Advanced Calculus: MATH 410
Uniform Convergence of Functions

Professor David Levermore

7 December 2017

12. Sequences of Functions

We now explore two notions of what it means for a sequence of functions $\{f_n\}_{n\in\mathbb{N}}$ to converge to a function f. The first notion, *pointwise convergence*, might seem natural at first, but we will see that it is not strong enough to do much. The second notion, *uniform convergence*, is strong enough to do many things, but might seem less natural at first. We will explore these notions through examples that show the superiority of uniform convergence.

12.1. **Pointwise Convergence.** We begin with the definition of pointwise convergent.

Definition 12.1. *Let $S \subset \mathbb{R}$. Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of real-valued functions that are each defined over S. The sequence $\{f_n\}_{n\in\mathbb{N}}$ is said to be* pointwise convergent *or to* converge pointwise *over S if there exists a function f defined over S such that*

$$\lim_{n\to\infty} f_n(x) = f(x) \qquad \text{for every } x \in S\,.$$

We say f_n converges to f pointwise *over S and call f the* pointwise limit *of the sequence $\{f_n\}_{n\in\mathbb{N}}$ over S. We denote this as*

$$f_n \to f \qquad \text{pointwise over } S\,.$$

Because every Cauchy sequence of real numbers has a unique limit, we have the following.

Proposition 12.1. *Let $S \subset \mathbb{R}$. Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of real-valued functions that are each defined over S such that*

$$\text{for every } x \in S \text{ the real sequence } \quad \{f_n(x)\}_{n\in\mathbb{N}} \quad \text{is Cauchy}\,.$$

Then there exists a unique real-valued function f defined over S such that

$$f_n \to f \qquad \text{pointwise } S\,. \tag{12.1}$$

Proof. For each $x \in S$ define $f(x)$ to be the unique limit of the Cauchy sequence of real numbers $\{f_n(x)\}_{n\in\mathbb{N}}$. □

The following natural questions arise.

- If $f_n \to f$ pointwise over $[a,b]$ and each f_n is continuous over $[a,b]$ then is f continuous over $[a,b]$?
- If $f_n \to f$ pointwise over $[a,b]$ and each f_n is Riemann integrable over $[a,b]$ then is f Riemann integrable over $[a,b]$ and does
$$\int_a^b f_n \to \int_a^b f\,?$$
- If $f_n \to f$ pointwise over $[a,b]$ and each f_n is differentiable over $[a,b]$ then is f differentiable over $[a,b]$ and does
$$f_n' \to f' \qquad \text{pointwise over } [a,b]\,?$$

For pointwise convergence the answer to each of these three questions is **NO, NOT ALWAYS!** To understand why, consider the following examples.

Example. Consider $f_n(x) = x^n$ over $[0,1]$. It can be shown that

$$f_n(x) \to f(x) = \begin{cases} 0 & \text{for } x \in [0,1)\,, \\ 1 & \text{for } x = 1\,. \end{cases}$$

However f is not continuous over $[0,1]$.

Exercise. Prove the claim in the example above.

Example. Consider $f_n(x) = 2nx(1-x^2)^{n-1}$ over $[0,1]$. It can be shown that

$$f_n(x) \to f(x) = 0 \qquad \text{pointwise over } [0,1]\,.$$

However, for every $n \in \mathbb{Z}_+$ we have

$$\int_0^1 f_n = 2n \int_0^1 x(1-x^2)^{n-1}\,\mathrm{d}x = -(1-x^2)^n\Big|_0^1 = 1\,,$$

so that

$$\int_0^1 f_n = 1 \not\to \int_0^1 f = 0\,.$$

Exercise. Prove the claim in the example above.

Example. Consider $f_n(x) = \frac{1}{n}\sin(nx)$ over $[-\pi,\pi]$. It is clear that

$$f_n(x) \to f(x) = 0 \qquad \text{pointwise over } [-\pi,\pi]\,.$$

Because $\sin(nx)$ is odd and $[-\pi,\pi]$ is symmetric, we have

$$\int_{-\pi}^{\pi} f_n = \frac{1}{n}\int_{-\pi}^{\pi} \sin(nx)\,\mathrm{d}x = 0\,.$$

Therefore

$$\int_{-\pi}^{\pi} f_n = 0 \to \int_{-\pi}^{\pi} f = 0\,.$$

However, it can be shown that

$$\lim_{n\to\infty} f_n'(x) = \lim_{n\to\infty} \cos(nx) \qquad \text{diverges for every nonzero } x \in [-\pi,\pi]\,.$$

Exercise. Prove the claim in the example above.

Exercise. Find a sequence $\{f_n\}_{n\in\mathbb{N}}$ of continuously differentiable functions over $[0,1]$ such that

$$f_n \to 0 \qquad \text{pointwise over } [0,1]\,, \quad \text{and} \quad f_n'(x) \text{ diverges for } x \in [0,1]\cap\mathbb{Q}\,.$$

12.2. **Uniform Convergence.** In the previous section we saw that the pointwise limit of a sequence of functions may not have many properties we might expect. In this section we introduce the *uniform limit* of a sequence of functions, which will behave better. We motivate this notion from the following characterization of pointwise convergence.

Proposition 12.2. *Let $S \subset \mathbb{R}$. Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of real-valued functions that are each defined over S. Let f be a real-valued functon that is defined over S. Then*

$$f_n \to f \qquad \text{pointwise over } S$$

if and only if for every $x \in S$ and every $\epsilon > 0$ there exists an $n_{x,\epsilon} \in \mathbb{N}$ such that

$$n \geq n_{x,\epsilon} \qquad \Longrightarrow \qquad |f_n(x) - f(x)| < \epsilon\,.$$

Proof. Exercise.

We now define *uniform convergence*, which is a stronger notion of convergence.

Definition 12.2. *Let $S \subset \mathbb{R}$. Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of real-valued functions that are each defined over S. The sequence $\{f_n\}_{n\in\mathbb{N}}$ is said to be* uniformly convergent *or to* converge uniformly *over S if there exists a function f defined over S such that for every ϵ there exists an $n_\epsilon \in \mathbb{N}$ such that for every $x \in S$*

$$n \geq n_\epsilon \qquad \Longrightarrow \qquad |f_n(x) - f(x)| < \epsilon\,.$$

We say f_n converges to f uniformly *over S and call f the* uniform limit *of the sequence $\{f_n\}_{n\in\mathbb{N}}$ over S. We denote this as*

$$f_n \to f \qquad \text{uniformly over } S\,.$$

It should be clear from this definition and from Proposition 12.2 that uniform convergence implies pointwise convergence.

Proposition 12.3. *Let $S \subset \mathbb{R}$. Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of real-valued functions that are each defined over S. Let f be a real-valued functon that is defined over S. If $f_n \to f$ uniformly over S then $f_n \to f$ pointwise over S.*

Proof. Exercise.

Remark. This is why we say uniform convergence is a stronger notion of convergence than pointwise convergence.

The first payoff of this stronger notion is the following.

Proposition 12.4. *Let $S \subset \mathbb{R}$. Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of real-valued functions that are each continuous over S. Let f be a real-valued functon that is defined over S. If $f_n \to f$ uniformly over S then f is continuous over S.*

Proof. Let $x \in S$ be arbitrary. We want to show that f is continuous at x. Let $\epsilon > 0$. Because $f_n \to f$ uniformly over S, there exists $n \in \mathbb{N}$ such that

$$|f_n(z) - f(z)| < \frac{\epsilon}{3} \qquad \text{for every } z \in S\,.$$

Because f_n is continuous over S there exists $\delta > 0$ such that for every $y \in S$

$$|y - x| < \delta \qquad \Longrightarrow \qquad |f_n(y) - f_n(x)| < \frac{\epsilon}{3}\,.$$

Then

$$|y - x| < \delta \quad \Longrightarrow \quad \begin{aligned} |f(y) - f(x)| &\leq |f(y) - f_n(y)| + |f_n(y) - f_n(x)| + |f_n(x) - f(x)| \\ &< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon\,. \end{aligned}$$

Therefore f is continuous at $x \in S$. Because $x \in S$ was arbitrary, we conclude that f is continuous over S. □

We can replace "continuous" by "uniformly continuous" in the foregoing proposition.

Proposition 12.5. *Let $S \subset \mathbb{R}$. Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of real-valued functions that are each uniformly continuous over S. Let f be a real-valued functon that is defined over S. If $f_n \to f$ uniformly over S then f is uniformly continuous over S.*

Proof. Exercise.

Uniform convergence behaves as we might hope for integrals.

Proposition 12.6. *Let $a < b$. Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of real-valued functions that are each Riemann integrable over $[a, b]$. Let f be a real-valued functon that is defined over $[a, b]$. If $f_n \to f$ uniformly over $[a, b]$ then f is Riemann integrable over $[a, b]$ and*

$$\lim_{n\to\infty} \int_a^b f_n = \int_a^b f\,.$$

Proof. Let $\epsilon > 0$. Because $f_n \to f$ uniformly over $[a, b]$ there exists an $n \in \mathbb{N}$ such that

$$|f_n(z) - f(z)| < \frac{\epsilon}{3(b-a)} \qquad \text{for every } z \in [a, b]\,.$$

Because f_n is Riemann integrable over $[a, b]$ there exists a partition P of $[a, b]$ such that

$$0 \leq U(f_n, P) - L(f_n, P) < \frac{\epsilon}{3}\,.$$

Because f_n is bounded over $[a, b]$ and because for every $z \in [a, b]$

$$f_n(z) - \frac{\epsilon}{3(b-a)} < f(z) < f_n(z) + \frac{\epsilon}{3(b-a)}\,,$$

we conclude that f is bounded over $[a, b]$ and that

$$L(f_n, P) - \frac{\epsilon}{3} < L(f, P) \leq U(f, P) < U(f_n, P) + \frac{\epsilon}{3}\,.$$

Then

$$\begin{aligned} 0 < U(f, P) - L(f, P) &< U(f_n, P) + \frac{\epsilon}{3} - L(f_n, P) + \frac{\epsilon}{3} \\ &< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon\,. \end{aligned}$$

Therefore f is Riemann integrable over $[a, b]$. □

The story for the convergence of derivatives is more complicated. For example, consider the sequence of functions over $\mathbb{R}$ given by

$$f_n(x) = \frac{1}{3^n} \sin(9^n x)\ .$$

It should be clear that $f_n \to 0$ uniformly over $\mathbb{R}$. Each f_n is differentiable over $\mathbb{R}$ with

$$f_n'(x) = 3^n \cos(9^n x)\ .$$

Notice that if $x = (k\pi)/9^m$ for some $k \in \mathbb{Z}$ and $m \in \mathbb{N}$ then for every $n > m$ we have

$$f_n'(x) = 3^n \cos\big(9^{n-m} k\pi\big) = (-1)^k 3^n \,.$$

Therefore the real sequence $\{f_n'(x)\}$ diverges to $+\infty$ when k is even and diverges to $-\infty$ when k is odd. But the sets

$$\left\{ \frac{k\pi}{9^m} \,:\, k \text{ is even and } m \in \mathbb{N} \right\} , \qquad \left\{ \frac{k\pi}{9^m} \,:\, k \text{ is odd and } m \in \mathbb{N} \right\}$$

are each dense in $\mathbb{R}$. It is clear that the sequence of functions $\{f_n'\}$ is not well behaved as $n \to \infty$.

The foregoing discussion shows that knowing a sequence of continuously differentiable functions $\{f_n\}$ converges uniformly to a continuously differentiable function f does not imply that the sequence of their derivatives $\{f_n'\}$ will converge pointwise to f', or even that it will converge pointwise at all! However, if we assume that the sequence derivatives $\{f_n'\}$ converges uniformly then we can use the Fundamental Theorems of Calculus and Propositions 12.4 and 12.6 to obtain the following useful theorem.

Proposition 12.7. *Let $a < b$. Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of real-valued continuously differentiable functions over $[a, b]$. Let f and g be real-valued functions over $[a, b]$ such that*

- *$f_n \to f$ pointwise over $[a, b]$;*
- *$f_n' \to g$ uniformly over $[a, b]$.*

Then g is continuous over $[a, b]$ and f is continuously differentiable over $[a, b]$ with $f' = g$.

Proof. Because each f_n is continuously differentiable over $[a, b]$, the First Fundamental Theorem of Calculus implies that

$$f_n(x) = f_n(a) + \int_a^x f_n' \qquad \text{for every } x \in [a, b] \,.$$

Because $f_n \to f$ pointwise over $[a, b]$ we have

$$\lim_{n\to\infty} f_n(x) = f(x) \qquad \text{for every } x \in [a, b] \,.$$

Because $f_n' \to g$ uniformly over $[a, b]$ and each f_n' is continuous, Proposition 12.4 implies that g is continuous over $[a, b]$. Moreover, Propostion 12.6 implies that

$$\lim_{n\to\infty} \int_a^x f_n' = \int_a^x g \qquad \text{for every } x \in [a, b] \,.$$

Therefore

$$f(x) = f(a) + \int_a^x g \qquad \text{for every } x \in [a, b] \,.$$

Because g is continuous, the Second Fundamental Theorem of Calculus implies that the above right-hand side is continuously differentiable over $[a, b]$ and that its derivative is g. Therefore f is continuously differentiable over $[a, b]$ and $f' = g$. $\square$

12.3. **Uniform Norms.** Uniform convergence is best studied with a tool called the *uniform norm*. Let $S \subset \mathbb{R}$. Let $\mathrm{B}(S)$ denote the set of all bounded functions $f : S \to \mathbb{R}$. Then for every $f \in \mathrm{B}(S)$ we define its uniform norm $\|f\|_{\mathrm{B}(S)}$ by

$$\|f\|_{\mathrm{B}(S)} = \sup\big\{|f(x)| \,:\, x \in S\big\}\,. \tag{12.2}$$

Clearly, $f \in \mathrm{B}(S)$ if and only if $\|f\|_{\mathrm{B}(S)} < \infty$.

Exercise. Let $S \subset \mathbb{R}$. Show that $f \in \mathrm{B}(S)$ if and only if $\|f\|_{\mathrm{B}(S)} < \infty$.

Remark. Uniform norms were introduced by Karl Weierstrass in the mid 1800s, but he neither call them that nor denoted them as we do here. He used the notation $M(f)$, which he called the *majorizer* of f. The modern name and notation used here where introduced in the 1900s, when it was realized that majorizers are a special case of the concept of a *norm*.

Before we show how the set $\mathrm{B}(S)$ and its uniform norm are connected to uniform convergence, we collect some of their basic properties.

Proposition 12.8. *Let $S \subset \mathbb{R}$. Then for every $\alpha \in \mathbb{R}$ and every $f, g \in \mathrm{B}(S)$ we have*

$$\alpha f \in \mathrm{B}(S)\,, \qquad f + g \in \mathrm{B}(S)\,, \qquad \text{and} \qquad fg \in \mathrm{B}(S)\,. \tag{12.3}$$

Moreover, for every $\alpha \in \mathbb{R}$ and every $f, g \in \mathrm{B}(S)$ the uniform norm $\|\cdot\|_{\mathrm{B}(S)}$ satisfies:

- $\|f\|_{\mathrm{B}(S)} \geq 0$, — *nonegativity;*
- $\|f\|_{\mathrm{B}(S)} = 0$ *if and only if* $f = 0$, — *definitness;*
- $\|\alpha f\|_{\mathrm{B}(S)} = |\alpha|\|f\|_{\mathrm{B}(S)}$, — *homogeneity;*
- $\|f + g\|_{\mathrm{B}(S)} \leq \|f\|_{\mathrm{B}(S)} + \|g\|_{\mathrm{B}(S)}$, — *triangle inequality;*
- $\|fg\|_{\mathrm{B}(S)} \leq \|f\|_{\mathrm{B}(S)}\|g\|_{\mathrm{B}(S)}$, — *product inequality.*

Proof. Exercise. □

Exercise. Prove Proposition 12.8.

Remark. The first two properties in (12.3) state that the set $\mathrm{B}(S)$ is a *linear space* over the reals. All three properties in (12.3) state that the set $\mathrm{B}(S)$ is an *algebra* over the reals. The first four properties of $\|\cdot\|_{\mathrm{B}(S)}$ listed in Proposition 12.8 are shared by all *norms*. You will meet these abstractions in later courses.

Remark. Proposition 12.8 shows that if f and g are functions over S such that $f - g \in \mathrm{B}(S)$ then we can think of $\|f - g\|_{\mathrm{B}(S)}$ as a measure of distance between f and g.

The connection of the set $\mathrm{B}(S)$ and its uniform norm with uniform convergence is provided by the following characterization, due to Weierstrass.

Proposition 12.9. *Let $S \subset \mathbb{R}$. Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of real-valued functions that are each defined over S. Let f be a real-valued functon that is defined over S. Then $f_n \to f$ uniformly over S as $n \to \infty$ if and only if*

$$f_n - f \in \mathrm{B}(S) \ \text{eventually} \qquad \text{and} \qquad \|f_n - f\|_{\mathrm{B}(S)} \to 0 \qquad \text{as } n \to \infty\,.$$

Proof. Exercise. □

Exercise. Prove Proposition 12.9.

12.4. **Uniformly Cauchy.** In 1841 Weierstrass gave a beautiful extension to sequences of functions of the Cauchy Criterion for convergence of sequences of real numbers. We begin by defining what it means for a sequence of functions to be uniformly Cauchy.

Definition 12.3. *Let $S \subset \mathbb{R}$. Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of real-valued functions that are each defined over S. We say that the sequence $\{f_n\}_{n\in\mathbb{N}}$ is* uniformly Cauchy *over S if for every $\epsilon > 0$ there exists $N_\epsilon \in \mathbb{N}$ such that*

$$m, n \geq N_\epsilon \quad \Longrightarrow \quad \|f_m - f_n\|_{\mathrm{B}(S)} < \epsilon \,. \tag{12.4}$$

Remark. Notice that this definition is analogous to the definition of what it means for a sequence of reals to be Cauchy, with the uniform norm playing the role of the absolute value.

The key fact about uniformly Cauchy sequences is the following.

Proposition 12.10. *Let $S \subset \mathbb{R}$. Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of real-valued functions that is* uniformly Cauchy *over S. Then for every $\epsilon > 0$ there exists $N_\epsilon \in \mathbb{N}$ such that for every $x \in S$ we have*

$$m, n \geq N_\epsilon \quad \Longrightarrow \quad |f_m(x) - f_n(x)| < \epsilon \,. \tag{12.5}$$

In particular, for every $x \in S$ the real sequence $\{f_n(x)\}_{n\in\mathbb{N}}$ is Cauchy.

Proof. The result follows directly from the definition of uniformly Cauchy (12.4) and the fact that for every $x \in S$ and every $m, n \in \mathbb{N}$ we have

$$|f_m(x) - f_n(x)| \leq \|f_m - f_n\|_{\mathrm{B}(S)} \,.$$

The details are left as an exercise. □

Exercise. Prove Proposition 12.10.

Weierstrass gave the following characterization of uniformly convergent sequences in terms of sequences being uniformly Cauchy.

Proposition 12.11. (Weierstrass Criterion) *Let $S \subset \mathbb{R}$. Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of real-valued functions that are each defined over S. Then the sequence $\{f_n\}_{n\in\mathbb{N}}$ is uniformly convergent over S if and only if it is uniformly Cauchy over S.*

Proof. The proof of the direction ($\Longrightarrow$) is similar to that for the Cauchy Criterion for real sequences, so we do it first.

($\Longrightarrow$) Let $\{f_n\}_{n\in\mathbb{N}}$ be uniformly convergent over S. This means there exists a function $f : S \to \mathbb{R}$ such that $f_n \to f$ uniformly over S. We must show that $\{f_n\}_{n\in\mathbb{N}}$ is uniformly Cauchy over S. Let $\epsilon > 0$. Because $f_n \to f$ uniformly over S, by Proposition 12.9 there exists an $N_\epsilon \in \mathbb{N}$ such that

$$n \geq N_\epsilon \quad \Longrightarrow \quad \|f_n - f\|_{\mathrm{B}(S)} < \frac{\epsilon}{2} \,.$$

Then for every $m, n \geq N_\epsilon$ we have

$$\begin{aligned} \|f_n - f_m\|_{\mathrm{B}(S)} &= \|(f_n - f) - (f_m - f)\|_{\mathrm{B}(S)} \\ &\leq \|f_n - f\|_{\mathrm{B}(S)} + \|f_m - f\|_{\mathrm{B}(S)} < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \,. \end{aligned}$$

Therefore $\{f_n\}_{n\in\mathbb{N}}$ is uniformly Cauchy over S.

The proof of the direction ($\Longleftarrow$) uses Proposition 12.10 and the Cauchy Criterion for real sequences to construct the limiting function f.

($\Longleftarrow$) Let $\{f_n\}_{n\in\mathbb{N}}$ be uniformly Cauchy over S. We want to show that there exists a function $f : S \to \mathbb{R}$ such that $f_n \to f$ uniformly over S. Proposition 12.10 implies that for every $x \in S$ the sequence of real numbers $\{f_n(x)\}_{n\in\mathbb{N}}$ is Cauchy, and thereby is convergent by the Cauchy Criterion. Therefore we can define $f : S \to \mathbb{R}$ by

$$f(x) = \lim_{n\to\infty} f_n(x) \qquad \text{for every } x \in S\,.$$

This shows that $f_n \to f$ pointwise over S. We still must show that $f_n \to f$ uniformly over S.

Let $\epsilon > 0$. Let $\eta_\epsilon \in (0, \epsilon)$. Because $\{f_n\}_{n\in\mathbb{N}}$ is uniformly Cauchy over S there exists $N_\epsilon \in \mathbb{N}$ such that

$$m, n \geq N_\epsilon \qquad \Longrightarrow \qquad \|f_n - f_m\|_{\mathrm{B}(S)} < \eta_\epsilon\,.$$

Therefore for every $x \in S$ we have

$$m, n \geq N_\epsilon \qquad \Longrightarrow \qquad |f_n(x) - f_m(x)| < \eta_\epsilon\,.$$

By letting $m \to \infty$ in the above we see that

$$n \geq N_\epsilon \qquad \Longrightarrow \qquad |f_n(x) - f(x)| \leq \eta_\epsilon < \epsilon\,,$$

which implies that

$$n \geq N_\epsilon \qquad \Longrightarrow \qquad \|f_n - f\|_{\mathrm{B}(S)} \leq \eta_\epsilon < \epsilon\,.$$

Therefore $f_n \to f$ uniformly over S. $\square$

13. Series of Functions

In the 1800s much mathematical attention was focused on series of simple functions that defined other functions. The two most important examples of such series are *power series* and *trigonometric series*.

Power series have the form

$$\sum_{k=0}^{\infty} c_k x^k \,. \tag{13.1}$$

These define functions as sums of monomials. They are natural extensions of polynomials that were studied long before the time of Newton and Liebniz. Newton used power series to express solutions of differential equations. This technique became wide spread in the 1800s.

Trigonometric series have the form

$$\tfrac{1}{2}a_0 + \sum_{k=1}^{\infty} \big(a_k \cos(kx) + b_k \sin(kx) \big) \,. \tag{13.2}$$

These define 2π-periodic functions as sums of basic trigonometric functions. They were used in the 1700s by D. Bernoulli and L. Euler to express certain continuous periodic functions. In 1807 J. Fourier claimed that periodic functions with jump discontinuities could also be expressed as such a sum. At that time this claim was controversial because then many (maybe most) mathematicians did not believe that summing nice functions like $\cos(kx)$ and $\sin(kx)$ could ever produce functions with jump discontinuities! This controversy drove much of the work in the 1800s to understand the sense in which such series converge. For example, Riemann developed his integral theory as part of this effort. Eventually Fourier was proven correct. Trigonometric series are also called Fourier series to recognize the importance of his work.

13.1. **Uniform Convergence for Series of Functions.**

Now consider a sequence $\{h_k\}_{k=0}^{\infty}$ of real-valued functions defined over a common domain $D \subset \mathbb{R}$. For each $x \in D$ consider the infinite series

$$\sum_{k=0}^{\infty} h_k(x) \,.$$

Let $S = \{x \in D : \text{the above series converges}\}$. Define a function $f : S \to \mathbb{R}$ by

$$f(x) = \sum_{k=0}^{\infty} h_k(x) \qquad \text{for every } x \in S \,. \tag{13.3}$$

This states that the associated sequence of partial sums converges to f pointwise over S. The set S is called the *domain of convergence* for the series.

The following notions of uniform convergence for series of functions are natural.

Definition 13.1. *The series of functions (13.3) is said to* converge uniformly *over a set S if the associated sequence of partial sums converges to f uniformly over S. It is said to* converge absolutely uniformly *over a set S if the sequence of partial sums associated with $|h_k|$ converges uniformly over S.*

With this definition of uniform convergence for series the following propositions are immediate corollaries of Propositions 12.4, 12.5, 12.6, and 12.7.

Proposition 13.1. *Let $S \subset \mathbb{R}$. Let $\{h_k\}_{k=0}^{\infty}$ be a sequence of functions in* $\mathrm{B}(S)$ *such that*

$$\sum_{k=0}^{\infty} h_k \qquad \text{converges uniformly over } S\,.$$

Let $f : S \to \mathbb{R}$ be defined by

$$f(x) = \sum_{k=0}^{\infty} h_k(x) \qquad \text{for every } x \in S\,.$$

Then we have the following.

- *If each h_k is continuous over S then f is continuous over S.*
- *If each h_k is uniformly continuous over S then f is uniformly continuous over S.*

Proof. Exercise.

The next proposition gives conditions that permit term-by-term integration of infinite series.

Proposition 13.2. *Let $[a, b] \subset \mathbb{R}$. Let $\{h_k\}_{k=0}^{\infty}$ be a sequence of functions in* $\mathrm{B}([a, b])$ *such that*

$$\sum_{k=0}^{\infty} h_k \qquad \text{converges uniformly over } [a, b]\,.$$

Let $f : [a, b] \to \mathbb{R}$ be defined by

$$f(x) = \sum_{k=0}^{\infty} h_k(x) \qquad \text{for every } x \in [a, b]\,.$$

If each h_k is Riemann integrable over $[a, b]$ then f is Riemann integrable over $[a, b]$ with

$$\int_a^b f = \sum_{k=0}^{\infty} \int_a^b h_k\,.$$

Proof. Exercise.

The next proposition gives conditions that permit term-by-term differentiation of infinite series.

Proposition 13.3. *Let $[a, b] \subset \mathbb{R}$. Let $\{h_k\}_{k=0}^{\infty}$ be a sequence of continuously differentiable functions over $[a, b]$ such that*

$$\sum_{k=0}^{\infty} h_k \qquad \text{and} \qquad \sum_{k=0}^{\infty} h_k' \qquad \text{converge uniformly over } [a, b]\,.$$

Let $f : [a, b] \to \mathbb{R}$ be defined by

$$f(x) = \sum_{k=0}^{\infty} h_k(x) \qquad \text{for every } x \in [a, b]\,.$$

Then f is continuously differentiable over $[a, b]$ with

$$f'(x) = \sum_{k=0}^{\infty} h_k'(x) \qquad \text{for every } x \in [a, b]\,.$$

Proof. Exercise.

Exercise. Prove Proposition 13.1.

Exercise. Prove Proposition 13.2.

Exercise. Prove Proposition 13.3.

In order to apply Propositions 13.1, 13.2, and 13.3, we need a useful criterion that tells us when a series of functions converges uniformly over a set S. Weierstrass gave such a criterion for series based upon Proposition 12.11, the Weierstrass Criterion for sequences of functions.

Proposition 13.4. (Weierstrass M-Test) *Let $S \subset \mathbb{R}$. Let $\{h_k\}_{k=0}^{\infty}$ be a sequence of functions in* $\mathrm{B}(S)$ *that satisfies*

$$\sum_{k=0}^{\infty} \|h_k\|_{\mathrm{B}(S)} < \infty \,. \tag{13.4}$$

Then

$$\sum_{k=0}^{\infty} h_k \qquad \text{converges absolutely uniformly over } S \,. \tag{13.5}$$

Remark. This is called the Weierstrass M-Test because he used the notation $M_k = \|h_k\|_{\mathrm{B}(S)}$ in (13.4). He used this notation because he called M_k the majorizer of h_k.

Proof. Let f_n be the n^{th} partial sum of (13.4), which for each $n \in \mathbb{N}$ is the function over S defined by

$$f_n(x) = \sum_{k=0}^{n} h_k(x) \qquad \text{for every } x \in S \,.$$

We will show that the sequence $\{f_n\}$ is uniformly Cauchy, and thereby is uniformly convergent by the Weierstrass Criterion, Proposition 12.11.

Let $\epsilon > 0$. By condition (13.4) there exists an $N_\epsilon \in \mathbb{N}$ such that

$$\sum_{k=N_\epsilon}^{\infty} \|h_k\|_{\mathrm{B}(S)} < \epsilon \,.$$

Let $m, n \geq N_\epsilon$. Without loss of generality we can assume that $m < n$. Then

$$\|f_n - f_m\|_{\mathrm{B}(S)} = \left\| \sum_{k=m+1}^{n} h_k \right\|_{\mathrm{B}(S)} \leq \sum_{k=m+1}^{n} \|h_k\|_{\mathrm{B}(S)} \leq \sum_{k=N_\epsilon}^{\infty} \|h_k\|_{\mathrm{B}(S)} < \epsilon \,.$$

Therefore the sequence $\{f_n\}$ is uniformly Cauchy, and thereby is uniformly convergent by the Weierstrass Criterion, Proposition 12.11.

Finally, because the sequence $\{h_k\}_{k=0}^{\infty}$ satisfies condition (13.4), the sequence $\{|h_k|\}_{k=0}^{\infty}$ also satisfies condition (13.4). Therefore, by what we have already proved, the series associated with $\{|h_k|\}_{k=0}^{\infty}$ is uniformly convergent over S. Therefore the series in (13.5) converges absolutely uniformly over S. $\square$

13.2. **Power Series.** We now apply the propositions from the previous section to power series. In order to apply the Root Test to the power series (13.1), we compute

$$\rho = \limsup_{k\to\infty} \left|c_k x^k\right|^{\frac{1}{k}} = |x| \limsup_{k\to\infty} |c_k|^{\frac{1}{k}} .$$

The Root Test states the series converges absolutely when $\rho < 1$ and diverges when $\rho > 1$. Therefore the series (13.1) converges absolutely when $|x| < R$ and diverges when $|x| > R$ where

$$\frac{1}{R} = \limsup_{k\to\infty} |c_k|^{\frac{1}{k}} , \tag{13.6}$$

with $R = 0$ when the lim sup is ∞, and $R = \infty$ when the lim sup is 0. The number R is called the *radius of convergence* for the power series (13.1).

The Root Test shows that the set of x for which the power series converges is an interval, the so-called *interval of convergence* of the power series (13.1).

- When $R = \infty$ the interval of convergence is $(-\infty, \infty) = \mathbb{R}$.
- When $R = 0$ the interval of convergence is $[0, 0]$, which is the singleton set $\{0\}$.
- When $0 < R < \infty$ the interval of convergence will be either

$$(-R, R) , \qquad [-R, R) , \qquad (-R, R] , \qquad \text{or} \qquad [-R, R] ,$$

depending upon the convergence or divergence of the series

$$\sum_{k=0}^{\infty} c_k (-R)^k \qquad \text{and} \qquad \sum_{k=0}^{\infty} c_k R^k .$$

These endpoint cases must be analyzed by other convergence tests.

The power series (13.1) converges to a function $f(x)$ over this interval of convergence and diverges outside of it. This result is often called the Cauchy-Hadamard Theorem. It was published by Cauchy in 1821 and by Hadamard in 1888.

We now consider those power series (13.1) with a radius of convergence $R > 0$ and interval of convergence I. We have the following.

Proposition 13.5. *Let the power series (13.1) have radius of convergence $R > 0$ and interval of convergence I. For every $x \in I$ let $f(x)$ be the sum of this series.*

Then the series converges absolutely to f uniformly over $[a, b]$ for every $[a, b] \subset (-R, R)$. The function f is uniformly continuous over every $[a, b] \subset (-R, R)$ and

$$\int_a^b f = \sum_{k=0}^{\infty} c_k \int_a^b x^k = \sum_{k=0}^{\infty} c_k \, \frac{b^{k+1} - a^{k+1}}{k+1} . \tag{13.7}$$

The function f is infinitely differentiable over $(-R, R)$ with $f^{(n)}$ for each $n \in \mathbb{N}$ given by

$$f^{(n)}(x) = \sum_{k=0}^{\infty} \frac{(k+n)!}{n!} \, c_{k+n} \, x^k \qquad \text{for every } x \in (-R, R) . \tag{13.8}$$

Remark. This result states that functions defined by power series with radius of convergence R are infinitely differentiable over $(-R, R)$ with derivatives that are also given by power series with the same radius of convergence. Moreover, formula (13.8) shows that the power series for the derivatives are obtained via term-by-term differentation of the power series for the function.

Proof. Let $[a,b] \subset (-R,R)$. Let $r \in (0,R)$ such that $[a,b] \subset [-r,r]$. Let $s \in (r,R)$. Because

$$\limsup_{k\to\infty} |c_k|^{\frac{1}{k}} = \frac{1}{R} < \frac{1}{s}\,.$$

It follows that

$$|c_k| < \frac{1}{s^k} \qquad \text{eventually as } k \to \infty\,. \tag{13.9}$$

This implies that

$$\|c_k x^k\|_{\mathrm{B}([a,b])} < \frac{r^k}{s^k} \qquad \text{eventually as } k \to \infty\,.$$

By the Direct Comparison Test, we conclude that

$$\sum_{k=0}^{\infty} \|c_k x^k\|_{\mathrm{B}([a,b])} < \infty\,.$$

Then by the Weierstrass M-Test, the series converges absolutely to f uniformly over $[a,b]$, where $[a,b] \subset (-R,R)$ was arbitrary.

Because $c_k x^k$ is uniformly continuous over every $[a,b] \subset (-R,R)$, Proposition 13.1 implies that the function f is uniformly continuous over every $[a,b] \subset (-R,R)$. Proposition 13.2 then implies that (13.7) holds.

In order to show that f is infinitely differentiable with $f^{(n)}$ given by (13.8) for every $n \in \mathbb{N}$, we first show that for every $[a,b] \subset (-R,R)$ we have

$$\sum_{k=0}^{\infty} \frac{(k+n)!}{n!} \|c_{k+n}\, x^k\|_{\mathrm{B}([a,b])} < \infty\,. \tag{13.10}$$

Let $r \in (0,R)$ such that $[a,b] \subset [-r,r]$. Let $s \in (r,R)$. By (13.9) and the fact that $|x^k| \leq r^k$ we have

$$\|c_{k+n} x^k\|_{\mathrm{B}([a,b])} < \frac{r^k}{s^{k+n}} \qquad \text{eventually as } k \to \infty\,.$$

Because $r < s$, the Direct Comparison Test implies that (13.10) holds because it can be shown for every $n \in \mathbb{N}$ that

$$\sum_{k=0}^{\infty} \frac{(k+n)!}{n!} \frac{r^k}{s^{k+n}} = \frac{n!\, s}{(s-r)^{n+1}} < \infty\,. \tag{13.11}$$

Therefore the Weierstrass M-Test, Proposition 13.4, implies that

$$\sum_{k=0}^{\infty} \frac{(k+n)!}{n!}\, c_{k+n}\, x^k \qquad \text{converges uniformly over } [a,b]\,.$$

Hence, if f is n-times continuously differentiable and $f^{(n)}$ is given by (13.8) for some $n \in \mathbb{N}$ then Proposition 13.3 implies that f is $(n+1)$-times continuously differentiable and that $f^{(n+1)}$ is given by (13.8). Therefore by induction on n we conclude that f is infinitely differentiable and that $f^{(n)}$ is given by (13.8) for every $n \in \mathbb{N}$. □

Exercise. Prove (13.11) by first showing that for every $\rho \in (0,1)$ and every $n \in \mathbb{N}$ we have

$$\sum_{k=0}^{\infty} \frac{(k+n)!}{n!}\, \rho^k = \frac{n!}{(1-\rho)^{n+1}}\,.$$

14. Everywhere Continuous, Nowhere Differentiable Functions

In 1872 Karl Weierstrass startled the mathematical world by publishing examples of functions that where continuous everywhere but differentiable nowhere. Specifically, he considered functions of the form

$$f(x) = \sum_{k=0}^{\infty} a^k \cos\big(b^k x\big) \,, \qquad \text{where } 0 < a < 1 < b \text{ with } ab \geq 1 \,. \tag{14.1}$$

It is clear by the Weierstrass M-test that this series converges uniformly over $\mathbb{R}$, and thereby defines a function f that is continuous over $\mathbb{R}$. It is equally clear that if $ab < 1$ then this function would be continuously differentiable over $\mathbb{R}$ with

$$f'(x) = -\sum_{k=0}^{\infty} a^k b^k \sin\big(b^k x\big) \,.$$

Therefore the condition $ab \geq 1$ is necessary for the function f defined by (14.1) to be nondifferentable somewhere in $\mathbb{R}$. Weierstrass showed f is not differentiable anywhere in $\mathbb{R}$ when b is an odd integer that satisfies $ab > 1 + \frac{3}{2}\pi$. These conditions on b were reduced to the necessary condition $ab \geq 1$ in 1916 when Godfrey Harold Hardy showed that every function f defined by (14.1) is not differentiable anywhere in $\mathbb{R}$!

We will not prove the theorem of Hardy. Rather, we will give a different construction of an everywhere continuous, nowhere differentiable function built up from so-called zigzag functions.

14.1. **Zigzag Functions.**

The basic building block of our construction will be the unit zigzag function z, which we define for every $x \in \mathbb{R}$ by

$$\mathrm{z}(x) = \begin{cases} 2n - x & \text{when } x \in [2n-1, 2n) \text{ for some } n \in \mathbb{Z} \,, \\ x - 2n & \text{when } x \in [2n, 2n+1) \text{ for some } n \in \mathbb{Z} \,. \end{cases} \tag{14.2}$$

Clearly z is a continuous, periodic function with

$$\mathrm{z}(x+2) = \mathrm{z}(x) \,, \qquad 0 \leq \mathrm{z}(x) \leq 1 \,, \qquad \text{for every } x \in \mathbb{R} \,. \tag{14.3}$$

It is also piecewise differentiable over $\mathbb{R}$ with

$$\mathrm{z}'(x) = \begin{cases} -1 & \text{when } x \in (2n-1, 2n) \text{ for some } n \in \mathbb{Z} \,, \\ 1 & \text{when } x \in (2n, 2n+1) \text{ for some } n \in \mathbb{Z} \,. \end{cases} \tag{14.4}$$

It is not differentiable for $x \in \mathbb{Z}$.

Let $\{\mathrm{z}_k\}_{k\in\mathbb{N}}$ be the sequence of zigzag functions defined by

$$\mathrm{z}_k(x) = 4^{-k}\,\mathrm{z}\big(4^k x\big) \,. \tag{14.5}$$

Clearly z_k is a continuous, periodic function with

$$\mathrm{z}_k(x + 2\cdot 4^{-k}) = \mathrm{z}_k(x) \,, \qquad 0 \leq \mathrm{z}_k(x) \leq 4^{-k} \,, \qquad \text{for every } x \in \mathbb{R} \,. \tag{14.6}$$

It is piecewise differentiable over $\mathbb{R}$ with

$$\mathrm{z}_k'(x) = \begin{cases} -1 & \text{when } x \in (\frac{2n-1}{4^k}, \frac{2n}{4^k}) \text{ for some } n \in \mathbb{Z} \,, \\ 1 & \text{when } x \in (\frac{2n}{4^k}, \frac{2n+1}{4^k}) \text{ for some } n \in \mathbb{Z} \,. \end{cases} \tag{14.7}$$

It is not differentiable at points $x \in \mathbb{R}$ such that $4^k x \in \mathbb{Z}$.

14.2. **An Everywhere Continuous Function.** Because the unit zigzag function z satisfies $0 \leq \mathrm{z}(x) \leq \mathrm{z}(1) = 1$ for every $x \in \mathbb{R}$, we see that it is in $\mathrm{B}(\mathbb{R})$ and satisfies

$$\|\mathrm{z}\|_{\mathrm{B}(\mathbb{R})} = 1\,.$$

Because $\mathrm{z}_k(x) = 4^{-k}\,\mathrm{z}(4^k x)$ for every $x \in \mathbb{R}$ and $k \in \mathbb{N}$, we see that $\mathrm{z}_k \in \mathrm{B}(\mathbb{R})$ for every $k \in \mathbb{N}$, and that

$$\|\mathrm{z}_k\|_{\mathrm{B}(\mathbb{R})} = 4^{-k} \qquad \text{for every } k \in \mathbb{N}\,. \tag{14.8}$$

The geometric series summation formula with $r = 4^{-1}$ then shows that

$$\sum_{k=0}^{\infty} \|\mathrm{z}_k\|_{\mathrm{B}(\mathbb{R})} = \sum_{k=0}^{\infty} 4^{-k} = \frac{1}{1 - 4^{-1}} = \frac{4}{3} < \infty\,.$$

The Weierstrass M-Test then implies that the series

$$\sum_{k=0}^{\infty} \mathrm{z}_k \qquad \text{converges absolutely uniformly over } \mathbb{R}\,. \tag{14.9}$$

Therefore we can define a function $h : \mathbb{R} \to \mathbb{R}$ by

$$h(x) = \sum_{k=0}^{\infty} \mathrm{z}_k(x) \qquad \text{for every } x \in \mathbb{R}\,. \tag{14.10}$$

By Proposition 13.1 the function h is continuous over $\mathbb{R}$. In the next section we will show that h is not differentiable at any point of $\mathbb{R}$.

Exercise. Prove that h defined by (14.10) is periodic with period 2.

Exercise. Prove that h defined by (14.10) is unifromly continuous over $\mathbb{R}$.

14.3. **That is Nowhere Differentiable.** We want to show that the function h defined by (14.10) is nowhere differentiable. This means that h is not differentiable at any point of $\mathbb{R}$. More specifically, we want to show for every $x \in \mathbb{R}$ that

$$\lim_{y \to x} \frac{h(y) - h(x)}{y - x} \qquad \text{diverges}\,. \tag{14.11}$$

It suffices to show for every $x \in \mathbb{R}$ that there exists a sequence $\{x_n\}_{n\in\mathbb{N}} \subset \mathbb{R} - \{x\}$ such that

$$\lim_{n\to\infty} x_n = x\,, \qquad \text{but} \qquad \lim_{n\to\infty} \frac{h(x_n) - h(x)}{x_n - x} \qquad \text{diverges}\,. \tag{14.12}$$

The construction of the sequence $\{x_n\}_{n\in\mathbb{N}}$ is based upon the observation that for every $x \in \mathbb{R}$ the unit zigzag function z either is monotonic on $[x, x + \frac{1}{2}]$ or is monotonic on $[x - \frac{1}{2}, x]$. The graph of z should make this clear. It follows that for every $x \in \mathbb{R}$ and every $n \in \mathbb{N}$ the zigzag function z_n either is monotonic on $[x, x + \frac{1}{2}\cdot 4^{-n}]$ or is monotonic on $[x - \frac{1}{2}\cdot 4^{-n}, x]$.

Now let $x \in \mathbb{R}$. By the paragraph above, we can define the sequence $\{x_n\}_{n\in\mathbb{N}}$ by

$$x_n = \begin{cases} x + \frac{1}{2}\cdot 4^{-n} & \text{if } \mathrm{z}_n \text{ is monotonic on } [x, x + \frac{1}{2}\cdot 4^{-n}]\,, \\ x - \frac{1}{2}\cdot 4^{-n} & \text{otherwise}\,. \end{cases} \tag{14.13}$$

Because $|x_n - x| = \frac{1}{2}\cdot 4^{-n}$, it is clear that

$$\lim_{n\to\infty} x_n = x\,.$$

In order to establish (14.12) it remains to prove that

$$\lim_{n\to\infty} \frac{h(x_n) - h(x)}{x_n - x} \qquad \text{diverges}\,. \tag{14.14}$$

The proof of (14.14) is based upon two more observations. First, if $k > n$ then $4^{k-n+1} \in \mathbb{Z}$ and, because z is 2-periodic, we have

$$\begin{aligned} \mathrm{z}_k(x_n) = 4^{-k}\,\mathrm{z}\big(4^k\left(x \pm \tfrac{1}{2}\cdot 4^{-n}\right)\big) &= 4^{-k}\,\mathrm{z}\big(4^k x \pm \tfrac{1}{2}\cdot 4^{k-n}\big) \\ &= 4^{-k}\,\mathrm{z}\big(4^k x \pm 2\cdot 4^{k-n-1}\big) = 4^{-k}\,\mathrm{z}\big(4^k x\big) = \mathrm{z}_k(x)\,. \end{aligned}$$

Second, if $k \leq n$ then:

- if z_n is monotonic on $[x, x+\frac{1}{2}\cdot 4^{-n}]$ then z_k is monotonic on $[x, x+\frac{1}{2}\cdot 4^{-n}]$;
- if z_n is monotonic on $[x-\frac{1}{2}\cdot 4^{-n}, x]$ then z_k is monotonic on $[x-\frac{1}{2}\cdot 4^{-n}, x]$.

This means that if $k \leq n$ then

$$\frac{\mathrm{z}_k(x_n) - \mathrm{z}_k(x)}{x_n - x} = \pm 1\,.$$

Upon combining definition (14.10) of h with these observations we see that

$$\frac{h(x_n) - h(x)}{x_n - x} = \sum_{k=0}^{\infty} \frac{\mathrm{z}_k(x_n) - \mathrm{z}_k(x)}{x_n - x} = \sum_{k=0}^{n} \frac{\mathrm{z}_k(x_n) - \mathrm{z}_k(x)}{x_n - x} = \sum_{k=0}^{n} \pm 1\,. \tag{14.15}$$

It follows that

$$\left| \frac{h(x_{n+1}) - h(x)}{x_{n+1} - x} - \frac{h(x_n) - h(x)}{x_n - x} \right| = 1\,,$$

whereby (14.14) is satisfied for this sequence. Therefore h is nowhere differentiable over $\mathbb{R}$. □

Remark. It follows from (14.15) that

$$\frac{h(x_n) - h(x)}{x_n - x} = \begin{cases} \text{odd} & \text{when } n \text{ is even}\,, \\ \text{even} & \text{when } n \text{ is odd}\,. \end{cases}$$

Remark. It turns out that most continuous functions are nowhere differentiable. This sense in which this is true is beyond what we can do here.

15. Weierstrass Approximation Theorem

In 1885 Karl Weierstrass showed that every continuous real-valued function f over a compact interval $[a,b]$ can be uniformly approximated by polynomials. Without loss of generality we can consider the case where $[a,b]=[0,1]$.

15.1. Bernstein Polynomial Approximation.

In 1912 Sergei Bernstein gave a proof of the Weierstrass Theorem that used an explicit sequence of approximating polynomials. Specifically, given any $f \in C([0,1])$ and $n \in \mathbb{Z}_+$ he defined the so-called n^{th} *Bernstein approximating polynomial* by

$$\mathcal{B}^n f(x) = \sum_{m=0}^{n} f\left(\tfrac{m}{n}\right) b_m^n(x) \,, \tag{15.1}$$

where $b_m^n(x)$ are the so-called Bernstein basis polynomials that for every $n \in \mathbb{Z}_+$ and every $m \in \{0, \cdots, n\}$ are defined by

$$b_m^n(x) = \frac{n!}{m!(n-m)!}\, x^m (1-x)^{n-m} \,. \tag{15.2}$$

Each $b_m^n(x)$ is a polynomial of degree n that is positive over the interval $(0,1)$ with a unique maximizer at $x = \frac{m}{n}$. For every $n \in \mathbb{Z}_+$ the Bernstein basis polynomials $\{b_m^n(x)\}_{m=0}^n$ form a basis for the linear space of polynomials of degree at most n. Therefore each $\mathcal{B}^n f(x)$ is a polynomial of degree at most n. Bernstein proved that these polynomials uniformly approximate f over $[0,1]$.

Exercise. Prove that for every $n \in \mathbb{Z}_+$ and for every $m \in \{0, \cdots, n\}$ the Bernstein basis polynomial $b_m^n(x)$ given by (15.2) has degree n and is positive over the interval $(0,1)$ with a unique maximizer at $x = \frac{m}{n}$.

Exercise. The Bernstein basis polynomials of degree 4 are

$$b_0^4(x) = (1-x)^4 \,,\; b_1^4(x) = 4x(1-x)^3 \,,\; b_2^4(x) = 6x^2(1-x)^2 \,,\; b_3^4(x) = 4x^3(1-x) \,,\; b_4^4(x) = x^4 \,.$$

Sketch these polynomials over $[0,1]$ on a single graph.

Exercise. Prove that for every $n \in \mathbb{Z}_+$ the Bernstein basis polynomials $\{b_m^n(x)\}_{m=0}^n$ form a basis for the linear space of polynomials of degree at most n. (Hint: Show they are linearly independent.)

15.2. Three Elementary Identities.

The proof of the Weierstrass Approximation Theorem requires three elementry identites, which we record in the following lemma.

Lemma 15.1. *For every $n \in \mathbb{Z}_+$ we have elementary identities*

$$\mathcal{B}^n 1 = 1 \,, \qquad \mathcal{B}^n x = x \,, \qquad \mathcal{B}^n x^2 = x^2 + \frac{x(1-x)}{n} \,. \tag{15.3}$$

Remark. The last identity in (15.3) shows that the Bernstein approximating polynomials $\mathcal{B}^n f(x)$ do not generally converge to $f(x)$ quickly. Specifically, it shows that $\mathcal{B}^n x^2$ converges to x^2 like $1/n$ everywhere in $(0,1)$. This rate of convergence is typical for the Bernstein approximating polynomials. This is why they are not used for building accurate approximations in practice. However, they have properties that make them useful for other tasks.

Proof. For every $x, y \in \mathbb{R}$ every and every $n \in \mathbb{Z}_+$ the binomial formula yields

$$\sum_{m=0}^{n} \frac{n!}{m!(n-m)!}\, x^m y^{n-m} = (x+y)^n .$$

By applying $x\partial_x$ and $(x\partial_x)^2$ to this identity we obtain

$$\sum_{m=0}^{n} \frac{n!}{m!(n-m)!}\, m\, x^m y^{n-m} = nx(x+y)^{n-1} ,$$

$$\sum_{m=0}^{n} \frac{n!}{m!(n-m)!}\, m^2 x^m y^{n-m} = n(n-1)x^2(x+y)^{n-2} + nx(x+y)^{n-1} .$$

By setting $y = 1 - x$ in the three foregoing identities and by using definition (15.2) of the Bernstein basis polynomials we see that

$$\sum_{m=0}^{n} b_m^n(x) = \sum_{m=0}^{n} \frac{n!}{m!(n-m)!}\, x^m(1-x)^{n-m} = 1 ,$$

$$\sum_{m=0}^{n} m\, b_m^n(x) = \sum_{m=0}^{n} \frac{n!}{m!(n-m)!}\, m\, x^m(1-x)^{n-m} = nx ,$$

$$\sum_{m=0}^{n} m^2 b_m^n(x) = \sum_{m=0}^{n} \frac{n!}{m!(n-m)!}\, m^2 x^m(1-x)^{n-m} = n(n-1)x^2 + nx .$$

Identites (15.4) follow from these and from definition (15.1) of $\mathcal{B}^n f(x)$ with $f(x) = 1$, $f(x) = x$, and $f(x) = x^2$ respectively. □

Remark. The identites (15.3) are the zeroth, first, and second moments of the binomial probability distribution for Bernoulli trials with probability of success x. Both the identities and the foregoing proof might look familiar to those who have studied those distributions. This connection with the binomial probability was at the heart of how Bernstein approached the problem.

Remark. We could continue in the style of the proof of Lemma 15.1 to derive expressions for $\mathcal{B}^n x^k$ for every $k \in \mathbb{N}$. These expressions show that $\mathcal{B}^n x^k$ converges to x^k like $1/n$ as $n \to \infty$. However, we need only the expressions in (15.3) to prove that $\mathcal{B}^n f(x)$ converges uniformly to $f(x)$ like $1/n$ as $n \to \infty$ for a much more general class of functions f.

15.3. **Bernstein Approximation Bound.** The key bound in the Bernstein proof of the Weierstrass Approximation Theorem is provided by the following lemma, which gives much more.

Lemma 15.2. *Let f be twice continuously differentiable over $[0, 1]$. Then for every $n \in \mathbb{Z}_+$ we have the pointwise bounds*

$$\underline{K}\, \frac{x(1-x)}{2n} \le \mathcal{B}^n f(x) - f(x) \le \overline{K}\, \frac{x(1-x)}{2n} \quad \text{for every } x \in [0,1] , \tag{15.4}$$

where $\underline{K}$ and $\overline{K}$ are given by

$$\underline{K} = \min\left\{ f''(x) \,:\, x \in [0,1] \right\} , \qquad \overline{K} = \max\left\{ f''(x) \,:\, x \in [0,1] \right\} . \tag{15.5}$$

Remark. Because $x(1-x) \leq \frac{1}{4}$ over $[0,1]$, the pointwise bounds in (15.4) imply the uniform bound

$$(15.6) \qquad \left|\mathcal{B}^n f(x) - f(x)\right| \leq \frac{K}{8\,n}\,, \quad \text{where} \quad K = \max\{\overline{K}, -\underline{K}\} = \max\left\{|f''(x)| \,:\, x \in [0,1]\right\}.$$

This uniform bound will be used in the proof of the Weierstrass Approximation Theorem. It shows that when f is twice continuously differentiable the Bernstein approximating polynomials $\mathcal{B}^n f$ converge uniformly to f over $[0,1]$ at least as fast as $1/n$ as $n \to \infty$.

Remark. Because f'' is continuous over $[0,1]$ the lower pointwise bound in (15.4) will be nonzero if and only if $f''(x)$ is either always positive or always negative over $[0,1]$. In this case the lower pointwise bound shows that the Bernstein approximating polynomials $\mathcal{B}^n f$ converge uniformly to f over $[0,1]$ no faster than $1/n$ as $n \to \infty$.

Remark. When f is a polynomial of degree two then $\underline{K} = \overline{K}$. In that case the lower and upper pointwise bounds in (15.4) are equal. They are consistent with the third identity in (15.3).

Proof. Let $n \in \mathbb{Z}_+$. Let $x \in [0,1]$. Because f is twice differentiable over $[0,1]$, the Lagrange Remainder Theorem states that for every $m \in \{0, \cdots, n\}$ there exists a point z_m between x and $\frac{m}{n}$ such that

$$f\big(\tfrac{m}{n}\big) = f(x) + f'(x)\big(\tfrac{m}{n} - x\big) + \tfrac{1}{2} f''(z_m)\big(\tfrac{m}{n} - x\big)^2 .$$

Then by the identities (15.3) we have

$$\begin{aligned}
\mathcal{B}^n f(x) - f(x) &= \sum_{m=0}^{n} \Big(f\big(\tfrac{m}{n}\big) - f(x)\Big)\, b^n_m(x) \\
&= \sum_{m=0}^{n} f'(x)\big(\tfrac{m}{n} - x\big)\, b^n_m(x) + \sum_{m=0}^{n} \tfrac{1}{2} f''(z_m)\big(\tfrac{m}{n} - x\big)^2\, b^n_m(x) \\
&= \sum_{m=0}^{n} \tfrac{1}{2} f''(z_m)\big(\tfrac{m}{n} - x\big)^2\, b^n_m(x) .
\end{aligned}$$

Because $b^n_m(x) \geq 0$, the definitions of $\underline{K}$ and $\overline{K}$ given by (15.5) implies that

$$\underline{K}\,\frac{1}{2} \sum_{m=0}^{n} \big(\tfrac{m}{n} - x\big)^2\, b^n_m(x) \leq \mathcal{B}^n f(x) - f(x) \leq \overline{K}\,\frac{1}{2} \sum_{m=0}^{n} \big(\tfrac{m}{n} - x\big)^2\, b^n_m(x) .$$

The identities (15.3) can be used to evaluate the sums above as

$$\begin{aligned}
\sum_{m=0}^{n} \big(\tfrac{m}{n} - x\big)^2\, b^n_m(x) &= \mathcal{B}^n x^2 - 2x\mathcal{B}^n x + x^2 \mathcal{B}^n 1 \\
&= \left(x^2 + \frac{x(1-x)}{n}\right) - 2x \cdot x + x^2 \cdot 1 = \frac{x(1-x)}{n} .
\end{aligned}$$

Therefore the pointwise bounds (15.4) hold for every $n \in \mathbb{Z}_+$. □

15.4. **Proof of the Weierstrass Approximation Theorem.** Lemma 15.2 shows that every twice continuously differentiable function over $[0,1]$ is uniformly approximated by its sequence of Bernstein approximating polynomials. The Weierstrass Approximation Theorem asserts that the same is true for every continuous function over $[0,1]$. In order to prove it we must show that for every $f \in C([0,1])$ and every $\epsilon > 0$ we can find $g \in C^2([0,1])$ such that $\|f - g\| < \frac{1}{3}\epsilon$. Given this fact, a proof of the Weierstrass Approximation Theorem goes as follows.

Proof. Let $f \in C([0,1])$. Let $\epsilon > 0$. We want to show that $\|\mathcal{B}^n f - f\| < \epsilon$ eventually as $n \to \infty$. Let $g \in C^2([0,1])$ such that $\|f - g\| < \frac{1}{3}\epsilon$. Then

$$\|\mathcal{B}^n f - \mathcal{B}^n g\| < \tfrac{1}{3}\epsilon \qquad \text{for every } n \in \mathbb{Z}_+ .$$

Let $K = \|g''\|$. Then for every $n \in \mathbb{Z}_+$ we have

$$\|\mathcal{B}^n f - f\| \le \|\mathcal{B}^n f - \mathcal{B}^n g\| + \|\mathcal{B}^n g - g\| + \|g - f\| < \tfrac{1}{3}\epsilon + \frac{K}{8n} + \tfrac{1}{3}\epsilon .$$

Therefore $\|\mathcal{B}^n f - f\| < \epsilon$ for every $n > \frac{3}{8}K/\epsilon$. □