_0\|_{\mathcal{C}(K,\mathbb{R})} < \epsilon$.

11. *Dini's Theorem Version 2*

    If $\{f_n\}$ is a monotonically increasing sequence of functions from $\mathcal{C}(K, \mathbb{R})$ that converges pointwise to a continuous function f, then the sequence converges uniformly to f.

In many of our proofs of the Weierstrass approximation theorem, we begin by approximating a basic function (eg. $f(t) = |t|$ or $f(t) = \sqrt{t}$) by a polynomial. Then we use this approximation of the nice function to obtain an approximation for an arbitrary continuous function in $\mathcal{C}(K, \mathbb{K})$.

Lemma 5.2. *For each $\epsilon > 0$, there is a polynomial p such that $||t| - p(t)| < \epsilon$ for all $t \in [-1, 1]$.*

Proof. Fix $\epsilon > 0$. The Maclaurin series [cf. [8], Remark II.7.26]

$$\sum_{n=0}^{\infty} \binom{\frac{1}{2}}{n} x^n = (1+x)^{\frac{1}{2}}$$

converges absolutely and uniformly for all $x \in [-1, 1]$. We can see that

$$|t| = \left(1 + t^2 - 1\right)^{\frac{1}{2}}.$$

Let $x = t^2 - 1$ so that

$$|t| = \left(1 + t^2 - 1\right)^{\frac{1}{2}} = \sum_{n=0}^{\infty} \binom{\frac{1}{2}}{n} \left(t^2 - 1\right)^n$$

for $|t| \leq \sqrt{2}$. Again this series converges absolutely and uniformly for $|t| \leq \sqrt{2}$. Now there exists a sufficiently large $N \in \mathbb{N}$ such that

$$\left| \sum_{n=0}^{\infty} \binom{\frac{1}{2}}{n} \left(t^2 - 1\right)^n - \sum_{n=0}^{N} \binom{\frac{1}{2}}{n} \left(t^2 - 1\right)^n \right| < \epsilon$$

for $t \in [-1, 1]$. Consider the polynomial

$$p(t) := \sum_{n=0}^{N} \binom{\frac{1}{2}}{n} \left(t^2 - 1\right)^n.$$

Then

$$||t| - p(t)| < \epsilon$$

for $t \in [-1, 1]$. □

Remark 5.3. We can extend Lemma 5.2 from all $t \in [-1, 1]$ to all $t \in [-B, B]$ by a standard rescaling.

Lemma 5.4. *There is a sequence of real-valued polynomials $\{p_n\}$ defined on $[0, 1]$ such that*

1. *$p_n(t) \leq p_{n+1}(t)$ for all $t \in [0, 1]$ and for all $n \in \mathbb{N}$ and*

2. $\{p_n\}$ *converges uniformly to* $\sqrt{t}$*.*

Proof. Let p_n be defined as follows. Let $p_1 := 0$.

$$p_{n+1} := p_n(t) + \frac{1}{2}\left(t - p_n^2(t)\right) \quad \text{for } n \geq 1. \tag{5.1}$$

Claim. Let p_n be as above, then

$$p_n(t) \leq \sqrt{t} \quad \text{and} \quad p_n(t) \leq p_{n+1}(t) \tag{5.2}$$

for all $t \in [0,1]$ and for all $n \in \mathbb{N}$.

Proof of Claim. We prove the claim by induction. Let $n = 1$. We know that $p_1(t) = 0$ therefore

$$p_1(t) \leq \sqrt{t} \ \text{ for } t \in [0,1].$$

Given $p_1 = 0$ then $p_2 = \frac{1}{2}t$. So we can see that

$$p_1(t) \leq p_2(t) \ \text{ for } t \in [0,1].$$

Therefore, equation (5.2) holds for $n = 1$. Now we assume equation (5.2) is true for $n \geq 1$. We want to show equation (5.2) is true for $n+1$. Observe that

$$\sqrt{t} - p_{n+1}(t) = \sqrt{t} - p_n(t) - \frac{1}{2}\left(t - p_n^2(t)\right).$$

So then

$$\sqrt{t} - p_{n+1}(t) = \left(\sqrt{t} - p_n(t)\right)\left(1 - \frac{1}{2}\left(\sqrt{t} + p_n(t)\right)\right).$$

Thus given our inductive hypothesis, $p_n(t) \leq \sqrt{t}$, we can see that

$$\sqrt{t} - p_{n+1}(t) \geq 0,$$

$$p_{n+1}(t) \leq \sqrt{t}. \tag{5.3}$$

By equation (5.1),

$$p_{n+2}(t) - p_{n+1}(t) = \frac{1}{2}\left(t - p_{n+1}^2(t)\right).$$

So the inequality (5.3) gives us that

$$p_{n+2}(t) - p_{n+1}(t) \geq 0,$$

$$p_{n+2}(t) \geq p_{n+1}(t).$$

Therefore by induction, our Claim holds for all $t \in [0,1]$ and for all $n \in \mathbb{N}$.

Now, for all $t \in [0,1]$, the sequence $\{p_n(t)\}$ is increasing and bounded, therefore $\{p_n(t)\}$ converges to a limit $f(t)$. We can see by equation (5.1) that

$$t - f^2(t) = 0.$$

So as $f(t) \geq 0$, then

$$f(t) = \sqrt{t}.$$

Therefore as f is continuous and the sequence $\{p_n\}$ is increasing, this convergence is a uniform convergence by Dini's Theorem 5.1 part (11). □

We now collect up some elementary facts about algebras, which we will repeatedly use in the following section containing our collection of Stone-Weierstrass proofs.

Lemma 5.5. *Let $\mathcal{A}$ be an algebra of $\mathcal{C}(K,\mathbb{R})$. Let $f_1, f_2 \in \mathcal{A}$ and fix $\epsilon > 0$. Then there exists $g_1, g_2 \in \mathcal{A}$ such that*

$$\|g_1 - f_1 \vee f_2\|_{\mathcal{C}(K,\mathbb{R})} < \epsilon \tag{5.4}$$

$$\|g_2 - f_1 \wedge f_2\|_{\mathcal{C}(K,\mathbb{R})} < \epsilon. \tag{5.5}$$

Proof. Note that

$$g - [(-f_1) \vee (-f_2)] = g + (f_1 \wedge f_2) = -[(-g) - (f_1 \wedge f_2)] \ .$$

Thus, if (5.4) holds for all $f_1, f_2 \in \mathcal{A}$, then (5.5) also holds for all $f_1, f_2 \in \mathcal{A}$. Thus is suffices to show just (5.4) .

Claim. It is enough to show that if $f \in \mathcal{A}$ and $\epsilon > 0$ then there exists an $g \in A$ such that

$$\|g - |f|\|_{\mathcal{C}(K,\mathbb{R})} < \epsilon \ .$$

Proof of Claim. Fix $f_1, f_2 \in \mathcal{A}$ and $\epsilon > 0$. Then let

$$f := f_1 - f_2.$$

Note $f \in \mathcal{A}$. By assumption of the Claim, there exists $g \in \mathcal{A}$ such that

$$\|g - |f|\|_{\mathcal{C}(K,\mathbb{R})} < 2\epsilon.$$

Set

$$g_1 := \frac{g + (f_1 + f_2)}{2} \ . \tag{5.6}$$

Since $\mathcal{A}$ is an algebra, $g_1 \in \mathcal{A}$. Recall that from Definition 1.7 that

$$f_1 \vee f_2 = \frac{f_1 + f_2 + |f_1 - f_2|}{2} \ .$$

Let us compute:

$$\begin{aligned} \|g_1 - f_1 \vee f_2\|_{\mathcal{C}(K,\mathbb{R})} &= \left\| g_1 - \frac{f_1 + f_2 + |f_1 - f_2|}{2} \right\|_{\mathcal{C}(K,\mathbb{R})} \\ &= \frac{1}{2} \left\| [2g_1 - (f_1 + f_2)] - |f_1 - f_2| \right\|_{\mathcal{C}(K,\mathbb{R})} \end{aligned}$$

now substitute in g_1 from (5.6)

$$\begin{aligned} &= \frac{1}{2} \|g - |f|\|_{\mathcal{C}(K,\mathbb{R})} \\ &< \frac{1}{2} 2\epsilon = \epsilon \ . \end{aligned}$$

So the Claim holds.

Now we need to prove the statement of the Claim. Let $f \in \mathcal{A}$ and $\epsilon > 0$. We want to show that there exists a $g \in \mathcal{A}$ such that

$$\|g - |f|\|_{\mathcal{C}(K,\mathbb{R})} < \epsilon \ .$$

Let

$$B := \sup_{x \in K} |f(x)| \,.$$

Since K is compact and f is continuous, B is finite. By Remark 5.3, there exists a polynomial $p : [-B, B] \to \mathbb{R}$ such that

$$\Big| |t| - p(t) \Big| < \epsilon$$

for all $t \in [-B, B]$. So for all $x \in K$, taking $t = f(x)$,

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

Since $\mathcal{A}$ is an algebra, $g := p \circ f \in \mathcal{A}$. Thus

$$\|g - |f|\|_{\mathcal{C}(K,\mathbb{R})} < \epsilon \ .$$

Therefore Lemma 5.5 holds. □

Next we extend Lemma 5.5 to a finite number of f_i's. For this we need the following simple observation. If $a, b, c \in \mathbb{R}$, then

$$\begin{aligned} |a \vee c - b \vee c| &= \left| \frac{a + c + |a - c|}{2} - \frac{b + c + |b - c|}{2} \right| \\ &\leq \frac{1}{2} \Big[\big|(a + c) - (b + c)\big| + \Big| |a - c| - |b - c| \Big| \Big] \\ &\leq \frac{1}{2} \Big[|a - b| + |a - b| \Big] \\ &= |a - b| \ . \end{aligned}$$

Therefore

$$|a \vee c - b \vee c| \leq |a - b| \,. \tag{5.7}$$

Lemma 5.6. *Let* $\{f_i\}_{i=1}^n$ *be a finite sequence from* $\mathcal{A}$ *and let* $\epsilon > 0$*. Then there exists* $g, \widetilde{g} \in \mathcal{A}$ *such that*

$$\left\| g - \bigvee_{i=1}^n f_i \right\|_{\mathcal{C}(K,\mathbb{R})} < \epsilon \tag{5.8}$$

$$\left\| \widetilde{g} - \bigwedge_{i=1}^n f_i \right\|_{\mathcal{C}(K,\mathbb{R})} < \epsilon \ . \tag{5.9}$$

Proof. Note that

$$g - \bigvee_{i=1}^{n}(-f_i) = g + \bigwedge_{i=1}^{n} f_i = -\left[(-g) - \bigwedge_{i=1}^{n} f_i\right] .$$

Thus if (5.8) holds for all finite sequences $\{f_i\}_{i=1}^n$ from $\mathcal{A}$, then (5.9) also holds for all finite sequences $\{f_i\}_{i=1}^n$ from $\mathcal{A}$. So it suffices to show just (5.8).

We shall show Lemma 5.6 by induction on n. Lemma 5.5 shows (5.8) holds for all finite sequences from $\mathcal{A}$ with $n = 2$ elements. So assume (5.8) holds for all finite sequences from $\mathcal{A}$ with n elements for some $n \geq 2$. Next fix a finite sequence $\{f_i\}_{i=1}^{n+1}$ from $\mathcal{A}$. By inductive hypothesis, there exists $g_1 \in \mathcal{A}$ such that

$$\left\| g_1 - \bigvee_{i=1}^{n} f_i \right\|_{\mathcal{C}(K,\mathbb{R})} < \frac{\epsilon}{2} .$$

By Lemma 5.5, there exists $g \in \mathcal{A}$ such that

$$\left\| g - (g_1 \bigvee f_{n+1}) \right\|_{\mathcal{C}(K,\mathbb{R})} < \frac{\epsilon}{2} .$$

So now let us calculate using the above observation:

$$\begin{aligned} \left\| g - \bigvee_{i=1}^{n+1} f_i \right\|_{\mathcal{C}(K,\mathbb{R})} &\leq \| g - (g_1 \vee f_{n+1}) \|_{\mathcal{C}(K,\mathbb{R})} + \left\| (f_{n+1} \vee g_1) - (f_{n+1} \vee \bigvee_{i=1}^{n} f_i) \right\|_{\mathcal{C}(K,\mathbb{R})} \\ &< \frac{\epsilon}{2} + \left\| g_1 - \bigvee_{i=1}^{n} f_i \right\|_{\mathcal{C}(K,\mathbb{R})} \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon . \end{aligned}$$

Therefore $g \in \mathcal{A}$ approximates $\bigvee_{i=1}^{n+1} f_i$ within ϵ .

Thus Lemma 5.6 holds. □

Lemma 5.7. *Let $\mathcal{A}$ be an algebra of functions such that if $f \in \mathcal{A}$ then $|f| \in \mathcal{A}$. Then $\mathcal{A}$ is a lattice.*

Proof. Let $\mathcal{A}$ be an algebra as described in the statement of Lemma 5.7. Let $f, g \in \mathcal{A}$. In order to prove that $\mathcal{A}$ is a lattice we need only to show that $f \vee g \in \mathcal{A}$ and $f \wedge g \in \mathcal{A}$. By Definition 1.7, we defined $f \vee g$ by the following equation:

$$f \vee g = \frac{f + g + |f - g|}{2},$$

thus

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

Now, by Definition 2.1, we can see

$$f+g \in \mathcal{A} \quad \text{and} \quad \frac{1}{2}(f+g) \in \mathcal{A}.$$

Also, by Definition 2.1, we see that $-g \in \mathcal{A}$. Thus

$$f-g \in \mathcal{A}.$$

By our assumption, $|f-g| \in \mathcal{A}$ since $f-g \in \mathcal{A}$. So

$$\frac{1}{2}|f-g| \in \mathcal{A}.$$

Hence $f \vee g \in \mathcal{A}$.

A similar argument can be used to prove $f \wedge g \in \mathcal{A}$. Therefore, $\mathcal{A}$ is a lattice. □

The above lemma has a version using the closure of $\mathcal{A}$. The following corollary states this version.

Corollary 5.8. *Let $\mathcal{A}$ be an algebra of functions such that if $f \in \overline{\mathcal{A}}$ then $|f| \in \overline{\mathcal{A}}$. Then $\overline{\mathcal{A}}$ is a lattice.*

Proof. Let $\mathcal{A}$ be an algebra. Then, by Lemma 2.3, $\overline{\mathcal{A}}$ is an algebra. Therefore, by Lemma 5.7 and our assumption on $\mathcal{A}$, we get that $\overline{\mathcal{A}}$ is also a lattice. □

Notice the similarities of the following proof and of the Mean Value Theorem that uses Rolle's Theorem.

Lemma 5.9. *Let $\mathcal{A} \subset \mathcal{C}(K, \mathbb{R})$ satisfy the following conditions.*

a) $\mathcal{A}$ separates points of K.

b) If $f \in \mathcal{A}$ and $c \in \mathbb{R}$ then $cf \in \mathcal{A}$ and $f + c \in \mathcal{A}$.

Then for each distinct points $y, z \in K$ and for each $a, b \in \mathbb{R}$ there is an $h \in \mathcal{A}$ such that

$$h(y) = a \quad \textit{and} \quad h(z) = b. \tag{5.10}$$

Note that we are not assuming that $\mathcal{A}$ is an algebra but rather that condition b) holds.

Proof. Let $\mathcal{A} \subset \mathcal{C}(K, \mathbb{R})$ satisfy the assumptions of Lemma 5.9. Fix distinct points $y, z \in K$ and fix $a, b \in \mathbb{R}$. Since $\mathcal{A}$ separates points in K there is an $h_0 \in \mathcal{A}$ such that $h_0(y) \neq h_0(z)$. We define h to be the following.

$$h(x) := (a - b)\frac{h_0(x) - h_0(z)}{h_0(y) - h_0(z)} + b \ . \tag{5.11}$$

By part (b) of Lemma 5.9, h belongs to $\mathcal{A}$. It is easy to see h satisfies (5.10). □

Lemma 5.10. *Let K be a compact set. Let $\mathcal{A} \subset \mathcal{C}(K, \mathbb{R})$ satisfy the following conditions.*

1. *$\mathcal{A}$ is an algebra.*
2. *$\mathcal{A}$ separates points of K.*
3. *$1_K \in \mathcal{A}$.*
4. *If $f, g \in \mathcal{A}$ then $f \vee g \in \mathcal{A}$ and $f \wedge g \in \mathcal{A}$.*

Then $\overline{\mathcal{A}} = \mathcal{C}(K, \mathbb{R})$.

Proof. Let $f \in \mathcal{C}(K, \mathbb{R})$. Fix $\epsilon > 0$. We want to find a function $g \in \mathcal{A}$ such that

$$f(y) - \epsilon < g(y) < f(y) + \epsilon$$

for all $y \in K$.

For each pair of points $x, y \in K$, there exists $h_{xy} \in \mathcal{A}$ so that

$$h_{xy}(x) = f(x) \quad \text{and} \quad h_{xy}(y) = f(y) \ . \tag{5.12}$$

Indeed, if $x = y$ we can take $h_{xy}(\cdot) = f(x)1_K(\cdot)$. If $x \neq y$, then this follows from Lemma 5.9.

For now, let's fix $x \in K$. For each $y \in K$, let

$$O_{xy} := \{z \in K \colon h_{xy}(z) - f(z) < \epsilon\} \ .$$

By our hypothesis, $x, y \in O_{xy}$. So

$$K = \bigcup_{y \in K} O_{xy} \ .$$

Since f and h_{xy} are continuous, O_{xy} is open. Since K is compact, there exists $y_1, y_2, ..., y_n \in K$ such that $\cup_{i=1}^n O_{xy_i}$ is a finite subcovering of K. Let

$$g_x(\cdot) := h_{xy_1}(\cdot) \wedge h_{xy_2}(\cdot) \wedge ... \wedge h_{xy_n}(\cdot) \, .$$

Then $g_x \in \mathcal{A}$ by condition 4 of Lemma 5.10. Clearly, $g_x(x) = f(x)$ by (5.12). Also

$$g_x(z) < f(z) + \epsilon \tag{5.13}$$

for all $z \in K$; indeed, each $z \in K$ is in some O_{xy_i}.

Now, for each $x \in K$, define V_x by

$$V_x := \{z \in K \colon f(z) - \epsilon < g_x(z)\} \ .$$

Note that

$$\bigcup_{x \in K} V_x = K$$

since, for each $x \in K$, we know that $f(x) = g_x(x)$ and so $x \in V_x$. Also, each V_x is open since f and g_x are continuous. Thus since K is compact, there exists $x_1, x_2, ..., x_m \in K$ such that $\cup_{j=1}^m V_{x_j}$ is a finite subcovering of K. Let

$$g(\cdot) := g_{x_1}(\cdot) \vee g_{x_2}(\cdot) \vee ... \vee g_{x_m}(\cdot) \, .$$

Then, by condition 4 of Lemma 5.10, $g \in \mathcal{A}$.

Consider a $z \in K$. Then z is in some V_{x_i} and so

$$f(z) - \epsilon < g_{x_i}(z) \le g(z) \ .$$

And by (5.13)

$$g(z) < f(z) + \epsilon \ .$$

Therefore $\|f - g\|_{\mathcal{C}(K,\mathbb{R})} < \epsilon$, i.e. $\overline{\mathcal{A}} = \mathcal{C}(K, \mathbb{R})$. □

The assumption on $\mathcal{A}$ in Lemma 5.10 can be weakened; indeed, the full strength of an algebra is not needed for the conclusion to hold, as we will show in the following theorem.

Theorem 5.11 (Kakutani-Krein Theorem)**.** *Let $\mathcal{L} \subset \mathcal{C}(K, \mathbb{R})$ satisfy the following conditions.*

a) $\mathcal{L}$ is a lattice.

b) $\mathcal{L}$ separates points of K.

c) $cf \in \mathcal{L}$ and $f + c \in \mathcal{L}$ provided $f \in \mathcal{L}$ and $c \in \mathbb{R}$.

Then $\overline{\mathcal{L}} = \mathcal{C}(K, \mathbb{R})$.

Proof. Let $\mathcal{L} \subset \mathcal{C}(K, \mathbb{R})$ satisfy the assumptions of Theorem 5.11. Fix $g \in \mathcal{C}(K, \mathbb{R})$ and define

$$\mathcal{L}_g := \{f \in \mathcal{L} : g(x) \le f(x) \text{ for each } x \in K\} \, .$$

We will use Dini's Theorem Version 1, as stated in Remark 5.1 part(10), applied to $F := \mathcal{L}_g$.

First note that $\mathcal{L}_g$ satisfies the condition (a) of Dini's Theorem Version 1; indeed, if $f_1, f_2 \in \mathcal{L}_g$, then $h := f_1 \wedge f_2 \in \mathcal{L}_g$. Condition (b) of Dini's Theorem Version 1 follows directly from the below Claim (with $f_0 := g$).

Claim. Fix $x \in K$. Then

$$g(x) = \inf \{f(x) : f \in \mathcal{L}_g\}.$$

Proof of Claim. Let $x \in K$ and $\epsilon > 0$. Now we will construct an $f_x \in \mathcal{L}_g$ such that

$$f_x(x) = g(x) + \epsilon. \tag{5.14}$$

First, consider the open set O defined by

$$O := \{y : g(y) < g(x) + \epsilon\}.$$

For each $z \in O^{\mathrm{c}}$, we can find an $h_z \in \mathcal{L}$, by Lemma 5.9, such that

$$h_z(z) = g(z) + \epsilon \quad \text{and} \quad h_z(x) = g(x) + \frac{\epsilon}{2}. \tag{5.15}$$

Now define V_z by

$$V_z := \{y \in K : h_z(y) > g(y)\}.$$

Note that V_z is open (since h_z and g are continuous) and that $z \in V_z$ (by (5.15)). Thus

$$O^{\mathrm{c}} \subseteq \cup_{z \in O^{\mathrm{c}}} V_z.$$

Now O^{c} is a compact set so there exists points $z_1, z_2, ...z_n \in O^{\mathrm{c}}$ such that $\cup_{j=1}^{n} V_{z_j}$ is a finite subcovering of O^{c}.

Let

$$f_x(\cdot) := (g(x) + \epsilon) 1_K(\cdot) \vee h_{z_1}(\cdot) \vee h_{z_2}(\cdot) \vee ... \vee h_{z_n}(\cdot).$$

By (5.15),

$$f_x(x) = (g(x) + \epsilon) \vee \left(g(x) + \frac{\epsilon}{2}\right) \vee ... \vee \left(g(x) + \frac{\epsilon}{2}\right) = g(x) + \epsilon.$$

Thus (5.14) holds.

Next we show that $f_x \in \mathcal{L}_g$. By part (c) of Theorem 5.11, $\mathcal{L}$ contains the constant function $(g(x) + \epsilon) 1_K(\cdot)$. Also recall $h_z \in \mathcal{L}$ and so by Definition 1.8 we see that

$f_x \in \mathcal{L}$. Now if $y \in O^c$, then $y \in V_{z_j}$ for some $j \in \{1, 2, ..., n\}$. Thus as a result, by the definition of V_{z_j}, we have

$$f_x(y) \geq h_{z_j}(y) > g(y).$$

And if $y \in O$, then by the definition of O,

$$f_x(y) \geq g(x) + \epsilon > g(y).$$

Thus $f_x \in \mathcal{L}_g$. So, by definition of infimum, $g(x) = \inf\{f(x) : f \in \mathcal{L}_g\}$. Thus the Claim holds.

So we can apply Dini's Theorem Version 1 to get

$$g \in \overline{\mathcal{L}_g}\ .$$

Therefore $\overline{\mathcal{L}} = \mathcal{C}(K, \mathbb{R})$. □

Theorem 5.12. *Let $f \in L_\infty(\mathbb{R}, \mathbb{K})$ be continuous on an open set U of $\mathbb{R}$. Let K be a compact subset of U. Let φ be an approximate identity and $\{\varphi_\epsilon\}_{\epsilon>0}$ be as in Remark 1.14. Then $f * \varphi_\epsilon$ converges uniformly on K to f as $\epsilon \to 0^+$, i.e.*

$$\lim_{\epsilon \to 0^+} \sup_{t \in K} |(f * \varphi_\epsilon)(t) - f(t)| = 0.$$

Proof. Let $f, U, K, \varphi, \varphi_\epsilon$ be as in the assumptions of Theorem 5.12. Thus

$$\varphi \in L_1(\mathbb{R}, \mathbb{K})$$
$$\int_{\mathbb{R}} \varphi(t)dt = 1$$
$$\varphi_\epsilon(t) := \frac{1}{\epsilon}\varphi\left(\frac{t}{\epsilon}\right)$$

for $\epsilon > 0$. Without loss of generality, $\|f\|_\infty \neq 0$. Fix $\eta > 0$. Since $\varphi \in L_1$, there exists $M > 0$ such that

$$\int_{|s|>M} |\varphi(s)|\, ds \leq \frac{\eta}{4\|f\|_\infty}. \tag{5.16}$$

Claim 1.

$$\operatorname{dist}(K, U^c) := \inf\{|k - v| : k \in K, v \in U^c\} > 0$$

Proof: Observe $K \subset U$ so $K \cap U^c = \varnothing$. We will prove Claim 1 by contradiction. Assume not. Then there exists a $k_n \in K$ and $v_n \in U^c$ such that $\lim_{n\to\infty} |k_n - v_n| = 0$. Now K is compact, so there exists a subsequence $\left\{k_{n_j}\right\}_j$ of $\{k_n\}_n$ and $k \in K$ such that $\lim_{j\to\infty} \left|k_{n_j} - k\right| = 0$. Then for all $j \in \mathbb{N}$

$$\left|k - v_{n_j}\right| \leq \left|k - k_{n_j}\right| + \left|k_{n_j} - v_{n_j}\right|.$$

So $\lim_{j\to\infty} \left|v_{n_j} - k\right| = 0$. Since $v_{n_j} \in U^c$ and U^c is closed because U is open, then $k \in U^c$. This is a contradiction. Therefore Claim 1 is true.

Fix $\epsilon_1 > 0$ such that

$$M\epsilon_1 < \operatorname{dist}(K, U^c).$$

Let's define K_1 by the following

$$K_1 := K + \overline{B_{M\epsilon_1}(0)} = \{k + z : k \in K, |z| \leq M\epsilon_1\}.$$

Clearly, $K \subseteq K_1$. Note also that $K_1 \subset U$. For indeed, let $k \in K$ and $z \in \overline{B_{M\epsilon_1}(0)}$. We want to show that $k + z \in U$. Assume $k + z \notin U$. Then

$$M\epsilon_1 < \operatorname{dist}(K, U^c) \leq |(k + z) - k| = |z| \leq M\epsilon_1.$$

This is a contradiction. Thus $k + z \in U$. Therefore $K_1 \subset U$. So $K \subseteq K_1 := K + \overline{B_{M\epsilon_1}(0)} \subseteq U$.

Claim 2. K_1 is compact.

Proof: We shall show, as is sufficient, that K_1 is sequentially compact. Fix $\{k_n + z_n\}_{n=1}^{\infty}$ from K_1 where $k_n \in K$ and $z_n \in \overline{B_{M\epsilon_1}(0)}$. Now K is compact, so there exists a $k \in K$ and a subsequence $\left\{k_{n_j}\right\}_{j=1}^{\infty}$ of $\{k_n\}_{n=1}^{\infty}$ such that $k_{n_j} \to k \in K$. Also, $\overline{B_{M\epsilon_1}(0)}$ is compact because $\overline{B_{M\epsilon_1}(0)}$ is closed and bounded. Thus there exists $z \in \overline{B_{M\epsilon_1}(0)}$ and a subsequence $\left\{z_{n_{j_i}}\right\}_{i=1}^{\infty}$ of $\left\{z_{n_j}\right\}_{j=1}^{\infty}$ such that $z_{n_{j_i}} \to z \in \overline{B_{M\epsilon_1}(0)}$.

Thus $k_{n_{j_i}} + z_{n_{j_i}} \to k + z$, and so $k + z \in K_1$. Therefore K_1 is compact. So Claim 2 holds.

Because K_1 is compact, f is uniformly continuous on K_1. So there exists $\epsilon_2 \in (0, \epsilon_1)$ such that

$$\sup\nolimits_{t\in K}\sup\nolimits_{|u|<M\epsilon_2} |f(t) - f(t-u)| < \frac{\eta}{2\|\varphi\|_1}. \tag{5.17}$$

Now, if $0 < \epsilon < \epsilon_2$, then

$$\begin{aligned}
&\sup_{t\in K} |f(t) - (f * \varphi_\epsilon)(t)| \\
&\qquad = \sup_{t\in K} \left| f(t) - \int_{\mathbb{R}} f\left(t - \widetilde{s}\right) \varphi_\epsilon\left(\widetilde{s}\right) d\widetilde{s} \right| \\
&\qquad = \sup_{t\in K} \left| \int_{\mathbb{R}} \left[f(t) - f\left(t - \widetilde{s}\right) \right] \frac{1}{\epsilon} \varphi\left(\frac{\widetilde{s}}{\epsilon}\right) d\widetilde{s} \right|
\end{aligned}$$

let $s = \frac{\widetilde{s}}{\epsilon}$

$$\begin{aligned}
&\le \sup_{t\in K} \int_{\mathbb{R}} |f(t) - f(t - \epsilon s)| \, |\varphi(s)| \, ds \\
&\le \sup_{t\in K} \left[2\|f\|_\infty \int_{|s|>M} |\varphi(s)| \, ds + \sup_{|s|\le M} |f(t) - f(t-\epsilon s)| \int_{|s|\le M} |\varphi(s)| \, ds \right] \\
&= 2\|f\|_\infty \int_{|s|>M} |\varphi(s)| \, ds + \sup_{t\in K} \sup_{|s|\le M} |f(t) - f(t-\epsilon s)| \int_{|s|\le M} |\varphi(s)| \, ds.
\end{aligned}$$

By inequality (5.16) and inequality (5.17)

$$\sup_{t\in K} |f(t) - (f * \varphi_\epsilon)(t)| \le 2\|f\|_\infty \frac{\eta}{4\|f\|_\infty} + \frac{\eta}{2\|\varphi\|_1} \cdot \|\varphi\|_1 = \eta.$$

Therefore $f * \varphi_\epsilon$ converges uniformly to f on K. □

Notice the similarities between the proofs of the above theorem and the following theorem. They both use the concept of splitting the domain into two pieces.

Theorem 5.13. *Let $f \in L_\infty(\mathbb{R}, \mathbb{K})$ be continuous on an open set U of $\mathbb{R}$ and K be a compact subset of U. Let $\{K_n\}_n$ be a Dirac sequence. Then $\{f * K_n\}_{n=1}^\infty$ converges to f uniformly on K.*

Proof. Let f, U, K, and $\{K_n\}_n$ be as in the assumption of Theorem 5.13.

As shown in Claim 1 of the proof of Theorem 5.12

$$0 < \mathrm{d}\left(K, U^{\mathrm{c}}\right) \ .$$

Fix $\delta_1 > 0$ such that

$$\delta_1 < \mathrm{d}\left(K, U^{\mathrm{c}}\right) \ .$$

Let's define K_1 by

$$K_1 := K + \overline{B_{\delta_1}(0)} = \{k + z : k \in K, |z| \leq \delta_1\} .$$

Clearly, $K \subseteq K_1$. Note also that $K_1 \subset U$. For indeed, let $k \in K$ and $z \in \overline{B_{\delta_1}(0)}$. We want to show that $k + z \in U$. Assume $k + z \notin U$. Then

$$\delta_1 < \mathrm{d}\left(K, U^{c}\right) \leq |(k+z) - k| = |z| \leq \delta_1 \ .$$

This is a contradiction. Thus $k + z \in U$.

So we now have

$$K \subseteq K_1 := K + \overline{B_{\delta_1}(0)} \subseteq U.$$

Also, as shown in Claim 2 of Theorem 5.12, K_1 is compact.

Fix $\epsilon > 0$. Using Definition 1.15 and Definition 1.12 property (2), we will observe that, for each $x \in \mathbb{R}$,

$$f(x) = f(x) \int_{\mathbb{R}} K_n(t)\, dt = \int_{\mathbb{R}} f(x)\, K_n(t)\, dt;$$

thus

$$(f * K_n)(x) - f(x) = \int_{\mathbb{R}} \left[f(x-t) - f(x)\right] K_n(t)\, dt. \tag{5.18}$$

Since f is uniformly continuous on K_1, there exists a $\delta \in (0, \delta_1)$ such that if $x \in K$ and $|t| < \delta$ then $x - t \in K_1$ and

$$|f(x-t) - f(x)| < \frac{\epsilon}{2}.$$

Without loss of generality, $\|f\|_\infty \neq 0$. By Definition 1.12 property (3), there exists $N \in \mathbb{N}$ such that if $n \geq N$,

$$\int_{|t|\geq\delta} K_n(t)\,dt < \frac{\epsilon}{4\|f\|_\infty}.$$

Now, by equation (5.18), we have for all $x \in K$

$$|(f * K_n)(x) - f(x)| \leq \int_{|t|\geq\delta} |f(x-t) - f(x)|\, K_n(t)\,dt + \int_{|t|<\delta} |f(x-t) - f(x)|\, K_n(t)\,dt.$$

Then, for each $n \geq N$ and $x \in K$,

$$\begin{aligned} |(f * K_n)(x) - f(x)| &\leq 2\|f\|_\infty \int_{|t|\geq\delta} K_n(t)\,dt + \frac{\epsilon}{2}\int_{|t|<\delta} K_n(t)\,dt \\ &< 2\|f\|_\infty \frac{\epsilon}{4\|f\|_\infty} + \frac{\epsilon}{2}\int_{\mathbb{R}} K_n(t)\,dt \\ &= \frac{\epsilon}{2} + \frac{\epsilon}{2}\cdot 1 = \epsilon. \end{aligned}$$

Therefore $f * K_n$ converges uniformly to f on K. □

The following two lemmas will be used in the Stone-Weierstrass Theorem to reduce the $\mathbb{K} = \mathbb{C}$ case to the $\mathbb{K} = \mathbb{R}$ case.

Lemma 5.14. *Let $\mathcal{A} \subset \mathcal{C}(K, \mathbb{C})$ be a self-conjugate algebra and let $\Re\mathcal{A}$ denote the set of real parts of functions in $\mathcal{A}$, i.e.*

$$\Re\mathcal{A} = \left\{ \frac{f + \overline{f}}{2} : f \in \mathcal{A} \right\}.$$

Then $\Re\mathcal{A} \subset \mathcal{A}$ and $\Re\mathcal{A}$ is an algebra in $\mathcal{C}(K, \mathbb{R})$ over $\mathbb{R}$.

Proof. Let $\mathcal{A} \subset \mathcal{C}(K, \mathbb{C})$ be a self-conjugate algebra. Let $h \in \Re\mathcal{A}$. Then

$$h = \frac{f + \overline{f}}{2}$$

for some $f \in \mathcal{A}$. Now, by Definition 2.1, $\frac{1}{2}f \in \mathcal{A}$. Also, by Definition 2.9 and Definition 2.1, $\frac{1}{2}\overline{f} \in \mathcal{A}$. Thus, $h \in \mathcal{A}$ and therefore $\Re\mathcal{A} \subset \mathcal{A}$. Now let $j, k \in \Re\mathcal{A}$ and $c \in \mathbb{R}$. Then for some $f, g \in \mathcal{A}$

$$j = \frac{f + \overline{f}}{2}$$

and

$$k = \frac{g + \overline{g}}{2}.$$

First, we need to show that $j + k \in \Re\mathcal{A}$. Now

$$\begin{aligned} j + k &= \frac{f + \overline{f}}{2} + \frac{g + \overline{g}}{2} \\ &= \frac{f + g + \overline{f} + \overline{g}}{2} \\ &= \frac{(f+g) + \left(\overline{f+g}\right)}{2}. \end{aligned}$$

Notice, $f + g \in \mathcal{A}$ by Definition 2.1. Thus $j + k \in \Re\mathcal{A}$. Second, we need to show that $jk \in \Re\mathcal{A}$.

$$\begin{aligned} jk &= \left(\frac{f + \overline{f}}{2}\right)\left(\frac{g + \overline{g}}{2}\right) \\ &= \frac{1}{2}\left(f\left(\frac{g + \overline{g}}{2}\right) + \overline{f}\left(\frac{g + \overline{g}}{2}\right)\right) \\ &= \frac{f\left(\frac{g+\overline{g}}{2}\right) + \overline{f\left(\frac{\overline{g}+g}{2}\right)}}{2}. \end{aligned}$$

Notice, $f\left(\frac{g+\overline{g}}{2}\right) \in \mathcal{A}$ by Definition 2.1 and the assumption that $\mathcal{A}$ is self-conjugate. Thus $jk \in \Re\mathcal{A}$. Lastly, we need to show that $cj \in \Re\mathcal{A}$.

$$\begin{aligned} cj &= c\left(\frac{f + \overline{f}}{2}\right) \\ &= \frac{cf + c\overline{f}}{2} \\ &= \frac{cf + \overline{cf}}{2}. \end{aligned}$$

Again, notice that $cf \in \mathcal{A}$ by Definition 2.1. Thus $cj \in \Re\mathcal{A}$. Therefore, by Definition 2.1, $\Re\mathcal{A}$ is an algebra in $\mathcal{C}(K, \mathbb{R})$. □

Lemma 5.15. *Let $\mathcal{A} \subset \mathcal{C}(K, \mathbb{C})$ satisfy the conditions of Theorem 4.5. Let $\Re\mathcal{A}$ denote the set of real parts of functions in $\mathcal{A}$, ie. $\Re\mathcal{A} = \left\{\frac{f+\overline{f}}{2} : f \in \mathcal{A}\right\}$. Then $\Re\mathcal{A}$ separates points of K.*

Proof. Let w, z be two distinct elements of K. By condition (2) of Theorem 4.5, there is an $f \in \mathcal{A}$ such that $f(w) \neq f(z)$. Then

$$\Re f(w) \neq \Re f(z) \quad \text{or} \quad \Im f(w) \neq \Im f(z) .$$

Since $if = i\Re f - \Im f$

$$-\Im f(w) = \Re(if(w)) \quad \text{and} \quad -\Im f(z) = \Re(if(z)) .$$

Now we have that

$$\Re f(w) \neq \Re f(z) \quad \text{or} \quad \Re(if(w)) \neq \Re(if(z)) .$$

By Definition 2.1, if $\mathcal{A}$ is an algebra, then $if \in \mathcal{A}$. Therefore $\Re\mathcal{A}$ separates points in K. □

The following lemma will be used to easily extend $K = [0, 1]$ in Theorem 4.1 to $K = [a, b]$ in Theorem 4.2.

Lemma 5.16. *Assume Theorem 4.1 holds. Then Theorem 4.2 holds for all intervals* $[a, b]$.

Proof. Assume that Theorem 4.1 holds and let $[a, b]$ be an interval of $\mathbb{R}$. Fix $f \in \mathcal{C}([a, b], \mathbb{R})$ and $\epsilon > 0$. We want to find a polynomial $p : [a, b] \to \mathbb{R}$ such that

$$\|f - p\|_{\mathcal{C}([a,b],\mathbb{R})} < \epsilon .$$

Define $h : [a, b] \to [0, 1]$ by

$$h(x) = \frac{x - a}{b - a} .$$

Clearly h is a well-defined continuous bijection and $f \circ h^{-1} \in \mathcal{C}([0, 1], \mathbb{R})$. By our assumption, there exists a polynomial $g \in \mathcal{C}\,[0, 1], \mathbb{R})$ such that

$$\left\|g - f \circ h^{-1}\right\|_{\mathcal{C}([0,1],\mathbb{R})} < \epsilon .$$

Let

$$p := g \circ h \ .$$

Then $p : [a,b] \to \mathbb{R}$ is a polynomial and

$$\begin{aligned}
\|p - f\|_{\mathcal{C}([a,b],\mathbb{R})} &= \left\|(g \circ h - f) \circ h^{-1}\right\|_{\mathcal{C}([0,1],\mathbb{R})} \\
&= \left\|g - f \circ h^{-1}\right\|_{\mathcal{C}([0,1],\mathbb{R})} \\
&< \epsilon \ .
\end{aligned}$$

So the Lemma holds. □

Chapter 6

A Collection of Approximation Theorem Proofs

Now we can begin proving the various versions of the Stone-Weierstrass Theorem. We start out with the Weierstrass Approximation Theorem as originally stated by Karl Weierstrass. (In the following section we will consider his original work.)

The following proof (cf. [9]) of Theorem 4.1 will use convolutions (see Definition 1.15) and Dirac sequences (see Definition 1.12).

The Weierstrass Approximation Theorem 4.1 (Case: $\mathcal{C}\left([0,1],\mathbb{R}\right)$)

Let f be a continuous real-valued function on the closed interval $[0,1]$. Then f can be uniformly approximated by polynomials on $[0,1]$. That is given any $\epsilon > 0$, there is a polynomial P such that $|f(x) - P(x)| < \epsilon$ for each $x \in [0,1]$.

Proof. Let $f \in \mathcal{C}\left([0,1],\mathbb{R}\right)$.

Claim 1. Let $h : [0,1] \to \mathbb{R}$ be given by

$$h(x) := f(x) - f(0) - x\left[f(1) - f(0)\right].$$

Then h is continuous and $h(0) = 0 = h(1)$. Also, if h can be uniformly approximated by polynomials, then f can be uniformly approximated by polynomials.

Proof: Clearly, h is continuous and $h(0) = 0 = h(1)$. Assume h can be uniformly approximated by polynomials. Notice

$$f(x) = h(x) + f(0) + xC$$

where $f(1) - f(0) := C$. Since h can be approximated by polynomials and $f(0)$ and C are constants, f can be uniformly approximated by polynomials on $[0,1]$. So Claim 1 holds.

By Claim 1, we can now assume that $f(0) = f(1) = 0$. For $n \in \mathbb{N}$, let $c_n > 0$ and define $K_n : \mathbb{R} \to \mathbb{R}$ by

$$K_n(t) := \frac{(1-t^2)^n}{c_n} \qquad \text{if } |t| \le 1$$

and

$$K_n(t) := 0 \qquad\qquad \text{if } |t| > 1.$$

Then $K_n(t) \ge 0$ for all t and K_n is continuous. Thus the sequence $\{K_n\}$ satisfies property (1) of Definition 1.12. Now pick c_n so that property (2) of Definition 1.12 is satisfied, that is

$$\int_{\mathbb{R}} K_n(t)\, dt = 1.$$

So the following hold.

$$\begin{aligned} \int_{[-1,1]} \frac{(1-t^2)^n}{c_n} dt &= 1 \\ \frac{1}{c_n} \int_{[-1,1]} \left(1-t^2\right)^n dt &= 1 \\ \int_{[-1,1]} \left(1-t^2\right)^n dt &= c_n \ . \end{aligned}$$

Note K_n is even.

Claim 2. Let $\{K_n\}_{n=1}^{\infty}$ be as above. Then $\{K_n\}_{n=1}^{\infty}$ satisfies property (3) of Definition 1.12.

Proof: Recall $c_n = \int_{[-1,1]} (1-t^2)^n \, dt$. So

$$\frac{c_n}{2} = \int_{[0,1]} \left(1-t^2\right)^n dt = \int_{[0,1]} (1+t)^n (1-t)^n \, dt \ge \int_{[0,1]} (1-t)^n \, dt = \frac{1}{n+1}.$$

Then $\frac{c_n}{2} \ge \frac{1}{n+1}$, thus $c_n \ge \frac{2}{n+1}$. Fix $\delta \in (0,1)$. Then

$$\begin{aligned} \int_{[\delta,1]} K_n(t)\, dt &= \int_{[\delta,1]} \frac{(1-t^2)^n}{c_n} dt \\ &\le \int_{[\delta,1]} \frac{n+1}{2} \left(1-\delta^2\right)^n dt \\ &\le \frac{n+1}{2} \left(1-\delta^2\right)^n (1-\delta) \\ &\le \frac{n+1}{2} \left(1-\delta^2\right)^n \ . \end{aligned}$$

Let $r = (1 - \delta^2)$. Observe $0 < r < 1$, so $\lim_{n\to\infty} \frac{n+1}{2} r^n = 0$. Therefore Claim 2 is true.

Thus $\{K_n\}_{n=1}^{\infty}$ is a Dirac sequence as in Definition 1.12. Extend f to the whole of $\mathbb{R}$ by setting f equal to 0 outside of $[0,1]$. By Theorem 5.13, $\{f * K_n\}$ converges to f uniformly on $[0,1]$. We only need to show that $(f * K_n)$ is a polynomial. Note that

$$(f * K_n)(x) = \int_{\mathbb{R}} f(t) K_n(x-t)\, dt = \int_{[0,1]} f(t) K_n(x-t)\, dt \; .$$

We can see that $K_n(x-t) = \frac{\left(1-(x-t)^2\right)^n}{c_n}$ is a polynomial in t and x and can be written in the form

$$K_n(x-t) = p_0(t) + p_1(t)\, x + ... + p_{2n}(t)\, x^{2n}$$

where p_0,...,p_{2n} are polynomials in t. Then

$$(f * K_n)(x) = a_0 + a_1 x + ... + a_{2n} x^{2n},$$

where the coefficients a_i are expressed as integrals of the form

$$a_i = \int_{[0,1]} f(t)\, p_i(t)\, dt$$

for $0 \leq i \leq 2n$. Therefore f can be uniformly approximated by polnomials on $[a,b]$. □

The following proof (cf. [7]) uses the Bernstein Polynomials as defined in Definition 1.17.

<u>**The Weierstrass Approximation Theorem 4.1**</u> (Case: $\mathcal{C}([0,1], \mathbb{R})$)

Let $f : [0,1] \to \mathbb{R}$ be a continuous function. Then for each $\epsilon > 0$, there exists a real-valued polynomial function P such that for all $x \in [0,1]$, $|f(x) - P(x)| < \epsilon$. Equivalently, for any such f, there exists a sequence $\{P_n\}_{n=1}^{\infty}$ of real-valued polynomials such that P_n converges to f uniformly on $[0,1]$.

Proof. Let $f \in \mathcal{C}([0,1], \mathbb{R})$ and let $\{B_n\}_{n=1}^{\infty}$ be the Bernstein polynomials as defined in Definition 1.17. We want to show that $\{B_n\}_{n=1}^{\infty}$ converges uniformly to f on $[0,1]$.

First, we need some preliminary computations. Let $p, q \in \mathbb{R}$ and $n \in \mathbb{N}$. Then by the binomial theorem

$$\sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} = (p+q)^n \ . \tag{6.1}$$

Differentiating with respect to p we obtain

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

Thus we have

$$\sum_{k=0}^{n} \frac{k}{n} \binom{n}{k} p^k q^{n-k} = p\,(p+q)^{n-1} \,. \tag{6.2}$$

Differentiating again

$$\sum_{k=0}^{n} \frac{k^2}{n} \binom{n}{k} p^{k-1} q^{n-k} = p\,(n-1)\,(p+q)^{n-2} + (p+q)^{n-1}$$

$$\sum_{k=0}^{n} \frac{k^2}{n^2} \binom{n}{k} p^k q^{n-k} = p^2 \left(1 - \frac{1}{n}\right) (p+q)^{n-2} + \frac{p}{n}\,(p+q)^{n-1} \,. \tag{6.3}$$

Next, let $x \in [0,1]$ and set $p = x, q = 1 - x$. Then by equation (6.1)

$$\sum_{k=0}^{n} \binom{n}{k} x^k \,(1-x)^{n-k} = 1. \tag{6.4}$$

Also, by equation (6.2),

$$\sum_{k=0}^{n} \frac{k}{n} \binom{n}{k} x^k \,(1-x)^{n-k} = x, \tag{6.5}$$

and by equation (6.3),

$$\sum_{k=0}^{n} \frac{k^2}{n^2} \binom{n}{k} x^k \,(1-x)^{n-k} = x^2 \left(1 - \frac{1}{n}\right) + \frac{x}{n} \ . \tag{6.6}$$

From these equations, we can simplify the following sum.

$$
\begin{aligned}
&\sum_{k=0}^{n}\left(\frac{k}{n}-x\right)^{2}\binom{n}{k}x^{k}\left(1-x\right)^{n-k}\\
&=\sum_{k=0}^{n}\left(\frac{k^{2}}{n^{2}}-2x\frac{k}{n}+x^{2}\right)\binom{n}{k}x^{k}\left(1-x\right)^{n-k}\\
&=\sum_{k=0}^{n}\frac{k^{2}}{n^{2}}\binom{n}{k}x^{k}\left(1-x\right)^{n-k}-2x\sum_{k=0}^{n}\frac{k}{n}\binom{n}{k}x^{k}\left(1-x\right)^{n-k}+x^{2}\sum_{k=0}^{n}\binom{n}{k}x^{k}\left(1-x\right)^{n-k}\\
&=x^{2}\left(1-\frac{1}{n}\right)+\frac{x}{n}-2x\cdot x+x^{2}\cdot 1\\
&=x^{2}-\frac{x^{2}}{n}+\frac{x}{n}-2x^{2}+x^{2}\\
&=x\frac{1-x}{n}\ .
\end{aligned}
$$

So

$$\sum_{k=0}^{n}\left(\frac{k}{n}-x\right)^{2}\binom{n}{k}x^{k}\left(1-x\right)^{n-k}=x\frac{1-x}{n}$$

for $x\in[0,1]$.

Since f is continuous and $[0,1]$ is compact, $f\left([0,1]\right)$ is compact. Thus, there exist a real number $M>0$ such that

$$\left|f\left(x\right)\right|\leq M$$

for all $x\in[0,1]$. Fix $\epsilon>0$. Since f is uniformly continuous on $[0,1]$ there exists $\delta>0$ such that if $|x-y|<\delta$ then

$$\left|f\left(x\right)-f\left(y\right)\right|<\frac{\epsilon}{2}$$

for all $x,y\in[0,1]$. Fix $x\in[0,1]$. Now observe, by equation (6.4) and Definition 1.17, that

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

When $\left|x - \frac{k}{n}\right| < \delta$ we have

$$\left|f(x) - f\left(\frac{k}{n}\right)\right| < \frac{\epsilon}{2}.$$

Now when $\left|x - \frac{k}{n}\right| \geq \delta$, then

$$\left(x - \frac{k}{n}\right)^2 \geq \delta^2$$
$$\frac{\left(x - \frac{k}{n}\right)^2}{\delta^2} \geq 1$$

and so

$$\left|f(x) - f\left(\frac{k}{n}\right)\right| \leq 2M \leq 2M \frac{\left(x - \frac{k}{n}\right)^2}{\delta^2}\ .$$

Therefore we continue with the above expression to see that

$$\begin{aligned}
|f(x) - B_n(f)(x)| &\leq \sum_{k=0}^{n} \left(\frac{\epsilon}{2} + 2M \frac{\left(x - \frac{k}{n}\right)^2}{\delta^2} \right) \binom{n}{k} x^k (1-x)^{n-k} \\
&= \frac{\epsilon}{2} + \frac{2M}{\delta^2} \sum_{k=0}^{n} \left(x - \frac{k}{n}\right)^2 \binom{n}{k} x^k (1-x)^{n-k} \\
&= \frac{\epsilon}{2} + \frac{2M}{\delta^2} \frac{x(1-x)}{n} \\
&< \frac{\epsilon}{2} + \frac{2M}{\delta^2 n}
\end{aligned}$$

because $x(1-x) < 1$. Now choose $n \in \mathbb{N}$ large enough such that $n \geq \dfrac{4M}{\delta^2 \epsilon}$. Then

$$|f(x) - B_n(f)(x)| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon\ .$$

Therefore the Bernstein polynomials, $\{B_n\}_{n=1}^{\infty}$, converge uniformly to f on $[0, 1]$. $\square$

The following proof (cf. [10]) of Theorem 4.3 will use the Kakutani-Krein Theorem, which we presented as Theorem 5.11.

<u>**Stone-Weierstrass Theorem 4.3**</u> (Case: $\mathcal{C}(K, \mathbb{R})$)

Let K be a compact Hausdorff space. Suppose that $\mathcal{A} \subset \mathcal{C}(K, \mathbb{R})$ satisfies the following conditions.

1. $\mathcal{A}$ is an algebra.

2. $\mathcal{A}$ separates points of K.

3. $1_K \in \mathcal{A}$.

Then $\overline{\mathcal{A}} = \mathcal{C}(K, \mathbb{R})$.

Proof. Let $\mathcal{A} \subset \mathcal{C}(K, \mathbb{R})$ satisfy the assumptions of Theorem 4.3. Since $\mathcal{A}$ is an algebra of functions, so is $\overline{\mathcal{A}}$, as proved in Lemma 2.3. If we can prove that $\overline{\mathcal{A}}$ is also a lattice, then the conclusion will hold as a result of Theorem 5.11. By Lemma 5.7, to prove $\overline{\mathcal{A}}$ is a lattice, it suffices to show that if

$$f \in \overline{\mathcal{A}} \quad \text{then} \quad |f| \in \overline{\mathcal{A}}.$$

Fix $f \in \overline{\mathcal{A}}$. If $f \equiv 0$, then the above is trivial. So assume $f \not\equiv 0$, then $\|f\|_{\mathcal{C}(K,\mathbb{R})} \neq 0$. Let g be the normalization of f, i.e.

$$g(\cdot) := \frac{f(\cdot)}{\|f\|_{\mathcal{C}(K,\mathbb{R})}} .$$

Note that $g \in \overline{\mathcal{A}}$. Fix $\epsilon > 0$. Now we can apply Lemma 5.2 to obtain a polynomial p such that

$$||t| - p(t)| < \epsilon$$

for all $t \in [-1, 1]$. For each $x \in K$ we have that $g(x) \in [-1, 1]$ and so

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

Thus

$$\||g| - p \circ g\|_{\mathcal{C}(K,\mathbb{R})} < \epsilon .$$

Since $\overline{\mathcal{A}}$ is an algebra containing the constant functions and $g \in \overline{\mathcal{A}}$, we get that $p \circ g \in \overline{\mathcal{A}}$. Thus, $|g| \in \overline{\overline{\mathcal{A}}} = \overline{\mathcal{A}}$. Because $|f|$ is a scalar multiple of $|g|$ and $\overline{\mathcal{A}}$ is an algebra, we get that $|f| \in \overline{\mathcal{A}}$. Thus $\overline{\mathcal{A}}$ is a lattice.

Therefore $\overline{\mathcal{A}} = \mathcal{C}(K, \mathbb{R})$ by Theorem 5.11. $\square$

The following proof (cf. [13]) of Theorem 4.4 uses the Definition 2.6.

Stone-Weierstrass Theorem 4.4 (Case: $\mathcal{C}(K, \mathbb{R})$)

Let $\mathcal{A}$ be a collection of real-valued functions on K that forms an algebra and strongly separates points. Then any continuous real-valued function on K can be uniformly approximated by functions in $\mathcal{A}$.

Proof. Let $f \in \mathcal{C}(K, \mathbb{R})$. Fix $\epsilon > 0$. We want to find a functions $g \in \mathcal{A}$ such that

$$f(y) - \epsilon < g(y) < f(y) + \epsilon$$

for all $y \in K$.

For each pair of points $x, y \in K$, there exists $h_{xy} \in \mathcal{A}$ so that

$$h_{xy}(x) = f(x) \quad \text{and} \quad h_{xy}(y) = f(y) \ . \tag{6.7}$$

Indeed, if $x = y$, by Definition 2.6, we can find $f_x \in \mathcal{A}$ such that $f_x(x) = 1$. So we can take $h_{xy}(\cdot) = f(x)f_x(\cdot)$. Also, if $x \neq y$, then (6.7) follows from Definition 2.6.

For now, let's fix $x \in K$. For each $y \in K$, let

$$O_{xy} := \left\{ z \in K \colon h_{xy}(z) - f(z) < \frac{\epsilon}{3} \right\} \ .$$

Note that $x, y \in O_{xy}$ by (6.7). So

$$K = \bigcup_{y \in K} O_{xy} \ .$$

Since f and h_{xy} are continuous, O_{xy} is open. Since K is compact, there exists $y_1, y_2, ..., y_n \in K$ such that $\cup_{i=1}^{n} O_{xy_i}$ is a finite subcovering of K. Let

$$g_x(\cdot) := h_{xy_1}(\cdot) \wedge h_{xy_2}(\cdot) \wedge ... \wedge h_{xy_n}(\cdot) \, .$$

Clearly, $g_x(x) = f(x)$ by (6.7). Also

$$g_x(z) < f(z) + \frac{\epsilon}{3} \tag{6.8}$$

for all $z \in K$; indeed, each $z \in K$ is in some O_{xy_i}.

Unfortunately, we do not know that $g_x \in \mathcal{A}$. So we now approximate g_x by an $H_x \in \mathcal{A}$. By Lemma 5.5, there exists a $H_x \in \mathcal{A}$ such that

$$|H_x(z) - g_x(z)| < \frac{\epsilon}{3}$$

for all $z \in K$. Thus

$$H_x(z) < g_x(z) + \frac{\epsilon}{3} < f(z) + \frac{2\epsilon}{3}$$

but $H_x(x) > f(x) - \frac{\epsilon}{3}$ since $g_x(x) = f(x)$.

Now, for each $x \in K$, define V_x by

$$V_x := \left\{ z \in K \colon f(z) - \frac{\epsilon}{3} < H_x(z) \right\} \ .$$

Note that

$$\bigcup_{x \in K} V_x = K$$

since, for each $x \in K$, we know that $H_x(x) > f(x) - \frac{\epsilon}{3}$ and so $x \in V_x$. Also, each V_x is open since f and H_x are continuous. Thus since K is compact, there exists $x_1, x_2, ..., x_m \in K$ such that $\cup_{j=1}^{m} V_{x_j}$ is a finite subcovering of K. Let

$$G(\cdot) := H_{x_1}(\cdot) \vee H_{x_2}(\cdot) \vee ... \vee H_{x_m}(\cdot) \ .$$

Consider a $z \in K$. Then z is in some V_{x_j} and so

$$f(z) - \frac{2\epsilon}{3} < f(z) - \frac{\epsilon}{3} < G(z) < f(z) + \frac{2\epsilon}{3} \ .$$

By Lemma 5.5, there exists $g \in \mathcal{A}$ such that

$$\|G - g\|_{\mathcal{C}(K,\mathbb{R})} < \frac{\epsilon}{3}$$

Now let us put it all together:

$$\begin{aligned} \|g - f\|_{\mathcal{C}(K,\mathbb{R})} &= \|g - G + G - f\|_{\mathcal{C}(K,\mathbb{R})} \\ &\leq \|g - G\|_{\mathcal{C}(K,\mathbb{R})} + \|G - f\|_{\mathcal{C}(K,\mathbb{R})} \\ &< \frac{\epsilon}{3} + \frac{2\epsilon}{3} = \epsilon \ . \end{aligned}$$

Therefore $\overline{\mathcal{A}} = \mathcal{C}(K, \mathbb{R})$. $\square$

The following proof of Theorem 4.5 will use Lemma 5.14 and Lemma 5.15.

<u>Stone-Weierstrass Theorem 4.5</u> (Case: $\mathcal{C}(K,\mathbb{C})$)

Suppose that $\mathcal{A} \subset \mathcal{C}(K,\mathbb{C})$ satisfies the following conditions.

1. $\mathcal{A}$ is an algebra.

2. $\mathcal{A}$ separates points of K.

3. $1_K \in \mathcal{A}$.

4. If $f \in \mathcal{A}$ then $\overline{f} \in \mathcal{A}$. In words, $\mathcal{A}$ is self-conjugate.

Then $\overline{\mathcal{A}} = \mathcal{C}(K,\mathbb{C})$.

Proof. Let $\mathcal{A} \subset \mathcal{C}(K,\mathbb{C})$ satisfy the assumptions of Theorem 4.5. Let

$$\begin{aligned} \Re\mathcal{A} &:= \{\Re f \colon f \in \mathcal{A}\} \equiv \left\{\frac{f+\overline{f}}{2} : f \in \mathcal{A}\right\} \\ \Im\mathcal{A} &:= \{\Im f \colon f \in \mathcal{A}\} \equiv \left\{\frac{f-\overline{f}}{2i} : f \in \mathcal{A}\right\} . \end{aligned}$$

Since $\Im f = \Re(-if)$ and $\mathcal{A}$ is an algebra,

$$\Re\mathcal{A} = \Im\mathcal{A} .$$

Then, by Lemma 5.14, $\Re\mathcal{A} \subset \mathcal{A}$ and $\Re\mathcal{A}$ is an algebra in $\mathcal{C}(K,\mathbb{R})$ over $\mathbb{R}$. Now by Lemma 5.15, $\Re\mathcal{A}$ separates points in K. Thus $\Re\mathcal{A}$ satisfies the conditions of Theorem 4.3. So we have $\overline{\Re\mathcal{A}} = \mathcal{C}(K,\mathbb{R})$.

Fix $\epsilon > 0$ and let $f \in \mathcal{C}(K,\mathbb{C})$. We can find $g, h \in \Re\mathcal{A}$ such that $\|\Re f - g\|_{\mathcal{C}(K,\mathbb{R})} < \frac{\epsilon}{2}$ and $\|\Im f - h\|_{\mathcal{C}(K,\mathbb{R})} < \frac{\epsilon}{2}$. Now

$$\begin{aligned} \|f-(g+ih)\|_{\mathcal{C}(K,\mathbb{C})} &= \|\Re f + i\Im f - g - ih\|_{\mathcal{C}(K,\mathbb{C})} \\ &\leq \|\Re f - g\|_{\mathcal{C}(K,\mathbb{C})} + \|\Im f - h\|_{\mathcal{C}(K,\mathbb{C})} \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \end{aligned}$$

Since $\Re\mathcal{A} \subset \mathcal{A}$, we have that $g + ih \in \mathcal{A}$.

Therefore $\overline{\mathcal{A}} = \mathcal{C}(K,\mathbb{C})$. ∎

Chapter 7

Weierstrass' Proof of His Approximation Theorem

In 1885, Karl Weierstrass proved that algebraic polynomials are dense in the class of continuous real-valued functions over a compact interval.

The idea of his proof, casted in the language of approximate identities (see Definition 1.13) and convolutions (see Definition 1.15), is as follows. Define $\varphi : \mathbb{R} \to \mathbb{R}$ by

$$\varphi(t) = \frac{1}{\sqrt{\pi}} \mathrm{e}^{-t^2} \ .$$

Then φ is an approximate identity. Form the family $\{\varphi_h\}_{h>0}$ where

$$\varphi_h(t) := \frac{1}{h} \varphi\left(\frac{t}{h}\right) = \frac{1}{h\sqrt{\pi}} \mathrm{e}^{-(\frac{t}{h})^2} \ .$$

Consider $f \in \mathcal{C}([a,b], \mathbb{R})$. Extend f to a continuous function (again called f) from $\mathbb{R}$ to $\mathbb{R}$ with support in $[a-1, b+1]$. It is (now) well-known, indeed we showed in Theorem 5.12, that under the conditions above, $f * \varphi_h$ converges uniformly to f as $h \to 0^+$. Note that

$$(f * \varphi_h)(x) = \frac{1}{h\sqrt{\pi}} \int_{-\infty}^{\infty} f(u) \, \mathrm{e}^{-\left(\frac{u-x}{h}\right)^2} du \ .$$

But, $f * \varphi_h$ need not be a polynomial so we use the Taylor expansion of φ, which converges uniformly to φ on bounded intervals, to uniformly approximate $(f * \varphi_h)$ by a polynomial.

We now give a self-contained proof (cf. [11]) of these ideas, which is Weierstrass' original proof (cf. [14]) cast in modern language, in order to pay Weierstrass homage.

Theorem 7.1. *Let $f : \mathbb{R} \to \mathbb{R}$ be a bounded uniformly continuous function. Then the function*

$$\mathbb{R} \ni x \;\to\; \frac{1}{h\sqrt{\pi}} \int_{-\infty}^{\infty} f(u)\, \mathrm{e}^{-\left(\frac{u-x}{h}\right)^2} du \in \mathbb{R}$$

converges uniformly to f as $h \to 0^+$.

Proof. Let $f : \mathbb{R} \to \mathbb{R}$. Fix $\epsilon > 0$. Then there exists $\delta > 0$ such that if $|x - y| < \delta$, then

$$|f(x) - f(y)| < \frac{\epsilon}{2}$$

for all $x, y \in \mathbb{R}$. Also, there exists a $M \in \mathbb{R}$ such that $|f(x)| \leq M$. We can see, using the fact that

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

that

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

Then we can write

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

Now let $0 < h_0 < \frac{\epsilon\delta\sqrt{\pi}}{4M}$, then

$$\begin{aligned}
\left| \frac{1}{h\sqrt{\pi}} \int_{-\infty}^{\infty} f(u)\, \mathrm{e}^{-\left(\frac{u-x}{h}\right)^2} du - f(x) \right| &\leq \frac{1}{h\sqrt{\pi}} \int_{-\infty}^{\infty} |f(u) - f(x)|\, \mathrm{e}^{-\left(\frac{u-x}{h}\right)^2} du \\
&\leq \frac{\epsilon}{2} + \frac{1}{h\sqrt{\pi}} \int_{|x-u|\geq\delta} |f(u) - f(x)|\, \mathrm{e}^{-\left(\frac{u-x}{h}\right)^2} du \\
&\leq \frac{\epsilon}{2} + \frac{2M}{h\sqrt{\pi}} \int_{|x-u|\geq\delta} \mathrm{e}^{-\left(\frac{u-x}{h}\right)^2} du \\
&= \frac{\epsilon}{2} + \frac{2M}{\sqrt{\pi}} \int_{|v|\geq\frac{\delta}{h}} \mathrm{e}^{-v^2} dv \\
&\leq \frac{\epsilon}{2} + \frac{2Mh}{\delta\sqrt{\pi}} \int_{|v|\geq\frac{\delta}{h}} |v|\, \mathrm{e}^{-v^2} dv
\end{aligned}$$

since $|v|\left(\frac{h}{\delta}\right) \geq 1$

$$\begin{aligned}
&\leq \frac{\epsilon}{2} + \frac{4Mh}{\delta\sqrt{\pi}} \int_0^\infty v\mathrm{e}^{-v^2} dv \\
&= \frac{\epsilon}{2} + \frac{2Mh}{\delta\sqrt{\pi}} \\
&< \frac{\epsilon}{2} + \frac{2M}{\delta\sqrt{\pi}} \left(\frac{\epsilon\delta\sqrt{\pi}}{4M}\right) = \epsilon
\end{aligned}$$

for all $0 < h \leq h_0$ and all $x \in \mathbb{R}$.

So the theorem holds. $\square$

Now we look at the original Weierstrass Approximation Theorem as he proved it.

Theorem 7.2 (Weierstrass Approximation Theorem)**.** *Let $f \in \mathcal{C}([a,b],\mathbb{R})$. Then f can be uniformly approximated by polynomials on $[a,b]$.*

Proof. Let $f \in \mathcal{C}([a,b],\mathbb{R})$. Now we extend f to a bounded uniformly continuous function, again called f, from $\mathbb{R}$ to $\mathbb{R}$ by

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

Observe there exists $J > 0$ such that $f(x) = 0$ for all $|x| > J$. Fix $\epsilon > 0$. There exists a $M \in \mathbb{R}$ such that $|f(x)| < M$ for all $x \in \mathbb{R}$. Now by Theorem 7.1 there exists a $h_0 > 0$ such that for all $x \in \mathbb{R}$,

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

Since $f(u) = 0$ for all $|u| > J$, we have

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

Now, on the interval $\left[\frac{-2J}{h_0}, \frac{2J}{h_0}\right]$ the power series of e^{-v^2} converges uniformly. So there exists N such that

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

for all $|x| \leq J$ and for all $|u| \leq J$, because in that case $|u-x| \leq 2J$. Thus we can see,

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

for all $|x| \leq J$. Now let $p(x)$ be defined as follows.

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

Observe that $p(x)$ is a polynomial in x of degree at most $2N$ such that

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

for all $|x| \leq J$. Thus, we see that

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

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

Therefore f can be approximated by polynomials on $[a, b]$. $\square$

Chapter 8

Stone's Approximation Theorem

Stone's approach to an approximation theorem is as follows. He started with a subset $\mathcal{X}_0$ of $\mathcal{C}(K, \mathbb{R})$. He then built a superset $\mathcal{U}_1(\mathcal{X}_0)$ of $\mathcal{X}_0$ and showed that, if $\mathcal{X}_0$ is sufficiently nice, then $\mathcal{U}_1(\mathcal{X}_0)$ is dense in $\mathcal{C}(K, \mathbb{R})$.

We start with some definitions and lemmas.

Definition 8.1. For an arbitrary subfamily $\mathcal{X}_0$ of $\mathcal{C}(K, \mathbb{R})$ we define the following sets.

a) $\mathcal{U}_1(\mathcal{X}_0)$ is the collection of all $f\colon K \to \mathbb{R}$ that can be obtained by applying lattice operations (i.e. $\vee$ and $\wedge$) to a finite number of elements for $\mathcal{X}_0$.

b) $\mathcal{U}_2(\mathcal{X}_0) := \{f\colon K \to \mathbb{R} \mid f \text{ is the uniform limit of some } \{f_n\}_{n=1}^{\infty} \subset \mathcal{U}_1(\mathcal{X}_0)\}$.

c) $\mathcal{U}(\mathcal{X}_0) := \{f\colon K \to \mathbb{R} \mid f \text{ is the uniform limit of some } \{f_n\}_{n=1}^{\infty} \subset \mathcal{U}_2(\mathcal{X}_0)\}$.

We also say that $\mathcal{X}_0$ is closed under:

d) lattice operations provided if $f, g \in \mathcal{X}_0$, then $f \vee g \in \mathcal{X}_0$ and $f \wedge g \in \mathcal{X}_0$

e) uniform limits provided if $\{f_n\}_{n=1}^{\infty} \subset \mathcal{X}_0$ converges uniformly to f, then $f \in \mathcal{X}_0$.

To clarify, a function f from $\mathcal{U}_1(\mathcal{X}_0)$ takes the form

$$f = f_1 \Diamond_1 f_2 \Diamond_2 \ldots f_{n-1} \Diamond_{n-1} f_n$$

where each $\Diamond_i$ is either $\vee$ or $\wedge$ and each f_i is of the form either

$$f_i = f_{i1} \wedge f_{i2} \wedge \ldots \wedge f_{in_i}$$

or

$$f_i = f_{i1} \vee f_{i2} \vee ... \vee f_{in_i}$$

where each $f_{ij} \in \mathcal{X}_0$

Remark 8.2. Throughout the remainder of this chapter, $\mathcal{X}_0$ will be a subset of $\mathcal{C}(K, \mathbb{R})$ and we will follow the notation set forth in Definition 8.1.

The goal of the next two lemmas is to show the following proposition, which we will use often. Proposition 8.3 follows directly from Lemmas 8.4 and 8.5.

Proposition 8.3. *The $\mathcal{C}(K, \mathbb{R})$-norm closure of $\mathcal{U}_1(\mathcal{X}_0)$ is $\mathcal{U}(\mathcal{X}_0)$. So, if $f \in \mathcal{C}(K, \mathbb{R})$, then $f \in \mathcal{U}(\mathcal{X}_0)$ if and only if for each $\epsilon > 0$ there exists $g \in \mathcal{U}_1(\mathcal{X}_0)$ such that $\|f - g\|_{\mathcal{C}(K,\mathbb{R})} < \epsilon$.*

Lemma 8.4. *The following holds: $\mathcal{X}_0 \subset \mathcal{U}_1(\mathcal{X}_0) \subset \mathcal{U}_2(\mathcal{X}_0) \subset \mathcal{U}(\mathcal{X}_0) \subset \mathcal{C}(K, \mathbb{R})$.*

Proof. Let $f \in \mathcal{X}_0$. Then $f \vee f \in \mathcal{U}_1(\mathcal{X}_0)$ by Definition 8.1. But $f \vee f = f$, so $f \in \mathcal{U}_1(\mathcal{X}_0)$. So the first set inclusion holds.

Now let $f \in \mathcal{U}_1(\mathcal{X}_0)$. For each $n \in \mathbb{N}$, let $f_n := f$. Then $\{f_n\}_{n=1}^{\infty} \subset \mathcal{U}_1(\mathcal{X}_0)$ which converges uniformly to f. So $\mathcal{U}_1(\mathcal{X}_0) \subset \mathcal{U}_2(\mathcal{X}_0)$.

Next let $f \in \mathcal{U}_2(\mathcal{X}_0)$. For each $n \in \mathbb{N}$, let $f_n := f$. Then $\{f_n\}_{n=1}^{\infty} \subset \mathcal{U}_2(\mathcal{X}_0)$ which converges uniformly to f. So $\mathcal{U}_2(\mathcal{X}_0) \subset \mathcal{U}(\mathcal{X}_0)$.

The last set inclusion follows directly from the fact that $\mathcal{C}(K, \mathbb{R})$ is closed under lattice operations and uniform limits. □

Lemma 8.5. *The family $\mathcal{U}_2(\mathcal{X}_0)$ is closed under lattice operations as well as uniform limits. Thus $\mathcal{U}_2(\mathcal{X}_0) = \mathcal{U}(\mathcal{X}_0)$.*

Proof. To see that $\mathcal{U}_2(\mathcal{X}_0)$ is closed under lattice operations, fix $f, g \in \mathcal{U}_2(\mathcal{X}_0)$. Then there exists sequences $\{f_n\}_n$ and $\{g_n\}_n$ from $\mathcal{U}_1(\mathcal{X}_0)$ such that $\{f_n\}$ converges uniformly to f and $\{g_n\}$ converges uniformly to g. The following statements follow from uniform convergence.

$$\{f_n \vee g_n\} \quad \text{converges uniformly to} \quad f \vee g.$$

$$\{f_n \wedge g_n\} \quad \text{converges uniformly to} \quad f \wedge g.$$

Now, $f_n \vee g_n, f_n \wedge g_n \in \mathcal{U}_1(\mathcal{X}_0)$ for each $n \in \mathbb{N}$. Thus $f \vee g, f \wedge g \in \mathcal{U}_2(\mathcal{X}_0)$. Therefore, by Definition 8.1, $\mathcal{U}_2(\mathcal{X}_0)$ is closed under lattice operations.

To see that $\mathcal{U}_2(\mathcal{X}_0)$ is closed under uniform limits, let f be the uniform limit of a sequence $\{f_n\}$ of functions from $\mathcal{U}_2(\mathcal{X}_0)$. Without loss of generality (pass to a subsequence if so needed),

$$\|f - f_n\|_{\mathcal{C}(K,\mathbb{K})} < \frac{1}{n} \ .$$

For each $n \in \mathbb{N}$, find $g_n \in \mathcal{U}_1(\mathcal{X}_0)$ so that

$$\|f_n - g_n\|_{\mathcal{C}(K,\mathbb{K})} < \frac{1}{n} \ .$$

Then $\{g_n\}$ converges uniformly to f since

$$\|f - g_n\|_{\mathcal{C}(K,\mathbb{K})} \leq \|f - f_n\|_{\mathcal{C}(K,\mathbb{K})} + \|f_n - g_n\|_{\mathcal{C}(K,\mathbb{K})} < \frac{1}{n} + \frac{1}{n} = \frac{2}{n} \ .$$

Thus $f \in \mathcal{U}_2(\mathcal{X}_0)$ and so $\mathcal{U}_2(\mathcal{X}_0)$ is closed under uniform limits.

We know from Lemma 8.4 that $\mathcal{U}_2(\mathcal{X}_0) \subset \mathcal{U}(\mathcal{X}_0)$. Next consider an $f \in \mathcal{U}(\mathcal{X}_0)$. Then f is a uniform limit of functions from $\mathcal{U}_2(\mathcal{X}_0)$. But we just showed that $\mathcal{U}_2(\mathcal{X}_0)$ is closed under uniform limits. So $f \in \mathcal{U}_2(\mathcal{X}_0)$. Thus $\mathcal{U}_2(\mathcal{X}_0) = \mathcal{U}(\mathcal{X}_0)$. □

Next we define a property of $\mathcal{X}_0$.

Definition 8.6. We say that $\mathcal{X}_0$ satisfies *Property SA* provided that for each

(1) $f \in \mathcal{C}(K, \mathbb{R})$

(2) pair of distinct points x and y in K

(3) $\epsilon > 0$

there exists a function f_{xy} in $\mathcal{U}_1(\mathcal{X}_0)$ such that

$$\begin{aligned} |f(x) - f_{xy}(x)| &< \epsilon \\ |f(y) - f_{xy}(y)| &< \epsilon. \end{aligned} \tag{8.1}$$

Note that for $\mathcal{X}_0$ to enjoy Property SA, each continuous function needs to be able to be approximated, arbitrarily close, by a function in $\mathcal{U}_1(\mathcal{X}_0)$, where the approximation is not uniform on the whole of K but just at *two* points in K. The goal of this chapter is to show the following theorem.

Theorem 8.7. *Let $\mathcal{X}_0$ have Property SA. Then $\mathcal{C}(K,\mathbb{R}) = \mathcal{U}(\mathcal{X}_0)$, in other words, each function $f \in \mathcal{C}(K,\mathbb{R})$ can be uniformly approximated, arbitrarily close in the $\mathcal{C}(K,\mathbb{R})$ norm, by a function in $\mathcal{U}_1(\mathcal{X}_0)$.*

Theorem 8.7 is a direct consequence of the following theorem, which shows even more.

Theorem 8.8 ([12])**.** *Let $\mathcal{X}_0$ be a subset of $\mathcal{C}(K,\mathbb{R})$. Let $f \in \mathcal{C}(K,\mathbb{R})$. Then the following are equivalent.*

(1) $f \in \mathcal{U}(\mathcal{X}_0)$.

(2) for each pair of points x and y in K and for each $\epsilon > 0$, there exists a function f_{xy} in $\mathcal{U}_1(\mathcal{X}_0)$ such that (8.1) *holds.*

(3) for each pair of distinct points x and y in K and for each $\epsilon > 0$, there exists a function f_{xy} in $\mathcal{U}_1(\mathcal{X}_0)$ such that (8.1) *holds.*

Proof. To show that (1) implies (2), fix $f \in \mathcal{U}(\mathcal{X}_0)$, a pair of points x and y in K, and a $\epsilon > 0$. By Proposition 8.3, there exists $g \in \mathcal{U}_1(\mathcal{X}_0)$ such that

$$\|f - g\|_{\mathcal{C}(K,\mathbb{R})} < \epsilon \ .$$

Set $f_{xy} := g$. Then

$$|f(x) - f_{xy}(x)| < \epsilon \quad \text{and} \quad |f(y) - f_{xy}(y)| < \epsilon \ .$$

Thus (2) holds.

Clearly (2) implies (3).

Next assume that (3) holds. We shall show that (1) holds. So fix an $\epsilon > 0$. By Proposition 8.3, it suffices to find a function $g \in \mathcal{U}_1(\mathcal{X}_0)$ such that

$$f(y) - \epsilon < g(y) < f(y) + \epsilon$$

for all $y \in K$.

By assumption, for each pair of distint points $x, y \in K$, there exists $h_{xy} \in \mathcal{U}_1(\mathcal{X}_0)$ so that

$$|f(x) - h_{xy}(x)| < \epsilon \quad \text{and} \quad |f(y) - h_{xy}(y)| < \epsilon \; . \tag{8.2}$$

For now, let's fix $x \in K$. For each $y \in K \setminus \{x\}$, let

$$O_{xy} := \{z \in K \colon h_{xy}(z) - f(z) < \epsilon\} \; .$$

Note that $x, y \in O_{xy}$ by (8.2). So

$$K = \bigcup_{y \in K \setminus \{x\}} O_{xy} \; .$$

Since f and h_{xy} are continuous, each O_{xy} is open. Since K is compact, there exists $y_1, y_2, ..., y_n$ in $K \setminus \{x\}$ such that $\cup_{i=1}^{n} O_{xy_i}$ is a finite subcovering of K. Since each y_i does not equal x, we can define $g_x \colon K \to \mathbb{R}$ by

$$g_x(\cdot) := h_{xy_1}(\cdot) \wedge h_{xy_2}(\cdot) \wedge ... \wedge h_{xy_n}(\cdot) \, .$$

Note that since each h_{xy_i} is in $\mathcal{U}_1(\mathcal{X}_0)$, we have that $g_x \in \mathcal{U}_1(\mathcal{X}_0)$. If $z \in K$, then z is in some O_{xy_i} and so

$$h_{xy_i}(z) < f(z) + \epsilon$$

which gives that

$$g_x(z) < f(z) + \epsilon \; . \tag{8.3}$$

Also note that

$$f(x) - \epsilon < g_x(x) \tag{8.4}$$

since, for each y_i, by (8.2)

$$f(x) - \epsilon < h_{xy_i}(x) \ .$$

Now, for each $x \in K$, define V_x by

$$V_x := \{z \in K \colon f(z) - \epsilon < g_x(z)\} \ .$$

Note that

$$\bigcup_{x \in K} V_x = K$$

since $x \in V_x$ by (8.4). Also, each V_x is open since f and g_x are continuous. Thus since K is compact, there exists $x_1, x_2, ..., x_m \in K$ such that $\cup_{j=1}^{m} V_{x_j}$ is a finite subcovering of K. Define $g \colon K \to \mathbb{R}$ by

$$g(\cdot) := g_{x_1}(\cdot) \vee g_{x_2}(\cdot) \vee ... \vee g_{x_m}(\cdot) \, .$$

Note that since each g_{x_i} is in $\mathcal{U}_1(\mathcal{X}_0)$, we have that $g \in \mathcal{U}_1(\mathcal{X}_0)$.

Consider a $z \in K$. Then z is in some V_{x_j} and so

$$f(z) - \epsilon < g_{x_j}(z) \leq g(z) \ .$$

And by (8.3)

$$g(z) < f(z) + \epsilon \ .$$

Therefore

$$f(z) - \epsilon < g(z) < f(z) + \epsilon \ ,$$

as desired. So (1) holds.

This completes the proof of Theorem 8.8. □

The following corollary follows from Theorem 8.7.

Corollary 8.9 ([12]). *Let $\mathcal{X}_0$ have the property that, for each pair $\alpha, \beta \in \mathbb{R}$ and each pair of distinct points $x, y \in K$, there exists a function $f \in \mathcal{X}_0$ for which*

$$f(x) = \alpha \quad \text{and} \quad f(y) = \beta \ .$$

Then

$$\mathcal{U}(\mathcal{X}_0) = \mathcal{C}(K, \mathbb{R}),$$

in other words, the $\mathcal{C}(K,\mathbb{R})$*-norm closure of* $\mathcal{U}_1(\mathcal{X}_0)$ *is* $\mathcal{C}(K,\mathbb{R})$.

Proof. Let $\mathcal{X}_0$ statisfy the assumptions of Corollary 8.9. Let $f \in \mathcal{C}(K,\mathbb{R})$ and x, y be a pair of distinct points from K. By the assumption, there is a function $f_{xy} \in \mathcal{X}_0$ such that

$$f_{xy}(x) = f(x) \quad \text{and} \quad f_{xy}(y) = f(y) \ .$$

Thus

$$|f(x) - f_{xy}(x)| \ = \ 0 \ = \ |f(y) - f_{xy}(y)| \ .$$

Thus $\mathcal{X}_0$ satisfies Property SA. Thus, by Theorem 8.7, we have that $\mathcal{U}(\mathcal{X}_0) = \mathcal{C}(K,\mathbb{R})$.

□

Next we give a proof Corollary 8.9 that is based on Theorem 5.11 rather than Theorem 8.7.

Another proof of Corollary 8.9. Let $\mathcal{X}_0$ statisfy the assumptions of Corollary 8.9.

Claim 1. $\mathcal{U}(\mathcal{X}_0)$ is a lattice, that is if $f, g \in \mathcal{U}(\mathcal{X}_0)$ then $f \vee g, f \wedge g \in \mathcal{U}(\mathcal{X}_0)$.

Proof of Claim 1. Let $f, g \in \mathcal{U}(\mathcal{X}_0)$ and $\epsilon > 0$. Then there exists $f_\epsilon, g_\epsilon \in \mathcal{U}_1(\mathcal{X}_0)$ such that

$$\|f - f_\epsilon\|_{\mathcal{C}(K,\mathbb{R})} < \frac{\epsilon}{2} \quad \text{and} \quad \|g - g_\epsilon\|_{\mathcal{C}(K,\mathbb{R})} < \frac{\epsilon}{2} \ .$$

Note that $f_\epsilon \vee g_\epsilon \in \mathcal{U}_1(\mathcal{X}_0)$ by Definition 8.1. Let's compute:

$$\begin{aligned}
\|f_\epsilon \vee g_\epsilon - f \vee g\|_{\mathcal{C}(K,\mathbb{R})} &= \left\| \frac{f_\epsilon + g_\epsilon + |f_\epsilon - g_\epsilon|}{2} - \frac{f + g + |f-g|}{2} \right\|_{\mathcal{C}(K,\mathbb{R})} \\
&= \frac{1}{2} \left\| f_\epsilon - f + g_\epsilon - g + |f_\epsilon - g_\epsilon| - |f - g| \right\|_{\mathcal{C}(K,\mathbb{R})} \\
&\le \frac{1}{2} \left[\|f_\epsilon - f\|_{\mathcal{C}(K,\mathbb{R})} + \|g_\epsilon - g\|_{\mathcal{C}(K,\mathbb{R})} + \| \ |f_\epsilon - g_\epsilon| - |g - f| \ \|_{\mathcal{C}(K,\mathbb{R})} \right] \\
&\le \frac{1}{2} \left[\|f_\epsilon - f\|_{\mathcal{C}(K,\mathbb{R})} + \|g_\epsilon - g\|_{\mathcal{C}(K,\mathbb{R})} + \|f_\epsilon - f\|_{\mathcal{C}(K,\mathbb{R})} + \|g_\epsilon - g\|_{\mathcal{C}(K,\mathbb{R})} \right] \\
&\le \frac{1}{2} \left(\frac{\epsilon}{2} + \frac{\epsilon}{2} + \frac{\epsilon}{2} + \frac{\epsilon}{2} \right) = \epsilon.
\end{aligned}$$

Thus $f \vee g \in \mathcal{U}(\mathcal{X}_0)$. A similar proof can be given to show that $f \wedge g \in \mathcal{U}(\mathcal{X}_0)$. Therefore $\mathcal{U}(\mathcal{X}_0)$ is a lattice. So Claim 1 holds.

Claim 2. $\mathcal{U}(\mathcal{X}_0)$ separates points of K.

Proof of Claim 2. Let $x, y \in K$ with $x \neq y$. By the assumption on $\mathcal{X}_0$, there exists a function $f \in \mathcal{X}_0$ such that $f(x) = 0$ and $f(y) = 17$. Now $f \in \mathcal{X}_0 \subset \mathcal{U}(\mathcal{X}_0)$. So there exists $f \in \mathcal{U}(\mathcal{X}_0)$ such that $f(x) \neq f(y)$. Therefore $\mathcal{U}(\mathcal{X}_0)$ separates points of K. Thus Claim 2 holds.

Claim 3. Let $f \in \mathcal{U}(\mathcal{X}_0)$ and $c \in \mathbb{R}$. Then $cf \in \mathcal{U}(\mathcal{X}_0)$ and $c + f \in \mathcal{U}(\mathcal{X}_0)$.

Proof of Claim 3. Let $f \in \mathcal{U}(\mathcal{X}_0)$ and $c \in \mathbb{R}$. Fix $\epsilon > 0$. Then there exists $f_\epsilon \in \mathcal{U}_1(\mathcal{X}_0)$ such that $\|f_\epsilon - f\|_{\mathcal{C}(K,\mathbb{R})} < \frac{\epsilon}{|c|+1}$. Now

$$\|cf_\epsilon - cf\|_{\mathcal{C}(K,\mathbb{R})} = |c| \, \|f_\epsilon - f\|_{\mathcal{C}(K,\mathbb{R})} < |c| \frac{\epsilon}{|c|+1} \leq \epsilon \, .$$

Also, there exists $g_\epsilon \in \mathcal{U}_1(\mathcal{X}_0)$ such that $\|g_\epsilon - f\|_{\mathcal{C}(K,\mathbb{R})} < \epsilon$. So

$$\|(g_\epsilon + c) - (f + c)\|_{\mathcal{C}(K,\mathbb{R})} = \|g_\epsilon - f\|_{\mathcal{C}(K,\mathbb{R})} < \epsilon \, .$$

Therefore, $cf \in \mathcal{U}(\mathcal{X}_0)$ and $c + f \in \mathcal{U}(\mathcal{X}_0)$. So Claim 3 holds.

So by Theorem 5.11, the $\mathcal{C}(K,\mathbb{R})$ norm closure of $\mathcal{U}(\mathcal{X}_0)$ is $\mathcal{C}(K,\mathbb{R})$. But in Lemma 8.5, we showed that $\mathcal{U}(\mathcal{X}_0)$ is norm closed. Therefore $\mathcal{U}(\mathcal{X}_0) = \mathcal{C}(K,\mathbb{R})$. □

We close with an application of Theorem 8.8.

Corollary 8.10 ([12]). *Let $f \in \mathcal{C}(K,\mathbb{R})$ be the pointwise limit of a monotonic sequence $\{f_n\}$ from $\mathcal{C}(K,\mathbb{R})$. Then the sequence $\{f_n\}$ converges uniformly to f.*

Proof. Let $f \in \mathcal{C}(K,\mathbb{R})$ be the pointwise limit of a monotonic sequence $\{f_n\}$ from $\mathcal{C}(K,\mathbb{R})$. Set $\mathcal{X}_0 := \{f_n : n \in \mathbb{N}\}$. Observe that, by monotonicity,

$$f_m \vee f_n \quad \text{is either} \quad f_m \text{ or } f_n$$

$$f_m \wedge f_n \quad \text{is either} \quad f_n \text{ or } f_m$$

and so

$$\mathcal{U}_1 \left(\mathcal{X}_0\right) = \mathcal{X}_0. \tag{8.5}$$

Consider points $x, y \in K$ and an $\epsilon > 0$. Since the sequence $\{f_n\}$ converges pointwise to f, there exists $n \in \mathbb{N}$ such that

$$|f(x) - f_n(x)| \ < \ \epsilon \quad \text{and} \quad |f(y) - f_n(y)| \ < \ \epsilon \ .$$

Thus, by Theorem 8.8

$$f \in \mathcal{U}\left(\mathcal{X}_0\right) \ . \tag{8.6}$$

So (8.5) and (8.6) give that f is the uniform limit of a sequence from $\mathcal{X}_0$. Again, fix an $\epsilon > 0$. Then there is a $N \in \mathbb{N}$ such that

$$\|f - f_N\|_{\mathcal{C}(K,\mathbb{R})} \ < \ \epsilon \ .$$

Let $n \geq N$. Since $\{f_n\}$ is monotonic, for each $x \in K$,

$$|f_n(x) - f(x)| \ \leq \ |f_N(x) - f(x)| \ .$$

Thus $\|f_n - f\|_{\mathcal{C}(K,\mathbb{R})} < \epsilon$.

Therefore $\{f_n\}$ converges unifromly to f. $\square$

Bibliography

[1] Heinrich Behnke, *Karl Weierstrass und seine Schule*, Festschr. Gedächtnisfeier K. Weierstrass, Westdeutscher Verlag, Cologne, 1966, pp. 13–40. 0202545 (34 #2412)

[2] Jöran Bergh and Jörgen Löfström, *Interpolation spaces. An introduction*, Springer-Verlag, Berlin, 1976, Grundlehren der Mathematischen Wissenschaften, No. 223. 0482275 (58 #2349)

[3] Felix E. Browder, *The Stone Age of mathematics on the Midway*, A century of mathematics in America, Part II, Hist. Math., vol. 2, Amer. Math. Soc., Providence, RI, 1989, pp. 191–193. 1003128 (91b:01090)

[4] J. Dieudonné, *Foundations of modern analysis*, Academic Press, New York, 1969, Enlarged and corrected printing, Pure and Applied Mathematics, Vol. 10-I. 0349288 (50 #1782)

[5] Donald Estep, *Practical analysis in one variable*, Undergraduate Texts in Mathematics, Springer-Verlag, New York, 2002. 1934646 (2003h:26001)

[6] Gerald B. Folland, *Real analysis*, second ed., Pure and Applied Mathematics (New York), John Wiley & Sons Inc., New York, 1999, Modern techniques and their applications, A Wiley-Interscience Publication. 1681462 (2000c:00001)

[7] Richard R. Goldberg, *Methods of real analysis*, second ed., John Wiley & Sons Inc., New York, 1976. 0473118 (57 #12796)

[8] Edwin Hewitt and Karl Stromberg, *Real and abstract analysis. A modern treatment of the theory of functions of a real variable*, Springer-Verlag, New York, 1965. 0188387 (32 #5826)

[9] Serge Lang, *Undergraduate analysis*, Undergraduate Texts in Mathematics, Springer-Verlag, New York, 1983. 704509 (85k:26001)

[10] John N. McDonald and Neil A. Weiss, *A course in real analysis*, Academic Press Inc., San Diego, CA, 1999, Biographies by Carol A. Weiss. 1680810 (2000g:00005)

[11] Anton R. Schep, *Weierstrass' Proof of the Weierstrass Approximation Theorem*, unpublished.

[12] M. H. Stone, *The generalized Weierstrass approximation theorem*, Math. Mag. **21** (1948), 167–184, 237–254. 0027121 (10,255a)

[13] Robert S. Strichartz, *The Way of Analysis*, Jones and Bartlett Books in Mathematics, Jones and Bartlett Publishers International, London, 1995.

[14] Karl Weierstrass, *Über die analytische Darstellbarkeit sogenannter willkürlicher Functionen einer reellen Veränderlichen*, Verl. d. Kgl. Akad. d. Wiss. **2** (1885), 633–639.

Appendix A

Biography of Karl Weierstrass

Karl Theodor Wilhelm Weierstrass was born October 31, 1815 in Ostenfelde, Westphalia (now Germany) to Wilhelm and Theodora Weierstrass. He was the oldest of four children. He died of pneumonia February 19, 1897 in Berlin, Germany.

At a young age, Karl frequently changed schools as his family moved around Prussia for his father's job. In 1827, his mother Theodora died. In 1829, he entered the Catholic Gymnasium in Paderborn. While there he excelled in mathematical competence. He graduated from the Gymnasium in 1834 and entered the University of Bonn where it was planned by his father for him to study law, finance and economics. He struggled with whether to obey his father's wishes or to study the subject he loved, mathematics. After four years of pretending he did not care about his studies and failing to study the subjects he was enrolled in, he left the university. In May of 1839, Weierstrass enrolled at the Academy in Munster. There he attended lectures by Gudermann on elliptic functions. In May 1840, Weierstrass presented his own research as an answer to a question on the representation of elliptic functions. He learned later in life that Gudermann read the paper and said that his work was 'of equal rank with the discoverers who were crowned with glory'. In April 1841, Weierstrass began a one year probation as a teacher at the Gymnasium in Munster. He wrote three short papers in 1841 and 1842 which are described in [1].

> "The concepts on which Weierstrass based his theory of functions of a complex variable in later years after 1857 are found explicitly in his unpublished works written in Munster from 1841 to 1842, while still under

> the influence of Gudermann. The transformation of his conception of an analytic function from a differentiable function to a function expansible into a convergent power series was made during this early period of Weierstrass' mathematical activity".

In 1842, he began his career as a qualified teacher of mathematics at the Pro-Gymnasium in Deutsch Krone in West Prussia (now Polland). In 1848, he moved to the Collegium Hoseanum in Braunsberg. From 1850 on, Weierstrass suffered from health problems. In 1854, he published *Zur Theorie der Abelschen Functionen in Crelle's Journal*. Before this publication, he was unnoticed for his work, but this publication was certainly noticed. After this publication, the University of Königsberg conferred an honorary doctor's degree on him on 31 March 1854. In 1856, he published a full version of his theory of inversion of hyperelliptic integrals. From 1856-1857 his lecture topics included the application of Fourier series and integrals to mathematical physics, an introduction to the theory of analytic functions (where he set out results he had obtained in 1841 but never published), the theory of elliptic functions (his main research topic), and applications to problems in geometry and mechanics. His lectures of 1859-1860 were an introduction to analysis where he tackled the foundations of the subject for the first time. In 1860-1861, he lectured on integral calculus. In 1863-1864, Weierstrass began to formulate his theory of the real numbers and in 1863 he proved that the complex numbers are the only commutative algebraic extension of the real numbers. In 1872, Weierstrass discovered a function, although continuous, had no derivative at any point. Many students benefited from Weierstrass' teaching. In 1870, Sofia Kovalevskaya came to Berlin and Weierstrass taught her privately since she was not allowed admission to the university. Through Weierstrass' efforts, Kovalevskaya received an honorary doctorate from Gottingen. Weierstrass strongly affected the future of mathematics by defining irrational numbers as limits of convergent series. It is said of him that

> "He is known as the father of modern analysis. He devised tests for the convergence of series, and he contributed to the theory of periodic functions, functions of real variables, elliptic functions, Abelian functions, converging infinite products and the calculus of variations. He also advanced the theory of bilinear and quadratic forms".

Of particular interest to us, when he was 70 years old, he published the Weierstrass approximation theorems. The first proved the density of algebraic polynomials in the space of continuous real-valued functions on a finite interval in the uniform norm, and the second proved the density of trigonometric polynomials in the space of 2π-periodic continuous real-valued functions on $\mathbb{R}$ in the uniform norm. (This is left for the reader to peruse at his or her inquisition.) These approximation theorems were in a sense the converse of Weierstrass' famous example of the existence of continuous nowhere differentiable functions. Every continuous function on $\mathbb{R}$ is a limit not only of infinitely differentiable or even analytic functions, but in fact of polynomials. Also, the limit is uniform if we restrict the approximation to any finite interval. Thus the set of continuous functions contains very many non-smooth functions, but they can each be approximated arbitrarily well by the ultimate in smooth functions. Before he died in 1897, Weierstrass helped with the publication of the first two volumes of his own complete works. Volumes three through seven have been published since then.

Appendix B

Biography of Marshall Stone

Marshall Harvey Stone was born April, 8, 1903 in New York to Harlan Fiske Stone and Agnes Harvey. His father was a distinguished lawyer and was on the Supreme Court for 21 years. Many thought that Marshall would follow his father's footsteps into a law career. He entered Harvard in 1919 intending to continue his studies at Harvard law school. He graduated in 1922 with a change of heart. His enthusiasm was now for mathematics. After a year as instructor in the Harvard Mathematics Department, he indeed decided to take his mathematical studies further.

He studied under the supervision of Birkhoff and, in 1926, he was awarded his doctorate for a thesis entitled *Ordinary Linear Homogeneous Differential Equations of Order n and the Related Expansion Problems*. In 1927, he was appointed to Harvard as an instructor. He also married Emmy Portman that year, but the marriage ended in divorce in 1962. Marshall and Emmy had three children.

Stone published eleven papers on the theory of orthogonal expansions between 1925 and 1928. In 1928, Stone was promoted to associate professor at Harvard. Before later returning to Harvard, he accepted a position as associate professor at Yale from 1931-1933. He went back to Harvard and was promoted to full professor in 1937. Beginning in 1929, Stone worked on self-adjoint operators in Hilbert space and included his findings in his 662 page book *Lindear transformations in Hilbert space and their applications to analysis* published in 1932. In 1934, he published two papers on Boolean algebras. Both appeared in the *Proceedings of the National Academy of Sciences*. During this period, Stone generalized Weierstrass' Approximation Theorem on uniform approximation of continuous functions by polynomials. This work is know as the Stone Weierstrass Theorem.

During World War II, Stone worked for the Office of Naval Operations from 1942-1943 and then the Office of the Chief of Staff at the War Department for the rest of World War II. In 1946, he left Harvard to become the mathematics department chair at the University of Chicago. Because of his 'forceful character and unquestioned integrity' [3] Stone returned this well-known research school to the prestige it had previously known. In 1952, he stepped down as chairman but remained at Chicago until he retired in 1968. He then accepted a full-time professorship at the University of Massachusetts until 1973 and then worked half-time until 1980.

Stone received much recognition for his outstanding achievements. In 1938, he was elected to the National Academy of Sciences. In 1939, he was American Mathematical Society Colloquium Lecturer and president of the Society in 1943-1944. In 1952-1954, he was elected president of the International Mathematical Union and was named Josiah Willard Gibbs lecturer for 1956. He was president of the International Committee of Mathematical Instruction from 1961-1967.

One of Stone's interests included he love to travel. He died on January 9, 1989 while on a trip Madras, India.