Analysis II
Supplementary course notes for Math 131B

Evan Randles

These are the supplementary course notes for Math 131B. The notes pick up at the end of our treatment of basic metric space topology from the courses' official textbook, Analysis II by T. Tao. I make no claim about the originality of these notes. Many of the proofs are taken (either directly or in spirit) from W. Rudin's classic text, Principles of Mathematical Analysis. In any case, these notes are a first approximation to my mind's ideal presentation of this material; I hope you enjoy them.

1 Uniform Convergence

Definition 1. *Let $(f^{(n)})_{n=1}^{\infty}$ be a sequence of functions from a metric space (X, d_X) to another, (Y, d_Y), and let $f : X \to Y$ be another function. For a subset E of X, we say that $(f^{(n)})_{n=1}^{\infty}$ converges uniformly to f on E if for all $\epsilon > 0$, there exists a natural number N for which*

$$d_Y(f^{(n)}(x), f(x)) < \epsilon \qquad \text{for all } n \geq N \text{ and } x \in E.$$

Remark 1. *Note that N does not depend on x; this is the difference between pointwise convergence and uniform convergence.*

The preceding definition is illustrated in Figure 1. In the figure, we see the graph of a real-valued function f (in black) in the center of a "band" of radius ϵ (in red). For a sequence of functions $(f^{(n)})_{n=1}^{\infty}$ to converge uniformly to f (on an interval) means that, for sufficiently large n, the graph of $f^{(n)}$ is completely contained in the band of radius ϵ surrounding f; the blue line is an example of the graph of one such $f^{(n)}$.

Example 1.1

For each natural number n, define $f^{(n)} : [0, 1] \to [0, 1]$ by $f^{(n)}(x) = x/n$ for $0 \leq x \leq 1$.

Claim 2. *The sequence $(f^{(n)})_{n=1}^{\infty}$ converges uniformly to the zero function on* $[0, 1]$.

Proof. Fix $\epsilon > 0$ and choose a natural number N such that $N > 1/\epsilon$. Then, simply observe that

$$d(f^{(n)}(x), \mathbf{0}(x)) = \left|\frac{x}{n} - 0\right| \leq \frac{1}{n} < \epsilon$$

for all $n \geq N$ and $x \in [0, 1]$. Here we have denoted the standard metric on $[0, 1]$ by d and the zero function by $\mathbf{0}$. $\square$

When the metric space (Y, d_Y) is complete, we obtain the following theorem.

Theorem 3. *Let $(f^{(n)})_{n=1}^{\infty}$ be a sequence of functions from a metric space (X, d_X) to a complete metric space (Y, d_Y). Let $E \subseteq X$. The sequence $(f^{(n)})_{n=1}^{\infty}$ converges uniformly (to some function f) on E if and only if the following condition is satisfied:*

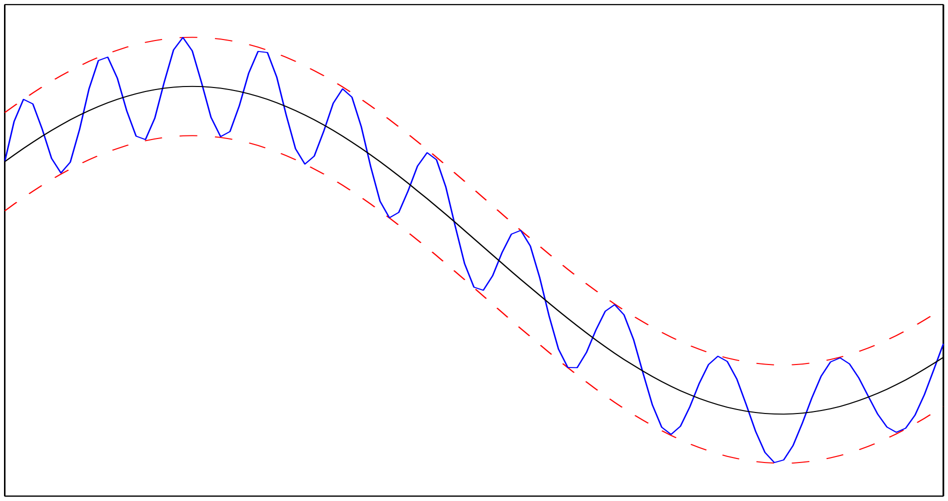

Figure 1: An illustration of uniform convergence

(UC) For all $\epsilon > 0$*, there exists a natural number* N *such that*

$$d_Y(f^{(n)}(x), f^{(m)}(x)) < \epsilon \qquad \textit{for all } m, n \geq N \textit{ and } x \in E.$$

Any sequence of functions $(f^{(n)})_{n=1}^{\infty}$ satisfying the condition (UC) (regardless of whether or not the space Y is complete) is said to be *uniformly Cauchy on* E.

Proof. Suppose that $(f^{(n)})_{n=1}^{\infty}$ converges uniformly to f on E. Then, given $\epsilon > 0$, there exists a natural number N such that

$$d_Y(f^{(n)}(x), f(x)) < \epsilon/2$$

for all $n \geq N$ and $x \in E$. Therefore

$$d_Y(f^{(n)}(x), f^{(m)}(x)) \leq d_Y(f^{(n)}(x), f(x)) + d_Y(f(x), f^{(m)}(x) < \epsilon/2 + \epsilon/2 = \epsilon$$

whenever $m, n \geq N$ and $x \in E$. Hence the sequence $(f^{(n)})_{n=1}^{\infty}$ is uniformly Cauchy on E.

Conversely, let's assume that $(f^{(n)})_{n=1}^{\infty}$ is uniformly Cauchy on E. Then, for each $x \in E$, the sequence of points $(f^{(n)}(x))_{n=1}^{\infty}$ is necessarily a Cauchy sequence in the space Y. Because Y is complete, this sequence converges and we will denote its limit by $f(x) \in Y$. In this way, we obtain a function $f : X \to Y$ to which our sequence $(f^{(n)})_{n=1}^{\infty}$ converges pointwise on E (and so f is our candidate for the uniform limit). We must show this convergence is, in fact, uniform on E.

Let $\epsilon > 0$ and choose a natural number N for which

$$d_Y(f^{(n)}(x), f^{(m)}(x)) < \epsilon/2$$

for all $n, m \geq N$ and $x \in E$. We observe that, for each $x \in E$ and $n \geq N$,

$$d_Y(f^{(n)}(x), f(x)) = \lim_{m\to\infty} d_Y(f^{(n)}(x), f^{(m)}(x)) \leq \epsilon/2 < \epsilon$$

(this uses the fact that $y \mapsto d_Y(f^{(n)}(x), y)$ is a continuous function from Y to $\mathbb{R}$). This is precisely what it means for $(f^{(n)})_{n=1}^{\infty}$ to converge uniformly to f on E. $\square$

Theorem 4. *Suppose that $(f^{(n)})_{n=1}^\infty$ is a sequence of functions from a metric space (X, d_X) to a complete metric space (Y, d_Y) which converges uniformly to a function $f : X \to Y$ on a subset E of X. Let x_0 be an adherent point of E and suppose that, for each natural number n,*

$$\lim_{x \to x_0; x \in E} f^{(n)}(x) = y^{(n)};$$

that is, the limit exists and is equal to $y^{(n)}$. Then

$$\lim_{x \to x_0; x \in E} f(x) = \lim_{n \to \infty} y^{(n)},$$

that is, both limits exists and are equal. In other words,

$$\lim_{x \to x_0; x \in E} \lim_{n \to \infty} f^{(n)}(x) = \lim_{n \to \infty} \lim_{x \to x_0; x \in E} f^{(n)}(x).$$

Proof. Let $\epsilon > 0$ be given. In view of the previous theorem, there is a natural number N for which

$$d_Y(f^{(n)}(x), f^{(m)}(x)) < \epsilon$$

for all $n, m \geq N$ and $x \in E$. In view of our hypothesis,

$$d_Y(y^{(n)}, y^{(m)}) = \lim_{x \to x_0; x \in E} d_Y(f^{(n)}(x), f^{(m)}(y)) \leq \epsilon.$$

whenever $n, m \geq N$. Consequently $(y^{(n)})_{n=1}^\infty$ is a Cauchy sequence in Y and because Y is complete it converges to some $y \in Y$. To prove the theorem, we need to show that $\lim_{x \to x_0; x \in E} f(x) = y$. Now, for any $x \in E$ and $n \in \mathbb{N}$,

$$d_Y(f(x), y) \leq d_Y(f(x), f^{(n)}(x)) + d_Y(f^{(n)}(x), y^{(n)}) + d_Y(y^{(n)}, y).$$

Invoking the uniform convergence of $(f^{(n)})$ choose $N_0 \in \mathbb{N}$ for which $d_Y(f(x), f^{(n)}(x)) < \epsilon/3$ for all $x \in E$ and $n \geq N_0$. Because $(y^{(n)})$ converges to y, we can select a natural number $N \geq N_0$ for which $d_Y(y^{(n)}, y) < \epsilon/3$ for all $n \geq N$. In particular, for $n = N$ (which we fix now), the preceding estimate guarantees that

$$d_Y(f(x), y) < \epsilon/3 + d_Y(f^{(N)}(x), y^{(n)}) + \epsilon/3 = 2\epsilon/3 + d_Y(f^{(N)}(x), y^{(n)})$$

for all $x \in E$. Using the hypothesis that $\lim_{x \to x_0; x \in E} f^N(x) = y^{(N)}$, choose $\delta > 0$ such that $d_Y(f^{(N)}(x), y^{(N)}) < \epsilon/3$ whenever $d_X(x, x_0) < \delta$ for $x \in E$. Thus, for any $x \in E$ for which $d_X(x, x_0) < \delta$, we have

$$d_Y(f(x), y) < 2\epsilon/3 + d_Y(f^{(N)}(x), y^{(N)}) < 2\epsilon/3 + \epsilon/3 = \epsilon.$$

Hence $\lim_{x \to x_0; x \in E} f(x) = y$ as claimed. $\square$

We obtain the following corollary:

Corollary 5. *Let $(f^{(n)})$ be a sequence of functions from a metric space (X, d_X) to a complete metric space (Y, d_Y) which converges uniformly to a function $f : X \to Y$ on X. If, for each n, $f^{(n)}$ is continuous at x_0, then f is continuous at x_0. If functions $(f^{(n)})$ are continuous on X, then f is also continuous on X.*

Proof. Assuming that each function $f^{(n)}$ is continuous at x_0, we have

$$\lim_{x \to x_0} f^{(n)}(x) = \lim_{x \to x_0; x \in X} f^{(n)}(x) = f^{(n)}(x_0).$$

Thus, by the preceding theorem,

$$\lim_{x \to x_0} f(x) = \lim_{n \to \infty} f^{(n)}(x_0) = f(x_0)$$

and hence f is continuous at x_0. If each function $f^{(n)}$ is continuous, then the preceding argument shows that f is continuous at every point in X and hence f is continuous. $\square$

2 The metric of uniform convergence

Definition 6. *Let (X,d) be a metric space and let $f : X \to \mathbb{R}$ be a function. We say that f is bounded on X if there exists a positive constant M such that $|f(x)| \leq M$ for all $x \in X$. In this case we set*

$$\|f\|_\infty = \sup_{x\in X} |f(x)|;$$

$\|f\|_\infty$ is called the sup norm of f. We denote by $B(X,\mathbb{R})$ or $B(X)$ the set of bounded functions $f : X \to \mathbb{R}$. The subset of $B(X)$ consisting of continuous functions is denoted by $C(X)$ or $C(X,\mathbb{R})$. In other words,

$$B(X) = \{f : X \to \mathbb{R} : f \text{ is bounded on } X\}$$

and

$$C(X) = \{f \in B(X) : f \text{ is continuous on } X\} = \{f : X \to \mathbb{R} : f \text{ is bounded and continuous on } X\}.$$

You shall prove the following proposition in Homework 4.

Proposition 7. *Let (X,d) be a metric space and $(f^{(n)})_{n=1}^\infty$ be a sequence of functions in $B(X)$. Let $f : X \to \mathbb{R}$ (not necessarily bounded). Then $(f^{(n)})_{n=1}^\infty$ converges uniformly to f on X if and only if $f \in B(X)$ and*

$$\lim_{n\to\infty} \|f^{(n)} - f\|_\infty = 0.$$

Given the sup norm $\|\cdot\|_\infty$ we can equip $B(X)$ with a metric as follows: For $f, g \in B(X)$, define

$$d_\infty(f,g) = \|f - g\|_\infty.$$

In Homework 4, you will verify that this is indeed a metric on $B(X)$. In light of Proposition 7, a sequence of functions $(f^{(n)})_{n=1}^\infty \in B(X)$ converges uniformly to $f \in B(X)$ if and only if $(f^{(n)})_{n=1}^\infty$ converges to f with respect to the metric d_∞. For this reason, we call d_∞ the *metric of uniform convergence*. As a subspace of $B(X)$, $C(X)$ becomes a metric space equipped with the induced metric. By an abuse of notation, we denote the induced metric on $C(X)$ by d_∞. Thus $(C(X), d_\infty)$ is a metric space. In fact, more is true:

Theorem 8. *$(C(X), d_\infty)$ is a complete metric space.*

Proof. Let $(f^{(n)})_{n=1}^\infty$ be a Cauchy sequence in $C(X)$. Thus, for each $\epsilon > 0$, there is a natural number N, for which

$$d_{\mathbb{R}}(f^{(n)}(x), f^{(m)}(x)) = |f^{(n)}(x) - f^{(m)}(x)| \leq \|f^{(n)} - f^{(m)}\|_\infty = d_\infty(f^{(n)}, f^{(m)}) < \epsilon$$

for all $n, m \geq N$ and $x \in X$. Hence, the sequence of functions $(f^{(n)})_{n=1}^\infty$ is uniformly Cauchy on X. By Theorem 3, it converges uniformly to some function $f : X \to \mathbb{R}$. By an appeal to Proposition 7, $f \in B(X)$ and $(f^{(n)})_{n=1}^\infty$ converges to f with respect to the sup norm metric. It remains to show that f is continuous, and so $f \in C(X)$. This however follows immediately from Corollary 5 because the target space $\mathbb{R}$ is complete and $(f^{(n)})_{n=1}^\infty$ converges uniformly to f on X. □

3 Uniform convergence and series

Definition 9. *Let (X,d) be a metric space and $(f^{(n)})_{n=1}^\infty$ a sequence of bounded real-valued functions on X. Let $f : X \to \mathbb{R}$ and let $E \subseteq X$. For each natural number $n \in N$, define $S^{(n)} : X \to \mathbb{R}$ by*

$$S^{(n)}(x) = \sum_{j=1}^n f^{(j)}(x)$$

for $x \in X$. We say that the series $\sum f^{(n)}$ converges pointwise on E to f *if*

$$f(x) = \lim_{n\to\infty} S^{(n)}(x) = \lim_{n\to\infty} \sum_{j=1}^n f^{(j)}(x)$$

for each $x \in E$. In this case we say that f is the sum of the series $\sum f^{(n)}$. If this convergence is uniform on E, i.e., $S^{(n)}$ converges uniformly to f on E, we say that that series $\sum f^{(n)}$ converges uniformly to f on E.

Example 3.1

For each n, let $f^{(n)} : \mathbb{R} \to \mathbb{R}$ be defined by $f^{(n)}(x) = x^n$ for $x \in \mathbb{R}$. Then $\sum f^{(n)}$ converges pointwise to $f(x) = x/(1-x)$ on $(-1, 1)$. It does not converge uniformly on $(-1, 1)$ as you will show in the homework.

Theorem 10 (Weierstrass M-test)**.** *Let (X, d) be a metric space and $(f^{(n)})_{n=1}^\infty$ be a sequence of bounded real-valued functions on X. For each n, set $M_n = \|f^{(n)}\|_\infty$. If the series $\sum_{n=1}^\infty M_n$ converges (note that this is just a series of non-negative numbers), then the series $\sum f^{(n)}$ converges uniformly to some function f on X. If in addition $(f^{(n)})_{n=1}^\infty \subseteq C(X)$, then f is continuous.*

Proof. Let $\epsilon > 0$. If the series $\sum M_n$ converges, in view of the Cauchy criterion for numerical series, there exists a natural number N for which

$$\left|S^{(n)}(x) - S^{(m)}(x)\right| = \left|\sum_{j=n+1}^{m} f^{(j)}(x)\right| \le \sum_{j=n}^{m} |f^{(j)}(x)| \le \sum_{j=n}^{m} M_j < \epsilon$$

for all $n, m \ge N$ and $x \in X$. Consequently $(S^{(n)})_{n=1}^\infty$ is uniformly Cauchy on X and so it converges uniformly to some function $f : X \to \mathbb{R}$ in view of Theorem 3. This is precisely what it means for the series $\sum f^{(n)}$ to converge uniformly to f. Finally, if each $f^{(j)} \in C(X)$, then, because the finite sum of continuous functions is continuous (we proved this), $S^{(n)} \in C(X)$ for each natural number n. An appeal to Corollary 5 then guarantees that $f \in C(X)$. $\square$

Example 3.2

For each natural number n, define $f^{(n)}(x) = (1/n^2)\cos(nx)$ for $x \in \mathbb{R}$. Observe that

$$\|f^{(n)}\|_\infty = \sup_{x \in \mathbb{R}} \left|\frac{1}{n^2}\cos(nx)\right| = \frac{1}{n^2} = M_n$$

for each n. Now, the series $\sum_{j=1}^\infty M_j = \sum_{j=1}^\infty (1/j^2)$ converges (In fact, finding its value was a long outstanding problem, called the Basel Problem. In 1734, Leonard Euler showed that it converges to $\pi^2/6$, making him instantly famous.) Therefore, by the Weierstrass M-test, the series

$$\sum_{j=1}^{\infty} f^{(j)} = \sum_{j=1}^{\infty} \frac{1}{j^2}\cos(jx)$$

converges uniformly on $\mathbb{R}$.

4 Uniform convergence and integration

4.1 A short recap on Riemann integration

Let $I = [a, b] \subseteq \mathbb{R}$. A partition P of I is a finite subset $P = \{x_0, x_1, x_2, \ldots, x_N\}$ of I such that

$$a = x_0 < x_1 < x_2 < \cdots < x_{N-1} < x_N = b.$$

Given a bounded function $f : I \to \mathbb{R}$ and a partition P of I, define

$$m_n = \inf_{x_{n-1} \le x \le x_n} f(x) \qquad \text{and} \qquad M_n = \sup_{x_{n-1} \le x \le x_n} f(x)$$

for each $n = 1, 2, \dots, N$. With these, we define the upper and lower sums of f with respect to the partition P respectively by

$$U(f,P) = \sum_{n=1}^{N} M_n(x_n - x_{n-1}) \qquad \text{and} \qquad L(f,P) = \sum_{n=1}^{N} m_n(x_n - x_{n-1}).$$

Further, we define the upper and lower Riemann sums of f over I by

$$\overline{\int_I} f(x)\,dx = U(f) = \inf_P U(f,P) \qquad \text{and} \qquad \underline{\int_I} f(x)\,dx = L(f) = \sup_P L(f,P)$$

where the supremum and infimum are taken over all partitions P of I. A function $f \in B(I)$ is said to be *Riemann-integrable* if its upper and lower Riemann sums are equal. In this case, the integral of f is defined to be the number

$$\int_I f(x)\,dx = \overline{\int_I} f(x)\,dx = \underline{\int_I} f(x)\,dx.$$

We have the following standard facts:

Fact 1. *For any $f \in B(I)$,*

$$\underline{\int_I} f(x)\,dx \le \overline{\int_I} f(x)\,dx.$$

If $f, g \in B(I)$ are such that $f(x) \le g(x)$ for all $x \in I$, then

$$\underline{\int_I} f(x)\,dx \le \underline{\int_I} g(x)\,dx \qquad \textit{and} \qquad \overline{\int_I} f(x)\,dx \le \overline{\int_I} g(x)\,dx.$$

Constant functions are Riemann-integrable and, for $\alpha \in \mathbb{R}$,

$$\int_I \alpha\,dx = \alpha(b-a).$$

Given Riemann-integrable functions f and g and $\alpha, \beta \in \mathbb{R}$, the linear combination $\alpha f + \beta g$ is Riemann-integrable and

$$\int_I (\alpha f + \beta g)(x)\,dx = \alpha \int_I f(x)\,dx + \beta \int_I g(x)\,dx.$$

Theorem 11. *Let $I = [a,b] \subseteq \mathbb{R}$ and, for each natural number n, let $f^{(n)} : I \to \mathbb{R}$ be Riemann-integrable. If the sequence $(f^{(n)})_{n=1}^{\infty}$ converges uniformly to a function $f : I \to \mathbb{R}$. Then f is Riemann-integrable on I and*

$$\int_I f(x)\,dx = \lim_{n\to\infty} \int_I f^{(n)}(x)dx.$$

In other words, when convergence is uniform, you can exchange integrals and limits.

Proof. We first note that f is necessarily bounded by Proposition 7 (and so its upper and lower Riemann sums exist). Let $\epsilon > 0$. Since $(f^{(n)})$ converges uniformly to f, there is a natural number N for which $d_{\mathbb{R}}(f^{(n)}(x), f(x)) = |f^{(n)}(x) - f(x)| < \epsilon$ for all $n \ge N$ and $x \in I$. in other words,

$$f^{(n)}(x) - \epsilon \le f(x) \le f^{(n)}(x) + \epsilon$$

for all $n \ge N$ and $x \in I$. By our fact above,

$$\underline{\int_I} (f^{(n)}(x) - \epsilon)\,dx \le \underline{\int_I} f(x)\,dx \le \overline{\int_I} f(x)\,dx \le \overline{\int_I} (f^{(n)}(x) + \epsilon)\,dx$$

for all $n \geq N$. Because each $f^{(n)}$ is Riemann-integrable, we have

$$\begin{aligned}\int_I f^{(n)}(x)\,dx - \epsilon(b-a) &= \int_I (f^{(n)}(x)-\epsilon)\,dx \leq \underline{\int_I} f(x)\,dx \\ \leq \quad \overline{\int_I} f(x)\,dx &\leq \int_I (f^{(n)}(x)+\epsilon)\,dx = \int_I (f^{(n)}(x)\,dx + \epsilon(b-a)\end{aligned}$$

for all $n \geq N$ in view of the preceding fact. In particular,

$$0 \leq \overline{\int_I} f(x)\,dx - \underline{\int_I} f(x)\,dx \leq 2\epsilon(b-a).$$

Now, because ϵ was arbitrary, the upper and lower Riemann sums of f must be equal and hence f is Riemann-integrable. Furthermore, the inequality above also shows that

$$\left|\int_I f^{(n)}(x)\,dx - \int_I f(x)\,dx\right| \leq 2\epsilon(b-a)$$

for all $n \geq N$. This shows, by perhaps selecting a cleaner ϵ, that

$$\int_I f(x)\,dx = \lim_{n\to\infty} \int_I f^{(n)}(x)\,dx$$

as desired. □

You will use the theorem above to prove the following corollary in Homework 4.

Corollary 12. *Let $(f^{(n)})_{n=1}^\infty$ be a sequence of Riemann-integrable functions on I. If the series $\sum f^{(n)}$ converges uniformly to f on I, then $f(x) = \sum_{j=1}^\infty f^{(j)}(x)$ is Riemann-integrable and*

$$\int_I f(x)\,dx = \int_I \sum_{j=1}^\infty f^{(j)}(x)\,dx = \sum_{j=1}^\infty \int_I f^{(j)}(x)\,dx.$$

5 Uniform convergence and differentiation

In this section, we explore the interplay between uniform convergence and differentiation. To keep our presentation simple, we will work on an open interval (a,b); the case of the closed interval $[a,b]$ can be treated similarly by considering left and right derivatives.

Theorem 13. *Let $(f^{(n)})_{n=1}^\infty$ be a sequence of functions mapping from the interval (a,b) to $\mathbb{R}$ with $-\infty < b < a < \infty$. Suppose that, for each n, $f^{(n)}$ is differentiable on (a,b) with derivative $\frac{df^{(n)}}{dx}$. Suppose that the sequence $(\frac{df^{(n)}}{dx})_{n=1}^\infty$ converges uniformly to a function g on (a,b). Suppose additionally, the sequence $(f^{(n)})$ converges at (at least) one point $x_0 \in X$. Then the sequence $(f^{(n)})_{n=1}^\infty$ converges uniformly on (a,b) to a function f, Moreover, f is differentiable on (a,b) with derivative $\frac{df}{dx} = g$.*

Before proving the theorem, let us discuss the (seemingly strange) condition that $(f^{(n)})_{n=1}^\infty$ converges at one point. To see that this condition is necessary, consider the sequence of constant functions defined by $f^{(n)}(x) = n$ for $n \in \mathbb{N}$ and $x \in \mathbb{R}$. Here, it's obvious that $\frac{df^{(n)}}{dx}(x) = 0$ for all $n \in \mathbb{N}$ and for all $x \in \mathbb{R}$. Trivially then, $(\frac{df^{(n)}}{dx})_{n=1}^\infty$ converges uniformly to the zero function. But, it is obvious that $(f^{(n)})$ doesn't converge (it diverges to infinity at every point). Thus the conclusion of the theorem doesn't hold.

Proof. Let $\epsilon > 0$. We suppose that $(f^{(n)})_{n=1}^\infty$ converges at $x_0 \in X$ and so it is a Cauchy sequence in $\mathbb{R}$. Consequently, there exists $N_0 \in \mathbb{N}$ for which

$$|f^{(n)}(x_0) - f^{(m)}(x_0)| < \epsilon/2$$

for all $n, m \geq N_0$. Further, let $N \geq N_0$ be a natural number for which

$$\left|\frac{df^{(n)}}{dx}(x) - \frac{df^{(m)}}{dx}(x)\right| < \frac{\epsilon}{2(b-a)}$$

for all $n, m \geq N$ and $a < x < b$. By invoking the mean value theorem on (a, b) (applied to the function $f^{(n)} - f^{(m)}$), we obtain

$$|f^{(n)}(x) - f^{(m)}(x) - (f^{(n)}(y) - f^{(m)}(y))| = \left|\frac{df^{(n)}}{dx}(c) - \frac{df^{(m)}}{dx}(c)\right| |x-y| < \frac{\epsilon|x-y|}{2(b-a)} \leq \epsilon/2 \tag{1}$$

for every $x, y \in (a, b)$ and $n, m \geq N$ (here, for each $a < x < y < b$, $x < c < y$ as guaranteed by the mean value theorem). In particular,

$$|f^{(n)}(x) - f^{(m)}(x) - (f^{(n)}(x_0) - f^{(m)}(x_0))| < \epsilon/2$$

for all $x \in (a, b)$ and $m, n \geq N$. Consequently, for any $x \in (a, b)$ and $n, m \geq N$,

$$|f^{(n)}(x) - f^{(m)}| \leq |f^{(n)}(x) - f^{(m)}(x) - (f^{(n)}(x_0) - f^{(m)}(x_0))| + |f^{(n)}(x_0) - f^{(m)}(x_0)| < \epsilon/2 + \epsilon/2 = \epsilon.$$

Thus we have shown that the sequence $(f^{(n)})$ is uniformly Cauchy on (a, b); it therefore converges to a function $f : (a, b) \to \mathbb{R}$ (because $Y = \mathbb{R}$ is complete) by virtue of Theorem 3. It remains to show that f is differentiable with derivative equal to g.

Let $\epsilon > 0$ and, for (fixed) $y \in (a, b)$ define

$$\psi^{(n)}(x) = \frac{f^{(n)}(x) - f^{(n)}(y)}{x - y} \text{ and } \psi(x) = \frac{f(x) - f(y)}{x - y}$$

for $x \in (a, b)$ such that $x \neq y$ and $n \in \mathbb{N}$. The hypothesis guarantees that

$$\lim_{x \to y; x \neq y} \psi^{(n)}(x) = \frac{df^{(n)}}{dx}(y).$$

We note that the inequality (1) guarantees that

$$|\psi^{(n)}(x) - \psi^{(m)}(x)| < \frac{\epsilon}{2(b-a)}$$

for all $n, m \geq N$ whenever $x \in (a, b)$ is such that $x \neq y$. Thus, $(\psi^{(n)})_{n=1}^{\infty}$ is a uniformly Cauchy sequence on the set $E = \{a < x < b : x \neq y\}$ and it therefore converges. Using the fact that $f^{(n)}$ converges to f, we have

$$\lim_{n \to \infty} \psi^{(n)}(x) = \lim_{n \to \infty} \frac{f^{(n)}(x) - f^{(n)}(y)}{x - y} = \psi(x)$$

for each $a < x < b$ such that $x \neq y$. An appeal to Theorem 4 shows that

$$\lim_{x \to y : x \neq y} \psi(x) = \lim_{n \to \infty} \frac{df^{(n)}}{dx}(y) = g(y).$$

This is precisely the statement that f is differentiable at y with derivative $g(y)$. $\square$

6 Weierstrass Approximation Theorem

Consider the space of continuous function $C(I)$ on a closed bounded (compact) interval $I = [a, b] \subseteq \mathbb{R}$ and the metric of uniform convergence d_∞. Also, consider the set $\mathcal{P}$ consisting of *polynomials* or *polynomial functions*; that is, functions of the form

$$P(x) = a_0 + a_1 x + a_2 x^2 + \cdots a_N x^N$$

defined for $x \in I$ where $N \in \mathbb{N}$ and $\{a_0, a_2, \dots, a_N\} \subseteq \mathbb{R}$. From your previous course in analysis, you know that $\mathcal{P} \subseteq C(I)$, i.e., polynomial functions are continuous. You should also note that $\mathcal{P} \neq C(I)$; for example, the exponential function $x \mapsto e^x$ is not a member of $\mathcal{P}$. This can be seen by observing that e^x has non-vanishing derivatives of all orders, a fact that is not true for any polynomial. The subject of this section is to prove the Weierstrass approximation theorem. This theorem is captured succinctly by the statement

$$\overline{\mathcal{P}} = C(I).$$

That is, with respect the the metric d_∞ of $C(I)$, every continuous function f is an adherent point of the set of polynomials $\mathcal{P}$.

This theorem is of significant practical importance. For one, the space of continuous functions includes many intractable members (e.g., continuous functions which are nowhere differentiable); however, polynomials are fantastically easy to work with and easy to compute. They are everywhere differentiable, they are specified by finitely many numbers ($\{a_0, a_1, \dots, a_N\}$-their coefficients) and evaluating them is cheap from a computational perspective. Without further ado, here is the Weierstrass approximation theorem, established by Karl Weierstrass in 1885:

Theorem 14 (Weierstrass approximation theorem)**.** *For any $f \in C(I)$, there exists a sequence of polynomials $(P^{(n)})_{n=1}^\infty \subseteq \mathcal{P}$ which converges uniformly to f on I. Equivalently, $(P^{(n)})_{n=1}^\infty$ converges to f with respect to the metric d_∞.*

The major hurdle in proving the Weierstrass approximation theorem is to establish the following technical lemma.

Lemma 15. *Let $f \in C([0,1])$ be such that $f(0) = f(1) = 0$. There exists a sequence of polynomials $(Q^{(n)})_{n=1}^\infty$ (on $[0,1]$) which converges uniformly to f on $[0,1]$.*

Proof of Lemma 15. In the flavor of Homework 3, we first extend f to the entire real line by putting $f(x) = 0$ for all $x \in \mathbb{R} \setminus [0,1]$; note that f is uniformly continuous on $\mathbb{R}$. Indeed, because it is continuous on the compact set $[0,1]$, it is uniformly continuous on $[0,1]$. Further, a simple verification shows that it is uniformly continuous of the complement of $[0,1]$ in $\mathbb{R}$ and hence uniformly continuous on $\mathbb{R}$.

For each positive natural number n, define

$$\rho^{(n)}(x) = c_n(1-x^2)^n$$

for $x \in \mathbb{R}$, where $c_n \in \mathbb{R}$ is chosen so that

$$\int_{-1}^{1} \rho^{(n)}(x)dx = 1. \tag{2}$$

Observe that

$$\begin{aligned}
\int_{-1}^{1} (1-x^2)^n \, dx &= 2\int_0^1 (1-x^2)^n \, dx \\
&\geq 2\int_0^{1/\sqrt{n}} (1-x^2)^n \, dx \\
&\geq 2\int_0^{1/\sqrt{n}} (1-nx^2) \, dx
\end{aligned}$$

where we have used the fact that $g(x) = (1-x^2)^n - (1-nx^2)$ is non-negative and non-decreasing on $[0,1]$ (you can check this with elementary calculus). Consequently

$$\int_{-1}^{1} (1-x^2)^n \, dx \geq \frac{4}{3\sqrt{n}}$$

and therefore

$$c_n = \left(\int_{-1}^{1} (1-x^2)^n \, dx\right)^{-1} \leq \frac{3\sqrt{n}}{4} < \sqrt{n}$$

for $n = 1, 2, \ldots, n$. Now, given any $\delta > 0$,

$$|\rho^{(n)}(x)| \leq \sqrt{n}(1-\delta^2)^n \tag{3}$$

for $\delta \leq |x| \leq 1$. Consequently $(\rho^{(n)})_{n=1}^{\infty}$ converges uniformly to the zero function on $\delta \leq |x| \leq 1$. Now, define

$$Q^{(n)}(x) = \int_{-1}^{1} f(x+t)\rho^{(n)}(t)\,dt$$

for $x \in [0,1]$ and observe that

$$Q^{(n)}(x) = \int_{-x}^{1-x} f(x+t)\rho^{(n)}(t)\,dt = \int_{0}^{1} f(t)\rho^{(n)}(t-x)\,dt$$

for $x \in [0,1]$ where we have used the fact that $f(x) = 0$ for all $x \in \mathbb{R} \setminus [0,1]$ and made a simple change of variables. The right hand side of this equation shows immediately that $Q^{(n)}(x)$ is a polynomial functions for the integral does nothing more than specify its coefficients. Hence $(Q^{(n)})_{n=1}^{\infty}$ is a sequence of polynomials.

Let $\epsilon > 0$ and choose $\delta > 0$ for which

$$|f(y) - f(x)| < \epsilon/2 \quad \text{whenever } |y - x| < \delta$$

in view of the uniform continuity of f and define $M = \sup_{x\in[0,1]} |f(x)|$. In view of (2) and (3) and the fact that $\rho^{(n)}(x) \geq 0$ for $x \in [0,1]$, we have

$$\begin{aligned}
|Q^{(n)}(x) - f(x)| &= \left| \int_{-1}^{1} f(x+t)\rho_n(t)\,dt - f(x)\int_{-1}^{1}\rho^{(n)}(t)\,dt \right| \\
&= \left| \int_{-1}^{1} \left(f(x+t) - f(x)\right)\rho^{(n)}(t)\,dt \right| \\
&\leq \int_{-1}^{1} |f(x+t) - f(x)|\,\rho^{(n)}(t)\,dt \\
&\leq 2M\left(\int_{-1}^{-\delta}\rho^{(n)}(t)\,dt + \int_{\delta}^{1}\rho^{(n)}(t)\,dt\right) + \frac{\epsilon}{2}\int_{-\delta}^{\delta}\rho^{(n)}(t)\,dt \\
&\leq 4M\sqrt{n}(1-\delta^2)^n + \epsilon/2.
\end{aligned}$$

Upon noting (again) that $\sqrt{n}(1-\delta^2)^n \to 0$ as $n \to \infty$, we can find $N \in \mathbb{N}$ for which $4M\sqrt{n}(1-\delta^2)^n < \epsilon/2$ for $n \geq N$. Consequently,

$$|Q^{(n)}(x) - f(x)| < \epsilon$$

for all $n \geq N$ and $x \in [0,1]$. □

We note that the formula by which $Q^{(n)}$ was defined above is intricately related to an operation on functions called *convolution*. This topic has important applications in many branches of mathematics, e.g., probability theory (the central limit theorem), Fourier analysis, harmonic analysis, partial differential equations, number theory, algebraic geometry, to name a few. We will revisit this topic (in the narrow context of this course) when our focus turns to Fourier series.

Proof of Theorem 14. The proof proceeds in two steps. We first show that any continuous function on the interval $[0,1]$ can be uniformly approximated by polynomials. In the second step, we treat the arbitrary compact interval $I = [a,b]$.

Step 1. Let $f : [0,1] \to \mathbb{R}$ be continuous (and not necessarily with $f(0) = f(1) = 0$). We claim that there exists a sequence of polynomials $(R^{(n)})_{n=1}^{\infty}$ which converge uniformly to f on $[0,1]$. To see this, let $\epsilon > 0$ and let $g : [0,1] \to \mathbb{R}$ be defined by

$$g(x) = f(x) - f(0) + (f(1) - f(0))\,x$$

for $x \in [0,1]$. We observe that g is a continuous function on $[0,1]$ which satisfies the conditions of Lemma 15 and hence, there exists a sequence of polynomials $(Q^{(n)})_{n=1}^{\infty}$ which converges to g uniformly on $[0,1]$. For each natural number n, set

$$R^{(n)}(x) = f(0) - (f(1) - f(0))x + Q^{(n)}(x)$$

for $x \in I$. It is clear that $(R^{(n)})_{n=1}^{\infty}$ is a sequence of polynomials (as they are simply obtained by adding a linear function to $Q^{(n)}$. Moreover, let $\epsilon > 0$ and using the uniform convergence of $(Q^{(n)})$ to g, select and $N > 0$ for which $|Q^{(n)}(x) - g(x)| < \epsilon$ for all $x \in [0,1]$ and $n \geq N$. Then we simply observe that

$$|f(x) - R^{(n)}(x)| = |f(x) - f(0) + (f(1) - f(0))x - Q^{(n)}(x)| = |g(x) - Q^{(n)}(x)| < \epsilon$$

for all $x \in [0,1]$ and $n \geq N$. This proves the claim.

Step 2. Let $f \in C(I) = C([a,b])$ and set $g(x) = f((b-a)x + a))$ for $x \in [0,1]$. We note that $g \in C([0,1]$ for it is nothing but the composition of the continuous function f with the continuous map $[0,1] \ni x \mapsto (b-a)x + a \in [a,b]$. By the claim above, there is a sequence of polynomials, $(R^{(n)})_{n=1}^{\infty}$ which converge uniformly to g on $[0,1]$. For each n, define $P^{(n)}(x) = R^{(n)}((x-a)/(b-a))$ for $x \in [a,b]$. Then, using the uniform continuity of the sequence $(R^{(n)})_{n=1}^{\infty}$ to g, for any $\epsilon > 0$ there exists a natural number N for which $|g(x) - R^{(n)}(x)| < \epsilon$ for all $x \in [0,1]$ and $n \geq N$. Consequently,

$$|f(x) = P^{(n)}(x)| = |g((x-a)/(b-a)) - R^{(n)}(x-a)/(b-a))| < \epsilon$$

for all $x \in I = [a,b]$ and $n \geq N$ (because $(x-a)/(b-a) \in [0,1]$ whenever $x \in I$. This proves the theorem. □

One important thing to note about the Weierstrass approximation theorem is the necessity for the underlying interval $I = [a,b]$ be compact. To see this, we consider the continuous function f on $(0,1]$ defined by $f(x) = 1/x$ for $x \in (0,1]$. Observe that, for any polynomial P on $(0,1]$,

$$\lim_{x\to 0; x>0} \left| \frac{1}{x} - P(x) \right| = \left| \lim_{x \to 0; x>0} \frac{1}{x} - P(0) \right| = \infty$$

and hence

$$\sup_{x\in\mathbb{R}} \left| \frac{1}{x} - P(x) \right| = \infty.$$

Consequently, there is no sequence of polynomials $(P^{(n)})_{n=1}^{\infty}$ which converges to $f(x) = 1/x$ uniformly on $(0,1]$.

There is another familiar theorem that helps to approximate "nice" functions by polynomials: Taylor's theorem. As you know from Math 131A, Taylor's theorem gives a way to approximate sufficiently differentiable functions by polynomial functions near an interior point of the function's domain. You may be led to ask: What is the difference between the Weirstrass approximation theorem and Taylor's theorem when applied to a sufficiently differentiable (and therefore continuous) functions f? Sometimes, there is no difference. For example, when f is chosen to be a polynomial itself, both theorems eventually yield the same approximant (*an element which approximates* f), f itself. However, in general, the theorems give very different approximants; this difference highlights an important distinction in analysis: global vs. local. Taylor's theorem gives a very good local approximation because it matches the function value at f along with its derivatives. So the approximants yielded by Taylor's theorem match f (its value) at a point and share the same same slope, curvature, etc. By contrast, the approximants yielded by the Weierstrass approximation theorem are good globally (this is the uniform result) but not generally good locally, e.g., nothing is said about the derivatives of the approximants found by Theorem 14.

In 1937 Marshall H. Stone extended the Weierstrass approximation theorem to a much more general context. Stone's result is commonly referred to as the Stone-Weierstrass theorem and it is used frequently in many branches of functional analysis. Following a couple of necessary definitions, we will state two versions of the Stone-Weierstrass theorem in the context of metric spaces. In the interest of time, however, we will not prove the Stone-Weierstrass theorem. For two particularly nice proofs, see [4] and [5].

Definition 16. *Let (X, d) be a metric space. A collection $\mathcal{A}$ of real-valued functions $f : X \to \mathbb{R}$ is said to be an algebra of real-valued functions on X if, for any $f, g \in \mathcal{A}$ and $c \in \mathbb{R}$, $f + g \in \mathcal{A}$, $fg \in \mathcal{A}$ and $cf \in \mathcal{A}$.*

A collection $\mathcal{A}$ of complex-valued functions $f : X \to \mathbb{C}$ is said to be an algebra of complex-valued functions on X if, for any $f, g \in \mathcal{A}$ and $c \in \mathbb{C}$, $f + g \in \mathcal{A}$, $fg \in \mathcal{A}$ and $cf \in \mathcal{A}$.

Definition 17. *Given an algebra $\mathcal{A}$ of (real or complex-valued) functions on a metric space (X, d), we say that $\mathcal{A}$ separates points if, for any distinct points $x, y \in X$, there exists $f \in \mathcal{A}$ for which $f(x) \neq f(y)$.*

Theorem 18 (Stone-Weierstrass theorem (real version))**.** *Let (X, d) be a compact metric space and let $\mathcal{A}$ be an algebra of continuous real-valued functions on X (this requires $\mathcal{A} \subseteq C(X)$). If $\mathcal{A}$ separates points and contains the set of constant functions,*

$$\overline{\mathcal{A}} = C(X)$$

where the closure (or set of adherent points of $\mathcal{A}$) is taken with respect to the metric of uniform convergence, d_∞, on $C(X)$. That is, for each $f \in C(X)$, there exists a sequence of functions $(h^{(n)})_{n=1}^\infty \subseteq \mathcal{A}$ for which

$$\lim_{n\to\infty} d_\infty(f, h^{(n)}) = \lim_{n\to\infty} \|f - h^{(n)}\|_\infty = 0.$$

Example 6.1

The set of polynomials $\mathcal{P}(I)$ is an algebra of real-valued functions on $I = [a, b]$. This is easy to see because the multiplication of any two polynomials is a polynomial, the sum of two polynomials in a polynomial and any constant multiple of a polynomial is a polynomial. As we discussed previously, $\mathcal{P}(I) \subseteq C(I)$ and, moreover, $\mathcal{P}$ contains the constant functions. Further, for any $x', y' \in I$ such that $x' \neq y'$, the first degree polynomial $P(x) = x$ has $P(x') = x' \neq y' = P(y')$ and so $\mathcal{P}(I)$ separates points. Consequently, the Stone-Weierstrass theorem says that

$$\overline{\mathcal{P}(I)} = C(I);$$

this is, however, the conclusion of Theorem 14. Thus we see that The Stone-Weierstrass theorem generalizes the Weierstrass approximation theorem.

Next, we will state the complex-valued version of the Stone-Weierstrass theorem. To this end, we note that the set of continuous functions $f : X \to \mathbb{C}$, denoted by $C(X; \mathbb{C})$, can, as in the real case, be equipped with the metric of uniform convergence: For $f, g \in C(X; \mathbb{C})$,

$$d_\infty(f, g) = \sup_{x\in X} |f(x) - g(x)|.$$

We note that, the only difference between this and that given for real-valued functions is that, here, $|\alpha| = \sqrt{a^2 + b^2}$ is the complex-modulus of a number $\alpha = a + bi \in \mathbb{C}$ whereas, in the real case, it is simply the absolute value. Equipped with this metric, it is quite easy to show (the proof is almost identical to that of the real case), that $(C(X, \mathbb{C}), d_\infty)$ is a complete metric space (this is the analogue of Theorem 8) and convergence in the metric d_∞ is equivalent to uniform convergence.

Definition 19. *Let $\mathcal{A}$ be an algebra of function on X. We say that $\mathcal{A}$ is self-adjoint if, for each $f \in \mathcal{A}$, $\overline{f} \in \mathcal{A}$. Here, for any complex-valued function f, $\overline{f}$, called the conjugate of f, is defined by*

$$\overline{f}(x) = \overline{f(x)} = \overline{\operatorname{Re} f(x) + i \operatorname{Im} f(x)} = \operatorname{Re} f(x) - i \operatorname{Im} f(x)$$

for all $x \in X$.

We can now state the complex-valued version of the Stone-Weierstrass theorem.

Theorem 20. *(Stone-Weierstrass theorem (complex version)) Let (X,d) be a compact metric space and let $\mathcal{A}$ be an algebra of continuous complex-valued functions on X (this requires $\mathcal{A} \subseteq C(X;\mathbb{C})$). If $\mathcal{A}$ is self-adjoint, separates points and contains the set of constant functions, then*

$$\overline{\mathcal{A}} = C(X;\mathbb{C})$$

where the closure (or set of adherent points of $\mathcal{A}$) is taken with respect to the metric of uniform convergence, d_∞, on $C(X;\mathbb{C})$. That is, for each $f \in C(X;\mathbb{C})$, there exists a sequence of functions $(h^{(n)})_{n=1}^\infty \subseteq \mathcal{A}$ for which

$$\lim_{n\to\infty} d_\infty(f, h^{(n)}) = 0.$$

7 Power Series and Real Analytic Functions

In this section, we study a particular (and very important) class of functions, the set of real-analytic functions (we will give a precise definition shortly). To introduce these functions, we must first introduce power series and (as usual) worry about convergence. To help us with this task, we recall two facts from Math 131A concerning numerical series.

Fact 2 (Root test)**.** *Let $(a_n)_{n=0}^\infty$ be a sequence of non-negative real numbers and set*

$$\rho = \limsup_{n\to\infty} (a_n)^{1/n}.$$

If $\rho < 1$, the series $\sum_{n=0}^\infty a_n$ converges. If $\rho > 1$, the series $\sum_{n=0}^\infty a_n$ diverges.

Fact 3. *Let $(a_n)_{n=0}^\infty$ be a sequence of real numbers. If the series $\sum_{n=0}^\infty a_n$ converges, then $\lim_{n\to\infty} a_n = 0$.*

We now define power series: Given $a \in \mathbb{R}$ and a sequence of real numbers $(c_n)_{n=0}^\infty$, we consider the formal expression

$$\sum_{n=0}^\infty c_n (x-a)^n;$$

this is a *power series*. By formal, we are not (yet) saying anything about convergence. Set

$$\gamma = \limsup_{n\to\infty} |c_n|^{1/n} \text{ and } R = 1/\gamma$$

with the convention that $R = \infty$ if $\gamma = 0$ and $R = 0$ if $\gamma = \infty$. We call R the *radius of convergence* of the power series $\sum_{n=0}^\infty c_n(x-a)^n$.

Theorem 21. *Let $\sum_{n=0}^\infty c_n(x-a)^n$ be a power series with radius of convergence R.*

1. *If $x \in \mathbb{R}$ is such that $|x-a| < R$, then the series (now viewed as a series of functions) converges at x.*
2. *If $x \in \mathbb{R}$ is such that $|x-a| > R$, then the series diverges at x.*

Proof. For x such that $|x-a| < R$

$$\limsup_{n\to\infty} (|c_n(x-a)^n|)^{1/n} = |x-a| \limsup_{n\to\infty} |c_n|^{1/n} < 1$$

and so the series converges (absolutely) by the root test. If x is such that $|x-a| > R$, then the root test guarantees (by an analogous computation), that the series $\sum_{n=0}^\infty c_n(x-a)^n$ diverges. □

In view of the above theorem, when $R > 0$ we define $f : (a-R, a+R) \to \mathbb{R}$ by

$$f(x) = \sum_{n=0}^\infty c_n (x-a)^n$$

for $x \in (a-R, a+R)$.

Theorem 22. *Let $\sum_{n=0}^{\infty} c_n(x-a)^n$ be a power series with radius of convergence R. Suppose that $R>0$ and let $0<r<R$.*

1. *For any $0<r<R$, the series $\sum_{n_0}^{\infty} c_n(x-a)^n$ converges uniformly to f on $[a-r,a+r]$. In particular, f is continuous on $(a-R,a+R)$.*
2. *The function f is differentiable on $(a-R,a+R)$. For any $0<r<R$, the series*
$$\sum_{n=1}^{\infty} nc_n(x-a)^{n-1}=\sum_{n=0}^{\infty}(n+1)c_{n+1}(x-a)^n$$
converges uniformly to f' on $[a-r,a+r]$.
3. *For any closed interval $I=[x_1,x_2]\subseteq(a-R,a+R)$,*
$$\int_I f(x)\,dx=\sum_{n=0}^{\infty} c_n\frac{(x_2-a)^{n+1}-(x_1-a)^{n+1}}{n+1}.$$

Proof. I will leave Items 1 and 3 for you to prove in your homework. Here, I will prove Item 2. We observe, for any $0<r<R$,
$$\left|\frac{d}{dx}c_n(x-a)^n\right|=|nc_n(x-a)^{n-1}|\leq nc_nr^{n-1}=:M_n$$
for all $x\in(a-r,a+r)$ and $n\in\mathbb{N}$. Now
$$\limsup_{n\to\infty} M_n^{1/n}=\limsup_{n\to\infty} r^{(n-1)/n}n^{1/n}|c_n|^{1/n}<1$$
because
$$\limsup_{n\to\infty} r^{(n-1)/n}=\lim_{n\to\infty} r^{1-1/n}=r<R$$
and
$$\limsup_{n\to\infty} n^{1/n}=\lim_{n\to\infty} n^{1/n}=\lim_{n\to\infty} e^{(\ln n/n)}=1.$$
Thus, by virtue of the root test, $\sum_{n=0}^{\infty} M_n$ converges, and so The Weierstrass M-test guarantees that $\sum_{n=1}^{\infty} nc_n(x-a)^{n-1}$ converges uniformly on $[a-r,a+r]$. The desired result follows directly from Theorem 13. ☐

8 Real Analytic Functions

Definition 23. *Let $E\subseteq\mathbb{R}$ and let $f:E\to\mathbb{R}$. If a is an interior point of E, we say that f is real analytic at a if there exists an open interval $(a-r,a+r)\subseteq E$ for some $r>0$ such that there exists a power series $\sum_{n=0}^{\infty} c_n(x-a)^n$ centered at a with radius of convergence $R\geq r$ which converges to f on $(a-r,a+r)$. In this case, the series $\sum_{n=0}^{\infty} c_n(x-a)^n$ is called the power series representation of f at a. If E is an open set and f is real analytic at every point a of E we say that f is real analytic on E.*

Essentially, the above definition says that a function f which is real analytic at a has a power series *representation* at centered at a, i.e., f is given by
$$f(x)=\sum_{n=0}^{\infty} c_n(x-a)^n$$
on a neighborhood of a.

Remark 2. *If f is real analytic at a, say with $(a-r,a+r)$, then the series representation $\sum_{n=0}^{\infty} c_n(x-a)^n$ must converge uniformly to f on every compact subinterval of $(a-r,a+r)$ because it converges pointwise (this is the definition) and we know the series is uniformly convergent in view of Theorem 22.*

Remark 3. *The notion defined above is closely related for a similar one for complex-valued functions. These notions should not be confused! Complex-analytically is a much stronger property and is treated in a course on complex analysis.*

Example 8.1

As we have seen previously, the function $f(x) = 1/(1-x)$ has a power series representation (geometric series) about 0. That is,

$$f(x) = \sum_{n=1}^{\infty} x^n$$

for $-1 < x < 1$. Thus f is real analytic at 0 with $r = R = 1$. and $c_n = 1$ for all n. You should note that f is defined for all $x \neq 1$ but the above power (on the right hand side) doesn't converge for any such x. However, f is still real analytic at $a = 2$ with power series representation

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

for $x \in (1, 3)$.

Definition 24. *Let $\mathcal{O} \subseteq \mathbb{R}$ be an open set and let $f : \mathcal{O} \to \mathbb{R}$. We say that f is once differentiable on $\mathcal{O}$ if it is differentiable on $\mathcal{O}$. We write $f^{(1)} := f'$. We say that f is twice differentiable on $\mathcal{O}$ if $f' = f^{(1)}$ is differentiable on $\mathcal{O}$ and we write $f^{(2)} = (f^{(1)})'$. For $k \geq 2$, we say that f is k-times differentiable on $\mathcal{O}$ if $f' = f^{(1)}$ is $k-1$ times differentiable. In this case, we write $f^{(k)} = (f^{(k-1)})'$. We say that f is infinitely differentiable on $\mathcal{O}$ if it is k-times differentiable for every k.*

Proposition 25. *Let E be a subset of $\mathbb{R}$, let a be an interior point of E and let $f : E \to \mathbb{R}$. If f is real-analytic at a with power series representation*

$$f(x) = \sum_{n=0}^{\infty} c_n (x-a)^n$$

for $x \in (a-r, a+r)$. Then f is infinitely differentiable on $(a-r, a+r)$ with, for any $k \geq 1$,

$$f^{(k)}(x) = \sum_{n=0}^{\infty} c_{n+k} \frac{(n+k)!}{n!}(x-a)^n$$

for $x \in (a-r, a+r)$.

Proof. We induct on k. First, in view of Theorem 22, f is once-differentiable on $(a-r, a+r)$ and

$$f'(x) = \sum_{n=0}^{\infty} c_{n+1}(n+1)(x-a)^n = \sum_{n=0}^{\infty} c_{n+1} \frac{(n+1)!}{n!}(x-a)^n$$

for $x \in (a-r, a+r)$. We note that the right hand side converges on $(a-R, a+R)$ where R is the radius of convergence of the original series. However, because (in view of real-analyticity of f) we are only assuming that this power series matches with f on the (possibly proper) subinterval $(a-r, a+r)$ and so our statement must be in terms of this subinterval. We have shown the statement is true for $k = 1$. Let's assume that it's true for $k \geq 1$. That is, f is k-times differentiable on $(a-r, a+r)$ and

$$f^{(k)}(x) = \sum_{n=0}^{\infty} c_{n+k} \frac{(n+k)!}{n!}(x-a)^n$$

for $x \in (a-r, a+r)$; here the series also has radius of convergence $R \geq r$. Applying Theorem 22 to the series above, we have $f^{(k)}$ is once differentiable and

$$f^{(k+1)}(x) = (f^{(k)})'(x) = \sum_{n=0}^{\infty} \left(c_{(n+1)+k} \frac{((n+1)+k)!}{(n+1)!} \right) (n+1)(x-a)^n = \sum_{n=0}^{\infty} c_{n+k+1} \frac{(n+k+1)!}{n!} (x-a)^n$$

for $x \in (a-r, a+r)$. Thus the statement is true for $k+1$. In particular, f is k-times differentiable for every k and so f is infinitely differentiable. ☐

Corollary 26. *If f is analytic on an open set $\mathcal{O} \subseteq \mathbb{R}$, then f is infinitely differentiable on $\mathcal{O}$.*

Proof. In view of the openness criterion for open sets, for any $a \in \mathcal{O}$ let $r > 0$ be such that $(a-r, a+r) \subseteq \mathcal{O}$ and f has a power series representation $\sum_{n=0}^{\infty} c_n (x-a)^n$ valid on the interval $(a-r, a+r)$. Then, in view of the preceding theorem, f is infinitely differentiable on $(a-r, a+r)$. Since a was arbitrary, this holds for every $a \in \mathcal{O}$. ☐

Corollary 27 (Taylor's formula)**.** *Let f be real analytic at a with power series representation*

$$f(x) = \sum_{n=0}^{\infty} c_n (x-a)^n$$

for $x \in (a-r, a+r)$ (where $r > 0$). Then, for any $k \in \mathbb{N}$,

$$f^{(k)}(a) = k! c_k.$$

In particular, we have

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!} (x-a)^n$$

for $x \in (a-r, a+r)$.

Proof. We simply apply the preceding theorem (and evaluate at a). This gives

$$f^{(k)}(a) = \sum_{n=0}^{\infty} c_{n+k} \frac{(n+k)!}{n!} (a-a)^n = c_{0+k} \frac{(0+k)!}{0!} + 0 + 0 \cdots = c_k k! \tag{4}$$

for each $k \in \mathbb{N}$. ☐

Given an infinitely differentiable function f, it is tempting to naively write down Taylor's formula (4) to represent f as a power series centered at a. However, you should be careful! The above corollary is valid only for real-analytic functions f not simply for infinitely differentiable ones. In your homework this week, you will examine a function which is infinitely differentiable on $\mathbb{R}$ but (4) does not hold. In you have studied a little complex-analysis, you should note that complex analyticity doesn't suffer the same drawbacks (think about why this might be true).

Corollary 28. *Let f be real analytic at $a \in \mathbb{R}$ and suppose that f has two power series expansions centered at a:*

$$f(x) = \sum_{n=0}^{\infty} c_n (x-a)^n$$

and

$$f(x) = \sum_{n=0}^{\infty} d_n (x-a)^n$$

both valid for $x \in (a-r, a+r)$. Then $c_n = d_n$ for all $n \in \mathbb{N}$. In other words, power series expansions are unique (if they exist).

Proof. In view the the preceding corollary, the coefficients of a power series of f are determined by the derivatives of f at a and hence $c_n = d_n = n! f^{(n)}(a)$ for all n. ☐

8.1 Multiplication of power series

In this subsection, we show that the product of real analytic functions is real analytic. This is captured in the following theorem, in which (as we've seen before) the procedure of (discrete) convolution appears naturally.

Theorem 29. *Let f and g be real analytic at a with power series representations*

$$f(x)=\sum_{n=0}^{\infty}c_n(x-a)^n$$

and

$$g(x)=\sum_{n=0}^{\infty}d_n(x-a)^n$$

valid for $x\in(a-r,a+r)$ (where $r>0$). Then the product fg is also real analytic at a with power series representation

$$fg(x)=f(x)g(x)=\sum_{n=0}^{\infty}e_n(x-a)^n$$

where

$$e_n=(c*d)_n=\sum_{m=0}^{n}c_m d_{n-m}$$

for each $n\in\mathbb{N}$.

Before proving the theorem, we have to prove a lemma which, under certain conditions, allows us to reverse the order of infinite sums. This lemma is a discrete (special) form of an important theorem in measure theory, known as Fubini's theorem.

Lemma 30 (A special case of Fubini's theorem)**.** *Given a double sequence $(a_{nm})_{n,m=0}^{\infty}$ of real numbers, suppose that, for each $n\in\mathbb{N}$,*

$$M_n=\sum_{m=0}^{\infty}|a_{nm}| \qquad \text{(i.e., this series converges)}$$

and $\sum_{n=0}^{\infty}M_n$ converges. Then

$$\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}a_{nm}=\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}a_{nm}.$$

Proof of Lemma 30. We closely follow the clever proof given in [5] which heavily relies on the Weierstrass M-test. We first construct a metric space. Consider the set of points

$$E=\{0\}\cup\{1/k:k=1,2,\dots\}\subseteq\mathbb{R}.$$

We set $x^*=0$ and $x_k=1/k$ for all k; thus $E=\{x^*\}\cup\{x_k:k\geq 1\}$ is a countable subset of $\mathbb{R}$. As a subset of $\mathbb{R}$, we can furnish E with the metric d inherited from $\mathbb{R}$ and, in this way, E becomes a metric space. A moment's thought shows that, $x^*\in E$ is the only point in E for which, for all $r>0$, $B(x^*,r)$ contains an element of E distinct from x^*. This is to say that E has only one *accumulation point*, namely x^*. Further, let's note that the sequence $(x_k)_{k=1}^{\infty}$ converges to x^* and has the property that $0<d(x^*,x_k)<d(x^*,x_n)$ whenever $k>n$. Also, note that, by taking points in E, there is only essentially one way to construct a sequence that converges to x^* and this is by building it with member of (x_k). More precisely, for any sequence $(y_k)\subseteq E$, the set of elements of this sequence satisfies $\{y_k:k\geq 1\}\subseteq\{x^*\}\cup\{x_k:k\geq 1\}$. This special property of E makes checking continuity particularly simple. In your homework you will prove that, given E with the aforementioned properties and $f:E\to\mathbb{R}$,

$$f\in C(E)\qquad\text{if and only if}\qquad \lim_{j\to\infty}f(x_j)=f(x^*).$$

Armed with this fact, we define a sequence of real valued functions $(f^{(n)})_{n=0}^{\infty}$ on E by

$$f^{(n)}(x^*) = \sum_{m=0}^{\infty} a_{nm} \qquad \text{and, for } k \geq 1, \qquad f^{(n)}(x_k) = \sum_{m=0}^{k} a_{nm}$$

for each $n \in \mathbb{N}$. Also, set

$$g(x) = \sum_{n=0}^{\infty} f^{(n)}(x) \tag{5}$$

for $x \in E$. Let's make note of a few things: For each n, our hypotheses guarantee that the series defining $f^{(n)}(x^*)$ converges and further $\lim_{k\to\infty} f^{(n)}(x_k) = f^{(n)}(x^*)$. Hence, in view of the discussion above $f^{(n)} \in C(E)$ for each n. Furthermore, $|f^{(n)}(x)| \leq M_n$ for each n and $x \in E$, and since the series $\sum M_n$ converges, the Weierstrass M-test guarantees that the series defining $g(x)$ converges uniformly on E and, moreover, $g \in C(E)$. Consequently

$$\begin{aligned}
\sum_{n=0}^{\infty}\sum_{m=0}^{\infty} a_{nm} &= \sum_{n=0}^{\infty} f^{(n)}(x^*) \\
&= g(x^*) \\
&= \lim_{k\to\infty} g(x_k) \\
&= \lim_{k\to\infty} \sum_{n=0}^{\infty} f^{(n)}(x_k) \\
&= \lim_{k\to\infty} \sum_{n=0}^{\infty}\sum_{m=0}^{k} a_{nm} \\
&= \lim_{k\to\infty} \sum_{m=0}^{k}\sum_{n=0}^{\infty} a_{nm} \\
&= \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} a_{nm};
\end{aligned}$$

here we note that the penultimate equality is valid because the finite sum of convergent series is the series of the finite sum of their summands. □

Proof of Theorem 29. We assume, without loss of generality, that $a = 0$. We show that the series $\sum_{n=0}^{\infty} e_n x^n$ converges to $f(x)g(x)$ for $x \in (-r, r)$. By our hypothesis, the power series representations

$$\sum_{m=0}^{\infty} c_m x^m \qquad \text{and} \qquad \sum_{n=0}^{\infty} d_n x^n$$

for f and g respectively have radii of convergence greater than or equal to r. Consequently, both series converge absolutely at this (fixed) x and we set

$$C = \sum_{n=0}^{\infty} |c_n x^n| \qquad \text{and} \qquad D = \sum_{n=0}^{\infty} |d_n x^n|$$

(which are both finite). For any $N \geq 0$, we consider

$$\sum_{n=0}^{N}\sum_{m=0}^{\infty} |d_n x^n c_m x^m|$$

which we can rewrite as

$$\sum_{n=0}^{N} |d_n x^n| \sum_{m=0}^{\infty} |c_m x^m| = \sum_{n=0}^{N} |d_n x^n| C.$$

Consequently, each such partial sum is bounded above by DC for all N and so it follows that

$$\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}|d_n x^n c_m x^m| = \lim_{N\to\infty}\sum_{n=0}^{N}\sum_{m=0}^{\infty}|d_n x^n c_m x^m|,$$

i.e., the (double) series of absolute values converges. This guarantees that

$$\sum_{n=0}^{\infty}\sum_{m=0}^{\infty} d_n x^n c_m x^m$$

is convergent. We compute this series two different ways: First, it is clear that (by properties of series) the above series is equal to

$$\sum_{n=0}^{\infty} d_n x^n \sum_{m=0}^{\infty} c_m x^m = \sum_{n=0}^{\infty} d_n x^n f(x) = f(x)g(x)$$

and so

$$f(x)g(x) = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} c_m x^m d_n x^n.$$

We can also write this as

$$f(x)g(x) = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} c_m d_n x^{n+m}.$$

We know however that the (double) series of absolute values converges and so we apply the lemma (Fubini) to write this as

$$f(x)g(x) = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} c_m d_n x^{n+m}.$$

For the inner series, we make the substitution (change of variables) $n' = n+m$ and so

$$f(x)g(x) = \sum_{m=0}^{\infty}\sum_{n'=m}^{\infty} c_m d_{n'-m} x^{n'}$$

and by adopting the convention that $d_n = 0$ whenever $n < 0$, this can be rewritten as

$$f(x)g(x) = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} c_m d_{n-m} x^n.$$

We now, apply the lemma again (which we can do because of the absolute series converges) to obtain

$$f(x)g(x) = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} c_m d_{n-m} x^n$$

which can be rewritten as

$$f(x)g(x) = \sum_{n=0}^{\infty} x^n \sum_{m=0}^{\infty} c_m d_{n-m}.$$

But again, noting that $d_j = 0$ when $j < 0$, this is

$$f(x)g(x) = \sum_{n=0}^{\infty} x^n \sum_{m=0}^{n} c_m d_{n-m}$$

or, equivalently,

$$f(x)g(x) = \sum_{n=0}^{\infty} x^n e_n = \sum_{n=0}^{\infty} e_n x^n.$$

$\square$

Our next theorem shows that if f is real analytic at a with power series representation valid on the interval $(a-r,a+r)$, then f, for any $b\in(a-r,a+r)$, f has a power series representation centered at b. In other words, f is analytic on the interval $(a-r,a+r)$.

Theorem 31. *Let f be real analytic at a with representation*

$$f(x)=\sum_{n=0}^{\infty}c_n(x-a)^n$$

for $x\in(a-r,a+r)$ (necessarily $r\leq R$ where R is the radius of convergence of the above series). Then for any $b\in(a-r,a+r)$, f is real analytic at b with

$$f(x)=\sum_{n=0}^{\infty}d_n(x-b)^n$$

for $x\in(b-s,b+s)\subseteq(a-r,a+r)$ where

$$d_n=\sum_{m=n}^{\infty}\frac{m!}{n!(m-n)!}(b-a)^{m-n}c_m.$$

Proof. For any $x\in(b-s,b+s)\subseteq(a-r,a+r)$,

$$\begin{aligned}
f(x)=\sum_{n=0}^{\infty}c_n(x-a)^n &= \sum_{n=0}^{\infty}c_n(x-b+(b-a))^n\\
&= \sum_{n=0}^{\infty}c_n\sum_{m=0}^{n}\frac{n!}{m!(n-m)!}(b-a)^{n-m}(x-b)^m\\
&= \sum_{n=0}^{\infty}\sum_{m=0}^{n}c_n\frac{n!}{m!(n-m)!}(b-a)^{n-m}(x-b)^m\\
&= \sum_{n=0}^{\infty}\sum_{m=0}^{\infty}e_{nm}(x-b)^m \qquad (6)
\end{aligned}$$

where

$$e_{nm}=\begin{cases}c_n\frac{n!}{m!(n-m)!}(b-a)^{n-m} & \text{if } m\leq n\\ 0 & \text{if } m>n\end{cases}$$

for $n,m\in\mathbb{N}$. Now, for each n,

$$\begin{aligned}
M_n &= \sum_{m=0}^{\infty}|e_{nm}(x-b)^m|\\
&= \sum_{m=0}^{n}|c_n|\frac{n!}{m!(n-m)!}|b-a|^{n-m}|x-b|^m=|c_n|(|x-b|+|b-a|)^n<|c_n|(s+r-s)^n=|c_n|r^n
\end{aligned}$$

because $|x-b|<s$ and $|b-a|<r-s$ because $(b-s,b+s)\subseteq(a-r,a+r)$ (you should check these basic inequalities). Also, because $r\leq R$ where R is the radius of convergence of the original series $\sum c_n(x-a)^n$, $\sum_n M_n$ must converge. Consequently, we have met the hypotheses of Fubini's theorem and therefore, in view of (6),

$$f(x)=\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}e_{mn}(x-b)^m=\sum_{m=0}^{\infty}\sum_{n=m}^{\infty}c_n\frac{n!}{m!(n-m)!}(b-a)^{n-m}(x-b)^m=\sum_{m=0}^{\infty}d_m(x-b)^m$$

as desired. ∎

Corollary 32. *If f is analytic at a with valid power series representation on the interval $(a-r, a+r)$. Then f is analytic on the interval $(a-r, a+r)$.*

Our final theorem in this section is Abel's theorem. Due to lack of time, we will not prove the theorem. For a proof, see your textbook (Analysis II by Tao).

Theorem 33 (Abel's theorem)**.** *Let f be analytic at a with power series representation*

$$f(x) = \sum_{n=0}^{\infty} c_n (x-a)^n$$

for $x \in (a-r, a+r)$. If $\sum_{n=0}^{\infty} c_n r^n$ converges, then f is continuous (from the left) at $x = r$ and

$$\lim_{x \mapsto r; x \in (a-r, a+r)} \sum_{n=0}^{\infty} c_n (x-a)^n = \sum_{n=0}^{\infty} c_n r^n.$$

Similarly, if $\sum_{n=0}^{\infty} c_n (-r)^n$ converges, then f is continuous (from the right) at $x = -r$ and

$$\lim_{x \mapsto -r; x \in (a-r, a+r)} \sum_{n=0}^{\infty} c_n (x-a)^n = \sum_{n=0}^{\infty} c_n (-r)^n.$$

8.2 Exponential and Log

Consider the following definition.

Definition 34. *Let $\exp : \mathbb{R} \to \mathbb{R}$ be defined by*

$$\exp(x) = \sum_{n=0}^{\infty} \frac{1}{n!} x^n$$

for each $x \in \mathbb{R}$. We call $\exp$ the exponential function.

In the homework you will prove that the radius of the defining series is infinite, i.e., $R = \infty$, and the series converges for every real number x. In particular, the definition above makes sense. In the homework you will also prove the following proposition.

Proposition 35. *The following properties for $\exp$ hold:*

1. *The function $\exp$ is differentiable and $\exp'(x) = \exp(x)$ for every $x \in \mathbb{R}$.*
2. *The function $\exp$ is continuous on $\mathbb{R}$ and, for each $a < b$,*
$$\int_{[a,b]} \exp(x)\, dx = \exp(b) - \exp(a).$$
3. *For every $x, y \in \mathbb{R}$, $\exp(x+y) = \exp(x)\exp(y)$.*
4. *We have $\exp(0) = 1$. Further, for each $x \in \mathbb{R}$, $\exp(x) > 0$ and $\exp(-x) = 1/\exp(x)$ (together with the previous item, this shows that $\exp$ is a homomorphism from the group $(\mathbb{R}, +)$ to $((0,\infty), \times)$.*
5. *The function $\exp$ is strictly increasing with range $(0,\infty)$ (hence, it's a bijection from $\mathbb{R}$ onto $(0,\infty)$).*

We define Euler's number e by

$$e = \exp(1) = \sum_{n=0}^{\infty} \frac{1}{n!} = \frac{1}{0!} + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \cdots.$$

With the help of the proposition above, we can prove the following corollary.

Corollary 36. *For any $x \in \mathbb{R}$,*

$$\exp(x) = e^x.$$

Proof. If $x = 1/n$ where n is a positive integer, $(\exp(1/n))^n = \exp(n/n) = \exp(1) = e$ and so, necessarily, $e^{1/n} = \exp(1/n)$. Now, if n is a positive integer and m is a non-negative integer,

$$\exp(m/n) = \exp\left(\sum_{j=1}^{m}(1/n)\right) = (\exp(1/n))^m = (e^{1/n})^m = e^{m/n}$$

where we have invoked Item 3 of the previous proposition. Finally, if n is a positive integer and m is a negative integer, we have $\exp(m/n) = 1/\exp(-m/n) = 1/e^{-m/n} = e^{m/n}$. Upon combining these results, we conclude that $e^x = \exp(x)$ for all $x \in \mathbb{Q}$. Since exp is continuous and strictly increasing function on $\mathbb{R}$, given any $x \in \mathbb{R} \setminus \mathbb{Q}$,

$$\exp(x) = \sup\{\exp(q) : q \in \mathbb{Q}, q \le x\} = \sup\{e^q : q \in \mathbb{Q}, q \le x\} = e^x$$

where the final equality holds because this is the way e^x is defined for irrational exponents. Hence, $\exp(x) = e^x$ for all $x \in \mathbb{R}$. $\square$

Looking back to Proposition 35, $\exp : \mathbb{R} \to (0, \infty)$ is a bijection which is everywhere differentiable with non-vanishing derivative. In particular, it makes sense to talk about its inverse function. Let's give it a definition.

Definition 37. *Let* $\ln : (0,\infty) \to \mathbb{R}$ *defined by* $\ln(x) = \exp^{-1}(x)$ *for* $x \in (0,\infty)$. *That is,* $\ln$ *is defined to be the inverse of the exponential function* exp. *We call* $\ln$ *the logarithm or the natural logarithm.*

Proposition 38. *The following statements hold for the logarithm.*

1. *For every $x \in (0,\infty)$, $\ln'(x) = 1/x$. In particular, by the fundamental theorem of calculus,*

$$\int_{[a,b]} \frac{1}{x}\, dx = \ln(b) - \ln(a)$$

   for each $0 < a < b$.

2. *For each $x, y \in (0,\infty)$, $\ln(xy) = \ln(x) + \ln(y)$.*

3. *For any $x \in (-1,1)$,*

$$\ln(1-x) = -\sum_{n=1}^{\infty} \frac{x^n}{n}.$$

   In particular, $\ln$ *is analytic at* 1 *with power series representation*

$$\ln(x) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}(x-1)^n$$

   for $x \in (0,2)$.

You will prove this proposition in Homework 6.

8.3 The complex exponential

In this subsection, we briefly discuss the complex exponential function. This is an extension of the function exp defined in the previous subsection. To study this function, we need to consider infinite series of complex numbers, and not simply real numbers as has previously been our main focus. By a close study, one notices that much of the preceding results about series hold true provided we replace the absolute value by the complex modulus [5].

We define

$$\exp(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!}$$

for $z \in \mathbb{C}$. Both the ratio and the root tests guarantee that the above series converges for every $z \in \mathbb{C}$. Moreover, given any $r > 0$, the above series converges uniformly on the closed disk $\overline{D}(0,r) = \{z \in \mathbb{C} : |z| \leq r\}$. Using the same argument you used for the real exponential (and in view of the validity of Theorem 29 for complex series), we have

$$\exp(z+w) = \exp(z)\exp(w)$$

for all $z, w \in \mathbb{C}$. Given any complex number z, we define

$$e^z = \exp(z).$$

We note that, in view of Corolary 36, this formula makes sense whenever z is real. This is why e^z is called the exponential function. Henceforth, we will use $\exp(z)$ and e^z interchangeably.

Theorem 39. *The complex exponential function,* exp*, is an extension of the real exponential function defined in Definition 34. The following properties hold:*

1. *For every complex z, $e^z \neq 0$.*
2. exp *is its own (complex) derivative. That is*

$$\exp'(z) = \lim_{w \to z} \frac{\exp(w) - \exp(z)}{w - z} = \exp(z)$$

   for each $z \in \mathbb{C}$.
3. *We have*

$$\exp(it) = \cos(t) + i\sin(t)$$

   for all $t \in \mathbb{R}$ where

$$\cos(t) = \sum_{k=0}^{\infty} \frac{(-1)^k t^{2k}}{(2k)!} \qquad \textit{and} \qquad \sin(t) = \sum_{k=0}^{\infty} \frac{(-1)^k t^{2k+1}}{(2k+1)!}$$

   for $t \in \mathbb{R}$.
4. *The map $t \mapsto e^{it}$ is a continuous surjection from $\mathbb{R}$ onto the unit circle $S^1 = \{z \in \mathbb{C} : |z| = 1\}$.*

For a proof of the above theorem, see [5].

9 Fourier Series

In this section, we study Fourier series. As the Weierstrass approximation theorem and Taylor's theorem give ways of approximating certain functions by polynomials, the goal of the study to which we now turn is to approximate certain functions with trigonometric polynomials. Approximation by trigonometric polynomials (linear combinations of sines, cosines, e^{it} and their powers) has significant advantages over other approximation methods. For one, it makes easy the task of solving ordinary and partial differential equations. In fact, the subject of Fourier series was initially developed by its namesake, Jean-Baptiste Fourier, in his study of heat diffusion and the heat equation. This pioneering investigation was presented in Fourier's main work, *Théorie analytique de la chaleur* (Analytic theory of heat), in 1822. What Fourier likely didn't know is that this subject would be the source of so much beautiful mathematics over the next 200 years.

Definition 40. *A trigonometric polynomial is a finite sum of the form*

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

for $x \in \mathbb{R}$ where $a_0, a_1, \dots, a_N, b_1, b_2, \dots, b_N$ are complex numbers. The collection of all trigonometric polynomials will be denoted by $\mathcal{P}(\mathbb{T})$.

Let's make a couple of quick observations regarding trigonometric polynomials:

1. Given $f \in \mathcal{P}(\mathbb{T})$, f is periodic of period 2π, i.e.,

$$f(x + 2\pi) = f(x)$$

   for all $x \in \mathbb{R}$.

2. Each $f \in \mathcal{P}(\mathbb{T})$ is continuous, i.e., $f \in C(\mathbb{R}; \mathbb{C})$.

3. Each $f \in \mathcal{P}(\mathbb{T})$ can be written uniquely in the form

$$f(x) = \sum_{n=-N}^{N} c_n e^{inx}$$

   for $x \in \mathbb{R}$.

In light of the above observations, let's denote the set of all continuous, complex-valued functions which are periodic of period 2π by $C(\mathbb{T})$, i.e., each $f \in C(\mathbb{T})$ belongs to $C(\mathbb{R}; \mathbb{C})$ and satisfies $f(x) = f(x + 2\pi)$ for all $x \in \mathbb{R}$. By our observations above, $\mathcal{P}(\mathbb{T}) \subseteq C(\mathbb{T})$. Though we haven't defined the symbol $\mathbb{T}$ independently, the reason for its appearance is because each $f \in C(\mathbb{T})$ has a unique representation as a continuous function on the 1-dimensional torus $\mathbb{T} = \mathbb{R}/(2\pi\mathbb{Z})$ (a quotient of groups). You will study (lightly) this representation in Homework 7. In these notes, we will not otherwise touch on this representation involving the torus; this perspective becomes very important when one wants to study the more general subject of harmonic analysis.

Our goal for the remainder of this introductory subsection is to prove a uniform approximation theorem, analogous to the Weierstrass approximation theorem, for trigonometric polynomials. It will be key to our arguments in following sections. Our approach is different from that taken in the textbook for this course. In the following lemma S^1 denotes the unit circle in the complex plane $\mathbb{C}$. That is,

$$S^1 = \{z \in \mathbb{C} : |z| = 1\}.$$

Taking $\mathbb{C}$ to be equipped with the standard metric, i.e., $d(z, w) = |z - w| = \sqrt{((\mathrm{Re}(z - w))^2 + (\mathrm{Im}(z - w))^2}$, we view S^1 as a metric space equipped with the induced metric from $\mathbb{C}$.

Lemma 41. *There is a one to one correspondence between elements of $C(\mathbb{T})$ and elements of $C(S^1; \mathbb{C})$. Specifically, to each $f \in C(\mathbb{T})$ there exists a unique $\tilde{f} \in C(S^1; \mathbb{C})$ (a continuous complex-valued function from S^1 to $\mathbb{C}$) which satisfies*

$$f(t) = \tilde{f}(e^{it})$$

for all $t \in \mathbb{R}$ and, further, every such element of $C(S^1; \mathbb{C})$ is of this form.

Proof. For any $f \in C(\mathbb{T})$, define $\tilde{f} : S^1 \to \mathbb{C}$ by

$$\tilde{f}(z) = f(t)$$

when $z = e^{it}$ for $t \in \mathbb{R}$. We must verify that $\tilde{f}$ is well defined. To this end, observe that, if $z = e^{it} = e^{is}$ for, possibly distinct $t, s \in \mathbb{R}$, then $e^{i(t-s)} = 1$ by the additive property of exp. Consequently $t = s + 2\pi k$ for some integer k and therefore $f(s) = f(s + 2\pi k) = f(t)$ where we have used the fact that f is periodic of period 2π.

To see that $\tilde{f}$ is continuous, let $w \in S^1$ and $(w_n)_{n=1}^\infty$ be a sequence of elements in S^1 which converges to w with respect to the metric inherited from $\mathbb{C}$. Because multiplication and division by non-zero numbers are continuous operations in $\mathbb{C}$, it follows that

$$\lim_{n\to\infty} z_n = \lim_{n\to\infty} w_n w^{-1} = ww^{-1} = 1$$

where $z_n = w_n w^{-1}$ for each n. We may assume without loss of generality (by possibly throwing out a finite number of terms) that $|\arg(z_n)| \leq \pi/2$ for all n. Consequently, for each n, there is a unique solution to the equation

$$z_n = e^{it} = \cos(t) + i\sin(t)$$

for $t \in [-\pi/2, \pi/2]$ and we will denote this by t_n. I claim that

$$\lim_{n\to\infty} t_n = 0.$$

If this wasn't the case, we could find a subsequence (t_{n_j}) of (t_n) which converges to some non-zero $\alpha \in [-\pi/2, \pi/2]$. In view of the continuity of the map $t \mapsto e^{it}$, we would then have

$$\lim_{j\to\infty} z_{n_j} = \lim_{j\to\infty} e^{it_{n_j}} = e^{i\alpha} \neq 1$$

but this would contradict our assumption that $z_n \to 1$. This proves the claim. Consequently,

$$\lim_{n\to\infty} e^{i(t_n+s)} = \lim_{n\to\infty} e^{it_n}e^{is} = \lim_{n\to\infty} z_n w = \lim_{n\to\infty} w_n = w$$

where $w = e^{is}$ for some $s \in (-\pi, \pi]$. In view of our mapping $f \mapsto \tilde{f}$,

$$\lim_{n\to\infty} \tilde{f}(w_n) = \lim_{n\to\infty} \tilde{f}(e^{i(t_n+s)}) = \lim_{n\to\infty} f(t_n + s) = f(s) = \tilde{f}(w)$$

because $t_n + s \to s$ and f is continuous on the real line. We have therefore shown that $\tilde{f}$ is continuous.

It remains to show that, given $h \in C(S^1; \mathbb{C})$, there exits a continuous function $f \in C(\mathbb{T})$ for which $\tilde{f} = h$. To this end, given $h \in C(S^1; \mathbb{C})$, define $f : \mathbb{R} \to \mathbb{C}$

$$f(t) = h(e^{it})$$

for $t \in \mathbb{R}$. It is clear that f continuous because it is the composition of continuous functions. Moreover, f is 2π-periodic because $t \mapsto e^{it}$ is 2π-periodic. It remains to show that $\tilde{f} = h$. Given any $w = e^{is} \in S^1$ (where $t \in \mathbb{R}$), we have

$$\tilde{f}(w) = f(s) = h(e^{is}) = h(w),$$

as desired. $\square$

Theorem 42. *Let $f \in C(\mathbb{T})$. For any $\epsilon > 0$, there exits $P \in \mathcal{P}(\mathbb{T})$ such that*

$$\sup_{t\in\mathbb{R}} |f(t) - P(t)| < \epsilon.$$

Proof. We consider the collection $\mathcal{A}$ of functions $p : S^1 \to \mathbb{C}$ of the form

$$p(z) = a_{-N}z^{-N} + \cdots a_{-1}z^{-1} + a_0 + a_1 z^1 + \cdots a_N z^N = \sum_{n=-N}^{N} a_n z^n \tag{7}$$

for $z \in S^1$ where $\{a_{-N}, \ldots, a_{-1}, a_0, a_1, \ldots, a_N\} \subseteq \mathbb{C}$ and N is a natural number. Because addition, multiplication and division (by non-zero complex numbers) are continuous operations on $\mathbb{C}$, it is clear that $\mathcal{A} \subseteq C(S^1; \mathbb{C})$. We claim that $\mathcal{A}$ is a self-adjoint algebra of continuous complex-valued functions on S^1 which separates points. First, it is clear that $\mathcal{A}$ forms an algebra because of the distributive laws of complex numbers and the fact that $z^m z^n = z^{m+n}$

for all integers m,n. To see that $\mathcal{A}$ separates points, simply note that, for any distinct points $z_1, z_2 \in S^1$, the function $q(z) = z$ is a member of $\mathcal{A}$ and has $q(z_1) = z_1 \neq z_2 = q(z_2)$. Finally, for any $p \in \mathcal{A}$ (of the form (7)),

$$\overline{p}(z) = \overline{p(z)} = \overline{\sum_{n=-N}^{N} a_n z^n} = \sum_{n=-N}^{N} \overline{a_n}\,\overline{z^n} = \sum_{n=-N}^{N} \overline{a_n} z^{-n} = \sum_{j=-N}^{N} \overline{a_{-j}} z^j$$

for all $z \in S^1$ where we have used the fact that $\overline{w} = \overline{e^{it}} = e^{-it} = (e^{-it})^{-1} = w^{-1}$ for each $w = e^{it} \in S^1$ in view of the properties of the complex exponential. Consequently, $\overline{p} \in \mathcal{A}$. Thus we have shown that $\mathcal{A}$ is a self-adjoint algebra of continuous complex valued functions which separates points. Moreover, since S_1 is the image of the compact set $[0, 2\pi]$ under the continuous function $t \mapsto e^{it}$, S^1 is compact. Thus, an appeal to the Stone-Weierstrass theorem shows that

$$\overline{\mathcal{A}} = C(S^1; \mathbb{C})$$

where the closure is taken with respect to the metric of uniform convergence.

Let $f \in C(\mathbb{T})$ and take $\epsilon > 0$. In view of the preceding lemma, there exists $\tilde{f} \in C(S^1; \mathbb{C})$ for which $f(t) = \tilde{f}(e^{it})$ for all $t \in \mathbb{R}$. By virtue of the preceding paragraph, let $p \in \mathcal{A}$ be such that

$$\sup_{z \in S^1} |\tilde{f}(z) - p(z)| = \sup_{t \in \mathbb{R}} |\tilde{f}(e^{it}) - p(e^{it})| < \epsilon$$

where we have used the fact that $t \mapsto e^{it}$ is a surjective map from $\mathbb{R}$ onto S^1. We notice however that

$$P(t) = p(e^{it}) = \sum_{n=-N}^{N} a_n (e^{it})^n = \sum_{n=-N}^{N} e^{int}$$

for $t \in \mathbb{R}$ defines an element $P \in \mathcal{P}(\mathbb{T})$. Consequently,

$$\sup_{t \in \mathbb{R}} |P(t) - f(t)| = \sup_{t \in \mathbb{R}} |p(e^{it}) - \tilde{f}(e^{it})| < \epsilon$$

as desired. $\square$

9.1 The L^2 theory

In the last subsection, we showed that continuous 2π-periodic functions can be approximated uniformly by trigonometric polynomials. In this subsection, we turn our focus to a larger class of functions, 2π-periodic and Riemann integrable functions, and a completely different mode of approximation, the L^2 approximation. It is my hope that you will see that the L^2 theory, outlined in this subsection, is the cleanest and best-adapted for approximation by trigonometric polynomials. In fact, convergence of Fourier series in L^2 will happen even when other forms of convergence fail.

Let $f : \mathbb{R} \to \mathbb{C}$. We say that f is Riemann integrable on an interval $[a, b]$ if its real and imaginary parts are Riemann integrable on $[a, b]$. In this case, we define the Riemann integral of f (a complex-valued function) on $[a, b]$ by

$$\int_{[a,b]} f = \int_{[a,b]} f(x)\,dx = \int_{[a,b]} \mathrm{Re}(f)(x)\,dx + i \int_{[a,b]} \mathrm{Im}(f)(x)\,dx.$$

Using the properties of the Riemann integral established for real-valued functions, it is easy to check that the Riemann integral, defined here for complex-valued functions, is also linear.

Now, given a 2π-periodic function $f : \mathbb{R} \to \mathbb{C}$, we say that f is Riemann integrable on $\mathbb{T}$ if f is Riemann integrable on $[-\pi, \pi]$ and in this case we write

$$\int_{\mathbb{T}} f = \int_{\mathbb{T}} f(x)\,dx = \int_{[-\pi,\pi]} f(x)\,dx.$$

When the context is clear, we will simply say that f is Riemann integrable. The set of all such functions will henceforth be denoted by $R(\mathbb{T})$. It is clear that $C(\mathbb{T}) \subseteq R(\mathbb{T})$.

Given $f, g \in R(\mathbb{T})$, we define the $L^2 = L^2(\mathbb{T})$ inner product by

$$\langle f, g\rangle = \frac{1}{2\pi}\int_{\mathbb{T}} f(x)\overline{g(x)}\,dx.$$

You should verify that $\langle \cdot, \cdot \rangle$ satisfies all of the properties of an inner product except positive definiteness. In particular, $\langle f, f\rangle \geq 0$ for all $f \in R(\mathbb{T})$ and so it makes sense to define the corresponding L^2 norm by

$$\|f\|_{L^2(\mathbb{T})} = \sqrt{\langle f, f\rangle} = \left(\frac{1}{2\pi}\int_{\mathbb{T}} |f(x)|^2\,dx\right)^{1/2};$$

this is also called the root-mean-square norm. We will often use the shorthand $\|f\|_2$ for $\|f\|_{L^2(\mathbb{T})}$.

As a quick remark, I should point out that $\|\cdot\|_2$ isn't a bona fide norm on the set $R(\mathbb{T})$. It satisfies all of the properties of a norm except positive definiteness. To see that positive definiteness fails, consider the function

$$f(x) = \begin{cases} 1 & \text{if } x \in 2\pi\mathbb{Z} \\ 0 & \text{otherwise.}\end{cases}$$

It is clear that $f \in R(\mathbb{T})$. Further,

$$\|f\|_2^2 = \frac{1}{2\pi}\int_{[-\pi,\pi]} f(x)^2\,dx = 0$$

however f is not the zero function. As it turns out $\|\cdot\|_2$ does become a norm on the Lebesgue space $L^2(\mathbb{T})$ which we do not discuss here. In fact, $L^2(\mathbb{T})$ is a complete (as a metric space) inner product space, such a space is called a Hilbert space. This is standard material in a course on measure theory.

Let us now observe that, for each $n, m \in \mathbb{Z}$,

$$\int_{\mathbb{T}} e^{inx}\overline{e^{imx}}\,dx = \int_{\mathbb{T}} e^{inx-imx}\,dx = \int_{\mathbb{T}} e^{i(n-m)x}\,dx = \int_{[-\pi,\pi]} \cos((n-m)x)\,dx + i\int_{[-\pi,\pi]} \sin((n-m)x)\,dx.$$

Consequently,

$$\langle e^{inx}, e^{imx}\rangle = \frac{1}{2\pi}\int_{\mathbb{T}} e^{inx}\overline{e^{imx}}\,dx = \begin{cases} 1 & \text{if } n = m \\ 0 & \text{if } n \neq m.\end{cases}$$

In view of the preceding calculation, the collection of functions $(e^{inx})_{n\in\mathbb{Z}}$ is called an *orthonormal system.* Let's make a further observation, given $P \in \mathcal{P}(\mathbb{T})$ of the form

$$P(x) = \sum_{n=-N}^{N} c_n e^{inx},$$

for each $-N \leq m \leq N$,

$$\langle P, e^{imx}\rangle = \sum_{n=-N}^{N} c_n\langle e^{inx}, e^{imx}\rangle = c_m$$

and $\langle P, e^{imx}\rangle = 0$ whenever $|m| > N$. In this way, we find a way to find the coefficients of a trigonometric polynomial by simply integrating against the elements of the system (e^{inx}). This is analogous to the way that the coefficients of an analytic function can be computed by taking derivatives. Taking our cues from the above computation, we make a definition.

Definition 43. *Let $f \in R(\mathbb{T})$. For each $n \in \mathbb{Z}$, define*

$$\hat{f}(n) = \langle f, e^{inx} \rangle = \frac{1}{2\pi} \int_{\mathbb{T}} f(x) e^{-inx} dx.$$

The collection of complex numbers $(\hat{f}(n))_{n\in\mathbb{Z}}$ are called the Fourier coefficients of f. Considered as a formal series, the series

$$\sum_{n\in\mathbb{Z}} \hat{f}(n) e^{inx}$$

is called the Fourier series for f and we write

$$f(x) \sim \sum_{n\in\mathbb{Z}} \hat{f}(n) e^{inx}$$

to indicate that the right hand side is the Fourier series for f.

In view of the discussion preceding the definition, we see that computing the Fourier series corresponding to a trigonometric polynomial P returns the polynomial P itself. Do you see an analogue with Taylor series here? Throughout the rest of this section, we begin to analyze the ways in which the Fourier series for a function f converges. As you will see, if the Fourier series is to converge, it will converge back to f. To this end, we will start talking about partial sums.

Let $f \in R(\mathbb{T})$ and $(\hat{f}(n))_{n\in\mathbb{Z}}$ be the Fourier coefficients of f. For each $N \in \mathbb{N}$, the Nth order trigonometric polynomial

$$S_N(x) = \sum_{n=-N}^{N} \hat{f}(n) e^{inx}$$

defined for $x \in \mathbb{R}$ is called the Nth partial sum of the Fourier series $\sum \hat{f}(n)e^{inx}$. Our first main result of the subsection says that, of all Nth order trigonometric polynomials, S_N is the best root-mean-square approximation to f.

Theorem 44. *Let $f \in R(\mathbb{T})$ and let $(\hat{f}(n))_{n\in\mathbb{Z}}$ be its Fourier coefficients. Given $N \in \mathbb{N}$ let $S_N(x)$ be the Nth order partial sum of the Fourier series for f and let $P_N \in \mathcal{P}(\mathbb{T})$ be another (possibly different) Nth order trigonometric polynomial of the form*

$$P_N(x) = \sum_{n=-N}^{N} c_n e^{inx}.$$

Then

$$\|f - S_N\|_2 \leq \|f - P_N\|_2$$

where equality holds if and only if $c_n = \hat{f}(n)$ for all n.

Proof. We first recall the basic linearity properties of the inner product: For $u, v, w \in R(\mathbb{T})$, and $a \in \mathbb{C}$,

$$\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle, \qquad \langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle$$

and

$$\langle au, v \rangle = a\langle u, v \rangle, \qquad \langle u, av \rangle = \overline{a}\langle u, v \rangle.$$

By virtue of these properties, we can write

$$\begin{aligned}
\|f-P_N\|_2^2 &= \langle f-P_N, f-P_N\rangle = \langle f,f\rangle - \langle f,P_N\rangle - \langle P_N,f\rangle + \langle P_N,P_N\rangle\\
&= \|f\|_2^2 - \sum_{n=-N}^{N}\overline{c_n}\langle f,e^{inx}\rangle - \sum_{n=-N}^{N} c_n\langle e^{inx},f\rangle + \sum_{n=-N}^{N}\sum_{m=-N}^{N} c_n\overline{c_m}\langle e^{inx},e^{imx}\rangle\\
&= \|f\|_2^2 - \sum_{n=-N}^{N}\hat{f}(n)\overline{c_n} - \sum_{n=-N}^{N} c_n\overline{\hat{f}(n)} + \sum_{n=-N}^{N} c_n\overline{c_n}\\
&= \|f\|_2^2 - \sum_{n=-N}^{N}|\hat{f}(n)|^2 + \sum_{n=-N}^{N}|c_n-\hat{f}(n)|^2
\end{aligned}$$

where we have used the fact that $\langle e^{inx},f\rangle = \overline{\langle f,e^{inx}\rangle} = \overline{\hat{f}(n)}$. Obviously, making the above computation when $c_n = \hat{f}(n)$ yields

$$\|f-S_N\|_2^2 = \|f\|_2^2 - \sum_{n=-N}^{N}|\hat{f}(n)|^2 \tag{8}$$

and therefore

$$\|f-P_N\|_2^2 = \|f-S_N\|_2^2 + \sum_{n=-N}^{N}|c_n-\hat{f}(n)|^2$$

from which we observe that

$$\|f-S_N\|_2 \le \|f-P_N\|_2$$

with equality if and only if $c_n = \hat{f}(n)$ for all n. □

Theorem 45 (Bessel's inequality)**.** *Let $f\in R(\mathbb{T})$ and let $(\hat{f}(n))_{n\in\mathbb{Z}}$ be the Fourier coefficients of f. Then*

$$\sum_{n\in\mathbb{Z}}|\hat{f}(n)|^2 \le \|f\|_2^2 = \frac{1}{2\pi}\int_{\mathbb{T}}|f(x)|^2\,dx.$$

In particular, the series

$$\sum_{n\in\mathbb{Z}}|\hat{f}(n)|^2 = \lim_{N\to\infty}\sum_{n=-N}^{N}|\hat{f}(n)|^2$$

converges.

Proof. In view of (8), we have

$$\sum_{n=-N}^{N}|\hat{f}(n)|^2 \le \|f\|_2^2$$

for all $N\in\mathbb{N}$. The desired result follows by taking the limit of the left hand side as $N\to\infty$ and noting that the partial sums, whose summands are all non-negative, form a non-decreasing sequence of non-negative numbers. □

Corollary 46 (The Riemann-Lebesgue lemma)**.** *Let $f\in R(\mathbb{T})$ and let $(\hat{f}(n))_{n\in[Z}$ be the Fourier coefficients of f. Then*

$$\lim_{n\to\infty}\hat{f}(n) = \frac{1}{2\pi}\int_{\mathbb{T}}f(x)e^{-inx}\,dx = 0$$

and

$$\lim_{n\to\infty}\hat{f}(-n) = \frac{1}{2\pi}\int_{\mathbb{T}}f(x)e^{inx}\,dx = 0.$$

Proof. The convergence of the series $\sum_{n=-\infty}^{\infty}|\hat{f}(n)|^2$ implies that the summands for sufficiently large and largely negative n converge to zero. □

Corollary 47 (A sharper form of the Riemann-Lebesgue lemma)**.** *Let $[a,b]\subseteq[-\pi,\pi]$ and suppose that f is Riemann integrable on $[a,b]$. Then*

$$\lim_{n\to\infty}\frac{1}{2\pi}\int_{[a,b]}f(x)\cos(nx)\,dx=0$$

and

$$\lim_{n\to\infty}\frac{1}{2\pi}\int_{[a,b]}f(x)\sin(nx)\,dx=0.$$

Proof. Given f as above, consider $g\in R(\mathbb{T})$ defined by

$$g(x)=\begin{cases}f(x) & x\in[a,b]\\ 0 & x\in[-\pi,\pi]\setminus[a,b]\end{cases}$$

and extended periodically to $\mathbb{R}$. Then, by the Riemann-Lebesgue lemma applied to g, we have

$$\lim_{n\to\infty}\frac{1}{2\pi}\int_{[a,b]}f(x)e^{-inx}\,dx=\lim_{n\to\infty}\frac{1}{2\pi}\int_{[-\pi,\pi]}g(x)e^{-inx}\,dx=\lim_{n\to\infty}\hat{g}(n)=0$$

and

$$\lim_{n\to\infty}\frac{1}{2\pi}\int_{[a,b]}f(x)e^{inx}\,dx=\lim_{n\to\infty}\frac{1}{2\pi}\int_{[-\pi,\pi]}g(x)e^{inx}\,dx=\lim_{n\to\infty}\hat{g}(-n)=0.$$

Consequently,

$$\begin{aligned}\lim_{n\to\infty}\frac{1}{2\pi}\int_{[a,b]}f(x)\cos(nx)\,dx &= \lim_{n\to\infty}\frac{1}{2\pi}\int_{[a,b]}f(x)\left(\frac{e^{inx}+e^{-inx}}{2}\right)dx\\ &= \frac{1}{2}\left(\lim_{n\to\infty}\frac{1}{2\pi}\int_{[a,b]}f(x)e^{inx}\,dx+\lim_{n\to\infty}\frac{1}{2\pi}\int_{[a,b]}f(x)e^{-inx}\,dx\right)\\ &= 0.\end{aligned}$$

The proof that $\lim_{n\to\infty}(2\pi)^{-1}\int_{[a,b]}f(x)\sin(nx)\,dx=0$ is similar. □

Our next result shows that the Fourier series for f converges to f with respect to the $L^2(\mathbb{T})$ norm. The result also shows that Bessel's inequality is, in fact, an equality. We first need the following simple lemma that you will prove in your Homework 7.

Lemma 48. *Let $f\in\mathbb{R}(\mathbb{T})$. For any $\epsilon>0$, there exists $g\in C(\mathbb{T})$ such that*

$$\|f-g\|_2<\epsilon.$$

Theorem 49 (Parseval's Theorem)**.** *Let $f\in R(\mathbb{T})$ and let $(\hat{f}(n))_{n\in\mathbb{Z}}$ the Fourier coefficients of f. For each natural number N, denote by S_N the Nth partial sum of the Fourier series for f. Then*

$$\lim_{N\to\infty}\|f-S_N\|_2=0;$$

this is the statement that the Fourier series $\sum_{n\in\mathbb{Z}}\hat{f}(n)e^{inx}$ converges to f with respect to the L^2 norm. Further

$$\|f\|_2^2=\sum_{n\in\mathbb{Z}}|\hat{f}(n)|^2.$$

Proof. Let $\epsilon>0$ and by an appeal to the lemma choose $h\in C(\mathbb{T})$ for which $\|f-h\|_2<\epsilon/2$. Now, in view of Theorem 42, Let $P\in\mathcal{P}(\mathbb{T})$ be a trigonometric polynomial of the form

$$P(x)=\sum_{n=-N}^{N}c_ne^{inx}$$

for which
$$|h(x) - P(x)| < \epsilon/2$$
for all $x \in \mathbb{R}$. From this it follows immediately that
$$\|f - P\|_2 \leq \|f - h\|_2 + \|h - P\|_2 < \epsilon/2 + \left(\frac{1}{2\pi}\int_{[-\pi,\pi]} (\epsilon/2)^2\, dx\right)^{1/2} = \epsilon.$$
Now, for any $M \geq N$, the Mth partial sum of the Fourier series for f, S_M can be compared with the trigonometric polynomial P (which can trivially be though of as of Mth degree by taking its coefficients to be zero for $M \leq |n| > N$. Thus, in view of Theorem 44
$$\|f - S_M\|_2 \leq \|f - P\|_2 < \epsilon.$$
Hence, for every $\epsilon > 0$, there exists and $N \in \mathbb{N}$ such that, for every $M \geq N$, $\|f - S_M\|_2 < \epsilon$ and therefore
$$\lim_{N\to\infty} \|f - S_N\|_2 = 0.$$
With this, (8) gives
$$\lim_{N\to\infty} \sum_{n=-N}^{N} |\hat{f}(n)|^2 = \|f\|_2^2 - \lim_{N\to\infty} \|f - S_N\|_2^2 = \|f\|_2^2 - 0$$
form which the desired result follows immediately. $\square$

Example 9.1

Consider the so-called sawtooth function defined by
$$f(x) = x \qquad -\pi < x \leq \pi$$
and extended 2π-periodically to $\mathbb{R}$. The graph of f is illustrated in Figure 2 (the vertical lines are not part of the graph; they are inserted automatically by Matlab).

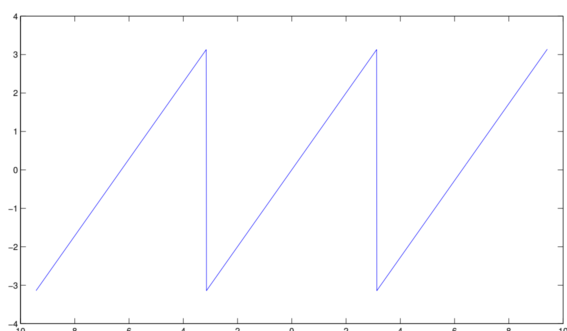

Figure 2: The graph of f for $-3\pi < x \leq 3\pi$.

Let's compute the Fourier coefficients of f. For $n = 0$, we have
$$\hat{f}(0) = \frac{1}{2\pi}\int_{\mathbb{T}} f(x)e^{-i\cdot 0\cdot x}\, dx = \frac{1}{2\pi}\int_{\mathbb{T}} x\, dx = \frac{1}{2\pi}\int_{[-\pi,\pi]} x\, dx = 0.$$

For $n \neq 0$,

$$\begin{aligned}
2\pi\hat{f}(n) &= \int_{\mathbb{T}} xe^{-inx}\,dx \\
&= \int_{[-\pi,\pi]} x\cos(-nx)\,dx + i\int_{[-\pi,\pi]} x\sin(-nx)\,dx \\
&= \int_{[-\pi,\pi]} x\cos nx\,dx - i\int_{[-\pi,\pi]} x\sin nx\,dx \\
&= -i\left(\frac{-1}{n}x\cos nx\Big|_{-\pi}^{\pi} - \frac{-1}{n}\int_{[-\pi,\pi]}\cos nx\,dx\right) \\
&= \frac{i}{n}(\pi\cos n\pi - (-\pi\cos(-n\pi)) - 0 \\
&= \frac{2\pi i}{n}\cos n\pi = \frac{2\pi i}{n}(-1)^n
\end{aligned}$$

where we have integrated by parts and used (heavily) the periodicity and odd and even properties of sine/cosine. Consequently,

$$\hat{f}(n) = \begin{cases} \frac{i(-1)^n}{n} & n \neq 0 \\ 0 & n = 0. \end{cases}$$

Observe that

$$\sum_{n\in\mathbb{Z}} |\hat{f}(n)|^2 = \sum_{n\in\mathbb{Z}\setminus\{0\}} \frac{1}{n^2} = 2\sum_{n=1}^{\infty}\frac{1}{n^2}.$$

Also,

$$\|f\|_2^2 = \frac{1}{2\pi}\int_{\mathbb{T}} |f(x)|^2\,dx = \frac{1}{2\pi}\int_{[-\pi,\pi]} x^2\,dx = \frac{1}{2\pi}\frac{x^3}{3}\Big|_{-\pi}^{\pi} = \frac{2\pi^3}{6\pi} = \frac{\pi^2}{3}$$

and so, by an application of Parseval's theorem, we obtain

$$\frac{\pi^2}{3} = \|f\|_2^2 = \sum_{x\in\mathbb{Z}} |\hat{f}(n)|^2 = 2\sum_{n=1}^{\infty}\frac{1}{n^2}.$$

It was long understood that the series $\sum_{n=1}^{\infty} 1/n^2$ converged. The Basel problem, posed by Pietro Mengoli around 1644, asks: What is the exact value of this series? In 1734, a mathematician named Leonhard Euler showed that the value of the series is exactly $\pi^2/6$. The solution made Euler instantly famous. By simply dividing the previous equation by 2 we obtain Euler's result (Theorem 50 below). We note that our approach (via Parseval's Theorem) is completely different than that of Euler. This should not be surprising as Fourier series wasn't discovered for nearly one hundred years after Euler presented his result.

Theorem 50 (Euler 1734)**.**

$$\sum_{n=1}^{\infty}\frac{1}{n^2} = \frac{\pi^2}{6}.$$

As our attention will soon turn to the investigation of pointwise convergence of Fourier series, let's continue our Fourier series analysis of the Sawtooth function. As we will see, this example contains a surprising number of the strange phenomena/pathologies found commonplace in the study of Fourier series, including the Gibb's phenomenon. For each N, the Nth Fourier polynomial for f is given by

$$\begin{aligned} S_N(x) &= \sum_{n=-N}^{N} \hat{f}(n)e^{inx} = \hat{f}(0) + \sum_{n=1}^{N} \hat{f}(n)e^{inx} + \sum_{n=1}^{N} \hat{f}(-n)e^{-inx} \\ &= \sum_{n=1}^{N} \frac{i(-1)^n}{n}e^{inx} + \frac{i(-1)^{-n}}{-n}e^{-inx} = \sum_{n=1}^{N} \frac{(-1)^n}{n} i(e^{inx} - e^{-inx}) \\ &= \sum_{n=1}^{N} \frac{(-1)^n}{n} i(2i\sin nx) = \sum_{n=1}^{N} \frac{2(-1)^{n+1}}{n} \sin nx \end{aligned}$$

for $x \in \mathbb{R}$. The Fourier polynomials of degrees $1, 2, 3, 4, 5$ and 40 are illustrated in Figure 3 for $-3\pi < x \leq 3\pi$. The reader should note that the partial sums appear to be converging nicely except for the overshoot near the points of discontinuity at $x = \pm 3\pi, \pm\pi$; this is the Gibb's phenomenon.

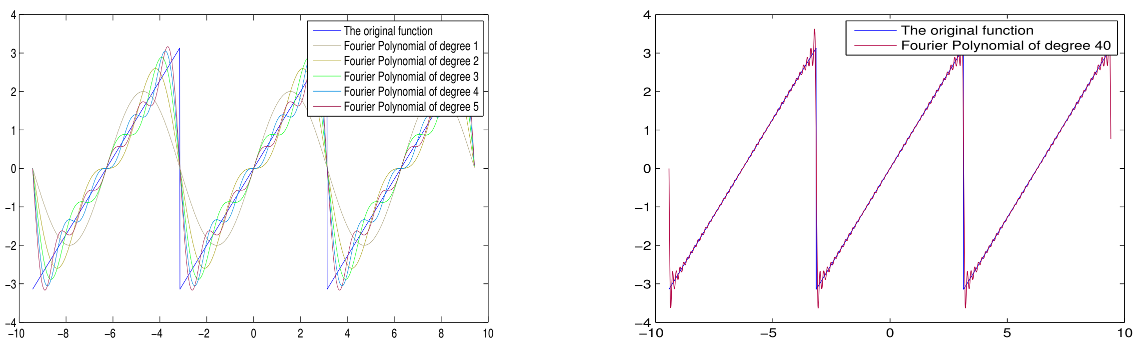

Figure 3: The graphs of f and P_n for $n = 1, 2, 3, 4, 5, 40$.

9.2 The pointwise and uniform theory theory

As we saw in the last subsection, the Fourier series for a function $f \in R(\mathbb{T})$ converges to f with respect to the L^2 norm. In this subsection, we investigate the same question from the perspective of pointwise (and uniform) convergence; as we have seen, both notions are stronger than L^2 convergence. When originally investigating Fourier series in its connection to the theory of heat, Jean-Baptiste Fourier claimed that, given $f \in C(\mathbb{T})$, the Fourier series $\sum_{n\in\mathbb{N}} \hat{f}(n)e^{inx}$ converges to $f(x)$ for all $x \in \mathbb{R}$. As it turns out, this isn't true.

Theorem 51 (Du Bois-Reymond, 1873)**.** *There exits $f \in C(\mathbb{T})$ whose Fourier series $\sum_{n\in\mathbb{Z}} \hat{f}(n)e^{inx}$ diverges at a point $x \in (-\pi, \pi]$. This means specifically that, for some $x \in (-\pi, \pi]$, the limit*

$$\lim_{N\to\infty} \sum_{n=-N}^{N} \hat{f}(n)e^{inx}$$

does not exist.

Following the discovery of this result, a frantic search began in mathematics to find precise conditions on a function f which would guarantee that its Fourier series converged. Throughout this search, it was discovered that the Riemann integral was insufficient for the needed purposes of investigation and this helped lead to a complete revolution in mathematics. Our of the revolution emerged the Lebesgue integral and Lebesgue's theory of integration. Any serious investigation of Fourier series (which ours is unfortunately not) requires one to understand the Lebesgue integral. In 1926 a brilliant young mathematician named Andre Kolmogorov showed that things were much worse than had been previously thought (and had been shown by Du Bois-Reymond).

Theorem 52 (Kolmogorov 1926). *There exists a Lebesgue integrable function f (we say $f \in L^1(\mathbb{T})$) whose Fourier series diverges at every point.*

We now begin our investigation of the uniform convergence of Fourier series.

Lemma 53. *Let $f, g \in C(\mathbb{T})$, if $\hat{f}(n) = \hat{g}(n)$ for all n, then $f = g$, i.e., $f(x) = g(x)$ for all $x \in \mathbb{R}$.*

Proof. As you show in your homework, $\widehat{(f-g)}(n) = \widehat{f}(n) - \widehat{g}(n)$ for all $n \in \mathbb{Z}$, i.e., the map $f \mapsto \hat{f}$ is linear. Thus, $\widehat{(f-g)}(n) = 0$ for all $n \in \mathbb{Z}$ and so, by an appeal to Parseval's theorem,

$$\|f - g\|_2^2 = \sum_{n \in \mathbb{Z}} |\widehat{(f-g)}(n)|^2 = 0.$$

Consequently,

$$\int_{\mathbb{T}} |f(x) - g(x)|^2 \, dx = 2\pi \|f - g\|_2^2 = 0.$$

It now follows, by the result I prove in class during Week 6, that $f(x) - g(x) = 0$ for all $x \in [-\pi, \pi]$. Because f and g are 2π periodic, we conclude that $f(x) = g(x)$ for all $x \in \mathbb{R}$. $\square$

With the help of the preceding lemma and the Weierstrass M-test, we can give a sufficient condition for the uniform convergence of Fourier series.

Theorem 54. *Let $f \in C(\mathbb{T})$ and let $(\hat{f}(n))_{n \in \mathbb{Z}}$ be the Fourier coefficients of f. If the series $\sum_{n \in \mathbb{Z}} |\hat{f}(n)|$ converges, then the Fourier series of f, $\sum_{n \in \mathbb{Z}} \hat{f}(n) e^{inx}$, converges uniformly to f.*

Proof. We observe that

$$|\hat{f}(n) e^{inx}| = |\hat{f}(n)||e^{inx}| = |\hat{f}(n)|$$

for all $n \in \mathbb{Z}$ and $x \in \mathbb{R}$. In other words, the summands $g_n(x) = \hat{f}(n) e^{inx}$ satisfy $|g_n(x)| \leq M_n = |\hat{f}(n)|$ for all $n \in \mathbb{Z}$ and $x \in \mathbb{R}$. Since the series of non-negative numbers $\sum |\hat{f}(n)| = \sum M_n$ converges, the Weierstrass M-test (applied to series of complex-valued functions) guarantees that the Fourier series converges uniformly to some complex-valued function g on $\mathbb{R}$. Further, because each summand $\hat{f}(n) e^{inx}$ is continuous, g is continuous. Also, observe that

$$g(x + 2\pi) = \lim_{N \to \infty} \sum_{n=-N}^{N} \hat{f}(n) e^{in(x+2\pi)} = \lim_{N \to \infty} \sum_{n=-N}^{N} \hat{f}(n) e^{inx} = g(x)$$

for all $x \in \mathbb{R}$ and so g is 2π-periodic. We therefore conclude that $g \in C(\mathbb{T})$. It remains to show that $g = f$. To this end, we compute the Fourier coefficients of g. Observe that, for any $m \in \mathbb{Z}$,

$$g(x) e^{-imx} = e^{-imx} \lim_{N \to \infty} \sum_{n=-N}^{N} \hat{f}(n) e^{inx} = \lim_{N \to \infty} \sum_{n=-N}^{N} \hat{f}(n) e^{inx} e^{-imx}$$

for all $x \in \mathbb{R}$. In fact, the Weierstrass M-text applied to the summands $n \mapsto \hat{f}(n) e^{inx} e^{-imx}$ shows that the partial sums on the right hand side converge uniformly to $g(x) e^{-mx}$. Since each summand is Riemann integrable (it is continuous) and the series converges uniformly, we may integrate term-by-term. For each $m \in \mathbb{Z}$, we have

$$\hat{g}(m) = \langle g, e^{imx} \rangle = \frac{1}{2\pi} \int_{\mathbb{T}} g(x) e^{-imx} \, dx = \sum_{n \in \mathbb{Z}} \frac{1}{2\pi} \int_{\mathbb{T}} (\hat{f}(n) e^{inx} e^{-imx}) \, dx = \sum_{n \in \mathbb{Z}} \hat{f}(n) \langle e^{inx}, e^{imx} \rangle = \hat{f}(m)$$

where we have used the fact that $\langle e^{inx}, e^{imx} \rangle = 1$ when $n = m$ and 0 otherwise. Therefore $\hat{f}(n) = \hat{g}(n)$ for all $n \in \mathbb{Z}$ and, in view of the preceding lemma $f = g$. In other words,

$$f(x) = \sum_{n \in \mathbb{Z}} \hat{f}(n) e^{inx}$$

where the convergence is uniform for $x \in \mathbb{R}$. $\square$

I've always found the result above somewhat unsatisfying, though it is powerful as we will shortly see. The reason I find it unsatisfying is because its hypotheses are stated in terms of the Fourier coefficients of f and so, to apply the theorem, one has to compute the Fourier coefficients of f and then ask if they are absolutely summable. One would like to instead have hypotheses stated in terms of f itself. In any case, the theorem allows us to prove the following result.

Corollary 55. *Let $f \in C(\mathbb{T})$. If f is twice continuously differentiable, i.e., f's real and imaginary parts are twice continuously differentiable on $\mathbb{R}$, then the Fourier series for f, $\sum_{n\in\mathbb{Z}} \hat{f}(n)e^{inx}$, converges uniformly to f on $\mathbb{R}$.*

Proof. Let $f(x) = u(x) + iv(x)$ where u and v are real-valued 2π periodic functions which are, by hypothesis, twice continuously differentiable on $\mathbb{R}$. We have $f'(x) = u'(x) + iv'(x)$ and $f''(x) = u''(x) + iv''(x)$ for $x \in \mathbb{R}$. I claim that, for all non-zero $n \in \mathbb{N}$,

$$\hat{f}(n) = \frac{-1}{2\pi n^2}\int_{\mathbb{T}} f''(x)e^{inx}\,dx.$$

To see this, first observe that

$$\begin{aligned}
\int_{\mathbb{T}} f''(x)e^{inx}\,dx &= \int_{\mathbb{T}} (u''(x) + iv''(x))(\cos nx + i\sin nx)\,dx\\
&= \int_{\mathbb{T}} (u''(x)\cos nx - v''(x)\sin nx) + i(u''(x)\sin nx + v''(x)\cos nx)\,dx\\
&= \int_{[-\pi,\pi]} (u''(x)\cos nx - v''(x)\sin nx)\,dx + i\int_{[-\pi,\pi]} (u''(x)\sin nx + v''(x)\cos nx)\,dx\\
&= \int_{[-\pi,\pi]} u''(x)\cos nx\,dx - \int_{[-\pi,\pi]} v''(x)\sin nx\,dx\\
&\qquad + i\int_{[-\pi,\pi]} u''(x)\sin nx\,dx + i\int_{[-\pi,\pi]} v''(x)\cos nx\,dx
\end{aligned}$$

for all $n \neq 0$. Let's analyze the first integral above. For $n \neq 0$, integration by parts gives

$$\begin{aligned}
\int_{[-\pi,\pi]} u''(x)\cos nx\,dx &= u'(x)\cos nx\big|_{-\pi}^{\pi} - \int_{[-\pi,\pi]} u'(x)(-n\sin nx)\,dx\\
&= (u'(\pi) - u'(-\pi))\cos(n\pi) - \left(u(x)(-n\sin nx)\big|_{-\pi}^{\pi} - \int_{[-\pi,\pi]} u(x)(-n^2\cos nx)\,dx\right)\\
&= -n^2\int_{[-\pi,\pi]} u(x)\cos nx\,dx
\end{aligned}$$

where we have used the fact that u' is 2π-periodic (as it must be because u is). By completely analogous calculations we obtain

$$\begin{aligned}
\int_{\mathbb{T}} f''(x)e^{inx}\,dx &= -n^2\int_{[-\pi,\pi]} u(x)\cos nx\,dx + n^2\int_{[-\pi,\pi]} v(x)\sin nx\,dx\\
&\qquad -in^2\int_{[-\pi,\pi]} u(x)\sin nx\,dx - in^2\int_{[-\pi,\pi]} v(x)\cos nx\,dx\\
&= -n^2\left(\int_{[-\pi,\pi]} (u(x)\cos nx - v(x)\sin nx)\,dx + i\int_{[-\pi,\pi]} (u(x)\sin nx + v(x)\cos nx)\,dx\right)\\
&= -n^2\int_{\mathbb{T}} f(x)e^{inx}\,dx = -2\pi n^2\hat{f}(n)
\end{aligned}$$

for each $n \neq 0$. This proves the claim.

Now observe that, for each non-zero $n \in \mathbb{Z}$,

$$|\hat{f}(n)| = \left|\frac{1}{2\pi n^2}\int_{\mathbb{T}} f''(x)e^{inx}\,dx\right| \leq \frac{1}{2\pi n^2}\int_{\mathbb{T}} |f''(x)e^{inx}|\,dx = \frac{1}{2\pi n^2}\int_{\mathbb{T}} |f''(x)|\,dx = \frac{M}{n^2}$$

where $M = (1/2\pi)\int_{\mathbb{T}} |f''(x)|\,dx$; M is finite because f'' is continuous and therefore bounded on $[-\pi,\pi]$. Consequently, $\sum_{n\in\mathbb{Z}} |\hat{f}(n)|$ converges and so, by an appeal the the preceding theorem, the Fourier series $\sum_{n\in\mathbb{Z}} \hat{f}(n)e^{inx}$ converges uniformly to f. $\square$

This concludes our investigation of the uniform convergence of Fourier series. We now move on to the theory of pointwise convergence. In contrast to our study of uniform convergence, the main object in our study of pointwise convergence of Fourier series revolves around a careful (analytical) study of the Dirichlet kernel, which we introduce now.

Proposition 56. *Let $f \in R(\mathbb{T})$ and let $(\hat{f}(n))_{n\in\mathbb{Z}}$ be its Fourier coefficients. For each N, let S_N denote the Nth Fourier polynomial of f, i.e.,*

$$S_N(x) = \sum_{n=-N}^{N} \hat{f}(n)e^{inx}.$$

Then, for each $N \in \mathbb{N}$ and $x \in \mathbb{R}$,

$$S_N(x) = \frac{1}{2\pi}\int_{\mathbb{T}} D_N(x-y)f(y)\,dy$$

where

$$D_N(y) := \begin{cases} \frac{\sin((N+1/2)y)}{\sin y/2} & y \neq 2\pi k, k \in \mathbb{Z} \\ 2N+1 & y = 2\pi k, k \in \mathbb{Z} \end{cases} = \sum_{n=-N}^{N} e^{iny};$$

this is called the Dirichlet kernel. For each N,

$$\frac{1}{2\pi}\int_{\mathbb{T}} D_N(y)\,dy = 1$$

and

$$\frac{1}{2\pi}\int_{[-\pi,0]} D_N(y)\,dy = \frac{1}{2\pi}\int_{[0,\pi]} D_N(y)\,dy = \frac{1}{2}.$$

Proof. By virtue of the linearity of the integral, it is evident that

$$\begin{aligned} S_N(x) &= \sum_{n=-N}^{N}\left(\frac{1}{2\pi}\int_{\mathbb{T}} f(y)e^{-iny}\,dy\right)e^{inx} \\ &= \frac{1}{2\pi}\int_{\mathbb{T}}\sum_{n=-N}^{N} f(y)e^{in(x-y)}\,dy \end{aligned}$$

for $N \in \mathbb{N}$ and $x \in \mathbb{R}$. We now show that

$$D_N(y) = \sum_{n=-N}^{N} e^{iny}$$

for $N \in \mathbb{N}$ and $x \in \mathbb{R}$. First, it is clear that, if $y = 2\pi k$ for $k \in \mathbb{Z}$, $D_N(y) = 2N+1 = \sum_{n=-N}^{N} 1 = \sum_{n=-N}^{N} e^{iny}$.

Thus, we assume without loss of generality that $y \neq 2\pi k$ for $k \in \mathbb{Z}$ and hence $\sin y/2 \neq 0$. Now,

$$\begin{aligned}
D_1(y) &= \frac{\sin((1+1/2)y)}{\sin(y/2)} = \frac{\sin y \cos(y/2) + \sin(y/2)\cos y}{\sin(y/2)} \\
&= \frac{\sin((1/2+1/2)y)\cos(y/2)}{\sin(y/2)} + \cos y = \frac{2\sin(y/2)\cos(y/2)\cos(y/2)}{\sin(y/2)} + \cos y \\
&= 2\cos^2(y/2) + \cos y = 1 + 2\cos y \\
&= e^{i\cdot 0\cdot y} + e^{iy} + e^{iy} = \sum_{n=-1}^{1} e^{iny}
\end{aligned}$$

where we have used the half-angle identity $2\cos^2(A/2) = 1 + \cos(A)$. Thus, the desired result is true for $N = 1$. Let's induct on N. Let's assume that the formula holds for $N \geq 1$, we will show it holds for $N+1$. We have

$$\begin{aligned}
D_{N+1}(y) &= \frac{\sin((N+1+1/2)y)}{\sin y/2} \\
&= \frac{\sin y \cos((N+1/2)y) + \cos y \sin((N+1/2)y}{\sin y/2} \\
&= 2\cos(y/2)\cos((N+1/2)) + \cos y D_N(y) \\
&= 2\cos(y/2)(\cos Ny \cos(y/2) - \sin Ny \sin(y/2)) + \cos y D_N(y) \\
&= (1+\cos y)\cos Ny - \sin Ny \sin y + \cos y D_N(y) \\
&= \cos Ny + (\cos y \cos Ny - \sin Ny \sin y) + \cos y D_N(y) \\
&= \cos Ny + \cos((N+1)y) + \cos y D_N(y) \\
&= \frac{e^{iNy} + e^{-iNy}}{2} + \frac{e^{i(N+1)y} + e^{-i(N+1)y}}{2} + \frac{e^{iy} + e^{-iy}}{2} D_N(y) \\
&= \frac{e^{iNy} + e^{-iNy}}{2} + \frac{e^{i(N+1)y} + e^{-i(N+1)y}}{2} + \frac{e^{iy} + e^{-iy}}{2} \sum_{n=-N}^{N} e^{iny} \\
&= \frac{1}{2}\left(e^{iNy} + e^{-iNy} + e^{i(N+1)y} + e^{-i(N+1)y} + \sum_{n=-N}^{N} e^{i(n+1)y} + \sum_{n=-N}^{N} e^{i(n-1)y}\right) \\
&= \frac{1}{2}\left(e^{iNy} + e^{-iNy} + e^{i(N+1)y} + e^{-i(N+1)y} + \sum_{n=-(N-1)}^{N+1} e^{iny} + \sum_{n=-(N+1)}^{N-1} e^{iny}\right) \\
&= \frac{1}{2}\left(\sum_{n=-(N+1)}^{N+1} e^{iny} + \sum_{n=-(N+1)}^{N+1} e^{iny}\right) \\
&= \sum_{n=-(N+1)}^{N+1} e^{iny}
\end{aligned}$$

where we have made use of the induction hypothesis and a tour de force of trigonometric identities.

Now, for any $N \in \mathbb{N}$, due to the periodicity of the functions $x \mapsto e^{inx}$ for $n \neq 0$ and their antiderivatives,

$$\frac{1}{2\pi}\int_{\mathbb{T}} D_N(y)\,dy = \frac{1}{2\pi}\int_{\mathbb{T}} \sum_{n=-N}^{N} e^{iny}\,dy = \frac{1}{2\pi}\int_{\mathbb{T}} e^{i0\cdot y}\,dy = \frac{1}{2\pi}\int_{[-\pi,\pi]} 1\,dy = 1.$$

Finally, by a quick examination, it is clear the $D_N(y)$ is an even function. Consequently

$$1 = 2\left(\frac{1}{2\pi}\int_{[0,\pi]} D_N(y)\,dy\right) = 2\left(\frac{1}{2\pi}\int_{[-\pi,0]} D_N(y)\,dy\right)$$

from which the final result follows. $\square$

The Dirichlet kernels D_5, D_{10} and D_{20} are illustrated in Figure 4.

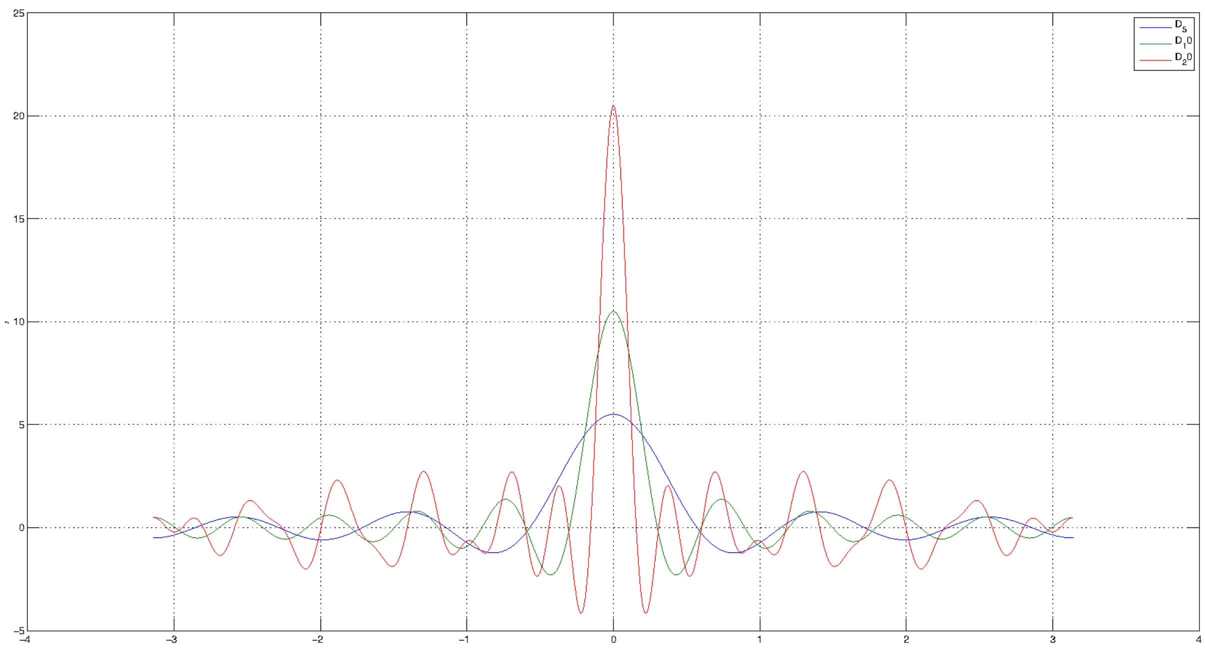

Figure 4: The graphs of D_5, D_{10} and D_20.

With the properties of the Dirichlet kernel established in the preceding proposition, we have our first (truely) pointwise result.

Theorem 57. *Let $f \in R(\mathbb{T})$ and assume that the $f = u + iv$ is piecewise differentiable, i.e., its real and imaginary parts u and v are continuously differentiable on $[-\pi, \pi]$ except possibly at a finite number of points where u, v, u' and v' have (at worse) removable or jump discontinuities. For any $x_0 \in \mathbb{R}$, define*

$$f(x_0^+) = \lim_{x \to x_0; x > x_0} f(x) \qquad \text{and} \qquad f(x_0^-) = \lim_{x \to x_0 : x < x_0} f(x).$$

Then, at each $x_0 \in \mathbb{R}$, the Fourier series for f converges and

$$\frac{f(x_0^+) + f(x_0^-)}{2} = \sum_{n \in \mathbb{Z}} \hat{f}(n) e^{inx_0} = \lim_{N \to \infty} \sum_{n=-N}^{N} \hat{f}(n) e^{inx_0}.$$

In particular, if f is continuous at x_0 (or has a removable discontinuity at x_0),

$$f(x_0) = \sum_{n \in \mathbb{Z}} \hat{f}(n) e^{inx_0};$$

otherwise, the Fourier series for f converges to the average of the left and right limits of f at x_0.

Before proving the theorem, let's illustrate its conclusion by revisiting the sawtooth function and its Fourier series.

Example 9.2

We recall the sawtooth function defined by $f(x) = x$ for $-\pi < x \leq \pi$ and extended periodically to $\mathbb{R}$. We previously studied this function in Example 9.1 and computed its Fourier series. We found

$$f(x) \sim \sum_{n\in\mathbb{Z}\setminus\{0\}} \frac{i(-1)^n}{n} e^{inx} = \sum_{n=1}^{\infty} \frac{2(-1)^{n+1}}{n} \sin nx$$

Now, by a straightforward computation, f is differentiable on the open set $\mathbb{R} \setminus \{(2k+1)\pi : k \in \mathbb{Z}\}$ (consisting of the entire real line except for the *breakpoints* $\pm\pi, \pm 3\pi, \pm 5\pi, \dots$) and, on this set, $f'(x) = 1$. Consequently, f is piecewise differentiable and so we may apply the theorem.

For $-\pi < x < \pi$, f is continuous and by virtue of the theorem we conclude that the Fourier series for f converges to $f(x) = x$ on the open set $(-\pi, \pi)$. At $x = \pi$, f has a discontinuity. Here we have

$$f(\pi_-) = \lim_{x\to\pi; x<\pi} f(x) = \lim_{x\to\pi; x<\pi} x = \pi$$

and

$$f(\pi_+) = \lim_{x\to\pi; x>\pi} f(x) = \lim_{x\to\pi; x>\pi} (x - 2\pi) = \pi - 2\pi = -\pi.$$

Consequently,

$$\sum_{n\in\mathbb{Z}} \hat{f}(n) e^{in\pi} = \lim_{N\to\infty} S_N(\pi) = \frac{1}{2}(\pi + -\pi) = 0;$$

in fact, this result holds at all the breakpoints $\pm\pi, \pm 3\pi, \dots,$. Appealing to the full scope of the theorem (or simply noting that the above conclusion extends by periodicity), we have

$$\sum_{n\in\mathbb{Z}} \hat{f}(n) e^{inx} = \lim_{N\to\infty} S_N(x) = \begin{cases} (x - 2\pi k) & (2k-1)\pi < x < (2k+1)\pi, k \in \mathbb{Z} \\ 0 & x = (2k+1)\pi, k \in \mathbb{Z} \end{cases}.$$

In other words, the Fourier series for f converges to f pointwise on the open set $\mathbb{R} \setminus \{(2k+1)\pi : k \in \mathbb{Z}\}$ and it converges to 0 elsewhere. We note that the series does not converge uniformly for, if this were the case, f would be continuous. This convergence is illustrated in the following Figure 5.

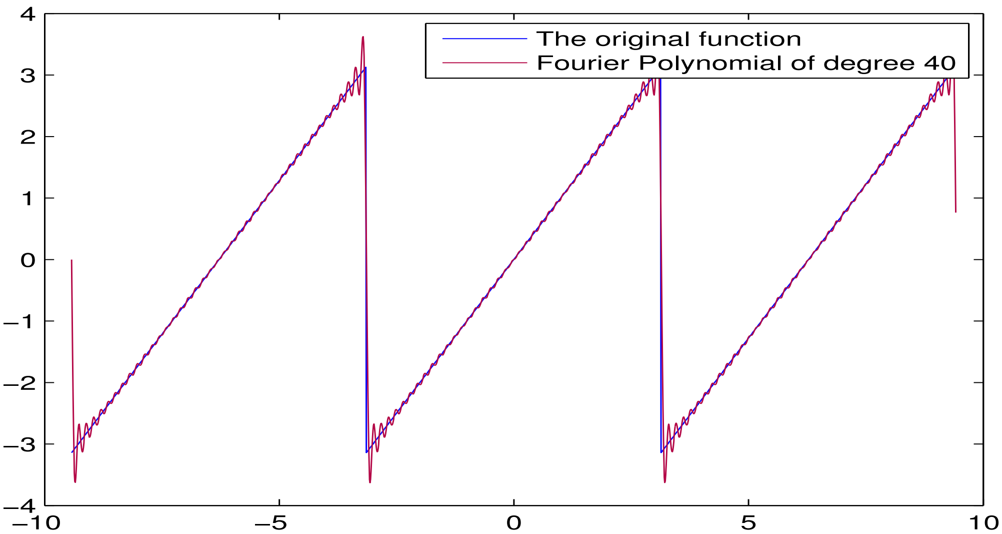

Figure 5: f and S_{40}

We now prove the theorem.

Proof of Theorem 57. Our aim is to show that

$$\lim_{N\to\infty}\left(S_N(x_0)-\frac{f(x_0^+)+f(x_0^-)}{2}\right)=0.$$

In view of Proposition 56, we have

$$\begin{aligned}\left(S_n(x_0)-\frac{f(x_0^+)+f(x_0^-)}{2}\right) &= \frac{1}{2\pi}\int_{\mathbb{T}} f(x_0-y)D_N(y)\,dy-\left(f(x_0^+)\frac{1}{2}+f(x_0^-)\frac{1}{2}\right)\\ &= \frac{1}{2\pi}\int_{[-\pi,0]} f(x_0-y)D_N(y)\,dy+\frac{1}{2\pi}\int_{[0,\pi]} f(x_0-y)D_N(y)\,dy\\ &\qquad -\left(f(x_0^+)\frac{1}{2\pi}\int_{[-\pi,0]} D_N(y)\,dy+f(x_0^-)\frac{1}{2\pi}\int_{[0,\pi]} D_N(y)\,dy\right)\\ &= \frac{1}{2\pi}\int_{[-\pi,0]}(f(x_0-y)-f(x_0^+))D_N(y)\,dy+\frac{1}{2\pi}\int_{[0,\pi]}(f(x_0-y)-f(x_0^-))D_N(y)\,dy\\ &=: I_+(N)+I_-(N).\end{aligned}$$

We consider the integrals I_+ and I_-. By virtue of Proposition 56,

$$\begin{aligned}
I_+(N) &= \frac{1}{2\pi}\int_{[-\pi,0]}(f(x_0-y)-f(x_0^+))\frac{\sin((N+1/2)y)}{\sin(y/2)}\,dy\\
&= \frac{1}{2\pi}\int_{[0,\pi]}(f(x_0+y)-f(x_0^+))\frac{\sin((N+1/2)y)}{\sin(y/2)}\,dy\\
&= \frac{1}{2\pi}\int_{[0,\pi]}(f(x_0+y)-f(x_0^+))\left(\cos(Ny)+\frac{\cos(y/2)\sin(Ny)}{\sin(y/2)}\right)\,dy\\
&= \frac{1}{2\pi}\int_{[0,\pi]}(f(x_0+y)-f(x_0^+))\cos(Ny)\,dy+\frac{1}{2\pi}\int_{[0,\pi]}\left(\frac{f(x_0+y)-f(x_0^+)}{y}\right)\left(2(y/2)\frac{\cos(y/2)}{\sin(y/2)}\right)\sin(Ny)\,dy\\
&= \frac{1}{2\pi}\int_{[0,\pi]}g(y)\cos(Ny)\,dy+\frac{1}{2\pi}\int_{[0,\pi]}h(y)\sin(Ny)\,dy
\end{aligned}$$

where

$$g(y)=f(x_0+y)-f(x_0^+)$$

and

$$h(y)=\left(\frac{f(x_0+y)-f(x_0^+)}{y}\right)\left(2(y/2)\frac{\cos(y/2)}{\sin(y/2)}\right).$$

It is clear that $g\in R(\mathbb{T})$ and thus, by Corollary 47,

$$\lim_{N\to\infty}\frac{1}{2\pi}\int_{[0,\pi]}g(y)\cos(Ny)\,dy=0.$$

Making the same conclusion concerning h isn't so straightforward. First, given our hypotheses concerning f, it is clear that h is piecewise continuous (continuous except at a finite number of points) on the interval $(0,\pi]$ and so it is Riemann-integrable on every compact subinterval of $(0,\pi]$. We must examine h near $y=0$ for the only possible impediment for Riemann integrability on $[0,\pi]$ is the behavior of h as $y\to 0$. First,

$$\lim_{y\to 0;y>0}2(y/2)\frac{\cos(y/2)}{\sin(y/2)}=\lim_{y\to 0;y>0}2\cos(y/2)\frac{(y/2)}{\sin(y/2)}=2.$$

Secondly, because f is piecewise differentiable on $[-\pi,\pi]$,

$$\lim_{y\to 0:y>0}\frac{f(x_0+y)-f(x_0^+)}{y}=\lim_{y\to 0:y>0}f'(x_0+y)=f'(x_0^+).$$

We remark that this is really a statement about exchanging limits and derivatives and its validity is far from obvious. A rigorous proof of this limit (which you should attempt), can be seen as an application of Theorem 13. Putting these two result together shows that

$$\lim_{h\to 0;h>0}h(y)=2f'(x_0^+).$$

In particular, h is bounded (and well-behaved) at 0 and so it follows that h is Riemann integrable on $[0,\pi]$. By an application of Corollary 47, we conclude that

$$\lim_{N\to\infty}\frac{1}{2\pi}\int_{[0,\pi]}h(y)\sin(Ny)\,dy=0.$$

Consequently,

$$\lim_{N\to\infty}I_+(N)=0.$$

By making completely analagous reasoning, it follows that

$$\lim_{N\to\infty}I_-(N)=0$$

and therefore

$$\lim_{N\to\infty}\left(S_N(x_0) - \frac{f(x_0^+)+f(x_0^-)}{2}\right) = \lim_{N\to\infty}\left(I_+(N)+I_-(N)\right) = 0.$$

□

Theorem 57 is the strongest result about the convergence of Fourier series we will prove in this course. I hope that you find the result satisfying, it encompasses most of the functions that you know about and can write down. The Nobel prize-winning physicist, Richard Feynman, was quite happy with this results (and ones like it) when made the following statement: “The mathematicians have shown, for a wide class of functions, in fact for all that are of interest to physicists, that if we can do the integrals we will get back $f(t)$.” He made this statement in the 1960’s in his famous lecture series at Caltech, just before Lennart Carleson (UCLA Professor Emeritus) completely solved the problem and determined the exact class of functions representable by their Fourier series [2, 3]. Carleson’s result, now known as Carleson’s theorem, was a long standing conjecture known as Luzin’s conjecture. For your cultural benefit, I will state it here; I’ll first need to make a definition.

Definition 58. *For any interval $I = [a, b] \subseteq \mathbb{R}$, we define*

$$\ell(I) = b - a$$

to be the length of I. Now, for any subset E of $\mathbb{R}$, we say that E is a set of measure zero (or a null set) if, for every $\epsilon > 0$, there is an infinite collection of intervals $\{I_n\}$ such that

$$E \subseteq \bigcup_{n=1}^{\infty} I_n$$

and

$$\sum_{n=1}^{\infty} \ell(I_n) < \epsilon.$$

You should think of a set of measure zero as an extremely small set. For instance, any finite (or countably infinite) collection of points is of measure zero. Using this notion of small sets we can state Carleson’s theorem in the context of Riemann integrable functions; the general result is formulated using the Lebesgue integral [1]. Here it is:

Theorem 59 (Carleson 1966)**.** *For any $f \in R(\mathbb{T})$, there exists a set E of measure zero such that*

$$f(x) = \sum_{n\in\mathbb{Z}} \hat{f}(n)e^{inx} = \lim_{N\to\infty} S_N(x)$$

for all $x \notin E$.

9.3 The Gibb’s phenomenon

In this final subsection, we will study the Gibb’s phenomenon. The Gibb’s phenomenon, named after J. Willard Gibb’s (yes, the free energy Gibbs), describes the pointwise convergence of Fourier series of a function with a jump discontinuity. Instead of working in the general setting, we will study the Gibb’s phenomenon as it occurs when we consider the Fourier series of the sawtooth function. Focusing on this specific case will allow us to very precisely see what’s going on. If you are worried about the general case, I’ll refer you to a very nice discussion by T. W. Körner in which he describes how to extend the results pertaining to this example to the general class of piecewise differentiable functions in $R(\mathbb{T})$ [3].

So let’s return to our favorite example, the sawtooth function f, defined by

$$f(x) = x$$

for all $-\pi < x \leq \pi$ and extended periodically to $\mathbb{R}$. We recall that

$$f(x) \sim \sum_{n\in\mathbb{Z}\setminus\{0\}} \frac{(-1)^n}{n} e^{inx} = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \sin(nx) \tag{9}$$

Let's again consider the graph of the Fourier polynomial S_{40}; this is illustrated in Figure 6.

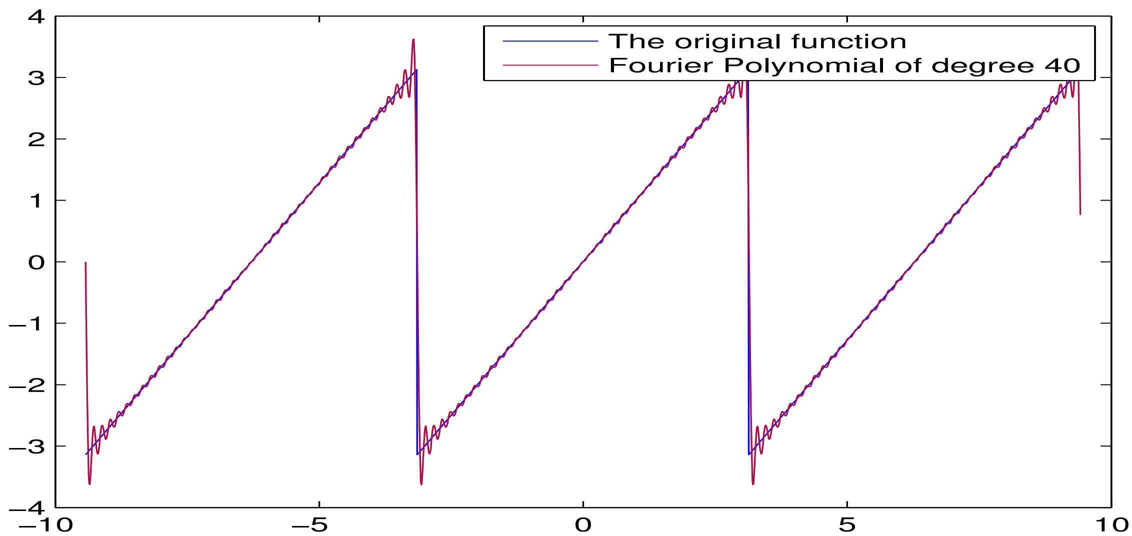

Figure 6: f and S_{40}

As we discussed in the last subsection, Theorem 57 guarantees that

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

for all $x \neq m\pi$ where $m \in \mathbb{Z}$ is odd. We also showed that, at any $x = m\pi$ where $m \in \mathbb{Z}$ is odd, $S_N(x) \to 0$ as $N \to \infty$. These two things should appear to be somewhat clear by looking at Figure 6. There is however one thing that should bother you: near (but not at) the breakpoints, $\{\dots, -3\pi, -\pi, \pi, 3\pi, \dots\}$, the graph of S_{40} seems to "overshoot" (or "undershoot") the graph of f. These are the spikes you see close to the discontinuity in f. This behavior is called the *Gibb's phenomenon.* Let's study this behavior more closely at, say, $x = \pi$; Figure 7 shows the graphs of $S_N(x)$ for $N = 25, 26, \dots, 50$.

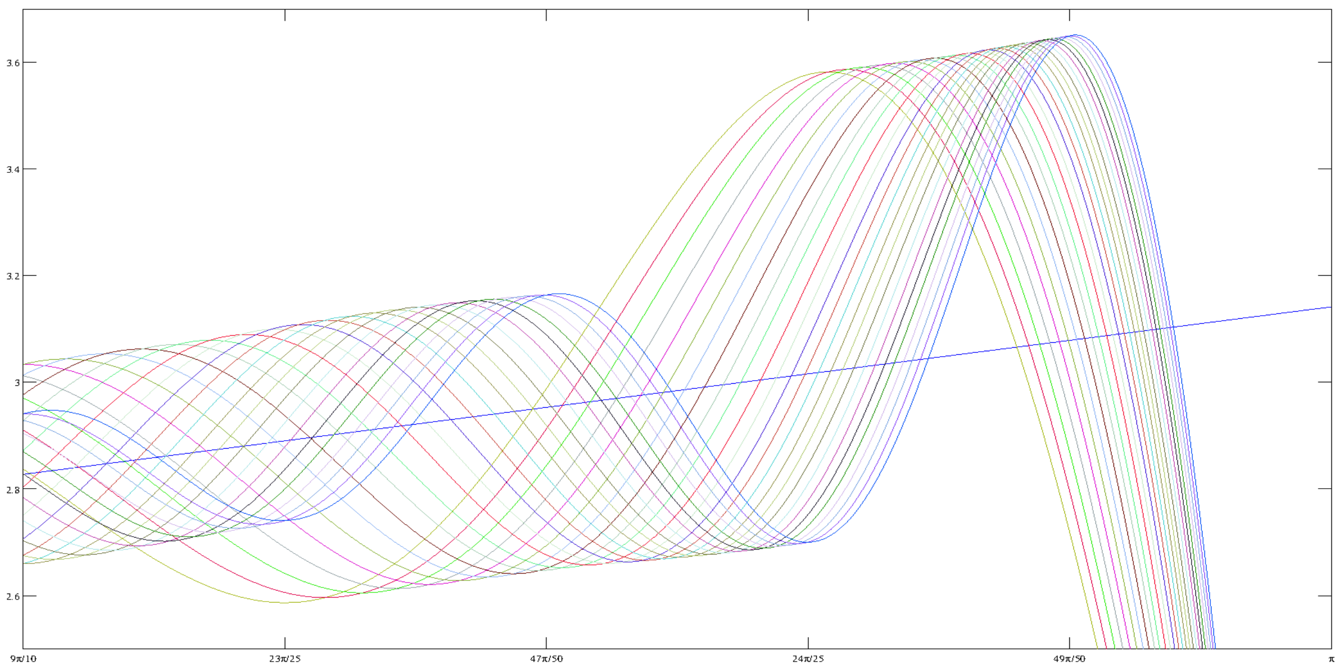

Figure 7: The graphs of $S_n(x)$ for $n = 25, 26, \dots 50$ and $f(x)$ for $9\pi/10 \le x \le \pi$.

Upon studying Figure 7 closely, we see that the overshoot is moving right as N increases. You might say: Theorem 57 guarantees that $S_N(x) \to f(x) = x$ for all $-\pi < x < \pi$ but, upon looking at the figure, $S_N(x)$ isn't converging to $f(x)$ for x very close to π. So where did we go wrong? The answer is that we haven't gone wrong at all, the apparent discrepancy can be understood by recognizing that pointwise convergence is weaker than convergence in the graph–this is the difference between pointwise convergence and uniform convergence. Remember, that for pointwise convergence, we first select x and ϵ and find a natural number $M = M(\epsilon, x)$ for which

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

for all $N \ge M$. In the case at hand, we can understand this notion as follows: If I select an $x < \pi$, but as close to π as I want, since the overshoot in the Fourier polynomials are moving to the right, I simply have to wait until they have moved so far right that they've passed x–this will determine N. After this, the Fourier polynomials evaluated at x will get much much closer to $f(x)$. Okay, so now you understand how we still get pointwise convergence. Let's now try to understand the overshoot.

Using the same numbers I've used to make the graphs in Figure 7, I can quantify this overshoot. For each $N \in \mathbb{N}$, denote by

$$M_N = \max_{9\pi/10 \le x\pi} S_N(x),$$

the maximum of the function $S_N(x)$ near π. We also denote by x_N the unique x near π for which

$$f(x_N) = M_N.$$

The following table shows M_N, M_N/π, x_N and $\pi - \pi/N$ to four decimal places for $n = 25, 30, \dots, 50$.

N	M_N	M_N/π	x_N	$\pi - \pi/N$
25	3.5822	1.403	3.0204	3.0159
30	3.6020	1.1465	3.0404	3.0369
35	3.6172	1.1514	3.0544	3.0518
40	3.6321	1.1561	3.0654	3.0631
45	3.6433	1.1597	3.0734	3.0718
50	3.6516	1.1624	3.0804	3.0788

Upon looking at the table, we see that, as N increases the x_N's are close to $\pi - \pi/N$ and the ratio M_N/π grows toward $1.17\dots$. So, by following the x at which $S_N(x)$ is maximized, the ratio M_N/π approaches some number A as $N \to \infty$ (note that π is half the gap of the discontinuity of f at π); this describes the overshoot. The following theorem formalizes it:

Theorem 60. *Let f, S_N be as above. Then*

$$\lim_{N\to\infty} S_N(\pi - \pi/N) = \pi A$$

where

$$A = 1.178979744447216727\dots$$

Thus, the $S_N(\pi-\pi/N)/\pi$ converges to A (called the Gibb's constant) times half of the gap of the jump discontinuity.

Proof. Using our trigonometric identities, we find that

$$\begin{aligned}
S_N(\pi - \pi/N) &= \sum_{n=1}^{N} \frac{2(-1)^{n+1}}{n} \sin(n\pi - n\pi/N) \\
&= \sum_{n=1}^{N} \frac{2(-1)^{n+1}}{n} (\sin(nx)\cos(n\pi/N) - \sin(n\pi/N)\cos(n\pi) \\
&= \sum_{n=1}^{N} \frac{2(-1)^{n+1}}{n} (0 - \sin(n\pi/N)\cos(n\pi) \\
&= \sum_{n=1}^{N} \frac{2(-1)^{n+1}}{n} ((-1)^{n+1}\sin(n\pi/N) \\
&= 2\sum_{n=1}^{N} \frac{\sin n\pi/N}{n\pi/N}\frac{\pi}{N}.
\end{aligned}$$

You should recognize that

$$\sum_{n=1}^{N} \frac{\sin n\pi/N}{n\pi/N}\frac{\pi}{N}$$

is a (right) Riemann sum for the integral

$$\int_0^\pi \frac{\sin(x)}{x}\,dx$$

and because $\sin x/x$ is Riemann integrable on $[0,\pi]$, we immediately conclude that

$$\lim_{N\to\infty} S_N(\pi - \pi/N) = \lim_{N\to\infty} 2\sum_{n=1}^{N} \frac{\sin n\pi/N}{n\pi/N}\frac{\pi}{N} = 2\int_0^\pi \frac{\sin x}{x}\,dx = \pi A$$

where

$$A = \frac{2}{\pi}\int_0^\pi \frac{\sin x}{x}\,dx.$$

It remains to compute A. We may easily compute the power series for $\sin x/x$ about 0:

$$\frac{\sin x}{x} = \frac{1}{x}\left(x - \frac{x^3}{3!} + \frac{x^5}{5!}\cdots\right) = 1 - \frac{x^2}{3!} + \frac{x^4}{5!} + \cdots.$$

It is easily verified that the radius of convergence of this series is $R = \infty$. By an application of Theorem 22, we conclude that

$$\begin{aligned} A &= \frac{2}{\pi}\int_0^\pi (1 - \frac{x^2}{3!} + \frac{x^4}{5!} + \cdots)\,dx = \frac{2}{\pi}\left(\int_0^\pi 1\,dx - \frac{1}{3!}\int_0^\pi x^2\,dx + \frac{1}{5!}\int_0^\pi x^4\,dx\right) \\ &= \frac{2}{\pi}\left(\pi - \frac{\pi^3}{3\cdot 3!} + \frac{\pi^5}{5\cdot 5!} + \cdots\right) = 2 - \frac{2\pi^2}{3\cdot 3!} + \frac{2\pi^4}{5\cdot 5!}\cdots \\ &= 1.178979744447216727\ldots] \end{aligned}$$

as desired. □

References

[1] Lennart Carleson. On convergence and growth of partial sums of Fourier series. *Acta Math.*, 116(1):135–157, dec 1966.

[2] Richard Phillips Feynman, Robert B Leighton, and Matthew L Sands. *The Feynman Lectures on Physics*, volume 1. 1963.

[3] Thomas William Körner. *Fourier Analysis.* Cambridge University Press, Cambridge, 1st edition, 1988.

[4] Serge Lang. *Analysis I.* Addison-Wesley Publishing Company, 1968.

[5] Walter Rudin. *Principles of Mathematical Analysis.* McGraw-Hill, Inc., 3rd edition, 1976.