MATH 3001
Real Analysis II
Series of Functions

Paul Skoufranis

April 5, 2022

Preface:

These are the first edition of these lecture notes for MATH 3001 (Real Analysis II). Consequently, there may be several typographical errors, missing exposition on necessary background, and more advance topics for which there will not be time in class to cover. Future iterations of these notes will hopefully be fairly self-contained provided one has the necessary background. If you come across any typos, errors, omissions, or unclear explanations, please feel free to contact me so that I may continually improve these notes.

Contents

Motivation for this Course

Analysis is the mathematics of approximation. The most simple object in analysis, the limit, says roughly that a sequence $(a_n)_{n\geq 1}$ converges to L provided if you go out far enough (i.e. $n \geq N$), then a_n is approximately L (i.e. $|a_n - L| < \epsilon$ for a fixed threshold $\epsilon > 0$). The idea of using ϵ-N and ϵ-δ to approximate objects as seen in previous courses may at first seem abstract, but are simple once one gets used to them and can be used to understand a plethora of analytical results.

Analysis is in stark contrast with other subjects, such as algebra. In algebra, one can only perform an operation a finite number of times. For example, one can only add a finite number of vectors in a vector space. However, with analysis, one can 'add' an infinite number of vectors by taking a limit of all of the finite sums. This leads to the concept of series that extends beyond real numbers to all vectors in vector spaces and thus, most importantly, continuous functions.

Series are incredibly useful not only for their abilities to add an infinite number of elements, but for their ability to approximate solutions. For example, to understand the behaviour of a vibrating string, the diffusion of heat, or quantum physics, one must understand the solutions to various differential equations. For quantum physics, the basic differential equation governing the behaviour of subatomic particles is the famous Schrödiner's equation. To understand the solutions to these differential equations, one must understand the basic solutions and then one can approximate all solutions via (potentially infinite) sums of these base solutions.

As an understanding of series of functions is vital for solving a vast number of problems in mathematics, physics, and beyond, this course will delve into an in-depth and rigorous study of series of functions. It is essential that the results stated in this course are precise and the proofs are complete as some of the biggest plunders in the history of mathematics came from mathematicians claiming a series of functions converged in a certain sense only for others to realizes decades or centuries later that this was not the case. Consequently, this course will focus on the various ways a series of functions can converge along with the pathological bad behaviours when series of functions converge in the incorrect way. Time permitting, we will look at some of the numerous applications of convergent series of functions.

Chapter 1

Series of Complex Numbers

Before we can delve into the study of series of functions and various ways these series can converge, it is important to first understand the behaviours and various ways series of scalars can converge. After all, a series of functions is a function of series of scalars. Thus thus chapter will thus focus on the various ways a series of scalars can converge, what can go right and wrong when one tries to rearrange the terms of a series, and how one can develop the exponential function.

As it is important for later chapters, we will not only deal with series of real numbers, but series of complex numbers. In particular, a reader that is unfamiliar or uncomfortable with complex numbers should refer to Appendix A for the basic and necessary facts.

1.1 Sequences of Complex Numbers

As a series will be a limit of finite sums, it is first useful to formally recall the definition of a limit. Since we will be dealing with complex numbers throughout the course, we will quickly generalize the basic properties of limits of sequences of real numbers seen in a previous course to the complex setting.

Definition 1.1.1. A sequence $(z_n)_{n\geq 1}$ of complex numbers is said to *converge to a complex number* L if for all $\epsilon > 0$ there exists an $N \in \mathbb{N}$ such that $|z_n - L| < \epsilon$ for all $n \geq N$. The complex number L is called a *limit of* $(z_n)_{n\geq 1}$ and is denoted by $\lim_{n\to\infty} z_n$.

The sequence $(z_n)_{n\geq 1}$ is said to *diverge* if $(z_n)_{n\geq 1}$ does not converge to any complex number.

Similarly to the definition for convergent sequences of real numbers, the 'less than ϵ" in Definition 1.1.1 can be replaced by "less than or equal to ϵ". As a reminder on how this works, consider the following.

Lemma 1.1.2. *Let $(z_n)_{n\geq 1}$ be a sequence of complex numbers and let $L \in \mathbb{C}$. Then $(z_n)_{n\geq 1}$ converges to L if and only if for all $\epsilon > 0$ there exists an $N \in \mathbb{N}$ such that $|z_n - L| \leq \epsilon$ for all $n \geq N$.*

Proof. Suppose $(z_n)_{n\geq 1}$ converges to L. To see the statement in the lemma is true, let $\epsilon > 0$ be arbitrary. Since $(z_n)_{n\geq 1}$ converges to L, Definition 1.1.1 implies there exists an $N \in \mathbb{N}$ such that $|z_n - L| < \epsilon$ for all $n \geq N$. As this implies $|z_n - L| \leq \epsilon$ for all $n \geq N$ and as $\epsilon > 0$ was arbitrary, one direction of the proof is complete.

For the other direction, assume for all $\epsilon > 0$ there exists an $N \in \mathbb{N}$ such that $|z_n - L| \leq \epsilon$ for all $n \geq N$. To see that $(z_n)_{n\geq 1}$ converges to L, let $\epsilon > 0$ be arbitrary. Let $\epsilon_0 = \frac{\epsilon}{2}$. Since $\epsilon_0 > 0$, the assumptions of this direction imply that there exists an $N \in \mathbb{N}$ such that $|z_n - L| \leq \epsilon_0$ for all $n \geq N$. Hence $|z_n - L| \leq \epsilon_0 < \epsilon$ for all $n \geq N$. As $\epsilon > 0$ was arbitrary, $(z_n)_{n\geq 1}$ converges to L by Definition 1.1.1. ∎

Remark 1.1.3. As the proof of Lemma 1.1.2 shows, when discussing convergence of sequences, one may always replace ϵ with any positive constant multiple of ϵ. Indeed the assumption "for all $\epsilon > 0$" is the same as "for all ϵ such that $k\epsilon > 0$" provided $k > 0$.

The convergence of a sequence of complex numbers is very similar to the convergence of a sequence of real numbers. In particular, the following shows that a sequence of complex numbers converges if and only if two sequences of real numbers converge!

Lemma 1.1.4. *Let $(z_n)_{n\geq 1}$ be a sequence of complex numbers and let $L \in \mathbb{C}$. Then $(z_n)_{n\geq 1}$ converges to L if and only if $(\mathrm{Re}(z_n))_{n\geq 1}$ and $(\mathrm{Im}(z_n))_{n\geq 1}$ converge to $\mathrm{Re}(L)$ and $\mathrm{Im}(L)$ respectively.*

Proof. Suppose $(z_n)_{n\geq 1}$ converges to L. To prove that $(\mathrm{Re}(z_n))_{n\geq 1}$ and $(\mathrm{Im}(z_n))_{n\geq 1}$ converge to $\mathrm{Re}(L)$ and $\mathrm{Im}(L)$ respectively, let $\epsilon > 0$ be arbitrary. Since $(z_n)_{n\geq 1}$ converges to L there exists an $N \in \mathbb{N}$ such that $|z_n - L| < \epsilon$ for all $n \geq N$. Hence for all $n \geq N$ we have that

$$|\mathrm{Re}(z_n) - \mathrm{Re}(L)| \leq |z_n - L| < \epsilon$$

and

$$|\mathrm{Im}(z_n) - \mathrm{Im}(L)| \leq |z_n - L| < \epsilon.$$

Therefore, as $\epsilon > 0$ was arbitrary, $(\mathrm{Re}(z_n))_{n\geq 1}$ and $(\mathrm{Im}(z_n))_{n\geq 1}$ converge to $\mathrm{Re}(L)$ and $\mathrm{Im}(L)$ respectively.

Conversely, suppose $(\mathrm{Re}(z_n))_{n\geq 1}$ and $(\mathrm{Im}(z_n))_{n\geq 1}$ converge to $\mathrm{Re}(L)$ and $\mathrm{Im}(L)$ respectively. To see that $(z_n)_{n\geq 1}$ converges to L, let $\epsilon > 0$ be arbitrary. Since $(\mathrm{Re}(z_n))_{n\geq 1}$ converges to $\mathrm{Re}(L)$, there exists an $N_1 \in \mathbb{N}$ such that

$$|\mathrm{Re}(z_n) - \mathrm{Re}(L)| < \frac{1}{\sqrt{2}}\epsilon$$

for all $n \geq N_1$. Similarly, since $(\text{Im}(z_n))_{n\geq 1}$ converges to $\text{Im}(L)$, there exists an $N_2 \in \mathbb{N}$ such that

$$|\text{Im}(z_n) - \text{Im}(L)| < \frac{1}{\sqrt{2}}\epsilon$$

for all $n \geq N_2$. Let $N = \max\{N_1, N_2\}$. Hence for all $n \geq N$, the above implies that

$$\begin{aligned}
|z_n - L| &= \sqrt{(\text{Re}(z_n) - \text{Re}(L))^2 + (\text{Im}(z_n) - \text{Im}(L))^2} \\
&< \sqrt{\left(\frac{1}{\sqrt{2}}\epsilon\right)^2 + \left(\frac{1}{\sqrt{2}}\epsilon\right)^2} \\
&= \sqrt{\frac{1}{2}\epsilon^2 + \frac{1}{2}\epsilon^2} \\
&= \sqrt{\epsilon^2} = \epsilon.
\end{aligned}$$

Therefore, as $\epsilon > 0$ was arbitrary, $(z_n)_{n\geq 1}$ converges to L. ∎

Lemma 1.1.4 allows one to quickly translate many results from convergent sequences of real numbers to the complex setting with ease.

Corollary 1.1.5. *Let $L, K \in \mathbb{C}$ and let $(z_n)_{n\geq 1}$ be a sequence of complex numbers. If L and K are limits of $(z_n)_{n\geq 1}$, then $L = K$.*

Proof. One quick way to prove this corollary is Lemma 1.1.4. Indeed if L and K are limits of $(z_n)_{n\geq 1}$, then Lemma 1.1.4 implies $\text{Re}(L)$ and $\text{Re}(K)$ are limits of $(\text{Re}(z_n))_{n\geq 1}$ and $\text{Im}(L)$ and $\text{Im}(K)$ are limits of $(\text{Im}(z_n))_{n\geq 1}$. Hence the corresponding result for sequences of real numbers implies that $\text{Re}(L) = \text{Re}(K)$ and $\text{Im}(L) = \text{Im}(K)$, so $L = K$ as desired.

Alternatively, it is not difficult to adapt the proof of the corresponding result for real sequences to complex sequences. Indeed the modified proof is below.

Suppose that $L \neq K$. Therefore, if

$$\epsilon = \frac{|L - K|}{2},$$

then $\epsilon > 0$. Since L is a limit of $(z_n)_{n\geq 1}$, Definition 1.1.1 implies that there exists an $N_1 \in \mathbb{N}$ such that $|z_n - L| < \epsilon$ for all $n \geq N_1$. Similarly, since K is a limit of $(z_n)_{n\geq 1}$, Definition 1.1.1 implies that there exists an $N_2 \in \mathbb{N}$ such that $|z_n - K| < \epsilon$ for all $n \geq N_2$.

Let $N = \max\{N_1, N_2\}$. Therefore, the above implies $|z_N - L| < \epsilon$ and $|z_N - K| < \epsilon$ so

$$|L - K| \leq |L - z_n| + |z_n - K| < \epsilon + \epsilon = 2\epsilon = |L - K|.$$

Since this is clearly a contradiction, it must be the case that $L = K$. ∎

Corollary 1.1.6. *If* $(z_n)_{n\geq 1}$ *is a convergent sequence of complex numbers, then* $(z_n)_{n\geq 1}$ *is bounded; that is,*

$$\sup(\{|z_n| \mid n \in \mathbb{N}\}) < \infty.$$

Proof. One quick way to prove this corollary is Lemma 1.1.4. Indeed if $(z_n)_{n\geq 1}$ converges to L, then Lemma 1.1.4 implies $(\mathrm{Re}(z_n))_{n\geq 1}$ and $(\mathrm{Im}(z_n))_{n\geq 1}$ converge to $\mathrm{Re}(L)$ and $\mathrm{Im}(L)$ respectively. Since convergent sequences of real numbers are bounded, there exists $M_1 > 0$ and $M_2 > 0$ such that $|\mathrm{Re}(z_n)| \leq M_1$ and $|\mathrm{Im}(z_n)| \leq M_2$ for all $n \in \mathbb{N}$. Hence for all $n \in \mathbb{N}$.

$$|z_n| = \sqrt{(\mathrm{Re}(z_n))^2 + (\mathrm{Im}(z_n))^2} \leq \sqrt{M_1^2 + M_2^2}.$$

Consequently, $(z_n)_{n\geq 1}$ is bounded.

Alternatively, it is not difficult to adapt the proof of the corresponding result for real sequences to complex sequences. Indeed the modified proof is below.

Suppose $(z_n)_{n\geq 1}$ converges to L. Let $\epsilon = 1$. By Definition 1.1.1, there exists an $N \in \mathbb{N}$ such that $|z_n - L| < \epsilon = 1$ for all $n \geq N$. Hence the reverse triangle inequality implies for all $n \geq N$ that

$$||z_n| - |L|| \leq |z_n - L| < 1$$

and thus $|z_n| < |L| + 1$ for all $n \geq N$.

Let

$$M = \max(\{|z_1|, |z_2|, \ldots, |z_{N-1}|, L + 1\}).$$

Clearly $M \in \mathbb{R}$ and $M \geq 0$. Moreover $|z_n| \leq M$ for all $n < N$ by construction. Furthermore, the above implies that $|z_n| < |L| + 1 \leq M$ for all $n \geq N$. Hence $|z_n| \leq M$ for all $n \in \mathbb{N}$. Therefore $(z_n)_{n\geq 1}$ is bounded. ∎

It is also possible to use Lemma 1.1.4 and the corresponding results for convergent sequences of real numbers to prove the following. However, proving the following result directly is both simpler and more instructive in ϵ-N arguments. Thus we only prove the following directly

Corollary 1.1.7. *Let* $L, K \in \mathbb{C}$ *and let* $(z_n)_{n\geq 1}$ *and* $(w_n)_{n\geq 1}$ *be sequences of complex numbers that converge to* L *and* K *respectively. Then the following are true:*

a) $(z_n + w_n)_{n\geq 1}$ *converges to* $L + K$.

b) $(z_n w_n)_{n\geq 1}$ *converges to* LK.

c) $(\alpha z_n)_{n\geq 1}$ *converges to* αL.

d) If $L \neq 0$ *and* $z_n \neq 0$ *for all* $n \in \mathbb{N}$, $\left(\frac{1}{z_n}\right)_{n\geq 1}$ *converges to* $\frac{1}{L}$.

e) $(\overline{z_n})_{n\geq 1}$ *converges to* $\overline{L}$.

f) $(|z_n|)_{n\geq 1}$ *converges to* $|L|$.

Proof. To see that part a) is true, let $\epsilon > 0$ be arbitrary. Since $(z_n)_{n\geq 1}$ converges to L, there exists an $N_1 \in \mathbb{N}$ such that $|z_n - L| < \frac{\epsilon}{2}$ for all $n \geq N_1$. Similarly, since $(w_n)_{n\geq 1}$ converges to K, there exists an $N_2 \in \mathbb{N}$ such that $|w_n - K| < \frac{\epsilon}{2}$ for all $n \geq N_2$. Let $N = \max\{N_1, N_2\}$. Thus for all $n \geq N$ we have that

$$|(z_n + w_n) - (L + K)| \leq |z_n - L| + |w_n - K| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Therefore, since $\epsilon > 0$ was arbitrary, $(z_n + w_n)_{n\geq 1}$ converges to $L + K$.

To see that part b) is true, let $\epsilon > 0$ be arbitrary. By Corollary 1.1.6 there exists an $M > 0$ such that $|z_n| \leq M$ for all $n \in \mathbb{N}$. Since $(z_n)_{n\geq 1}$ converges to L, there exists an $N_1 \in \mathbb{N}$ such that

$$|z_n - L| < \frac{\epsilon}{2(|K| + 1)}$$

for all $n \geq N_1$. Similarly, since $(w_n)_{n\geq 1}$ converges to K, there exists an $N_2 \in \mathbb{N}$ such that

$$|w_n - K| < \frac{\epsilon}{2(M + 1)}$$

for all $n \geq N_2$. Let $N = \max\{N_1, N_2\}$. Thus for all $n \geq N$ we have that

$$\begin{aligned}
|z_n w_n - LK| &= |z_n w_n - z_n K + z_n K - LK| \\
&\leq |z_n w_n - z_n K| + |z_n K - LK| \\
&= |z_n||w_n - K| + |z_n - L||K| \\
&\leq M\frac{\epsilon}{2(M+1)} + \frac{\epsilon}{2(|K|+1)}|K| \\
&< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon
\end{aligned}$$

Therefore, since $\epsilon > 0$ was arbitrary, $(z_n w_n)_{n\geq 1}$ converges to LK.

Next, note part c) follows directly from part b) using the constant sequence $(\alpha)_{n\geq 1}$ in place of $(w_n)_{n\geq 1}$.

To see that part d) is true, first note since $(z_n)_{n\geq 1}$ converges to L and since $L \neq 0$ and thus $|L| \neq 0$ that there exists an $N_1 \in \mathbb{N}$ such that $|z_n - L| < \frac{|L|}{2}$ for all $n \geq N_1$. Hence the reverse triangle inequality implies for all $n \geq N_1$ that

$$|L| - |z_n| < \frac{|L|}{2}$$

and thus $\frac{|L|}{2} < |z_n|$ for all $n \geq N_1$. Therefore $\frac{1}{|z_n|} \leq \frac{2}{|L|}$ for all $n \geq N_1$.

To see that $\left(\frac{1}{z_n}\right)_{n\geq 1}$ converges to $\frac{1}{L}$, let $\epsilon > 0$ be arbitrary. Since $|L| \neq 0$ and since $(z_n)_{n\geq 1}$ converges to L, there exists an $N_2 \in \mathbb{N}$ such that

$|z_n - L| < \frac{|L|^2}{2}\epsilon$ for all $n \geq N_2$. Let $N = \max\{N_1, N_2\}$. Hence for all $n \geq N$ we have that

$$\begin{aligned}\left|\frac{1}{z_n} - \frac{1}{L}\right| &= \left|\frac{L - z_n}{z_n L}\right| \\ &= |L - z_n| \frac{1}{|z_n|} \frac{1}{|L|} \\ &< \left(\frac{|L|^2}{2}\epsilon\right)\left(\frac{2}{|L|}\right)\frac{1}{|L|} = \epsilon.\end{aligned}$$

Therefore, as $\epsilon > 0$ was arbitrary, $\left(\frac{1}{z_n}\right)_{n\geq 1}$ converges to $\frac{1}{L}$.

To see that part e) is true, note by Lemma 1.1.4 that if $(z_n)_{n\geq 1}$ converges to L, then $(\mathrm{Re}(z_n))_{n\geq 1}$ and $(\mathrm{Im}(z_n))_{n\geq 1}$ converge to $\mathrm{Re}(L)$ and $\mathrm{Im}(L)$ respectively. Hence part c) implies that $(-\mathrm{Im}(z_n))_{n\geq 1}$ converge to $-\mathrm{Im}(L)$. Hence, by the definition of the complex conjugate, Lemma 1.1.4 implies $(\overline{z_n})_{n\geq 1}$ converges to $\overline{L}$ as desired.

Finally, to see that part f) is true, let $\epsilon > 0$ be arbitrary. Since $(z_n)_{n\geq 1}$ converges to L, there exists an $N \in \mathbb{N}$ such that $|z_n - L| < \epsilon$ for all $n \geq N$. Therefore for all $n \geq N$ we have by the reverse triangle inequality that

$$||z_n| - |L|| \leq |z_n - L| < \epsilon.$$

Therefore, as $\epsilon > 0$ was arbitrary, $(|z_n|)_{n\geq 1}$ converges to $|L|$. ∎

One difficulty in verifying that a sequence of complex numbers converges is that one must first guess the limit and then prove the limit is indeed the limit. As with sequences of real numbers, there is an alternative property of a sequence that is easier to check and is equivalent to convergence.

Definition 1.1.8. A sequence $(z_n)_{n\geq 1}$ of complex numbers is said to be *Cauchy* if for all $\epsilon > 0$ there exists an $N \in \mathbb{N}$ such that $|z_n - z_m| < \epsilon$ for all $n, m \geq N$.

Of course, by the same arguments as used in Lemma 1.1.2, one can easily replace the "less than ϵ" with "less than or equal to ϵ" in Definition 1.1.8.

As with the real numbers, the complex numbers are *complete*:

Theorem 1.1.9. *A sequence of complex numbers converges if and only if it is Cauchy.*

Proof. Let $(z_n)_{n\geq 1}$ be a sequence of complex numbers. Suppose $(z_n)_{n\geq 1}$ converges to a complex number L. To see that $(z_n)_{n\geq 1}$ is Cauchy, let $\epsilon > 0$ be arbitrary. Since $(z_n)_{n\geq 1}$ converges to L, there exists an $N \in \mathbb{N}$ such that $|z_n - L| < \frac{\epsilon}{2}$ for all $n \geq N$. Therefore, for all $n, m \geq N$ we have that

$$|z_n - z_m| = |z_n - L + L - z_m| \leq |z_n - L| + |L - z_m| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Therefore, as $\epsilon > 0$ was arbitrary, $(z_n)_{n\geq 1}$ is Cauchy.

Conversely, suppose $(z_n)_{n\geq 1}$ is Cauchy. To see that $(z_n)_{n\geq 1}$ converges, we first claim that $(\mathrm{Re}(z_n))_{n\geq 1}$ and $(\mathrm{Im}(z_n))_{n\geq 1}$ are Cauchy sequences of real numbers. To see this, let $\epsilon > 0$ be arbitrary. Since $(z_n)_{n\geq 1}$ is Cauchy, there exists an $N \in \mathbb{N}$ such that $|z_n - z_m| < \epsilon$ for all $n, m \geq N$. Thus for all $n, m \geq N$, we have that

$$|\mathrm{Re}(z_n) - \mathrm{Re}(z_m)| = |\mathrm{Re}(z_n - z_m)| \leq |z_n - z_m| < \epsilon$$

and

$$|\mathrm{Im}(z_n) - \mathrm{Im}(z_m)| = |\mathrm{Im}(z_n - z_m)| \leq |z_n - z_m| < \epsilon.$$

Therefore, as $\epsilon > 0$ was arbitrary, $(\mathrm{Re}(z_n))_{n\geq 1}$ and $(\mathrm{Im}(z_n))_{n\geq 1}$ are Cauchy sequences of real numbers.

Since every Cauchy sequence of real numbers converges, there exists $a, b \in \mathbb{R}$ such that $(\mathrm{Re}(z_n))_{n\geq 1}$ and $(\mathrm{Im}(z_n))_{n\geq 1}$ converge to a and b respectively. Hence $(z_n)_{n\geq 1}$ converges to $a + bi$ by Lemma 1.1.4. ■

1.2 Convergence of Series

With our understanding of convergent sequences of complex numbers, we can turn our attention to convergent series of complex numbers. In particular, a convergent series of complex numbers is just a particular form of convergent sequence of complex numbers.

Definition 1.2.1. Let $(z_n)_{n\geq 1}$ be a sequence of complex numbers and for each $N \in \mathbb{N}$ let $S_N = \sum_{k=1}^{N} z_k$. The series $\sum_{n=1}^{\infty} z_n$ is said to *converge to* $L \in \mathbb{C}$, denoted $\sum_{n=1}^{\infty} z_n = L$, if the sequence $(S_N)_{N\geq 1}$ converges to L. The term S_N is called the N^{th} *partial sum of the series*.

The series $\sum_{n=1}^{\infty} z_n$ is said to *diverge* if the sequence $(S_N)_{N\geq 1}$ diverges.

Remark 1.2.2. When trying to determine whether or not a series converges, the size of the "first few" terms of the series do not matter. To be specific, $\sum_{n=1}^{\infty} z_n$ converges if and only if $\sum_{n=K}^{\infty} z_n$ converges for some $K > 1$ in which case $\sum_{n=1}^{\infty} z_n = \sum_{n=1}^{K-1} z_n + \sum_{n=K}^{\infty} z_n$. To see this, note the partial sums of the first series are partial sums for the second series plus $\sum_{n=1}^{K-1} z_n$ and adding a constant to a sequence does not affect whether or not the sequence converges. We call $\sum_{n=K}^{\infty} z_n$ a *tail of the series* so, when using this fact in the future, we will state "only the tail of the series matters".

It is useful to begin our study of convergent series of complex numbers with a very important exampled. Throughout these notes, the convention $z^0 = 1$ for all $z \in \mathbb{C}$ will be used.

Example 1.2.3. Let $z \in \mathbb{C}$ be such that $|z| < 1$. We claim that

$$\sum_{n=0}^{\infty} z^n = \frac{1}{1-z}.$$

To see this, let $N \in \mathbb{N}$ be arbitrary. Notice that the N^{th} partial sum is

$$S_N = \sum_{k=0}^{N} z^k$$

so

$$zS_N = \sum_{k=1}^{N+1} z^k.$$

Hence $zS_N - S_N = z^{N+1} - 1$ so

$$S_N = \frac{z^{N+1} - 1}{z - 1}$$

since $z \neq 1$. Therefore, since $\left|z^{N+1}\right| = |z|^{N+1}$ is easily seen to converge to 0 as N tends to infinity as $|z| < 1$, we see that $\lim_{N\to\infty} z^{N+1} = 0$ so

$$\lim_{N\to\infty} S_N = \lim_{N\to\infty} \frac{z^{N+1} - 1}{z - 1} = \frac{0 - 1}{z - 1} = \frac{1}{1 - z}.$$

Thus $\sum_{n=0}^{\infty} z^n = \frac{1}{1-z}$. This series is called a *geometric series*.

Unsurprisingly, convergent series of complex numbers behave well in regards to addition and scalar multiplication as convergent sequences behave well due to Lemma 1.1.7.

Lemma 1.2.4. *Let $\sum_{n=1}^{\infty} z_n$ and $\sum_{n=1}^{\infty} w_n$ be convergent series of complex numbers. Then $\sum_{n=1}^{\infty} z_n + w_n$ converges and*

$$\sum_{n=1}^{\infty} z_n + w_n = \sum_{n=1}^{\infty} z_n + \sum_{n=1}^{\infty} w_n.$$

Moreover, for all $\alpha \in \mathbb{C}$, the series $\sum_{n=1}^{\infty} \alpha z_n$ converges and

$$\sum_{n=1}^{\infty} \alpha z_n = \alpha \sum_{n=1}^{\infty} z_n.$$

Proof. For all $N \in \mathbb{N}$, consider the N^{th} partial sums

$$S_N = \sum_{k=1}^{N} z_k, \qquad T_N = \sum_{k=1}^{N} w_k,$$

$$R_N = \sum_{k=1}^{N} z_k + w_k, \qquad \text{and} \qquad U_N = \sum_{k=1}^{N} \alpha z_k.$$

Since $\sum_{n=1}^{\infty} z_n$ and $\sum_{n=1}^{\infty} w_n$ converge, we know that $(S_N)_{N\geq 1}$ and $(T_N)_{N\geq 1}$ converge and

$$\lim_{N\to\infty} S_N = \sum_{n=1}^{\infty} z_n \quad \text{and} \quad \lim_{N\to\infty} T_n = \sum_{n=1}^{\infty} w_n.$$

Since $R_N = S_N + T_N$ and $U_n = \alpha S_N$, Lemma 1.1.7 implies $(R_N)_{N\geq 1}$ and $(U_N)_{N\geq 1}$ converge and

$$\lim_{N\to\infty} R_N = \lim_{N\to\infty} S_N + \lim_{N\to\infty} T_N \text{ and}$$
$$\lim_{N\to\infty} U_N = \alpha \lim_{N\to\infty} S_N.$$

Hence $\sum_{n=1}^{\infty} z_n + w_n$ and $\sum_{n=1}^{\infty} \alpha z_n$ converge and

$$\sum_{n=1}^{\infty} z_n + w_n = \sum_{n=1}^{\infty} z_n + \sum_{n=1}^{\infty} w_n \text{ and}$$
$$\sum_{n=1}^{\infty} \alpha z_n = \alpha \sum_{n=1}^{\infty} z_n$$

as desired. ∎

Like with convergent sequences of complex numbers, one can test whether a series of complex numbers converges via an alternate criterion that bypasses the need for determining and verifying a specific limit.

Theorem 1.2.5 (Cauchy Criterion). *A series $\sum_{n=1}^{\infty} z_n$ converges if and only if for all $\epsilon > 0$ there exists an $N \in \mathbb{N}$ such that $|\sum_{k=N}^{m} z_k| < \epsilon$ for all $m \geq N$.*

Proof. For each $N \in \mathbb{N}$, let $S_N = \sum_{k=1}^{N} z_k$.

Suppose $\sum_{n=1}^{\infty} z_n$ converges. To see the desired statement holds, let $\epsilon > 0$ be arbitrary. Since $\sum_{n=1}^{\infty} z_n$ converges, $(S_N)_{N\geq 1}$ converges and thus is Cauchy by Theorem 1.1.9. Hence there exists an $N_0 \in \mathbb{N}$ such that $|S_m - S_k| < \epsilon$ for all $m, k \geq N_0$. Therefore, if $N = N_0 + 1$, we see for all $m \geq N > N_0$ that

$$\left|\sum_{k=N}^{m} z_k\right| = \left|\sum_{k=1}^{m} z_k - \sum_{k=1}^{N-1} z_k\right| = |S_m - S_{N-1}| < \epsilon.$$

Therefore, since $\epsilon > 0$ was arbitrary, the desired statement from the theorem holds.

Conversely, suppose for all $\epsilon > 0$ there exists an $N \in \mathbb{N}$ such that $|\sum_{k=N}^{m} z_k| < \epsilon$ for all $m \geq N$. We claim that $(S_N)_{N\geq 1}$ is Cauchy. To see that $(S_N)_{N\geq 1}$ is Cauchy, let $\epsilon > 0$ be arbitrary. By the assumption of this direction of the proof, there exists an $N \in \mathbb{N}$ such that $|\sum_{k=N}^{m} z_k| < \frac{\epsilon}{2}$ for all

$m \geq N$. Hence, for all $n \geq m \geq N$, we see that

$$\begin{aligned} |S_n - S_m| &= \left| \sum_{k=1}^{n} z_k - \sum_{k=1}^{m} z_k \right| \\ &= \left| \sum_{k=N}^{n} z_k - \sum_{k=N}^{m} z_k \right| && \text{cancel the first } (N-1) \text{ terms} \\ &\leq \left| \sum_{k=N}^{n} z_k \right| + \left| \sum_{k=N}^{m} z_k \right| \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \end{aligned}$$

Therefore, as $\epsilon > 0$ was arbitrary, $(S_N)_{N\geq 1}$ is Cauchy.

Since $(S_N)_{N\geq 1}$ is Cauchy, $(S_N)_{N\geq 1}$ converges by Theorem 1.1.9. Hence $\sum_{n=1}^{\infty} z_n$ converges by definition. ∎

Immediately the Cauchy Criterion (Theorem 1.2.5) can be used to show a required property for series to converge.

Corollary 1.2.6. *If a series $\sum_{n=1}^{\infty} z_n$ of complex numbers converge, then* $\lim_{n\to\infty} z_n = 0$.

Proof. Let $\epsilon > 0$ be arbitrary. By the Cauchy Criterion (Theorem 1.2.5) there exists an $N \in \mathbb{N}$ such that

$$\left| \sum_{k=N}^{m} z_k \right| < \frac{\epsilon}{2}$$

for all $m \geq N$. Therefore, for all $n \geq N+1$ we have $n, n-1 \geq N$ so that

$$|z_n| = \left| \sum_{k=N}^{n} z_k - \sum_{k=N}^{n-1} z_k \right| \leq \left| \sum_{k=N}^{n} z_k \right| + \left| \sum_{k=N}^{n-1} z_k \right| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Therefore, as $\epsilon > 0$ was arbitrary, $\lim_{n\to\infty} z_n = 0$. ∎

Remark 1.2.7. It is important to point out that the converse of Corollary 1.2.6 is false; that is, there exist sequences $(z_n)_{n\geq 1}$ of complex numbers such that $\lim_{n\to\infty} z_n = 0$ but $\sum_{n=1}^{\infty} z_n$ does not converge. Some examples of this will be pointed out shortly (see Corollary 1.2.15).

Using Corollary 1.2.6, we can complete Example 1.2.3.

Example 1.2.8. Let $z \in \mathbb{C}$ be such that $|z| \geq 1$. We claim that the geometric series $\sum_{n=0}^{\infty} z^n$ does not converge. To see this, note that $|z^n - 0| = |z|^n$ for all $n \in \mathbb{N}$ so either $\lim_{n\to\infty} |z^n - 0|$ diverges to infinity or is equal to 1. In either case, $(z^n)_{n\geq 1}$ does not converge to 0 as n tends to infinity, so Corollary 1.2.6 implies $\sum_{n=0}^{\infty} z^n$ cannot converge.

Although the Cauchy Criterion (Theorem 1.2.5) helps us with a theoretical tool for determining when a series of complex numbers converges, it is still quite difficult to use. To aid us in our ability to determine whether or not a series converges, let us deal with the specific case of series of non-negative real numbers.

Theorem 1.2.9. *Let $(a_n)_{n\geq 1}$ be a sequence of real numbers with $a_n \geq 0$ for all $n \in \mathbb{N}$. The series $\sum_{n=1}^{\infty} a_n$ converges if and only if there exists an $M \in \mathbb{R}$ such that $\sum_{k=1}^{N} a_k \leq M$ for all $N \in \mathbb{N}$. Moreover, if $\sum_{k=1}^{N} a_k \leq M$ for all $N \in \mathbb{N}$, then $\sum_{n=1}^{\infty} a_n \leq M$.*

Finally, if $\sum_{n=1}^{\infty} a_n$ converges, then for all $\epsilon > 0$ there exists a $N_0 \in \mathbb{N}$ such that $\sum_{k=N}^{\infty} a_k < \epsilon$ for all $N > N_0$.

Proof. For every $N \in \mathbb{N}$, let $S_N = \sum_{k=1}^{N} a_k$. Since $a_{N+1} \geq 0$ for all $N \in \mathbb{N}$, we see that

$$S_{N+1} = \sum_{k=1}^{N+1} a_k = a_{N+1} + \sum_{k=1}^{N} a_k = a_{N+1} + S_N \geq S_N.$$

Hence $(S_N)_{N\geq 1}$ is a non-decreasing sequence of real numbers. Therefore, the Monotone Convergence Theorem implies $(S_N)_{N\geq 1}$ converges if and only if $(S_N)_{N\geq 1}$ is bounded. Moreover, as any upper bound of $(S_N)_{N\geq 1}$ must be greater than or equal to $\lim_{N\to\infty} S_N$, the first part of the statement is true.

To see the second part of the statement is true, let $\epsilon > 0$ be arbitrary. By the Cauchy Criterion (Theorem 1.2.5), there exists an $N_0 \in \mathbb{N}$ such that

$$\sum_{k=N_0}^{m} a_k = \left| \sum_{k=N_0}^{m} a_k \right| < \epsilon.$$

Hence, by taking the limit as m tends to infinity, we obtain that $\sum_{k=N_0}^{\infty} a_k < \epsilon$. Since $\sum_{k=N}^{\infty} a_k$ converges as only the tail matters, and since clearly the partial sums of $\sum_{k=N}^{\infty} a_k$ are bounded above by the partial sums of $\sum_{k=N_0}^{\infty} a_k$ for all $N \geq N_0$, the result follows. ∎

Using the idea that it should be easier to determine when series of non-negative real numbers converge, we consider the following idea of converting a series of complex numbers into a series of non-negative real numbers.

Definition 1.2.10. A series $\sum_{n=1}^{\infty} z_n$ of complex numbers is said to converge *absolutely* if $\sum_{n=1}^{\infty} |z_n|$ converges.

Clearly a series of non-negative real numbers converges if and only if the series converges absolutely by definition. Fortunately, we are in luck as the important direction holds for series of complex numbers thereby allowing us an alternate way to verify certain series of complex numbers converge.

Theorem 1.2.11. *If $\sum_{n=1}^{\infty} z_n$ is an absolutely convergent series of complex numbers, then $\sum_{n=1}^{\infty} z_n$ converges. Moreover*

$$\left|\sum_{n=1}^{\infty} z_n\right| \leq \sum_{n=1}^{\infty} |z_n|.$$

Proof. To see that $\sum_{n=1}^{\infty} z_n$ converges, let $\epsilon > 0$ be arbitrary. Since $\sum_{n=1}^{\infty} z_n$ converges absolutely, we know that $\sum_{n=1}^{\infty} |z_n|$ converges. Hence the Cauchy Criterion (Theorem 1.2.5) implies there exists an $N \in \mathbb{N}$ such that

$$\left|\sum_{k=N}^{m} |z_k|\right| < \epsilon$$

for all $m \geq N$. Hence for all $m \geq N$ we have that

$$\left|\sum_{k=N}^{m} z_k\right| \leq \sum_{k=N}^{m} |z_k| = \left|\sum_{k=N}^{m} |z_k|\right| < \epsilon.$$

Therefore, as $\epsilon > 0$ was arbitrary, the Cauchy Criterion (Theorem 1.2.5) implies $\sum_{n=1}^{\infty} z_n$ converges.

Moreover, by Corollary 1.1.7 part (f), we see that

$$\left|\sum_{n=1}^{\infty} z_n\right| = \lim_{N\to\infty} \left|\sum_{k=1}^{N} z_k\right| \leq \liminf_{N\to\infty} \sum_{k=1}^{N} |z_k| = \sum_{n=1}^{\infty} |z_n|$$

as desired. ∎

Remark 1.2.12. It is important to point out that the converse of Theorem 1.2.11 is false; that is, there exist series $\sum_{n=1}^{\infty} z_n$ of complex numbers such that $\sum_{n=1}^{\infty} z_n$ converges but $\sum_{n=1}^{\infty} z_n$ does not converge absolutely. Some examples of this will be pointed out shortly (see Example 1.2.23).

Since absolutely convergent series converge, it is useful to develop a collection of 'tests' that will aid us in determining whether a series of non-negative real numbers converges (thereby aiding in determining whether a series of complex numbers converges absolutely).

Theorem 1.2.13 (Comparison Test). *Let $(a_n)_{n\geq 1}$ and $(b_n)_{n\geq 1}$ be sequences of real numbers such that $0 \leq a_n \leq b_n$ for all $n \in \mathbb{N}$. Then*

a) If $\sum_{n=1}^{\infty} b_n$ converges, then $\sum_{n=1}^{\infty} a_n$ converges.

b) If $\sum_{n=1}^{\infty} a_n$ diverges, then $\sum_{n=1}^{\infty} b_n$ diverges.

Proof. As these two statements are contrapositives of each other, it suffices to prove the first statement. To see that the first statement is true, suppose $\sum_{n=1}^{\infty} b_n$ converges. Hence Theorem 1.2.9 implies that there exists an $M \in \mathbb{R}$

such that $\sum_{k=1}^N b_k \leq M$ for all $N \in \mathbb{N}$. Since $0 \leq a_n \leq b_n$ for all $n \in \mathbb{N}$, we obtain for all $N \in \mathbb{N}$ that

$$\sum_{k=1}^N a_k \leq \sum_{k=1}^N b_k \leq M.$$

Hence Theorem 1.2.9 implies $\sum_{n=1}^\infty a_n$ converges. ∎

Theorem 1.2.14 (Integral Test). *If $f : [1, \infty) \to [0, \infty)$ be a non-increasing function and $a_n = f(n)$ for all $n \in \mathbb{N}$, then $\sum_{n=1}^\infty a_n$ converges if and only if $\int_1^\infty f(x)\, dx$ converges.*

Proof. Using the definition of Riemann sums, it is not difficult to see for all $N \in \mathbb{N}$ that

$$\sum_{k=2}^N a_k \leq \int_1^N f(x)\, dx \leq \sum_{k=1}^N a_k.$$

Therefore, since f is non-increasing and non-negative and thus $a_n \geq 0$ for all $n \in \mathbb{N}$, we see that

$$\left\{ \sum_{k=1}^N a_k \,\middle|\, N \in \mathbb{N} \right\}$$

is bounded if and only if

$$\left\{ \int_1^b f(x)\, dx \,\middle|\, b \geq 1 \right\}$$

is bounded. Hence the result follows from Theorem 1.2.9 and the analogous result for improper integrals of non-increasing non-negative functions. ∎

The Integral Test (Theorem 1.2.14) is quite useful when determining whether series related to nice functions converge.

Corollary 1.2.15. *The series $\sum_{n=1}^\infty \frac{1}{n^p}$ converges if and only if $p > 1$.*

Proof. First, consider the case $p \leq 0$. In this case the sequence $\left(\frac{1}{n^p}\right)_{n\geq 1}$ does not converge to zero and thus $\sum_{n=1}^\infty \frac{1}{n^p}$ does not converge by Corollary 1.2.6.

Otherwise $p > 0$. Notice that the function $f : [1, \infty) \to (0, \infty)$ defined by $f(x) = \frac{1}{x^p}$ is well-defined. Moreover, since $f'(x) = \frac{-p}{x^{p+1}} < 0$ for all $x \in [1, \infty)$, we obtain that f is non-decreasing. Furthermore, notice for all $b > 1$ that

$$\int_1^b f(x)\, dx = \begin{cases} \ln(b) & \text{if } p = 1 \\ \frac{1}{p-1} - \frac{1}{(p-1)b^{p-1}} & \text{if } p \neq 1 \end{cases}.$$

Hence we easily see that

$$\lim_{b\to\infty} \int_1^b f(x)\, dx$$

exists if and only if $p > 1$ as desired. ∎

Remark 1.2.16. Note Corollary 1.2.15 immediately implies that $\sum_{n=1}^{\infty} \frac{1}{n}$ diverges. Therefore, since $\lim_{n\to\infty} \frac{1}{n} = 0$, we have an example to show that the converse of Corollary 1.2.6 is false.

Of course, with the convergence of these series, one fundamental question would be:

Question 1.2.17. *What are the values of* $\sum_{n=1}^{\infty} \frac{1}{n^p}$ *for all* $p > 1$*?*

Of course, the proof of the Integral Test (Theorem 1.2.14) can be used to show that

$$\int_1^\infty \frac{1}{x^p}\, dx \leq \sum_{n=1}^{\infty} \frac{1}{n^p} \leq \frac{1}{1^p} + \int_1^\infty \frac{1}{x^p}\, dx,$$

so

$$\frac{1}{p-1} \leq \sum_{n=1}^{\infty} \frac{1}{n^p} \leq 1 + \frac{1}{p-1} = \frac{p}{p-1},$$

but this is rather wide interval.

Using our knowledge of convergent series, we can also construct the following test to aid us in determine when some series converge.

Theorem 1.2.18 (Ratio Test). *Let* $(a_n)_{n\geq 1}$ *be a sequence of real numbers such that* $a_n > 0$ *for all* $n \in \mathbb{N}$*. Suppose* $r = \lim_{n\to\infty} \frac{a_{n+1}}{a_n}$ *exists. Then:*

a) $\sum_{n=1}^{\infty} a_n$ *converges if* $r < 1$*.*

b) $\sum_{n=1}^{\infty} a_n$ *diverges if* $r > 1$*.*

Proof. First suppose $r < 1$. Our goal for this direction is to show that the tail of the series is bounded above by a convergent geometric series.

Let $\epsilon = \frac{1-r}{2} > 0$. Since $r = \lim_{n\to\infty} \frac{a_{n+1}}{a_n}$, there exists an $K \in \mathbb{N}$ such that

$$\left| \frac{a_{k+1}}{a_k} - r \right| < \epsilon$$

for all $k \geq K$. Therefore

$$\frac{a_{k+1}}{a_k} < r + \epsilon = r + \frac{1-r}{2} = \frac{1+r}{2}$$

for all $k \geq K$ so

$$a_{k+1} < \left(\frac{1+r}{2}\right) a_k$$

for all $k \geq K$. A simple induction argument then implies

$$a_k \leq \left(\frac{1+r}{2}\right)^{k-K} a_K$$

for all $k \geq K$. Notice since $\frac{1+r}{2} < \frac{1+1}{2} = 1$ that the geometric series

$$\sum_{n=1}^{\infty} \left(\frac{1+r}{2}\right)^n$$

converges by Example 1.2.3. Thus

$$\sum_{n=K}^{\infty} \left(\frac{1+r}{2}\right)^{k-K} a_K$$

converges so $\sum_{n=K}^{\infty} a_n$ converges by the Comparison Test (Theorem 1.2.13). Hence $\sum_{n=1}^{\infty} a_n$ converges as only the tail of the series matters.

Now suppose $r > 1$. Our goal for this direction is to show that the tail of the series is bounded below by a divergent geometric series.

Let $\epsilon = \frac{r-1}{2} > 0$. Since $r = \lim_{n\to\infty} \frac{a_{n+1}}{a_n}$, there exists an $K \in \mathbb{N}$ such that

$$\left|\frac{a_{k+1}}{a_k} - r\right| < \epsilon$$

for all $k \geq K$. Therefore

$$\frac{a_{k+1}}{a_k} > r - \epsilon = r - \frac{r-1}{2} = \frac{r+1}{2}$$

for all $k \geq K$ so

$$a_{k+1} > \left(\frac{1+r}{2}\right) a_k$$

for all $k \geq K$. A simple induction argument then implies

$$a_k \geq \left(\frac{1+r}{2}\right)^{k-K} a_K$$

for all $k \geq K$. Notice since $\frac{1+r}{2} > \frac{1+1}{2} = 1$ that the geometric series

$$\sum_{n=1}^{\infty} \left(\frac{1+r}{2}\right)^n$$

diverges by Example 1.2.8. Thus

$$\sum_{n=K}^{\infty} \left(\frac{1+r}{2}\right)^{k-K} a_K$$

diverges (since $a_K \neq 0$) so $\sum_{n=K}^{\infty} a_n$ diverges by the Comparison Test (Theorem 1.2.13). Hence $\sum_{n=1}^{\infty} a_n$ diverges as only the tail of the series matters. ∎

Remark 1.2.19. In the context of the Ratio Test (Theorem 1.2.18), when $r = 1$ we obtain no information since we cannot bound the ratio away from 1 in either direction to obtain a comparison. In fact, when $r = 1$, the series could converge or diverge. In particular, by Corollary 1.2.15, we know that the series

$$\sum_{n=1}^{\infty} \frac{1}{n^2} \qquad \text{and} \qquad \sum_{n=1}^{\infty} \frac{1}{n}$$

converge and diverge respectively, but both have $r = 1$.

Using the Ratio Test (Theorem 1.2.18), we can prove a vital series converges.

Example 1.2.20. We claim that the series $\sum_{n=0}^{\infty} \frac{1}{n!} z^n$ converges absolutely for all $z \in \mathbb{C}$. To see this, fix $z \in \mathbb{C}$ and for all $n \in \mathbb{N}$ let

$$a_n = \left| \frac{1}{n!} z^n \right| = \frac{|z|^n}{n!}.$$

Note for all $n \in \mathbb{N}$ that

$$\left| \frac{a_{n+1}}{a_n} \right| = \left| \frac{|z|^{n+1}}{(n+1)!} \frac{n!}{|z|^{n+1}} \right| = \frac{|z|}{n+1}.$$

Since

$$\lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| = \lim_{n\to\infty} \frac{|z|}{n+1} = 0,$$

we obtain that $\sum_{n=1}^{\infty} a_n$ converges by the Ratio Test (Theorem 1.2.18) and thus $\sum_{n=0}^{\infty} \frac{1}{n!} z^n$ converges absolutely for all $z \in \mathbb{C}$.

A sometimes useful alternative to the Ratio Test (Theorem 1.2.18) is the following, which is proved in a similar manner.

Theorem 1.2.21 (Root Test). *Let $(a_n)_{n\geq 1}$ be a sequence of real numbers such that $a_n > 0$ for all $n \in \mathbb{N}$. Suppose $r = \lim_{n\to\infty} \sqrt[n]{a_n}$ exists. Then:*

a) $\sum_{n=1}^{\infty} a_n$ converges if $r < 1$.

b) $\sum_{n=1}^{\infty} a_n$ diverges if $r > 1$.

Proof. First suppose $r < 1$. Our goal for this direction is to show that the tail of the series is bounded above by a convergent geometric series.

Let $\epsilon = \frac{1-r}{2} > 0$. Since $r = \lim_{n\to\infty} \sqrt[n]{a_n}$, there exists an $K \in \mathbb{N}$ such that

$$|\sqrt[k]{a_k} - r| < \epsilon$$

for all $k \geq K$. Therefore

$$\sqrt[k]{a_k} < r + \epsilon = r + \frac{1-r}{2} = \frac{1+r}{2}$$

for all $k \geq K$ so

$$a_k < \left(\frac{1+r}{2}\right)^k$$

for all $k \geq K$. Notice since $\frac{1+r}{2} < \frac{1+1}{2} = 1$ that the geometric series

$$\sum_{n=K}^{\infty} \left(\frac{1+r}{2}\right)^n$$

converges by Example 1.2.3 so $\sum_{n=K}^{\infty} a_n$ converges by the Comparison Test (Theorem 1.2.13). Hence $\sum_{n=1}^{\infty} a_n$ converges as only the tail of the series matters.

Now suppose $r > 1$. Our goal for this direction is to show that the tail of the series is bounded below by a divergent geometric series.

Let $\epsilon = \frac{r-1}{2} > 0$. Since $r = \lim_{n\to\infty} \sqrt[n]{a_n}$, there exists an $K \in \mathbb{N}$ such that

$$|\sqrt[k]{a_k} - r| < \epsilon$$

for all $k \geq K$. Therefore

$$\sqrt[k]{a_k} > r - \epsilon = r - \frac{r-1}{2} = \frac{r+1}{2}$$

for all $k \geq K$ so

$$a_k > \left(\frac{1+r}{2}\right)^k$$

for all $k \geq K$. Notice since $\frac{1+r}{2} > \frac{1+1}{2} = 1$ that the geometric series

$$\sum_{n=1}^{\infty} \left(\frac{1+r}{2}\right)^n$$

diverges by Example 1.2.8 so $\sum_{n=K}^{\infty} a_n$ diverges by the Comparison Test (Theorem 1.2.13). Hence $\sum_{n=1}^{\infty} a_n$ diverges as only the tail of the series matters. ∎

Of course, there are other ways to ensure that a series converges. One method for obtaining a convergent series is the following.

Theorem 1.2.22 (Alternating Series Test (Leibniz's Theorem)). *Let $(a_n)_{n\geq 1}$ be a non-increasing sequence of non-negative real numbers such that $\lim_{n\to\infty} a_n = 0$. Then $\sum_{n=1}^{\infty}(-1)^{n+1}a_n$ converges.*

Proof. For each $N \in \mathbb{N}$, let $S_N = \sum_{k=1}^{N}(-1)^{k+1}a_k$. First, we claim for all $N \in \mathbb{N}$ that

$$S_{2N} \leq S_{2N+2} \leq S_{2N+3} \leq S_{2N+1}.$$

To see the first inequality, notice since $(a_n)_{n\geq 1}$ is a non-increasing sequence of non-negative real numbers that $a_{2N+1} - a_{2N+2} \geq 0$ for all $N \in \mathbb{N}$ so

$$\begin{aligned} S_{2N} &\leq (a_{2N+1} - a_{2N+2}) + S_{2N} \\ &= (-1)^{(2N+1)+1} a_{2N+1} + (-1)^{(2N+2)+1} a_{2N+2} + \sum_{k=1}^{2N} a_k \\ &= \sum_{n=1}^{2N+2} (-1)^{k+1} a_k = S_{2N+2}. \end{aligned}$$

Similarly, since $(a_n)_{n\geq 1}$ is a non-increasing sequence of non-negative real numbers that $a_{2N+3} - a_{2N+2} \leq 0$ for all $N \in \mathbb{N}$ so

$$\begin{aligned} S_{2N+3} &= \sum_{n=1}^{2N+3} (-1)^{k+1} a_k \\ &= (-1)^{(2N+3)+1} a_{2N+3} + (-1)^{(2N+2)+1} a_{2N+2} + \sum_{k=1}^{2N+1} a_k \\ &= (a_{2N+3} - a_{2N+2}) + S_{2N+1} \leq S_{2N+1}. \end{aligned}$$

Finally, since $a_{2N+3} \geq 0$, we obtain that

$$S_{2N+2} = \sum_{k=1}^{2N+2} a_k \leq a_{2N+3} + \sum_{k=1}^{2N+2} a_k = \sum_{k=1}^{2N+3} a_k = S_{2N+3}$$

as desired.

Notice the inequality proved in the above claim shows that $(S_{2N})_{N\geq 1}$ is a non-decreasing sequence and $(S_{2N+1})_{N\geq 1}$ is non-increasing sequence both of which are bounded below by S_2 and bounded above by S_1. Hence the Monotone Converge Theorem implies $(S_{2N})_{N\geq 1}$ and $(S_{2N+1})_{N\geq 1}$ converge.

Let

$$L = \lim_{N\to\infty} S_{2N} \qquad \text{and} \qquad K = \lim_{N\to\infty} S_{2N+1}.$$

Notice that

$$K - L = \lim_{N\to\infty} S_{2N+1} - S_{2N} = \lim_{N\to\infty} \sum_{k=1}^{2N+1} a_k - \sum_{k=1}^{2N} a_k = \lim_{N\to\infty} a_{2N+1} = 0.$$

Hence $L = K$. Therefore, since $(S_{2N})_{N\geq 1}$ and $(S_{2N+1})_{N\geq 1}$ both converge to L, $(S_N)_{N\geq 1}$ converges to L. Therefore $\sum_{n=1}^{\infty} (-1)^{n+1} a_n$ converges by definition. ■

The Alternating Series Test (Theorem 1.2.22) is a quick way to construct series that converge but do not converge absolutely.

Example 1.2.23. We claim that the series $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}$ converges but does not converge absolutely. To see that this series converges, note that $\lim_{n\to\infty} \frac{1}{n} = 0$. Hence the Alternating Series Test (Theorem 1.2.22) implies that $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}$. However, since

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

diverges by Corollary 1.2.15, $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}$ does not converge absolutely.

It is useful to have a name for such series.

Definition 1.2.24. A series $\sum_{n=1}^{\infty} z_n$ of complex numbers is said to converge *conditionally* if $\sum_{n=1}^{\infty} z_n$ converges but does not converge absolutely.

Of course, one natural question we can ask is:

Question 1.2.25. *What are the values of* $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}$*?*

In an attempt to solve this question, consider the following: Let

$$S = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}.$$

Notice that

$$\begin{aligned} S &= 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} + \cdots \\ &= \left(1 - \frac{1}{2}\right) - \frac{1}{4} + \left(\frac{1}{3} - \frac{1}{6}\right) - \frac{1}{8} + \left(\frac{1}{5} - \frac{1}{10}\right) - \frac{1}{12} + \cdots \\ &= \frac{1}{2} - \frac{1}{4} + \frac{1}{6} - \frac{1}{8} + \frac{1}{10} - \frac{1}{12} + \cdots \\ &= \frac{1}{2}\left(1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} + \cdots\right) \\ &= \frac{1}{2}S. \end{aligned}$$

Therefore $S = 0$. However, notice that

$$\begin{aligned} S &= \left(1 - \frac{1}{2}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \left(\frac{1}{5} - \frac{1}{6}\right) + \left(\frac{1}{7} - \frac{1}{8}\right) + \cdots \\ &\geq \frac{1}{2} + 0 + 0 + \cdots = \frac{1}{2}. \end{aligned}$$

How is this possible? Did we just break mathematics?

1.3 Rearrangements of Series

The issue with the above computation is that when we were trying to evaluate S, we rearranged the order of the terms in the series. This may seem valid in the sense that if we are adding up a only finite number of scalars, then we know we can rearrange the order of the terms in the sum due to the associative and commutative properties of addition. However, in Definition 1.2.1, the partial sums are formed by adding up the terms of the series in a very specific order and then taking the limit of the partial sums. By rearranging the terms in an infinite series, we are in a sense transforming the series as we are changing the partial sums and thereby modifying the value the partial sums converges to. In fact, when dealing with a conditional convergent series, we can reorder the series to make the value of the series anything we want!

Theorem 1.3.1. *Let* $\sum_{n=1}^{\infty} a_n$ *be a conditionally convergent series of real numbers. For any* $L \in \mathbb{R}$ *there exists a bijection* $\sigma : \mathbb{N} \to \mathbb{N}$ *such that* $\sum_{n=1}^{\infty} a_{\sigma(n)} = L$.

Proof. For each $n \in \mathbb{N}$, let

$$a_n^+ = \begin{cases} a_n & \text{if } a_n \geq 0 \\ 0 & \text{if } a_n < 0 \end{cases} \qquad \text{and} \qquad a_n^- = \begin{cases} 0 & \text{if } a_n \geq 0 \\ a_n & \text{if } a_n < 0 \end{cases}.$$

Hence for all $n \in \mathbb{N}$

$$a_n = a_n^+ + a_n^- \qquad \text{and} \qquad |a_n| = a_n^+ - a_n^-.$$

If both $\sum_{n=1}^{\infty} a_n^+$ and $\sum_{n=1}^{\infty} a_n^-$ converged, then $\sum_{n=1}^{\infty} |a_n|$ would converge since

$$\sum_{k=1}^{N} |a_k| = \sum_{k=1}^{N} a_k^+ - \sum_{k=1}^{N} a_k^-$$

for all $N \geq 1$. However, since $\sum_{n=1}^{\infty} |a_n|$ does not converge, it must be the case that least one of $\sum_{n=1}^{\infty} a_n^+$ and $\sum_{n=1}^{\infty} a_n^-$ diverges. Moreover, since $\sum_{n=1}^{\infty} a_n$ converges and

$$\sum_{k=1}^{N} a_k = \sum_{k=1}^{N} a_k^+ + \sum_{k=1}^{N} a_k^-$$

for all $N \in \mathbb{N}$, if one of $\sum_{n=1}^{\infty} a_n^+$ and $\sum_{n=1}^{\infty} a_n^-$ converged then both would need to converge thereby contradicting what was just demonstrated. Hence both $\sum_{n=1}^{\infty} a_n^+$ and $\sum_{n=1}^{\infty} a_n^-$ diverge.

Let $(p_n)_{n\geq 1}$ denote sequence of all non-negative terms from $(a_n)_{n\geq 1}$ listed in the same order they appear, and let $(q_n)_{n\geq 1}$ denote sequence of all negative terms from $(a_n)_{n\geq 1}$ listed in the same order they appear. Since

$$\sum_{n=1}^{\infty} p_n \qquad \text{and} \qquad \sum_{n=1}^{\infty} q_n$$

diverge by what was demonstrated above, Theorem 1.2.9 implies that

$$\sup\left(\left\{\sum_{k=1}^{N} p_k \;\middle|\; N\in\mathbb{N}\right\}\right)=\infty \quad\text{and}\quad \inf\left(\left\{\sum_{k=1}^{N} q_k \;\middle|\; N\in\mathbb{N}\right\}\right)=-\infty.$$

Fix $L \in \mathbb{R}$. To find a bijection $\sigma : \mathbb{N} \to \mathbb{N}$ so that $\sum_{n=1}^{\infty} a_{\sigma(n)} = L$, our goal is to add p_ks up to the point where we obtain a number just larger than L, then add (thereby decreasing the value) q_ks up to the point where we obtain a number just smaller than L, then add p_ks up to the point where we obtain a number just larger than L, then add (thereby decreasing the value) q_ks up to the point where we obtain a number just smaller than L, and so on. This procedure will create a rearrangement of the series so that partial sums will be within a_n of L for some increasingly large n. Therefore since $\sum_{n=1}^{\infty} a_n$ converges so $\lim_{n\to\infty} a_n = 0$ by Corollary 1.2.6, this rearrangement will converge to L. Hopefully that is clear as the formal write-up as seen below is rather technical and possibly not illuminating.

First note since the above supremum is infinity that there exists an $N \in \mathbb{N}$ such that

$$\sum_{k=1}^{N} p_k > L.$$

Choose $N_1 \in \mathbb{N}$ to be the smallest $N \in \mathbb{N}$ such that $\sum_{k=1}^{N} p_k > L$. Therefore, since $p_k \geq 0$ for all $k \in \mathbb{N}$, this implies that

$$\sum_{k=1}^{N_1} p_k > L \geq \sum_{k=1}^{N} p_k$$

for all $N < N_1$. Let

$$T_1 = \sum_{k=1}^{N_1} p_k$$

and note $0 < T_1 - L$.

Next, since the above infimum is negative infinity, there exists $N \in \mathbb{N}$ such that

$$\sum_{k=1}^{N} q_k < L - T_1.$$

Choose $M_1 \in \mathbb{N}$ to be the smallest $N \in \mathbb{N}$ such that $\sum_{k=1}^{N} q_k < L - T_1$. Therefore, since $q_k \leq 0$ for all $k \in \mathbb{N}$, this implies that

$$\sum_{k=1}^{M_1} q_k < L - T_1 \leq \sum_{k=1}^{N} q_k$$

for all $N < M_1$. Let

$$R_1 = \sum_{k=1}^{M_1} q_k.$$

Note the above inequalities imply that

$$0 < L - T_1 - R_1 = (L - T_1) - R_1 \leq \sum_{k=1}^{M_1 - 1} q_k - \sum_{k=1}^{M_1} q_k = -q_{M_1}.$$

Next, since the above supremum is infinity that there exists an $N \in \mathbb{N}$ such that

$$\sum_{k=N_1+1}^{N} p_k > L - T_1 - R_1.$$

Choose $N_2 \in \mathbb{N}$ to be the smallest $N \in \mathbb{N}$ such that $N > N_1$ and $\sum_{k=N_1+1}^{N} p_k > L - T_1 - R_1$. Therefore, since $p_k \geq 0$ for all $k \in \mathbb{N}$, this implies that

$$\sum_{k=N_1+1}^{N_2} p_k > L - T_1 - R_1 \geq \sum_{k=N_1+1}^{N} p_k.$$

for all $N \in \{N_1 + 1, \ldots, N_2 - 1\}$. Let

$$T_2 = \sum_{k=N_1+1}^{N_2} p_k.$$

Notice that the above inequalities and the fact that $p_k \geq 0$ for all $k \in \mathbb{N}$ imply that

$$0 \leq L - T_1 - R_1 - \sum_{k=N_1+1}^{N} p_k \leq L - T_1 - R_1 \leq -q_{M_1}$$

for all $N \in \{N_1 + 1, \ldots, N_2 - 1\}$ and

$$0 < T_2 - (L - T_1 - R_1) \leq \sum_{k=N_1+1}^{N_2} p_k - \sum_{k=N_1+1}^{N_2 - 1} p_k = p_{N_2}.$$

Once more for clarity, since $T_1 + R_1 + T_2 - L > 0$ and the above infimum is negative infinity, there exists an $N \in \mathbb{N}$ such that

$$\sum_{k=M_1+1}^{N} q_k < L - T_1 - R_1 - T_2.$$

Choose $M_2 \in \mathbb{N}$ to be the smallest $N \in \mathbb{N}$ such that $N > M_1$ and $\sum_{k=M_1+1}^{N} q_k < L - T_1 - R_1 - T_2$. Therefore, since $q_k \leq 0$ for all $k \in \mathbb{N}$, this implies that

$$\sum_{k=M_1+1}^{M_2} q_k < L - T_1 - R_1 - T_2 \leq \sum_{k=M_1+1}^{N} q_k.$$

for all $N \in \{M_1 + 1, \ldots, M_2 - 1\}$. Let

$$R_2 = \sum_{k=M_1+1}^{M_2} q_k.$$

Notice that the above inequalities and the fact that $q_k \leq 0$ for all $k \in \mathbb{N}$ imply that

$$0 \geq L - T_1 - R_1 - T_2 - \sum_{k=M_1+1}^{N} q_k \geq L - T_1 - R_1 - T_2 > -p_{N_2}$$

for all $N \in \{M_1 + 1, \ldots, M_2 - 1\}$ and

$$0 < (L - T_1 - R_1 - T_2) - R_2 \leq \sum_{k=M_1+1}^{M_2-1} q_k - \sum_{k=M_1+1}^{M_2} q_k = -q_{M_2}.$$

By repeating this procedure, there exist strictly increasing sequences $(N_j)_{j\geq 1}$ and $(M_j)_{j\geq 1}$ so that if

$$T_j = \sum_{k=N_{j-1}+1}^{N_j} p_k \qquad \text{and} \qquad R_j = \sum_{k=M_{j-1}+1}^{M_j} q_k,$$

then for all $\ell \geq 1$ we have that

$$0 \leq L - \sum_{j=1}^{\ell}(T_j + R_j) - \sum_{k=N_j+1}^{N} p_k \leq -q_{M_\ell}$$

for all $N \in \{N_\ell + 1, \ldots, N_{\ell+1} - 1\}$,

$$0 < -L + T_{\ell+1} + \sum_{j=1}^{\ell}(T_j + R_j) \leq p_{N_{\ell+1}},$$

and

$$0 \geq L - T_{\ell+1} - \sum_{j=1}^{\ell}(T_j + R_j) - \sum_{k=M_j+1}^{N} q_k > -p_{N_{\ell+1}}$$

for all $N \in \{M_\ell + 1, \ldots, M_{\ell+1} - 1\}$, and

$$0 < L - \sum_{j=1}^{\ell+1}(T_j + R_j) \leq -q_{M_{\ell+1}}.$$

Since $(N_j)_{j\geq 1}$ and $(M_j)_{j\geq 1}$ are strictly increasing sequences, we see that

$$p_1, \ldots, p_{N_1}, q_1, \ldots, q_{M_1}, p_{N_1+1}, \ldots, p_{N_2}, q_{M_1+1}, \ldots, q_{M_2},$$
$$p_{N_2+1}, \ldots, p_{N_3}, q_{M_2+1}, \ldots, q_{M_3}, \ldots$$

is a rearrangement of $\sum_{n=1}^{\infty} a_n$. Moreover, by construction, the partial sums of this rearrangement are within either p_{N_ℓ} or $-q_{M_\ell}$ of L for progressively large ℓ at ever step of the construction. Since $\sum_{n=1}^{\infty} a_n$ converges, Corollary 1.2.6 implies that $\lim_{n\to\infty} a_n = 0$. Therefore $\lim_{n\to\infty} p_n = \lim_{n\to\infty} q_n = 0$, so the partial sums of this rearrangement converge to L as desired. ∎

The main appeal of absolutely convergent series over conditionally convergent series is the dichotomy between the following and Theorem 1.3.1.

Theorem 1.3.2. *Let $\sum_{n=1}^{\infty} z_n$ be an absolutely convergent series of complex numbers. For all bijections $\sigma : \mathbb{N} \to \mathbb{N}$, the series $\sum_{n=1}^{\infty} z_{\sigma(n)}$ converges absolutely and $\sum_{n=1}^{\infty} z_{\sigma(n)} = \sum_{n=1}^{\infty} z_n$.*

Proof. Let $L = \sum_{n=1}^{\infty} z_n$ and fix a bijection $\sigma : \mathbb{N} \to \mathbb{N}$. For all $N \in \mathbb{N}$, let

$$S_N = \sum_{k=1}^{N} z_k \qquad \text{and} \qquad T_N = \sum_{k=1}^{N} z_{\sigma(k)}.$$

To see that $(T_N)_{N\geq 1}$ converges to L, let $\epsilon > 0$ be arbitrary. Since $L = \sum_{n=1}^{\infty} z_n$, there exists an $N_1 \in \mathbb{N}$ such that

$$|S_N - L| < \frac{\epsilon}{2}$$

for all $N \geq N_1$. Moreover, since $\sum_{n=1}^{\infty} z_n$ converges absolutely, Theorem 1.2.9 implies there exists an $N_2 \in \mathbb{N}$ such that

$$\sum_{k=N}^{\infty} |z_k| < \frac{\epsilon}{2}$$

for all $N \geq N_2$.

Let $N_0 = \max\{N_1, N_2\}$. Since $\sigma : \mathbb{N} \to \mathbb{N}$ is a bijection, there exists an $M_0 \in \mathbb{N}$ such that

$$\{1, 2, \ldots, N_0\} \subseteq \{\sigma(j) \mid j \in \{1, \ldots, M_0\}\}.$$

Therefore, for all $M \geq M_0$ we see that

$$\begin{aligned} |T_M - S_{N_0}| &= \left| \sum_{k=1}^{M} z_{\sigma(k)} - \sum_{k=1}^{N_0} z_k \right| \\ &= \left| \sum_{\substack{k\in\sigma(\{1,\ldots,M\}) \\ \text{and } k\notin\{1,\ldots,N_0\}}} z_k \right| && \text{i.e. we made } M \text{ so large to contain the first } N_0 \text{ terms of the initial series} \\ &\leq \sum_{\substack{k\in\sigma(\{1,\ldots,M\}) \\ \text{and } k\notin\{1,\ldots,N_0\}}} |z_k| \\ &\leq \sum_{k=N_0+1}^{\infty} |z_k| < \frac{\epsilon}{2}. \end{aligned}$$

Hence, for all $M \geq M_0$ we see that

$$|T_M - L| \leq |T_M - S_{N_0}| + |S_{N_0} - L| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Therefore, since $\epsilon > 0$ was arbitrary, $\sum_{n=1}^{\infty} z_{\sigma(n)}$ converges to L.

To see that $\sum_{n=1}^{\infty} z_{\sigma(n)}$ converges absolutely, note since $\sum_{n=1}^{\infty} z_n$ converges absolutely that $\sum_{n=1}^{\infty} |z_n|$ converges and thus $\sum_{n=1}^{\infty} |z_{\sigma(n)}|$ converges by the first part of the proof. Hence $\sum_{n=1}^{\infty} z_{\sigma(n)}$ converges absolutely. ∎

1.4 Double Indexed Series

Given that rearranging series may or may not change the value of the series based on the type of convergence, perhaps it is useful to ponder other similar situations that may arise. One such situation is when dealing with a sum over two indices. To be specific, given $z_{n,m} \in \mathbb{C}$ for all $n, m \in \mathbb{N}$, is

$$\sum_{n=1}^{\infty} \sum_{m=1}^{\infty} z_{n,m} = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} z_{n,m}?$$

For those that have taken multivariate calculus, this is very reminiscent of Fubini's Theorem that lets one exchange the order of a double integral for continuous functions.

To understand our question for sums, note that

$$\begin{aligned}
\sum_{n=1}^{\infty} \sum_{m=1}^{\infty} z_{n,m} &= \lim_{N\to\infty} \sum_{n=1}^{N} \sum_{m=1}^{\infty} z_{n,m} \\
&= \lim_{N\to\infty} \sum_{n=1}^{N} \lim_{M\to\infty} \sum_{m=1}^{M} z_{n,m} \\
&= \lim_{N\to\infty} \lim_{M\to\infty} \sum_{n=1}^{N} \sum_{m=1}^{M} z_{n,m}.
\end{aligned}$$

whereas

$$\begin{aligned}
\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} z_{n,m} &= \lim_{M\to\infty} \sum_{m=1}^{M} \sum_{n=1}^{\infty} z_{n,m} \\
&= \lim_{M\to\infty} \sum_{m=1}^{M} \lim_{N\to\infty} \sum_{n=1}^{N} z_{n,m} \\
&= \lim_{M\to\infty} \lim_{N\to\infty} \sum_{m=1}^{M} \sum_{n=1}^{N} z_{n,m}.
\end{aligned}$$

Thus is it always possible to interchange the limits? No!

Example 1.4.1. For all $n, m \in \mathbb{N}$, let

$$z_{n,m} = \begin{cases} 1 & \text{if } n = m \\ -1 & \text{if } n = m+1 \\ 0 & \text{otherwise} \end{cases}.$$

Notice that $\sum_{n=1}^{\infty} z_{n,m} = 0$ for all $m \in \mathbb{N}$ since the sequence $(z_{n,m})_{n\geq 1}$ always contains a single -1 followed by a single 1 and all other terms are 0. However notice that

$$\sum_{m=1}^{\infty} z_{n,m} = \begin{cases} 0 & \text{if } n \geq 2 \\ 1 & \text{if } n = 1 \end{cases}$$

since the sequence $(z_{1,m})_{m\geq 1}$ is the sequence $(1, 0, 0, \ldots)$ whereas for $n \geq 2$ the sequence $(z_{n,m})_{m\geq 1}$ always contains a single -1 followed by a single 1 and all other terms are 0. Hence

$$\sum_{n=1}^{\infty} \sum_{m=1}^{\infty} z_{n,m} = 1$$

whereas

$$\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} z_{n,m} = 0$$

so

$$\sum_{n=1}^{\infty} \sum_{m=1}^{\infty} z_{n,m} \neq \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} z_{n,m}.$$

As we will see via the following two results, this example is only possible since $z_{n,m}$ takes both positive and negative values and

$$\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} |z_{n,m}| = \sum_{m=1}^{\infty} 2$$

does not converge.

For certain series, we do not run into the same pathological behaviour as Example 1.4.1. In particular, like with Section 1.2, for series of non-negative real numbers, things work out as beautifully as possible.

Theorem 1.4.2 (Tonelli's Theorem for Sums). *For all $n, m \in \mathbb{N}$, let $a_{n,m} \geq 0$. The following are equivalent:*

i) $\sum_{n=1}^{\infty} a_{n,m}$ *converges for all* $m \in \mathbb{N}$ *and* $\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} a_{n,m}$ *converges.*

ii) $\sum_{m=1}^{\infty} a_{n,m}$ *converges for all* $n \in \mathbb{N}$ *and* $\sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{n,m}$ *converges.*

iii) $S = \sup\left(\left\{\sum_{n=1}^{N} \sum_{m=1}^{M} a_{n,m} \;\middle|\; N, M \in \mathbb{N}\right\}\right) < \infty.$

Moreover, in the case that one and thus all of the above are true, we have

$$\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_{n,m} = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} a_{n,m} = S.$$

Proof. Only the equivalence of (ii) and (iii) will be demonstrated as the equivalence of (i) and (iii) is nearly identical.

Suppose that (ii) holds and let

$$L = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} a_{n,m}.$$

Since $\sum_{m=1}^{\infty} a_{n,m} \geq 0$ for all $n \in \mathbb{N}$ as $a_{n,m} \geq 0$ for all $n, m \in \mathbb{N}$, we see if

$$S_N = \sum_{n=1}^{N}\sum_{m=1}^{\infty} a_{n,m}$$

for all $N \in \mathbb{N}$, then $(S_N)_{n\geq 1}$ is a non-decreasing sequence of partial sums that converge to L so

$$\sum_{n=1}^{N}\sum_{m=1}^{\infty} a_{n,m} \leq L$$

for all $N \in \mathbb{N}$. Similarly, since $a_{n,m} \geq 0$ for all $n, m \in \mathbb{N}$ we see that

$$\sum_{m=1}^{M} a_{n,m} \leq \sum_{m=1}^{\infty} a_{n,m}$$

for all $M, n \in \mathbb{N}$ so

$$\sum_{n=1}^{N}\sum_{m=1}^{M} a_{n,m} \leq \sum_{n=1}^{N}\sum_{m=1}^{\infty} a_{n,m} \leq L$$

for all $N, M \in \mathbb{N}$. Hence $S \leq L < \infty$ so (iii) holds.

Conversely, suppose that (iii) holds. Since

$$\sum_{m=1}^{M} a_{n,m} \leq S$$

for all $n, M \in \mathbb{N}$, Theorem 1.2.9 implies that $\sum_{m=1}^{\infty} a_{n,m}$ converges for all $n \in \mathbb{N}$. To see that $\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} a_{n,m}$ converges, fix $N \in \mathbb{N}$ and consider the partial sum

$$S_N = \sum_{n=1}^{N}\sum_{m=1}^{\infty} a_{n,m}.$$

Notice for all $M \in \mathbb{N}$ that

$$\sum_{n=1}^{N}\sum_{m=1}^{M} a_{n,m} \leq S.$$

Hence

$$\begin{aligned}
S_N &= \sum_{n=1}^{N} \lim_{M\to\infty} \sum_{m=1}^{M} a_{n,m} \\
&= \lim_{M\to\infty} \sum_{n=1}^{N} \sum_{m=1}^{M} a_{n,m} \quad \text{(since the sum over } n \text{ is finite)} \\
&\leq S.
\end{aligned}$$

Therefore, since $\sum_{m=1}^{\infty} a_{n,m} \geq 0$ for all $n \in \mathbb{N}$, Theorem 1.2.9 implies that $\sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{n,m}$ converges and

$$L = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{n,m} \leq S.$$

Hence, as we have shown that $S \leq L$ and $L \leq S$ when (ii) and (iii) are true, the result follows. ∎

Returning to series of complex numbers, again the key to being able to 'rearrange' doubly indexed series is absolute convergence.

Theorem 1.4.3 (Fubini's Theorem for Sums). *For all $n, m \in \mathbb{N}$, let $z_{n,m} \in \mathbb{C}$. Suppose that either*

- $\sum_{n=1}^{\infty} |z_{n,m}|$ *converges for all $m \in \mathbb{N}$, and*
- $\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} |z_{n,m}|$ *converges*

or

- $\sum_{m=1}^{\infty} |z_{n,m}|$ *converges for all $n \in \mathbb{N}$, and*
- $\sum_{n=1}^{\infty} \sum_{m=1}^{\infty} |z_{n,m}|$ *converges.*

Then

a) $\sum_{n=1}^{\infty} z_{n,k}$ *and* $\sum_{m=1}^{\infty} z_{k,m}$ *converge absolutely for all $k \in \mathbb{N}$,*

b) $\sum_{m=1}^{\infty} \left(\sum_{n=1}^{\infty} z_{n,m}\right)$ *and* $\sum_{n=1}^{\infty} \left(\sum_{m=1}^{\infty} z_{n,m}\right)$ *converge absolutely, and*

c) $\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} z_{n,m} = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} z_{n,m}$.

Proof. Note that if either conditions hold, then Tonelli's Theorem (Theorem 1.4.2) implies that

- $\sum_{n=1}^{\infty} |z_{n,m}|$ converges for all $m \in \mathbb{N}$,
- $\sum_{m=1}^{\infty} |z_{n,m}|$ converges for all $n \in \mathbb{N}$, and
- $\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} |z_{n,m}| = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} |z_{n,m}| < \infty$.

Hence a) is clearly true. Moreover, Theorem 1.2.11 then implies that

$$\sum_{m=1}^{\infty}\left|\sum_{n=1}^{\infty} z_{n,m}\right| \leq \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} |z_{n,m}| < \infty$$

and

$$\sum_{n=1}^{\infty}\left|\sum_{m=1}^{\infty} z_{n,m}\right| \leq \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} |z_{n,m}| < \infty,$$

so Theorem 1.2.9 implies b) is true.

To see that c) is true, note by part b) that

$$L_1 = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} z_{n,m} \qquad \text{and} \qquad L_2 = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} z_{n,m}$$

exist. To see that $L_1 = L_2$, let $\epsilon > 0$ be arbitrary. Since

$$\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} |z_{n,m}| < \infty,$$

Theorem 1.2.9 implies that there exists an $M_0 \in \mathbb{N}$ such that

$$\sum_{m=M_0+1}^{\infty}\sum_{n=1}^{\infty} |z_{n,m}| < \frac{\epsilon}{4}.$$

Similarly, since

$$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} |z_{n,m}| < \infty,$$

Theorem 1.2.9 implies that there exists an $N_0 \in \mathbb{N}$ such that

$$\sum_{n=N_0+1}^{\infty}\sum_{m=1}^{\infty} |z_{n,m}| < \frac{\epsilon}{4}.$$

Now notice that

$$
\begin{aligned}
&\left|L_1 - \sum_{m=1}^{M_0}\sum_{n=1}^{N_0} z_{n,m}\right| \\
&= \left|\left(\sum_{m=M_0+1}^{\infty}\sum_{n=1}^{\infty} z_{n,m}\right) + \left(\sum_{m=1}^{M_0}\sum_{n=1}^{\infty} z_{n,m} - \sum_{m=1}^{M_0}\sum_{n=1}^{N_0} z_{n,m}\right)\right| && \text{value of series is value of tail plus first few terms} \\
&\leq \left|\sum_{m=M_0+1}^{\infty}\sum_{n=1}^{\infty} z_{n,m}\right| + \left|\sum_{m=1}^{M_0}\sum_{n=1}^{\infty} z_{n,m} - \sum_{m=1}^{M_0}\sum_{n=1}^{N_0} z_{n,m}\right| && \text{triangle inequality} \\
&= \left|\sum_{m=M_0+1}^{\infty}\sum_{n=1}^{\infty} z_{n,m}\right| + \left|\sum_{m=1}^{M_0}\left(\sum_{n=1}^{\infty} z_{n,m} - \sum_{n=1}^{N_0} z_{n,m}\right)\right| && \text{finite sum} \\
&= \left|\sum_{m=M_0+1}^{\infty}\sum_{n=1}^{\infty} z_{n,m}\right| + \left|\sum_{m=1}^{M_0}\sum_{n=N_0+1}^{\infty} z_{n,m}\right| && \text{value of series is value of tail plus first few terms} \\
&\leq \sum_{m=M_0+1}^{\infty}\left|\sum_{n=1}^{\infty} z_{n,m}\right| + \sum_{m=1}^{M_0}\left|\sum_{n=N_0+1}^{\infty} z_{n,m}\right| && \text{Theorem 1.2.9 and finite sum} \\
&\leq \sum_{m=M_0+1}^{\infty}\sum_{n=1}^{\infty} |z_{n,m}| + \sum_{m=1}^{M_0}\sum_{n=N_0+1}^{\infty} |z_{n,m}| && \text{finite sum and Theorem 1.2.9} \\
&= \sum_{m=M_0+1}^{\infty}\sum_{n=1}^{\infty} |z_{n,m}| + \sum_{n=N_0+1}^{\infty}\sum_{m=1}^{M_0} |z_{n,m}| && \text{finite sum} \\
&\leq \sum_{m=M_0+1}^{\infty}\sum_{n=1}^{\infty} |z_{n,m}| + \sum_{n=N_0+1}^{\infty}\sum_{m=1}^{\infty} |z_{n,m}| && \text{Theorem 1.2.9} \\
&< \frac{\epsilon}{4} + \frac{\epsilon}{4} = \frac{\epsilon}{2}.
\end{aligned}
$$

A similar argument reversing the roles of n and m show that

$$\left|L_2 - \sum_{m=1}^{M_0}\sum_{n=1}^{N_0} z_{n,m}\right| = \left|L_2 - \sum_{n=1}^{N_0}\sum_{m=1}^{M_0} z_{n,m}\right| < \frac{\epsilon}{2}.$$

Hence

$$|L_1 - L_2| \leq \left|L_1 - \sum_{m=1}^{M_0}\sum_{n=1}^{N_0} z_{n,m}\right| + \left|\sum_{m=1}^{M_0}\sum_{n=1}^{N_0} z_{n,m} - L_2\right| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Therefore, as $\epsilon > 0$ was arbitrary, $|L_1 - L_2| = 0$. Thus $L_1 = L_2$ as desired. ∎

1.5 The Complex Exponential

Using our knowledge of absolutely convergent series and doubly indexed series, we can derive the complex exponential function along with many of its well-known properties. Said function will be of vital importance in Chapter 3. In particular, it should be pointed out that the results of this section also apply to the real exponential function and provide one of the most rigorous methods for constructing it and deriving its properties. Finally, it is possible to generalize this definition to object other than complex numbers, such as matrices. However, to do so requires technology from MATH 4011.

Definition 1.5.1. For $z \in \mathbb{C}$, the *complex exponential of* z, denoted e^z, is

$$e^z = \sum_{n=0}^{\infty} \frac{1}{n!} z^n.$$

Recall from Example 1.2.20 that $\sum_{n=0}^{\infty} \frac{1}{n!} z^n$ converges absolutely and so e^z is well-defined for all $z \in \mathbb{C}$. Also note that if $a \in \mathbb{R}$ then $e^a \in \mathbb{R}$ and that if $a > 0$, then

$$e^a = \sum_{n=0}^{\infty} \frac{1}{n!} a^n > 0.$$

Finally, it is elementary to see that

$$e^0 = \sum_{n=0}^{\infty} \frac{1}{n!} 0^n = 1.$$

Using just Definition 1.5.1 along with our knowledge of series, we can demonstrate that the complex exponential function does behave like an exponential function.

Theorem 1.5.2. *For all $z, w \in \mathbb{C}$,*

$$e^{z+w} = e^z e^w.$$

Proof. Fix $z, w \in \mathbb{C}$ and for all $n, m \in \mathbb{N}$, let

$$z_{n,m} = \begin{cases} \frac{1}{m!} \frac{1}{(n-m)!} z^{n-m} w^m & \text{if } n \geq m \\ 0 & \text{otherwise} \end{cases}.$$

Notice for all $m \in \mathbb{N}$ that

$$\begin{aligned} \sum_{n=0}^{\infty} |z_{n,m}| &= \sum_{n=m}^{\infty} \frac{1}{m!} \frac{1}{(n-m)!} |z|^{n-m} |w|^m \\ &= \sum_{k=0}^{\infty} \frac{1}{m!} \frac{1}{k!} |z|^k |w|^m \\ &= \frac{1}{m!} |w|^m \sum_{k=0}^{\infty} \frac{1}{k!} |z|^k \\ &= \frac{1}{m!} |w|^m e^{|z|} < \infty \end{aligned}$$

and

$$\begin{aligned}\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}|z_{n,m}| &= \sum_{m=0}^{\infty}\frac{1}{m!}|w|^m e^{|z|}\\ &= e^{|z|}\sum_{m=0}^{\infty}\frac{1}{m!}|w|^m\\ &= e^{|z|}e^{|w|} < \infty.\end{aligned}$$

Therefore, we can apply Fubini's Theorem (Theorem 1.4.3) to obtain that

$$\begin{aligned}e^{z+w} &= \sum_{n=0}^{\infty}\frac{1}{n!}(z+w)^n\\ &= \sum_{n=0}^{\infty}\sum_{m=0}^{n}\frac{1}{n!}\binom{n}{m}z^{n-m}w^m\\ &= \sum_{n=0}^{\infty}\sum_{m=0}^{n}\frac{1}{n!}\frac{n!}{m!((n-m)!)}z^{n-m}w^m\\ &= \sum_{n=0}^{\infty}\sum_{m=0}^{n}\frac{1}{m!}\frac{1}{(n-m)!}z^{n-m}w^k\\ &= \sum_{m=0}^{\infty}\sum_{n=m}^{\infty}\frac{1}{m!}\frac{1}{(n-m)!}z^{n-m}w^m \quad \text{by Fubini's Theorem}\\ &= \sum_{m=0}^{\infty}\sum_{k=0}^{\infty}\frac{1}{m!}\frac{1}{k!}z^k w^m\\ &= \sum_{m=0}^{\infty}\frac{1}{m!}w^m e^z\\ &= e^z e^w\end{aligned}$$

as desired. ∎

Using Theorem 1.5.2 and some ingenuity, other essential properties of the complex exponential function can be demonstrated.

Corollary 1.5.3. *For all $z \in \mathbb{C}$, $e^z \neq 0$. Moreover $e^{-z} = \frac{1}{e^z}$ for all $z \in \mathbb{C}$. Finally $e^x > 0$ for all $x \in \mathbb{R}$.*

Proof. Let $z \in \mathbb{C}$ be arbitrary. By Theorem 1.5.2 we see that

$$1 = e^0 = e^{z+(-z)} = e^z e^{-z}.$$

Hence $e^z \neq 0$ and $e^{-z} = \frac{1}{e^z}$ as desired.

Now let $x \in \mathbb{R}$ be arbitrary. If $x = 0$, then $e^x = 1 > 0$. Next, $x > 0$, then clearly

$$e^x = \sum_{n=0}^{\infty}\frac{1}{n!}x^n > 0$$

being the sum of positive terms. Finally, if $x < 0$, then $e^x = \frac{1}{e^{-x}} > 0$ as desired. ∎

Proposition 1.5.4. *For all $z \in \mathbb{C}$, $\overline{e^z} = e^{\overline{z}}$.*

Proof. Let $z \in \mathbb{C}$ be arbitrary. By Corollary 1.1.7, we obtain that

$$\begin{aligned}
\overline{e^z} &= \overline{\lim_{N\to\infty} \sum_{k=0}^{N} \frac{1}{k!} z^k} \\
&= \lim_{N\to\infty} \overline{\sum_{k=0}^{N} \frac{1}{k!} z^k} \\
&= \lim_{N\to\infty} \sum_{k=0}^{N} \frac{1}{k!} \overline{z}^k \\
&= e^{\overline{z}}
\end{aligned}$$

as desired. ∎

Our next result should not be a surprise to those that have studied complex numbers and Euler's formula.

Corollary 1.5.5. *For all $\theta \in \mathbb{R}$, $|e^{i\theta}| = 1$.*

Proof. Let $\theta \in \mathbb{C}$ be arbitrary. Then

$$\begin{aligned}
|e^{i\theta}|^2 &= e^{i\theta}\overline{e^{i\theta}} \\
&= e^{i\theta}e^{-i\theta} \quad \text{by Proposition 1.5.4} \\
&= e^{i\theta+(-i\theta)} \quad \text{by Theorem 1.5.2} \\
&= e^0 = 1.
\end{aligned}$$

Hence $|e^{i\theta}| = 1$ as desired. ∎

In fact, the best way to obtain Euler's Formula is to define it to be true and derive from the definition all of the desired properties of the resulting functions.

Definition 1.5.6. For all $\theta \in \mathbb{R}$, the *cosine and sine of* θ, denoted $\cos(\theta)$ and $\sin(\theta)$ respectively, are

$$\cos(\theta) = \mathrm{Re}\left(e^{i\theta}\right) = \frac{e^{i\theta} + e^{-i\theta}}{2} \qquad \text{and} \qquad \sin(\theta) = \mathrm{Im}\left(e^{i\theta}\right) = \frac{e^{i\theta} - e^{-i\theta}}{2i}.$$

Note $\cos(\theta), \sin(\theta) \in \mathbb{R}$ for all $\theta \in \mathbb{R}$.

Using the above definition, we obtain many of the base results of the cosine and sine functions, along with Euler's Formula.

Corollary 1.5.7. *The following are true:*

a) (Euler's Formula) $e^{i\theta} = \cos(\theta) + i\sin(\theta)$ *for all* $\theta \in \mathbb{R}$.

b) $\cos^2(\theta) + \sin^2(\theta) = 1$ *for all* $\theta \in \mathbb{R}$.

c) $\cos(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n}$ *for all* $x \in \mathbb{R}$ *with the series converging absolutely.*

d) $\sin(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1}$ *for all* $x \in \mathbb{R}$ *with the series converging absolutely.*

e) $\cos(0) = 1$ *and* $\sin(0) = 0$.

f) $\cos(-\theta) = \cos(\theta)$ *and* $\sin(-\theta) = -\sin(\theta)$ *for all* $\theta \in \mathbb{R}$.

g) $\cos(\theta + \varphi) = \cos(\theta)\cos(\varphi) - \sin(\theta)\sin(\varphi)$ *for all* $\theta, \varphi \in \mathbb{R}$.

h) $\sin(\theta + \varphi) = \sin(\theta)\cos(\varphi) + \cos(\theta)\sin(\varphi)$ *for all* $\theta, \varphi \in \mathbb{R}$.

Proof. Since $\cos(\theta) = \mathrm{Re}(e^{i\theta})$ and $\sin(\theta) = \mathrm{Im}(e^{i\theta})$, a) trivially follows.

To see that b) is true, note by Corollary 1.5.5 that for all $\theta \in \mathbb{R}$

$$1 = \left|e^{i\theta}\right|^2 = |\cos(\theta) + i\sin(\theta)|^2 = \cos^2(\theta) + \sin^2(\theta)$$

as desired.

To see that c) and d) are true, fix $x \in \mathbb{R}$. We will first show the series converge absolutely. For all $n \in \mathbb{N}$, let

$$a_n = \frac{(-1)^n}{(2n)!} x^{2n} \qquad \text{and} \qquad b_n = \frac{(-1)^n}{(2n+1)!} x^{2n+1}.$$

Since

$$\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = \lim_{n\to\infty} \frac{x^{2n+2}}{(2n+2)!} \frac{(2n)!}{x^{2n}} = \lim_{n\to\infty} \frac{x^2}{(2n+2)(2n+1)} = 0$$

and

$$\lim_{n\to\infty} \left|\frac{b_{n+1}}{b_n}\right| = \lim_{n\to\infty} \frac{x^{2n+3}}{(2n+3)!} \frac{(2n+1)!}{x^{2n+1}} = \lim_{n\to\infty} \frac{x^2}{(2n+3)(2n+2)} = 0,$$

the Ratio Test (Theorem 1.2.18) implies that both series converge absolutely.

To see the series converge to $\cos(x)$ and $\sin(x)$ respectively, notice since

$$i^k = \begin{cases} 1 & \text{if } k \equiv 0 \mod 4 \\ i & \text{if } k \equiv 1 \mod 4 \\ -1 & \text{if } k \equiv 2 \mod 4 \\ -i & \text{if } k \equiv 3 \mod 4 \end{cases}$$

that

$$\begin{aligned}\cos(x) + i\sin(x) &= e^{ix} \\ &= \lim_{N\to\infty} \sum_{k=0}^{4N-1} \frac{1}{k!}(ix)^k \\ &= \lim_{N\to\infty} \sum_{k=0}^{2N-1} \frac{(-1)^k}{(2k)!}x^{2k} + i\sum_{k=0}^{2N-1} \frac{(-1)^k}{(2k+1)!}x^{2k+1} \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!}x^{2n} + i\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!}x^{2n+1}.\end{aligned}$$

Therefore, by equating the real and imaginary parts, c) and d) follow.

Note e) clearly follows from c) and d). To see that f) holds, notice that

$$\cos(-\theta) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!}(-\theta)^{2n} = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!}\theta^{2n} = \cos(\theta)$$

and

$$\sin(-\theta) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!}(-\theta)^{2n+1} = \sum_{n=0}^{\infty}(-1)\frac{(-1)^n}{(2n+1)!}\theta^{2n+1} = -\sin(\theta)$$

as desired.

To see that g) and h) hold, let $\theta, \varphi \in \mathbb{R}$ be arbitrary. Then

$$\begin{aligned}&\cos(\theta+\varphi) + i\sin(\theta+\varphi) \\ &= e^{i(\theta+\varphi)} \\ &= e^{i\theta}e^{i\varphi} \quad \text{by Theorem 1.5.2} \\ &= (\cos(\theta) + i\sin(\theta))(\cos(\varphi) + i\sin(\varphi)) \\ &= (\cos(\theta)\cos(\varphi) - \sin(\theta)\sin(\varphi)) + i(\cos(\theta)\sin(\varphi) + \sin(\theta)\cos(\varphi)).\end{aligned}$$

Therefore, by equating the real and imaginary parts, g) and h) follow. ∎

Using the above, we obtain most of the essential elementary facts about cosine and sine. However, how do we obtain the traditional 'special angle' values of cosine and sine? Of course, if we could obtain $\cos(\pi) = -1$ and $\sin(\pi) = 0$, or $\cos(\frac{\pi}{2}) = 0$ and $\sin(\frac{\pi}{2}) = 1$, we could use Corollary 1.5.7 parts b), f), g), and h) to obtain all of the special angles. However, how do we obtain the above values? Moreover, can we obtain the usual facts about the continuity and derivatives of the exponential and trigonometric functions?

Chapter 2

Series of Functions

As seen in Chapter 1, for every $x \in \mathbb{R}$ the numbers e^x, $\cos(x)$, and $\sin(x)$ can all be defined via absolutely convergent series. In particular,

$$\begin{aligned} e^x &= \sum_{n=0}^{\infty} \frac{1}{n!} x^n, \\ \cos(x) &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n}, \text{ and} \\ \sin(x) &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1}. \end{aligned}$$

Viewing each of e^x, $\cos(x)$, and $\sin(x)$ as a function of x yields each function as an infinite series of functions (namely polynomials) in x. In this chapter, we will focus on some of the fundamental problems surrounding series of functions.

To begin, we know (but should show based on the above definitions) that cosine, sine, and the exponential functions are all continuous, differentiable functions. However, is it true that a series of continuous functions is continuous? Is it true that a series of differentiable functions is differentiable? If so, is the derivative of the series the infinite sum of the derivatives? Moreover, if we want to integrate a series of functions, can we integrate them term-by-term? Finally can we approximate every continuous function with a series or sequence of 'nicer' functions? If so, how can we find these approximating functions?

Understanding these behaviours and what can and cannot be done is essential to applying these results to various real-world applications. Thus we will delve into a precise, in-depth discussion of these questions in this chapter. In particular, we will show what results are true and give several examples illustrating the irredeemable behaviour if assumptions are dropped.

2.1 Continuity of Complex-Valued Functions

In order to discuss series of continuous functions, it is useful to recall the notion of a continuous function. We do so in the situation of functions on the complex numbers for use later in the course.

Definition 2.1.1. Let $\Omega \subseteq \mathbb{C}$. A function $f : \Omega \to \mathbb{C}$ is said to be *continuous at a point* $w \in \Omega$ if for all $\epsilon > 0$ there exists a $\delta > 0$ such that if $z \in \Omega$ and $|z - w| < \delta$ then $|f(z) - f(w)| < \epsilon$.

Moreover, f is said to be *continuous on* Ω if f is continuous at every point in Ω.

Remark 2.1.2. Again, just as Lemma 1.1.2 shows that the '$< \epsilon$' in the definition of a limit of a sequence can be replaced with '$\leq \epsilon$', one can replace one or both of '$< \delta$' and '$< \epsilon$' with '$\leq \delta$' and '$\leq \epsilon$' respectively in Definition 2.1.1 without modifying the notion of continuity.

Unsurprisingly, the notion of continuity for continuous functions on the complex numbers can also be characterized via limits of convergent sequences.

Lemma 2.1.3. *Let $\Omega \subseteq \mathbb{C}$, let $w \in \Omega$, and let $f : \Omega \to \mathbb{C}$. Then f is continuous at w if and only if whenever $(z_n)_{n\geq 1}$ is a sequence in Ω that converge to w, we have $\lim_{n\to\infty} f(z_n) = f(w)$.*

Proof. First suppose f is continuous at w. To see the desired result is true, let $(z_n)_{n\geq 1}$ be a sequence in Ω that converge to w. To see that $\lim_{n\to\infty} f(z_n) = f(w)$, let $\epsilon > 0$ be arbitrary. Since f is continuous at w, there exists a $\delta > 0$ such that if $z \in \Omega$ and $|z - w| < \delta$ then $|f(z) - f(w)| < \epsilon$. Since $(z_n)_{n\geq 1}$ converges to w, there exists an $N \in \mathbb{N}$ such that $|z_n - w| < \delta$ for all $n \geq N$. Hence for all $n \geq N$ we have that $|z_n - w| < \delta$ so $|f(z_n) - f(w)| < \epsilon$. Therefore, since $\epsilon > 0$ was arbitrary, $\lim_{n\to\infty} f(z_n) = f(w)$.

Conversely, suppose that f is not continuous at w. Therefore there exists an $\epsilon_0 > 0$ such that for all $\delta > 0$ there exists an $z \in \Omega$ such that $|z - w| < \delta$ but $|f(z) - f(w)| \geq \epsilon_0$. Hence for all $n \in \mathbb{N}$ there exists a $z_n \in \Omega$ such that $|z_n - w| < \frac{1}{n}$ but $|f(z_n) - f(w)| \geq \epsilon_0$. Thus $(z_n)_{n\geq 1}$ is a sequence in Ω that converges to w such that $(f(z_n))_{n\geq 1}$ does not converge to $f(w)$ since $|f(z_n) - f(w)| \geq \epsilon_0$ for all $n \in \mathbb{N}$ so Definition 1.1.1 fails for $\epsilon = \epsilon_0$. ∎

Consequently, a composition of continuous functions is continuous.

Proposition 2.1.4. *Let $\Omega \subseteq \mathbb{C}$, let $z_0 \in \Omega$, let $f : \Omega \to \mathbb{C}$, and let $g : Range(f) \to \mathbb{C}$. If f is continuous at z_0 and g is continuous at $f(z_0)$, then $g \circ f$ is continuous at z_0.*

Proof. Let $(z_n)_{n\geq 1}$ be an arbitrary sequence in Ω that converges to z_0. Since f is continuous at z_0 and $(z_n)_{n\geq 1}$ converges to z_0, $(f_n(z_0))_{n\geq 1}$ converges to $f(z_0)$ by Lemma 2.1.3. Similarly, since g is continuous at $f(z_0)$ and $(f_n(z_0))_{n\geq 1}$

converges to $f(z_0)$, $(g(f(z_n)))_{n\geq 1}$ converges to $g(f(z_0))$ by Lemma 2.1.3. Therefore, since $(z_n)_{n\geq 1}$ was arbitrary, Lemma 2.1.3 implies that $g \circ f$ is continuous at z_0. ∎

In addition, the following shows that the set of continuous functions on the complex numbers is a vector subspace of the vector space of all complex-valued functions.

Lemma 2.1.5. *Let $\Omega \subseteq \mathbb{C}$ and let $f : \Omega \to \mathbb{C}$ and $g : \Omega \to \mathbb{C}$ be continuous functions. Then the following are true:*

a) The function $f + g : \Omega \to \mathbb{C}$ defined by

$$(f + g)(z) = f(z) + g(z)$$

for all $z \in \Omega$ is continuous.

b) The function $fg : \Omega \to \mathbb{C}$ defined by

$$(fg)(z) = f(z)g(z)$$

for all $z \in \Omega$ is continuous.

c) For all $\alpha \in \mathbb{C}$, the function $\alpha f : \Omega \to \mathbb{C}$ defined by

$$(\alpha f)(z) = \alpha f(z)$$

for all $z \in \Omega$ is continuous.

d) If $f(z) \neq 0$ for all $z \in \Omega$, the function $\frac{1}{f} : \Omega \to \mathbb{C}$ defined by

$$\left(\frac{1}{f}\right)(z) = \frac{1}{f(z)}$$

for all $z \in \Omega$ is continuous.

e) The function $\overline{f} : \Omega \to \mathbb{C}$ defined by

$$\overline{f}(z) = \overline{f(z)}$$

for all $z \in \Omega$ is continuous.

f) The function $|f| : \Omega \to \mathbb{C}$ defined by

$$|f|(z) = |f(z)|$$

for all $z \in \Omega$ is continuous.

Proof. This immediately follows from Lemma 2.1.3 by using Corollary 1.1.7. ∎

As with complex numbers, the ability to take the real and imaginary part of complex-valued functions is important.

Definition 2.1.6. Let $\Omega \subseteq \mathbb{C}$ and let $f : \Omega \to \mathbb{C}$. The *real and imaginary parts of f* are the functions $\mathrm{Re}(f), \mathrm{Im}(f) : \Omega \to \mathbb{R}$ respectively where

$$(\mathrm{Re}(f))(x) = \mathrm{Re}(f(x)) = \frac{f(x) + \overline{f(x)}}{2} \quad \text{and}$$
$$(\mathrm{Im}(f))(x) = \mathrm{Im}(f(x)) = \frac{f(x) - \overline{f(x)}}{2i}$$

for all $x \in \Omega$.

Notice if $f : \Omega \to \mathbb{C}$, then $f = \mathrm{Re}(f) + iIm(f)$. Moreover

$$\mathrm{Re}(f) = \frac{f + \overline{f}}{2} \quad \text{and} \quad \mathrm{Im}(f) = \frac{f - \overline{f}}{2i}$$

Thus Lemma 2.1.5 implies the following.

Lemma 2.1.7. *Let $\Omega \subseteq \mathbb{C}$ and let $f : \Omega \to \mathbb{C}$. Then f is continuous if and only if $\mathrm{Re}(f)$ and $\mathrm{Im}(f)$ are continuous real-valued functions.*

Proof. If $\mathrm{Re}(f)$ and $\mathrm{Im}(f)$ are continuous, then $= \mathrm{Re}(f) + iIm(f)$ is continuous by Lemma 2.1.5, parts a) and c).

Similarly, if f is continuous, then

$$\mathrm{Re}(f) = \frac{f + \overline{f}}{2} \quad \text{and} \quad \mathrm{Im}(f) = \frac{f - \overline{f}}{2i}$$

are continuous by Lemma 2.1.5, parts a), c), and e). ∎

2.2 Continuity of Series of Functions

With our reminders of continuous functions out of the way, let us examine the question of whether a series of continuous functions is continuous. To answer this question we note since a series is a limit of its partial sums that a series of functions is a limit of a sequences of functions. Consequently, it is necessary to discuss limits of sequences of functions and whether a limit of continuous functions is continuous. We begin with the most obvious way to define the limit of a function and its corresponding restriction to a series of functions.

Definition 2.2.1. Let $\Omega \subseteq \mathbb{C}$. A sequence $(f_n)_{n \geq 1}$ of complex-valued functions on Ω is said to *converge pointwise on* Ω to $f : \Omega \to \mathbb{C}$ if

$$f(z) = \lim_{n \to \infty} f_n(z)$$

for all $z \in \Omega$.

Definition 2.2.2. Let $\Omega \subseteq \mathbb{C}$ and for each $n \in \mathbb{N}$, let $f_n : \Omega \to \mathbb{C}$. The series $\sum_{n=1}^{\infty} f_n$ is said to *converge pointwise on* Ω if $\sum_{n=1}^{\infty} f_n(z)$ converges for each $z \in \Omega$. Moreover the function $f : \Omega \to \mathbb{C}$ defined by

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

for all $z \in \mathbb{C}$ is called the *(pointwise) sum of* $(f_n)_{n\geq 1}$ and is denoted by $\sum_{n=1}^{\infty} f_n$.

Unfortunately, pointwise limits are not the limits we are looking for as a pointwise limit of continuous functions is not continuous. Thus a pointwise convergent series of continuous functions need not be continuous.

Example 2.2.3. For each $n \in \mathbb{N}$, define $f_n : [0,1] \to [0,1]$ by $f_1(x) = x$ and, for $n \geq 2$,

$$f_n(x) = x^n - x^{n-1}$$

for all $x \in [0,1]$. Clearly $(f_n)_{n\geq 1}$ is a sequence of continuous functions on $[0,1]$.

We claim the series $\sum_{n=1}^{\infty} f_n$ converges pointwise on $[0,1]$ to the function $f : [0,1] \to [0,1]$ defined by

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

As f is clearly not continuous at $x = 1$, this provides an example of a series of continuous functions that converges pointwise to a function that is discontinuous at a point.

To see the above, for each $N \in \mathbb{N}$ let $S_N = \sum_{k=1}^{N} f_k$. Notice that $S_n(x) = x^N$ for all $N \in \mathbb{N}$. Note if $x = 1$ then

$$\lim_{N\to\infty} S_N(x) = \lim_{N\to\infty} 1^N = 1$$

whereas if $x \in [0,1)$ then

$$\lim_{N\to\infty} S_N(x) = \lim_{N\to\infty} x^N = 0.$$

Hence the example is complete. Note this also show that there exists a sequence (namely $(x^n)_{n\geq 1}$) of continuous functions that converge pointwise to a function that is discontinuous at a point.

In order to rectify this situation, we simply need to require a stronger form of limit of functions.

Definition 2.2.4. Let $\Omega \subseteq \mathbb{C}$. A sequence $(f_n)_{n\geq 1}$ of complex-valued functions on Ω is said to *converge uniformly on* Ω to $f : \Omega \to \mathbb{C}$ if for all $\epsilon > 0$ there exists an $N \in \mathbb{N}$ such that if $n \geq N$ then $|f_n(z) - f(z)| \leq \epsilon$ for all $z \in \Omega$; that is,

$$\sup(\{|f_n(z) - f(z)| \mid z \in \Omega\}) \leq \epsilon$$

for all $n \geq N$.

Again, using arguments similar to those used in Lemma 1.1.2 the '$\leq \epsilon$' in the Definition 2.2.4 can be replaced with '$< \epsilon$'.

Remark 2.2.5. The notion of uniform convergence of sequences of functions can be rephrased in an analogous way to how the convergence of sequences of complex numbers was defined. Indeed instead of using the absolute value as a notion of distance, if $\mathcal{F}(\Omega)$ denotes the set of all bounded complex-valued functions on a set $\Omega \subseteq \mathbb{C}$, one can define a distance function (or metric) $d : \mathcal{F}(\Omega) \times \mathcal{F}(\Omega) \to [0, \infty)$ by

$$d(f, g) = \sup(\{|f(z) - g(z)| \mid z \in \Omega\})$$

for all $f, g \in \mathcal{F}(\Omega)$. Then a sequence $(f_n)_{n\geq 1}$ in $\mathcal{F}(\Omega)$ converges uniformly to $f \in \mathcal{F}(\Omega)$ if for all $\epsilon > 0$ there exists an $N \in \mathbb{N}$ such that $d(f_n, f) < \epsilon$ for all $n \geq N$. This notion of a distance function and convergence of sequences with respect to a distance function can be further generalized. But that is a topic for a future course (i.e. MATH 4011 - Analysis IIIA: Metric Spaces).

Remark 2.2.6. It is important to point out the difference between pointwise convergence and uniform convergence. The main difference is, given an $\epsilon > 0$, pointwise convergence simply lets us find for each $z \in \Omega$ an $N_Z \in \mathbb{N}$ that depends on z such that $|f_n(z) - f(z)| \leq \epsilon$ for all $n \geq N_Z$ whereas uniform convergence lets us find an $N \in \mathbb{N}$ that works for every $z \in \Omega$; that is, $|f_n(z) - f(z)| \leq \epsilon$ for all $n \geq N$ and $z \in \Omega$. More elegantly said uniform convergence lets us find one N to rule all N_z; that is

$$N = \sup(\{N_z \mid z \in \Omega\}) < \infty.$$

Note this clearly implies if a sequence of functions converges uniformly to f, then they converge pointwise to f. However, if a sequence of functions converges pointwise, it need not converge uniformly as the following example shows.

Example 2.2.7. Recall from Example 2.2.3 that if $f_n : [0, 1] \to [0, 1]$ is defined by $f_n(x) = x^n$ for all $x \in [0, 1]$ and $n \in \mathbb{N}$, the sequence $(f_n)_{n\geq 1}$ converges pointwise to $f : [0, 1] \to [0, 1]$ defined by

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

We claim that $(f_n)_{n\geq 1}$ does not converge uniformly to f on $[0, 1]$. To see this, let $\epsilon = \frac{1}{2} > 0$. Suppose there exists an $N \in \mathbb{N}$ such that $|f_n(x) - f(x)| < \frac{1}{2}$ for all $n \geq N$ and $x \in [0, 1]$. Since f_N is continuous at 1, there exists a $\delta > 0$ such that if $x \in [0, 1]$ and $|x - 1| < \delta$, then $|f_N(x) - f_N(1)| < \frac{1}{2}$. Hence, for any $x_0 \in \left(1 - \frac{\delta}{2}, 1\right)$ we have that

$$|x_0^N - 1| = |f_N(x) - f_N(1)| < \frac{1}{2}$$

whereas the above implies that

$$|x_0^N| = |x_0^N - f(x_0)| < \frac{1}{2}.$$

Since the first implies $x_0^N \in \left(\frac{1}{2}, \frac{3}{2}\right)$ whereas the second implies $x_0^N \in \left(-\frac{1}{2}, \frac{1}{2}\right)$, we have a contradiction. Hence $(f_n)_{n\geq 1}$ does not converge uniformly to f on $[0, 1]$.

However, the only pathology in the above example surrounds the value of the functions at $x = 1$.

Example 2.2.8. Fix $b \in [0, 1)$ and consider the functions $f_n : [0, b] \to [0, b]$ is defined by $f_n(x) = x^n$ for all $x \in [0, b]$ and $n \in \mathbb{N}$. We claim that $(f_n)_{n\geq 1}$ converge uniformly to 0 (the zero function) on $[0, b]$. To see this, let $\epsilon > 0$ be arbitrary. Since $\lim_{n\to\infty} b^n = 0$, there exists an $N \in \mathbb{N}$ such that $|b^n| < \epsilon$ for all $n \geq N$. Hence for all $x \in [0, b]$ and $n \geq N$ we see that

$$|f_n(x) - 0| = x^n \leq b^n < \epsilon.$$

Therefore, as $\epsilon > 0$ was arbitrary, $(f_n)_{n\geq 1}$ converge uniformly to 0 on $[0, b]$.

As promised, the following result shows continuity behaves properly when uniform limits are used. It is also useful to point out that the proof of the following result is a very common and incredibly useful argument in analysis known as a three-ϵ argument.

Theorem 2.2.9. *Let $\Omega \subseteq \mathbb{C}$, let $w \in \Omega$, and let $(f_n)_{n\geq 1}$ be a sequence of complex-valued functions on Ω that converge uniformly on Ω to $f : \Omega \to \mathbb{C}$. If each f_n is continuous at w, then f is continuous at w.*

Consequently, a uniform limit of continuous functions is continuous!

Proof. To see that f is continuous at w, let $\epsilon > 0$ be arbitrary. Since $(f_n)_{n\geq 1}$ converges to f uniformly on Ω, there exists an $N \in \mathbb{N}$ such that

$$|f_n(z) - f(z)| < \frac{\epsilon}{3}$$

for all $n \geq N$ and $z \in \Omega$. Since f_N is continuous at w, there exists a $\delta > 0$ such that if $z \in \Omega$ and $|z - w| < \delta$, then

$$|f_N(z) - f_N(w)| < \frac{\epsilon}{3}.$$

Hence, for all $z \in \Omega$ such that $|z - w| < \delta$, we have

$$\begin{aligned} |f(z) - f(w)| &\leq |f(z) - f_N(z)| + |f_N(z) - f_N(w)| + |f_N(w) - f(w)| \\ &< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon \end{aligned}$$

(where the first and last terms are less than $\frac{\epsilon}{3}$ by uniform convergence and the middle term is less than $\frac{\epsilon}{3}$ by continuity of f_N). Therefore, as $\epsilon > 0$ was arbitrary, f is continuous at w. ∎

To further emphasize the benefits of uniformly convergent continuous functions, we desire an analogue of 'every convergent sequence is bounded'. To do so, we define the following notion of boundedness for a sequence of functions.

Definition 2.2.10. Let $\Omega \subseteq \mathbb{C}$ and let $(f_n)_{n\geq 1}$ be a sequence of complex-valued functions on Ω. It is said that $(f_n)_{n\geq 1}$ is *uniformly bounded on* Ω if there exists an $M \in \mathbb{R}$ such that

$$|f_n(z)| \leq M$$

for all $z \in \Omega$ and $n \in \mathbb{N}$.

Proposition 2.2.11. *Let I be a bounded closed interval in $\mathbb{R}$ and let $(f_n)_{n\geq 1}$ be a sequence of complex-valued continuous functions on I that converge uniformly on I to $f : I \to \mathbb{C}$. Then $(f_n)_{n\geq 1}$ is uniformly bounded.*

Proof. Since I is a closed interval, the Extreme Value Theorem implies there exists an $M_n \in \mathbb{N}$ such that

$$|f_n(x)| \leq M_n$$

for all $x \in I$. Moreover, since $(f_n)_{n\geq 1}$ converge uniformly to f on I, Theorem 2.2.9 implies that f is continuous on I. Hence the Extreme Value Theorem implies there exists an $M_0 \in \mathbb{R}$ such that $|f(x)| \leq M_0$ for all $x \in I$.

Since $(f_n)_{n\geq 1}$ converges uniformly to f on I, there exists an N such that

$$|f_n(x) - f(x)| \leq 1$$

for all $n \geq N$ and $x \in I$ and thus

$$|f_n(x)| \leq M_0 + 1$$

for all $n \geq N$ and $x \in I$. Therefore, if

$$M = \max\{M_0 + 1, M_1, M_2, \ldots, M_N\},$$

then

$$|f_n(x)| \leq M$$

for all $x \in I$ and $n \in \mathbb{N}$. Hence $(f_n)_{n\geq 1}$ is uniformly bounded. ∎

Moreover, uniform convergence behaves well with respect to the operations of addition and scalar multiplication. The proofs of these facts are similar to the proof of Corollary 1.1.7 once combined with Proposition 2.2.11.

Proposition 2.2.12. *Let I be a bounded closed interval in $\mathbb{R}$ and let $(f_n)_{n\geq 1}$ and $(g_n)_{n\geq 1}$ be sequences of complex-valued continuous functions on I that converge uniformly on I to $f : I \to \mathbb{C}$ and $g : I \to \mathbb{C}$ respectively. Then the following are true:*

a) $(f_n + g_n)_{n\geq 1}$ converges uniformly to $f + g$ on I.

b) $(f_n g_n)_{n\geq 1}$ converges uniformly to fg on I.

c) If $(\alpha_n)_{n\geq 1}$ is a sequence of complex numbers that converges to $\alpha \in \mathbb{C}$, then $(\alpha_n f_n)_{n\geq 1}$ converges uniformly to αf on I.

d) $\left(\overline{f_n}\right)_{n\geq 1}$ converges uniformly to $\overline{f}$ on I.

Proof. Exercise. ∎

As Theorem 2.2.9 shows that uniform convergence is the limit we are looking for in order to preserve continuity of functions, we now discuss uniform convergence in the context of series.

Definition 2.2.13. Let $\Omega \subseteq \mathbb{C}$ and let $(f_n)_{n\geq 1}$ be a sequence of complex-valued functions on Ω. For each $N \in \mathbb{N}$, define $S_N : \Omega \to \mathbb{C}$ by

$$S_N(z) = \sum_{k=1}^{N} f_k(z).$$

The series $\sum_{n=1}^{\infty} f_n$ is said to *converge uniformly on* Ω to a function $f : \Omega \to \mathbb{C}$ if $(S_N)_{N\geq 1}$ converges uniformly to f on Ω.

Corollary 2.2.14. *Let $\Omega \subseteq \mathbb{C}$ and let $(f_n)_{n\geq 1}$ be a sequence of continuous complex-valued functions on Ω. If the series $\sum_{n=1}^{\infty} f_n$ converges uniformly on Ω to a function $f : \Omega \to \mathbb{C}$, then f is continuous.*

Proof. Since f_n is continuous for all $n \in \mathbb{N}$, the partial sums of $\sum_{n=1}^{\infty} f_n$ are continuous by Lemma 2.1.5. Hence f is continuous by Theorem 2.2.9 being the uniform limit of continuous functions. ∎

The simplest way to verify many series of functions converges uniformly is the following 'test', which should remind the reader of absolutely convergence of series of scalars. In particular, the following result shows that if a series of functions 'uniformly converges pointwise absolutely', then the series converges uniformly.

Theorem 2.2.15 (Weierstrass M-Test). *Let $\Omega \subseteq \mathbb{C}$ and let $(f_n)_{n\geq 1}$ be a sequence of complex-valued functions on Ω. For each $n \in \mathbb{N}$ suppose*

$$0 \leq M_n = \sup(\{|f_n(z)| \mid z \in \Omega\}) < \infty.$$

Furthermore, suppose $\sum_{n=1}^{\infty} M_n$ converges. Then $\sum_{n=1}^{\infty} f_n(z)$ converges absolutely for all $z \in \Omega$ and if $f : \Omega \to \mathbb{C}$ is defined by

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

for all $z \in \Omega$, then $\sum_{n=1}^{\infty} f_n$ converges uniformly to f.

Proof. To see that $\sum_{n=1}^{\infty} f_n(z)$ converges absolutely for all $z \in \Omega$, note for all $z \in \Omega$ that

$$|f_n(z)| \leq M_n.$$

Since $\sum_{n=1}^{\infty} M_n$ converges, $\sum_{n=1}^{\infty} f_n(z)$ converges absolutely for all $z \in \Omega$ by the Comparison Test (Theorem 1.2.13).

To see that $\sum_{n=1}^{\infty} f_n$ converges uniformly to f, let $\epsilon > 0$ be arbitrary. Since $\sum_{n=1}^{\infty} M_n$ converges, Theorem 1.2.9 implies there exists an $N_0 \in \mathbb{N}$ such that

$$\sum_{n=N_0}^{\infty} M_n < \epsilon.$$

Hence for all $N \geq N_0$ and $z \in \Omega$, we see that

$$\begin{aligned}
\left| \sum_{k=1}^{N} f_k(z) - \sum_{k=1}^{\infty} f_k(z) \right| &= \left| \sum_{k=N+1}^{\infty} f_k(z) \right| \\
&\leq \sum_{k=N+1}^{\infty} |f_k(z)| && \text{Theorem 1.2.11} \\
&\leq \sum_{k=N+1}^{\infty} M_k && \text{assumptions} \\
&\leq \sum_{k=N_0}^{\infty} M_k < \epsilon.
\end{aligned}$$

Therefore, since $\epsilon > 0$ was arbitrary, $\sum_{n=1}^{\infty} f_n$ converges uniformly to f. ∎

To demonstrate the power of the Weierstrass M-Test (Theorem 2.2.15), we demonstrate the following series of functions from Chapter 1 converge uniformly and thus define continuous functions.

Example 2.2.16. For each $n \in \mathbb{N} \cup \{0\}$, let $f_n : \mathbb{C} \to \mathbb{C}$ be defined by

$$f_n(z) = \frac{1}{n!} z^n$$

for all $z \in \mathbb{C}$. Clearly each f_n is a continuous function such that $\sum_{n=0}^{\infty} f_n(z) = e^z$ for all $z \in \mathbb{C}$.

For each $M \in \mathbb{N}$, let $\Omega_M = \{z \in \mathbb{C} \mid |z| \leq M\}$. Since

$$|f_n(z)| \leq \frac{1}{n!} M^n$$

for all $z \in \Omega_M$, and since $\sum_{n=0}^{\infty} \frac{1}{n!} M^n < \infty$ by Example 1.2.20, the Weierstrass M-Test (Theorem 2.2.15) implies that

$$e^z = \sum_{n=0}^{\infty} f_n(z)$$

is continuous on Ω_M. Since $M \in \mathbb{N}$ was arbitrary, e^z is a continuous function on $\bigcup_{M=1}^{\infty} \Omega_M = \mathbb{C}$.

Example 2.2.17. Consider the function $f : \mathbb{R} \to \mathbb{C}$ defined by

$$f(x) = e^{ix}$$

for all $x \in \mathbb{R}$. Since f is the restriction to $\mathbb{R}$ of a continuous function on $\mathbb{C}$ by Example 2.2.16, we obtain that f is continuous on $\mathbb{R}$. Since

$$(\mathrm{Re}(f))(x) = \cos(x) \qquad \text{and} \qquad (\mathrm{Im}(f))(x) = \sin(x)$$

for all $x \in \mathbb{R}$, we obtain that cos and sin are continuous functions by Lemma 2.1.7.

2.3 Continuous, Nowhere Differentiable Functions

Now that we know a pointwise series of continuous functions need not be continuous and a uniform series of continuous functions is continuous, we turn our attention to differentiability of series of differentiable functions. Since a differentiable function is automatically continuous, we know a pointwise limit of differentiable functions need not be differentiable. However, is a uniform convergent series of differentiable functions differentiable?

It turns out that the answer to this question is no. In particular, in this section we will give a family of examples of uniformly convergent series of differentiable functions that are continuous but not differentiable at any point in $\mathbb{R}$! Such function are said to be *nowhere differentiable*.

Before we get to the example of a uniformly convergent series of differentiable functions that is nowhere differentiable, we will use simply the Weierstrass M-Test (Theorem 2.2.15) to construct a continuous but nowhere differentiable function thereby motivating the procedure we will use for the primary examples of this section. To construct this example, we require the following.

Definition 2.3.1. A function $f : \mathbb{R} \to \mathbb{R}$ is said to be *periodic with period* $c > 0$ if $f(x + c) = f(x)$ for all $x \in \mathbb{R}$.

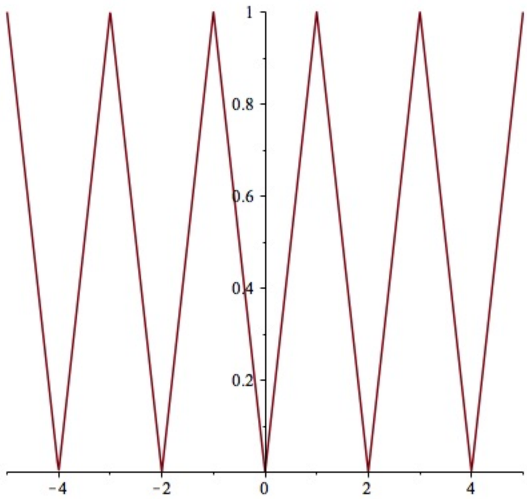
Remark 2.3.2. Fix $a, b \in \mathbb{R}$ with $a < b$. Note for each $x \in \mathbb{R}$ there exists a unique $y_x \in [a, b)$ and $n_x \in \mathbb{Z}$ such that

$$x = y_x + (b - a)n_x.$$

Therefore, if $f : [a, b) \to \mathbb{R}$, and we define $g : \mathbb{R} \to \mathbb{R}$ by $g(x) = f(y_x)$ for all $x \in \mathbb{R}$ where y_x is as above, then g is periodic with period $b - a$ and $g(x) = f(x)$ for all $x \in [a, b]$. We call g the *periodic extension of* f.

Let us now construct our first continuous but nowhere differentiable function.

Example 2.3.3. Let $h : [-1, 1] \to [0, 1]$ be defined by $h(x) = |x|$ for all $x \in [-1, 1]$. Since $h(-1) = h(1)$, we can define the periodic extension $g : \mathbb{R} \to [0, 1]$ of h and g is a continuous function with period 2. A portion of the graph of g can be seen in the following figure.



For any $n \in \mathbb{Z}$, it is not difficult to see that g is a line on $[n, n + 1]$ with either slope 1 or -1. In addition, we claim for all $x, y \in \mathbb{R}$ that

$$|g(x) - g(y)| \leq |x - y|.$$

To see this, first note if $|x - y| \geq 2$ then

$$|g(x) - g(y)| \leq |g(x)| + |g(y)| \leq 1 + 1 = 2 \leq |x - y|.$$

Otherwise, if $|x - y| < 2$, since g is 2-periodic, we can assume that $x, y \in [-1, 1]$. Therefore

$$|g(x) - g(y)| = ||x| - |y|| \leq |x - y|$$

by the reverse triangle inequality. Hence the claim has been demonstrated.

For all $n \in \mathbb{N}$ and $x \in \mathbb{R}$, notice that

$$\left|\left(\frac{3}{4}\right)^n g\left(4^n x\right)\right| \leq \left(\frac{3}{4}\right)^n.$$

Therefore, since g is a continuous function and the geometric series

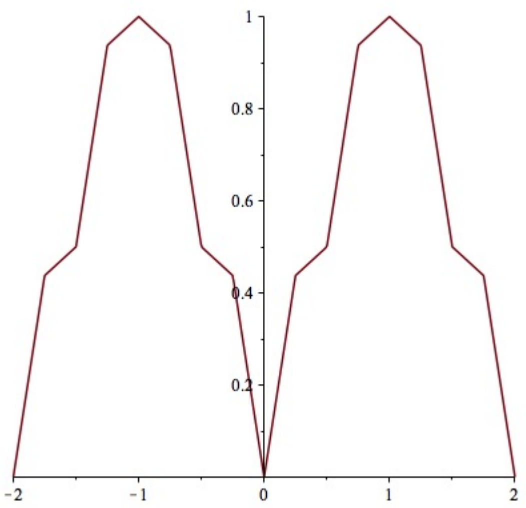
$$\sum_{n=0}^{\infty} \left(\frac{3}{4}\right)^n$$

converges, the Weierstrass M-Test (Theorem 2.2.15) implies that the function $f : \mathbb{R} \to \mathbb{R}$ defined by

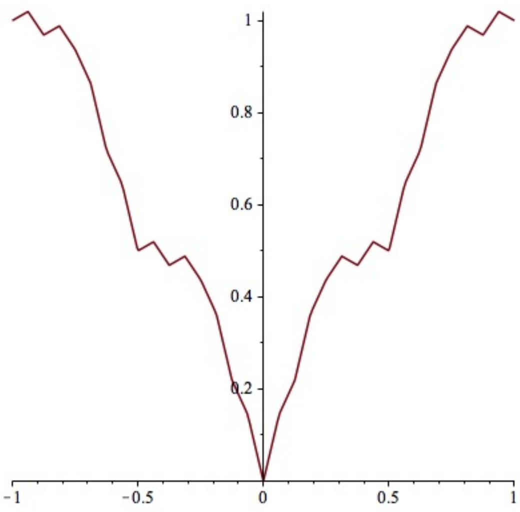
$$f(x) = \sum_{n=0}^{\infty} \left(\frac{3}{4}\right)^n g\left(4^n x\right)$$

for all $x \in \mathbb{R}$ exists, is continuous, and this sum converges uniformly.

To get an idea of what the graph of f looks like, the following is a portion of the graph of $g(x) + \frac{3}{4}g\left(4x\right)$



and the following is a portion of the graph of $g(x) + \frac{3}{4}g\left(4x\right) + \frac{9}{16}g(16x)$.



To see that f is nowhere differentiable, fix $x_0 \in \mathbb{R}$. To show that f is not differentiable at x_0, we will show that

$$\lim_{h\to\infty} \left| \frac{f(x_0+h) - f(x_0)}{h} \right|$$

does not exist by constructing a sequence $(h_m)_{m\geq 1}$ of non-zero numbers such that $\lim_{m\to\infty} h_m = 0$ yet

$$\lim_{m\to\infty} \left| \frac{f(x_0+h_m) - f(x_0)}{h_m} \right| = \infty.$$

For all $m \in \mathbb{N}$, notice that $\left[4^m x_0 - \frac{1}{2}, 4^m x_0 + \frac{1}{2}\right]$ has length 1. Therefore, only one of $\left(4^m x_0 - \frac{1}{2}, 4^m x_0\right)$ and $\left(4^m x_0, 4^m x_0 + \frac{1}{2}\right)$ can contain an integer. Let

$$h_m = \begin{cases} \frac{1}{2}4^{-m} & \text{if } \left(4^m x_0, 4^m x_0 + \frac{1}{2}\right) \cap \mathbb{Z} = \emptyset \\ -\frac{1}{2}4^{-m} & \text{if } \left(4^m x_0 - \frac{1}{2}, 4^m x_0\right) \cap \mathbb{Z} = \emptyset \end{cases}$$

To obtain a lower bound on $\left|\frac{f(x_0+h_m)-f(x_0)}{h_m}\right|$, for all $n, m \in \mathbb{N}$ let

$$D_{n,m} = \left| \frac{g\left(4^n(x_0+h_m)\right) - g\left(4^n x_0\right)}{h_m} \right|.$$

Let us compute some estimates on the value of $D_{n,m}$ depending on the relative sizes of m and n.

Case 1: $n > m$. In this case, notice that

$$4^n(x_0+h_m) = 4^n x_0 \pm \frac{1}{2}4^{n-m}.$$

Since $n > m$, $\frac{1}{2}4^{n-m}$ is an integer multiple of 2. Therefore, since g has period 2, we see that

$$g\left(4^n(x_0 + h_m)\right) = g\left(4^n x_0\right)$$

and thus $D_{n,m} = 0$ in this case.

Case 2: $n = m$. In this case, we see that

$$4^n(x_0 + h_m) = 4^n x_0 \pm \frac{1}{2}.$$

Therefore, since h_m was chosen so that there are no integers strictly between $4^n x_0$ and $4^n(x_0 + h_m)$, we obtain that g is a line with slope 1 or -1 between $4^n x_0$ and $4^n(x_0 + h_m)$. Hence

$$D_{n,m} = \left| \frac{g\left(4^n(x_0 + h_m)\right) - g\left(4^n x_0\right)}{h_m} \right| = \frac{|4^n(x_0 + h_m) - 4^n x_0|}{|h_m|} = \frac{\frac{1}{2}}{\frac{1}{2}4^{-n}} = 4^n.$$

Case 3: $n < m$. In this case, since $|g(x)| \leq 1$ for all $x \in \mathbb{R}$, we see that

$$D_{n,m} = \left| \frac{g\left(4^n(x_0 + h_m)\right) - g\left(4^n x_0\right)}{h_m} \right| \leq \frac{|4^n(x_0 + h_m) - 4^n x_0|}{|h_m|} = \frac{\frac{1}{2}4^{n-m}}{\frac{1}{2}4^{-m}} = 4^n.$$

Using the above three cases, we see for all $m \in \mathbb{N}$ that

$$\begin{aligned}
\left| \frac{f(x_0 + h_m) - f(x_0)}{h_m} \right| &= \left| \sum_{n=0}^{\infty} \left(\frac{3}{4}\right)^n \frac{g\left(4^n(x_0 + h_m)\right) - g\left(4^n x_0\right)}{h_m} \right| \\
&= \left| \sum_{n=0}^{m} \left(\frac{3}{4}\right)^n \frac{g\left(4^n(x_0 + h_m)\right) - g\left(4^n x_0\right)}{h_m} \right| \\
&\geq \left(\frac{3}{4}\right)^m D_{m,m} - \sum_{n=0}^{m-1} \left(\frac{3}{4}\right)^n D_{n,m} \\
&\geq \left(\frac{3}{4}\right)^m 4^m - \sum_{n=0}^{m-1} \left(\frac{3}{4}\right)^n 4^n \\
&= 3^m - \sum_{n=0}^{m-1} 3^n \\
&= 3^m - \frac{3^m - 1}{3 - 1} \\
&= \frac{3^m + 1}{2}.
\end{aligned}$$

Therefore

$$\lim_{m \to \infty} \left| \frac{f(x_0 + h_m) - f(x_0)}{h_m} \right| = \infty,$$

so f cannot be differentiable at x_0. Therefore, since x_0 was arbitrary, f is continuous but nowhere differentiable.

With the above example of a continuous but nowhere differentiable function, we return to our primary example of the section of a family of uniformly convergent series of differentiable functions that are nowhere differentiable. These famous examples were first demonstrated by Weierstrass in 1872.

The only tools we need to construct this family of functions is the Weierstrass M-Test (Theorem 2.2.15), the elementary properties of cosine and sine from Corollary 1.5.7, and knowledge about the derivatives of cosine and sine. The last of these three requirements will be shown in Example 2.6.6. Thus, for logical consistency, perhaps this examples should be discussed after Example 2.6.6. However, for educational purposes, it is best to introduce this example now in order to motivate the subsequent sections and enhance the previous section (and it is expected the reader is already familiar with the derivatives of the basic trigonometric functions and thus will excuse this minor logical gap).

Example 2.3.4 (Weierstrass, 1872). Fix $a, b \in \mathbb{R}$ such that a is a positive odd integer, $0 < b < 1$, and $ab > 1 + \frac{3}{2}\pi$ (e.g. $a = 13$ and $b = \frac{1}{2}$). For all $n \in \mathbb{N}$ and $x \in \mathbb{R}$, notice that

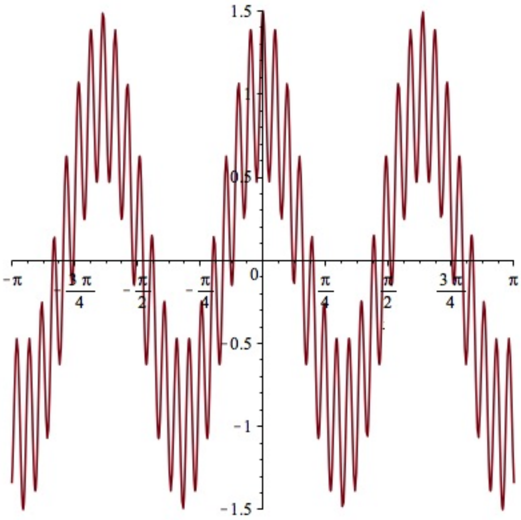
$$|b^n \cos(\pi a^n x)| \leq b^n.$$

Therefore, since $x \mapsto b^n \cos(\pi a^n x)$ is a continuous function and the geometric series $\sum_{n=0}^{\infty} b^n$ converges since $0 < b < 1$, the Weierstrass M-Test (Theorem 2.2.15) implies that the function $W : \mathbb{R} \to \mathbb{R}$ defined by

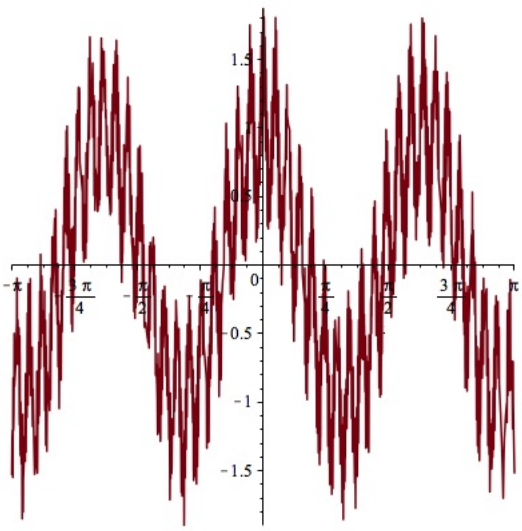
$$W(x) = \sum_{n=0}^{\infty} b^n \cos(\pi a^n x)$$

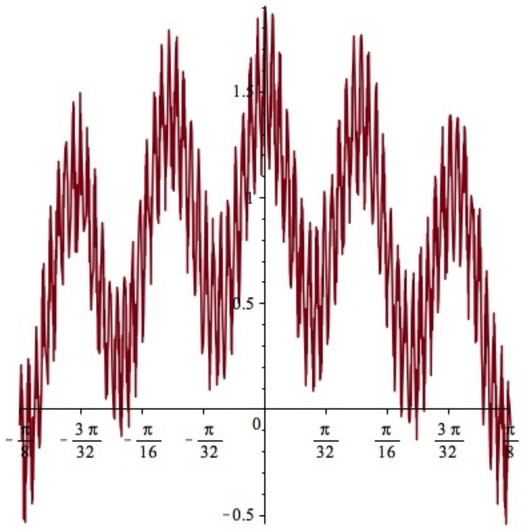
for all $x \in \mathbb{R}$ exists, is continuous, and this sum converges uniformly. The function W is known as a *Weierstrass function*.

To get an idea of what the graph of W looks like, the following is a portion of the graph of $\sum_{n=0}^{1} b^n \cos(\pi a^n x)$ when $a = 13$ and $b = \frac{1}{2}$



and the following is a portion of the graphs of $\sum_{n=0}^{10} b^n \cos(\pi a^n x)$ when $a = 13$ and $b = \frac{1}{2}$.





We claim that although W is continuous and a uniform limit of (infinitely) differentiable functions, W is not differentiable at any point in $\mathbb{R}$. This then completes the claim that there are uniformly convergent series of differentiable functions that are nowhere differentiable.

To see that W is nowhere differentiable, fix $x_0 \in \mathbb{R}$. To show that W is not differentiable at x_0, we will show that

$$\lim_{x \to x_0} \left| \frac{W(x) - W(x_0)}{x - x_0} \right|$$

does not exist by constructing a sequence $(x_m)_{m \geq 1}$ of real numbers such that $x_m \neq x_0$ for all m, $\lim_{m\to\infty} x_m = x_0$, and

$$\lim_{m \to \infty} \left| \frac{W(x_m) - W(x_0)}{x_m - x_0} \right| = \infty.$$

For all $m \in \mathbb{N}$, notice that the interval $\left[a^m x_0 + \frac{1}{2}, a^m x_0 + \frac{3}{2}\right)$ has length exactly 1 and therefore there exists a unique integer

$$\ell_m \in \mathbb{Z} \cap \left[a^m x_0 + \frac{1}{2}, a^m x_0 + \frac{3}{2}\right).$$

Notice this implies

$$\frac{1}{2} \leq \ell_m - a^m x_0 < \frac{3}{2}.$$

Since $0 < b < 1$ and $ab > 1 + \frac{3}{2}\pi$, we know that $a > 1$. Hence if $x_m = \frac{\ell_m}{a^m}$ we have that

$$\frac{1}{2a^m} \leq x_m - x_0 < \frac{3}{2a^m}$$

and thus $x_m > x_0$. Moreover, since $\lim_{m\to\infty} \frac{1}{2a^m} = 0 = \lim_{m\to\infty} \frac{3}{2a^m}$, the above inequality and the Squeeze Theorem imply that $\lim_{m\to\infty} x_m = x_0$.

Next notice for all $m \in \mathbb{N}$ that

$$\begin{aligned}
&\frac{W(x_m) - W(x_0)}{x_m - x_0} \\
&= \frac{1}{x_m - x_0} \sum_{n=0}^{\infty} b^n \cos(\pi a^n x_m) - \frac{1}{x_m - x_0} \sum_{n=0}^{\infty} b^n \cos(\pi a^n x_0) \\
&= \frac{1}{x_m - x_0} \left(\sum_{n=0}^{m-1} b^n \cos(\pi a^n x_m) + \sum_{n=m}^{\infty} b^n \cos(\pi a^n x_m) \right) \\
&\quad - \frac{1}{x_m - x_0} \left(\sum_{n=0}^{m-1} b^n \cos(\pi a^n x_0) + \sum_{n=m}^{\infty} b^n \cos(\pi a^n x_0) \right) && \text{only the tails matter} \\
&= \frac{1}{x_m - x_0} \sum_{n=0}^{m-1} b^n \left(\cos(\pi a^n x_m) - \cos(\pi a^n x_0) \right) \\
&\quad + \frac{1}{x_m - x_0} \sum_{n=m}^{\infty} b^n \left(\cos(\pi a^n x_m) - \cos(\pi a^n x_0) \right) && \text{adding finite and convergent series.}
\end{aligned}$$

For each $m \in \mathbb{N}$, let

$$\begin{aligned}
P_m &= \frac{1}{x_m - x_0} \sum_{n=0}^{m-1} b^n \left(\cos(\pi a^n x_m) - \cos(\pi a^n x_0) \right) \\
S_m &= \frac{1}{x_m - x_0} \sum_{n=m}^{\infty} b^n \left(\cos(\pi a^n x_m) - \cos(\pi a^n x_0) \right),
\end{aligned}$$

which are (uniformly) convergent series since $W(x_m)$ and $W(x_0)$ converge uniformly such that

$$\frac{W(x_m) - W(x_0)}{x_m - x_0} = P_m + S_m.$$

We desire bounds on $|P_m|$ and $|S_m|$.

Bound for $|P_m|$: Since the derivative of cosine is negative sine, the Mean Value Theorem implies for all $n \in \{0, 1, \ldots, m-1\}$ there exists a $c_{n,m}$ between $\pi a^n x_m$ and $\pi a^n x_0$ such that

$$\left| \frac{\cos(\pi a^n x_m) - \cos(\pi a^n x_0)}{\pi a^n x_m - \pi a^n x_0} \right| = |-\sin(c_{n,m})| = |\sin(c_{n,m})|.$$

Therefore, since $|\sin(x)| \leq 1$ for all $x \in \mathbb{R}$, we obtain that

$$\begin{aligned}
|P_m| &\leq \sum_{n=0}^{m-1} b^n \left| \frac{\cos(\pi a^n x_m) - \cos(\pi a^n x_0)}{x_m - x_0} \right| \\
&= \sum_{n=0}^{m-1} b^n \pi a^n \left| \frac{\cos(\pi a^n x_m) - \cos(\pi a^n x_0)}{\pi a^n x_m - \pi a^n x_0} \right| \\
&\leq \sum_{n=0}^{m-1} b^n \pi a^n \\
&= \pi \frac{(ab)^m - 1}{ab - 1} < \pi \frac{(ab)^m}{ab - 1}.
\end{aligned}$$

Bound for $|S_m|$: We will consider $(-1)^{\ell_m} S_m$. For a fixed $n \geq m$, notice that

$$\pi a^n x_m = \pi a^{n-m} \ell_m.$$

However, since a is an odd integer, we see that $\pi a^{n-m} \ell_m$ is an odd integer multiple of π when ℓ_m is odd and $\pi a^{n-m} \ell_m$ is an even integer multiple of π when ℓ_m is even. Hence

$$(-1)^{\ell_m} \cos(\pi a^n x_m) = 1.$$

Moreover, since

$$\sin(\pi a^n x_m) = \sin(\pi a^{n-m} \ell_m) = 0$$

as $\pi a^{n-m} \ell_m$ is an integer multiple of π, we see that

$$\begin{aligned}
\cos(\pi a^n x_0 - \pi a^n x_m) &= \cos(\pi a^n x_0) \cos(\pi a^n x_m) + \sin(\pi a^n x_0) \sin(\pi a^n x_m) \\
&= \cos(\pi a^n x_0)(-1)^{\ell_m} + \sin(\pi a^n x_0)(0) \\
&= (-1)^{\ell_m} \cos(\pi a^n x_0).
\end{aligned}$$

Therefore, since the above holds for all $n \geq m$, we obtain that

$$\begin{aligned}
(-1)^{\ell_m} S_m &= \sum_{n=m}^{\infty} b^n \frac{(-1)^{\ell_m} \cos(\pi a^n x_m) - (-1)^{\ell_m} \cos(\pi a^n x_0)}{x_m - x_0} \\
&= \sum_{n=m}^{\infty} b^n \frac{1 - \cos(\pi a^n x_0 - \pi a^n x_m)}{x_m - x_0} \\
&= \sum_{n=m}^{\infty} b^n \frac{1 - \cos(\pi a^{n-m}(a^m x_0 - \ell_m))}{x_m - x_0}.
\end{aligned}$$

Notice since $\frac{1}{2} \leq \ell_m - a^m x_0 < \frac{3}{2}$ that

$$\cos(\pi a^{m-m}(a^m x_0 - \ell_m)) \leq 0.$$

Moreover, we know that

$$1 - \cos(\pi a^{n-m}(a^m x_0 - \ell_m)) \geq 0$$

for all $n > m$. Finally, since $x_m > x_0$ for all $m \in \mathbb{N}$, all terms in the above series expression for $(-1)^{\ell_m} S_m(x)$ are non-negative. Hence $(-1)^{\ell_m} S_m(x)$ is at least the first term in the series so

$$\begin{aligned}
(-1)^{\ell_m} S_m &\geq b^m \frac{1 - \cos(\pi(a^m x_0 - \ell_m))}{x_m - x_0} \\
&\geq b^m \frac{1}{x_m - x_0} \\
&\geq b^m \frac{1}{\frac{3}{2a^m}} && \text{since } \tfrac{3}{2a^m} \geq x_m - x_0 \\
&\geq \frac{2}{3} a^m b^m.
\end{aligned}$$

Hence $|S_m(x)| \geq \frac{2}{3} a^m b^m$.

Using the above two bounds, we obtain for all $m \in \mathbb{N}$ that

$$\begin{aligned}
\left| \frac{W(x_m) - W(x_0)}{x_m - x_0} \right| &= |P_m + S_m| \\
&\geq |S_m| - |P_m| && \text{by the reverse triangle inequality} \\
&\geq \frac{2}{3} a^m b^m - \pi \frac{(ab)^m}{ab - 1} \\
&= (ab)^m \left(\frac{2}{3} - \frac{\pi}{ab - 1} \right).
\end{aligned}$$

However, since $ab > 1 + \frac{3}{2}\pi$, we see that $\frac{2}{3} - \frac{\pi}{ab-1} > 0$ and thus

$$\liminf_{m \to \infty} \left| \frac{W(x_m) - W(x_0)}{x_m - x_0} \right| \geq \liminf_{m \to \infty} (ab)^m \left(\frac{2}{3} - \frac{\pi}{ab - 1} \right) = \infty.$$

Hence W cannot be differentiable at x_0. Therefore, since x_0 was arbitrary, W is continuous but nowhere differentiable.

2.4 Integration of Series of Functions

The existence of the Weierstrass functions puts the idea that series of differentiable functions can be differentiable into great jeopardy. To rectify this situation, we ignore this question and turn to integration. This may seem odd to the reader in that the natural progression of calculus is to first introduce derivatives and then more onto integration since differentiation is easier computationally than integration. It turns out that integration actually behaves better than integration and we will be able to get at the desired differentiation results via integration and the Fundamental Theorem of Calculus. In addition, integration behaves far better in regards to limits of continuous functions than differentiation does. Well, except for the following examples that is (but of course these examples are with respect to pointwise convergence, which we know is not the right type of convergence to look at).

Example 2.4.1. We claim that there exists a sequence $(f_n)_{n\geq 1}$ of real-valued continuous functions on $[0, 1]$ that converge pointwise to a continuous function $f : [0, 1] \to \mathbb{R}$ such that

$$\int_0^1 f(x)\, dx \neq \lim_{n\to\infty} \int_0^1 f_n(x)\, dx.$$

To see this, for each $n \in \mathbb{N}$, let $f_n : [0, 1] \to \mathbb{R}$ be defined by

$$f_n(x) = \begin{cases} 2n^2 x & \text{if } 0 \leq x \leq \frac{1}{2n} \\ 2n - 2n^2 x & \text{if } \frac{1}{2n} \leq x \leq \frac{1}{n} \\ 0 & \text{if } \frac{1}{n} \leq x \leq 1 \end{cases}.$$

In particular, the graph of f_n creates an isosceles triangle with base $\left[0, \frac{1}{n}\right]$ with height n, and otherwise is 0. Thus f_n continuous and

$$\int_0^1 f_n(x)\, dx = \frac{1}{2}$$

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

We claim that $(f_n)_{n\geq 1}$ converges pointwise to 0 on $[0, 1]$. This will complete the example since

$$\int_0^1 0\, dx = 0 \neq \frac{1}{2} = \lim_{n\to\infty} \int_0^1 f_n(x)\, dx.$$

To see $(f_n)_{n\geq 1}$ converges pointwise to 0, let $x \in [0, 1]$ be arbitrary. If $x = 0$, then since $f_n(0) = 0$ for all $n \in \mathbb{N}$ we clearly see that $(f_n(x))_{n\geq 1}$ converges to 0. Otherwise, assume $x > 0$. Since $\lim_{n\to\infty} \frac{1}{n} = 0$, there exists an $N \in \mathbb{N}$ such that $\frac{1}{n} < x$ for all $n \geq N$. Thus the definition of f_n implies that $f_n(x) = 0$ for all $n \geq N$ and thus $(f_n(x))_{n\geq 1}$ converges to 0.

Example 2.4.2. We claim that there exists a sequence $(f_n)_{n\geq 1}$ of real-valued Riemann integrable functions on $[0, 1]$ that converge pointwise to a function $f : [0, 1] \to \mathbb{R}$ that is bounded but not Riemann integrable. To see this, recall that $\mathbb{Q}$ is a countable set. Hence we can write $\mathbb{Q} \cap [0, 1] = \{r_n \mid n \in \mathbb{N}\}$. Define $f_n : [0, 1] \to [0, 1]$ by

$$f_n(x) = \begin{cases} 1 & \text{if } x \in \{r_1, r_2, \dots, r_n\} \\ 0 & \text{otherwise} \end{cases}$$

for all $x \in [0,1]$ and define $f : [0,1] \to [0,1]$ by

$$f(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q} \cap [0,1] \\ 0 & \text{otherwise} \end{cases}.$$

We claim that $(f_n)_{n\geq 1}$ converges pointwise to f. To see this, we note that $f_n(x) = 0 = f(x)$ for all $x \in [0,1] \setminus \mathbb{Q}$. Otherwise, if $x \in \mathbb{Q} \cap [0,1]$, then $x = r_N$ for some $N \in \mathbb{N}$ and $f_n(x) = 1 = f(x)$ for all $n \geq N$. Hence $(f_n)_{n\geq 1}$ converges pointwise to f.

Next, we claim that f_n is Riemann integrable for all $n \in \mathbb{N}$. To see this, fix $n \in \mathbb{N}$ and let $\epsilon > 0$ be arbitrary. Let $\mathcal{P}_\epsilon$ be the partition of $[0,1]$ formed by taking the end points of the open intervals of length $\frac{\epsilon}{n}$ centred at each r_k for $k \leq n$. For any interval in $\mathcal{P}$ that does not contain an r_k for $k \leq n$, the maximal and minimal values of f_n on this interval are both 0. Moreover, the interval of $\mathcal{P}$ containing an r_k with $k \leq n$ is of length at most $\frac{\epsilon}{n}$ and the difference between the maximal and minimal values of f_n on this interval is at most 1. Therefore, as there are n possible r_k for $k \leq n$, we obtain that

$$U(f_n, \mathcal{P}_\epsilon) - L(f_n, \mathcal{P}\epsilon) \leq n(1-0)\frac{\epsilon}{n} = \epsilon.$$

Therefore, as $\epsilon > 0$ was arbitrary, f_n is Riemann integrable for all $n \in \mathbb{N}$.

However, notice for any partition of $[0,1]$ that

$$U(f, \mathcal{P}) = 1 \qquad \text{and} \qquad L(f, \mathcal{P}) = 0.$$

Hence f is not integrable. Therefore, the example is complete.

Remark 2.4.3. Example 2.4.1 is pathological for any conceivable notion of integral one would want to work with that models the area under a curve. However, Example 2.4.2 is more a pathology of the Riemann integral in that the function that is 1 on the rationals and 0 on the irrationals is not Riemann integral even through the rational numbers are quite 'meagre' with respect to the irrational numbers. In particular, there are ways to extend the Riemann integral to a better notion that will remove this pathology. However, that is a topic for MATH 4012.

Of course the real problem with the above two examples is that pointwise convergence generally yields no analytical information about the limit function. As we have seen with continuity, it is uniform convergence we should consider in analysis. The following results further emphasizes this point.

Theorem 2.4.4. *Let $(f_n)_{n\geq 1}$ be a sequence of real-valued, Riemann integrable functions on a closed interval $[a,b]$. If $(f_n)_{n\geq 1}$ converges uniformly on $[a,b]$ to $f : [a,b] \to \mathbb{R}$, then f is Riemann integrable and*

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

Proof. To see that f is Riemann integrable, let $\epsilon > 0$ be arbitrary. Since $(f_n)_{n\geq 1}$ converges to f uniformly, there exists an $N \in \mathbb{N}$ such that

$$|f_N(x) - f(x)| < \frac{\epsilon}{4(b-a)}$$

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

$$f_N(x) - \frac{\epsilon}{4(b-a)} < f(x) < f_N(x) + \frac{\epsilon}{4(b-a)}$$

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

Since f_N is Riemann integrable, there exists a partition $\mathcal{P}$ of $[a, b]$ such that

$$U(f_N, \mathcal{P}) - L(f_N, \mathcal{P}) < \frac{\epsilon}{2}.$$

Write $\mathcal{P} = \{t_k\}_{k=0}^{\ell}$ where $a = t_0 < t_1 < \cdots < t_\ell = b$. Then, if

$$M_k = \sup\left(\{f_N(x) \mid x \in [t_{k-1}t_k]\}\right) \quad \text{and}$$
$$m_k = \inf\left(\{f_N(x) \mid x \in [t_{k-1}t_k]\}\right),$$

we know by the definition of the upper and lower Riemann sums that

$$U(f_N, \mathcal{P}) = \sum_{k=1}^{\ell} M_k(t_k - t_{k-1}) \quad \text{and} \quad L(f_N, \mathcal{P}) = \sum_{k=1}^{\ell} m_k(t_k - t_{k-1}).$$

Notice for all $x \in [t_{k-1}, t_k]$ that

$$m_k - \frac{\epsilon}{4(b-a)} \leq f_N(x) - \frac{\epsilon}{4(b-a)} < f(x) < f_N(x) + \frac{\epsilon}{4(b-a)} \leq M_k + \frac{\epsilon}{4(b-a)}.$$

Therefore

$$\begin{aligned} U(f, \mathcal{P}) &\leq \sum_{k=1}^{\ell} \left(M_k + \frac{\epsilon}{4(b-a)}\right)(t_k - t_{k-1}) \\ &= \sum_{k=1}^{\ell} M_k(t_k - t_{k-1}) + \sum_{k=1}^{\ell} \frac{\epsilon}{4(b-a)}(t_k - t_{k-1}) \\ &= U(f_N, \mathcal{P}) + \frac{\epsilon}{4} \end{aligned}$$

and

$$\begin{aligned} L(f, \mathcal{P}) &\geq \sum_{k=1}^{\ell} \left(m_k - \frac{\epsilon}{4(b-a)}\right)(t_k - t_{k-1}) \\ &= \sum_{k=1}^{\ell} m_k(t_k - t_{k-1}) - \sum_{k=1}^{\ell} \frac{\epsilon}{4(b-a)}(t_k - t_{k-1}) \\ &= L(f_N, \mathcal{P}) - \frac{\epsilon}{4}. \end{aligned}$$

Hence

$$\begin{aligned} U(f,\mathcal{P}) - L(f,\mathcal{P}) &\leq \left(U(f_N,\mathcal{P}) + \frac{\epsilon}{4}\right) - \left(L(f_N,\mathcal{P}) - \frac{\epsilon}{4}\right) \\ &= (U(f_N,\mathcal{P}) - L(f_N,\mathcal{P})) + \frac{\epsilon}{2} \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \end{aligned}$$

Therefore, since $\epsilon > 0$ was arbitrary, f is Riemann integrable.

To see that

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

let $\epsilon > 0$ be arbitrary. Since $(f_n)_{n\geq 1}$ converges to f uniformly, there exists an $N \in \mathbb{N}$ such that

$$|f_n(x) - f(x)| < \frac{\epsilon}{b-a}$$

for all $n \geq N$ and $x \in [a,b]$. Therefore, for all $n \geq N$ we have

$$\begin{aligned} \left|\int_a^b f_n(x)\,dx - \int_a^b f(x)\right| &= \left|\int_a^b f_n(x) - f(x)\,dx\right| \\ &\leq \int_a^b |f_n(x) - f(x)|\,dx \\ &\leq \int_a^b \frac{\epsilon}{b-a}\,dx \qquad \text{since } |f_n(x)-f(x)|<\tfrac{\epsilon}{b-a} \text{ for all } x\in[a,b] \\ &= \epsilon. \end{aligned}$$

Therefore, since $\epsilon > 0$ was arbitrary, the result is complete. ∎

Theorem 2.4.4 immediately allows us to integrate uniformly convergent series of Riemann integrable functions term-by-term!

Corollary 2.4.5. *Let $(f_n)_{n\geq 1}$ be a sequence of real-valued Riemann integrable functions on $[a,b]$. If $\sum_{n=1}^\infty f_n$ converges uniformly to $f : [a,b] \to \mathbb{R}$, then f is Riemann integrable and*

$$\int_a^b f(x)\,dx = \sum_{n=1}^\infty \int_a^b f_n(x)\,dx.$$

Proof. For each $N \in \mathbb{N}$, let define $S_N : [a,b] \to \mathbb{R}$ by

$$S_N(x) = \sum_{k=1}^N f_k(x)$$

for all $x \in [a,b]$. Since f_n is Riemann integrable for all n, S_N is Riemann integrable. Moreover, since $(S_N)_{n\geq 1}$ converges uniformly to f by assumption, Theorem 2.4.4 implies that f is Riemann integrable and

$$\begin{aligned}\int_a^b f(x)\,dx &= \lim_{N\to\infty}\int_a^b S_N(x)\,dx \\ &= \lim_{N\to\infty}\int_a^b \sum_{k=1}^N f_k(x)\,dx \\ &= \lim_{N\to\infty}\sum_{k=1}^N \int_a^b f_k(x)\,dx \\ &= \sum_{n=1}^\infty \int_a^b f_n(x)\,dx\end{aligned}$$

as desired. ∎

Remark 2.4.6. Using Corollary 2.4.5, we can obtain some new series of functions and potentially find the values of some of the series of real numbers from Chapter 1. Indeed, for any $0 < b < 1$, since

$$|x^n| \leq b^n$$

for all $x \in [-b, b]$ and all $n \in \mathbb{N}$, and since $\sum_{n=1}^\infty b^n$ converges, the Weierstrass M-Test (Theorem 2.2.15) implies that

$$\frac{1}{1-x} = \sum_{n=0}^\infty x^n$$

with the convergence being uniform on $[-b, b]$. Hence Corollary 2.4.5 implies that

$$\begin{aligned}-\ln(1-x) &= \int_0^x \frac{1}{1-r}\,dr \\ &= \sum_{n=0}^\infty \int_0^x r^n\,dr \\ &= \sum_{n=0}^\infty \frac{1}{n+1}x^{n+1} \\ &= \sum_{n=1}^\infty \frac{1}{n}x^n\end{aligned}$$

for all $x \in [-b, b]$. As this holds for all $b \in (0, 1)$, we have that

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

for all $x \in (-1, 1)$.

Of course, this does not let us evaluate

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

(after all, this series only converges conditionally and can be rearranged to obtain any value). However, one may be tempted to try a similar idea to compute $\sum_{n=1}^{\infty} \frac{1}{n^2} x^n$ for any $x \in (-1, 1)$ and then take a limit as x tends to 1 to obtain the value of

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

Indeed, we note the above series implies that

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

In addition, we can again use the Weierstrass M-Test (Theorem 2.2.15) to show that this series converges uniformly on any closed subinterval of $(0, 1)$ and thus obtain by Corollary 2.4.5 that

$$\sum_{n=1}^{\infty} \frac{1}{n^2}(b^n - a^n) = \sum_{n=1}^{\infty} \int_a^b \frac{1}{n} x^{n-1}\, dx = \int_a^b -\frac{\ln(1-x)}{x}\, dx$$

for all $0 < a < b < 1$. This clearly poses some problems in that :

1. What is the value of this integral?
2. Can we actually take the limit as a tends to 0 and b tends to 1? After all, we have seen exchanging limits is problematic.
3. We have yet to actually define the natural logarithm.

Of course, the last question is the easiest to solve once we obtain some information about differentiation of series.

2.5 Differentiation of Series of Functions

Of course, as Example 2.3.4 shows, there exist uniformly convergent series of (infinitely) differentiable functions that are nowhere differentiable. However, some simple additional requirements can be added to resolve this problem. To be specific, provided the partial sums of the derivatives are Riemann integrable and converge uniformly, the series will be differentiable and the derivative can be obtained by summing the derivatives of the individual terms in the series. To obtain this result, we simply prove the following result pertaining to limits of differentiable functions.

Throughout these notes, given a closed interval $[a, b]$ and a function $f : [a, b] \to \mathbb{R}$, we say that f is differentiable on $[a, b]$ if f is continuous on $[a, b]$ and differentiable on (a, b).

Theorem 2.5.1. *Let $(f_n)_{n\geq 1}$ be a sequence of real-valued differentiable functions on a closed interval $[a,b]$ that converge pointwise to a function $f : [a,b] \to \mathbb{R}$. If*

- *f_n' is Riemann integrable (e.g. continuous) for all $n \in \mathbb{N}$, and*
- *$(f_n')_{n\geq 1}$ converges uniformly on $[a,b]$ to a continuous function $g : [a,b] \to \mathbb{R}$,*

then f is differentiable on $[a,b]$ and $f' = g$; that is, $(f_n')_{n\geq 1}$ converges uniformly to f' on $[a,b]$.

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

$$\begin{aligned} f(x) &= \lim_{n\to\infty} f_n(x) && \text{since } (f_n)_{n\geq 1} \text{ converges pointwise to } f \\ &= \lim_{n\to\infty} \int_a^x f_n'(r)\, dr && \text{by the Fundamental Theorem of Calculus} \\ &= \int_a^x g(r)\, dr && \begin{array}{c}\text{since Theorem 2.4.4 and since} \\ (f_n)_{n\geq 1} \text{ converges uniformly to } g.\end{array} \end{aligned}$$

Therefore, by the Fundamental Theorem of Calculus, f is differentiable on $[a,b]$ and $f' = g$. Hence $(f_n')_{n\geq 1}$ converges uniformly to f' on $[a,b]$. ∎

Of course, we immediately obtain the analogue for series.

Corollary 2.5.2. *Let $(f_n)_{n\geq 1}$ be a sequence of real-valued differentiable functions on $[a,b]$. If*

- *$\sum_{n=1}^\infty f_n(x)$ converges for all $x \in [a,b]$,*
- *f_n' is Riemann integrable (e.g. continuous) for all $n \in \mathbb{N}$, and*
- *$\sum_{n=1}^\infty f_n'$ converges uniformly on $[a,b]$ to a continuous function $g : [a,b] \to \mathbb{R}$,*

then the function $f : [a,b] \to \mathbb{R}$ defined by

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

for all $x \in [a,b]$ is differentiable and

$$\sum_{n=1}^\infty f_n'$$

converges uniformly to f' on $[a,b]$.

Proof. For all $N \in \mathbb{N}$, define $S_N : [a, b] \to \mathbb{R}$ by

$$S_N(x) = \sum_{k=1}^{n} f_k(x)$$

for all $x \in [a, b]$. By the first assumption, $(S_N)_{n\geq 1}$ converges pointwise to f. Moreover, by the second assumption

$$S_N' = \sum_{k=1}^{N} f_k'$$

is Riemann integrable on $[a, b]$ for all $N \in \mathbb{N}$ and, by the third assumption, $(S_N')_{N\geq 1}$ converges uniformly to a continuous function on $[a, b]$, Theorem 2.5.1 implies that f is differentiable and

$$\sum_{n=1}^{\infty} f_n'$$

converges uniformly to f' on $[a, b]$. ∎

Of course the challenge in using Corollary 2.5.2 is the requirement that the series of derivatives converges uniformly, but again we can often bypass this issue by using the Weierstrass M-Test (Theorem 2.2.15).

Before moving on to using and examples of Corollary 2.5.2, it is useful to note the following two things.

Remark 2.5.3. Corollary 2.5.2 is not in contradiction with the Weierstrass functions being nowhere differentiable in Example 2.3.4. Indeed, recall if $a, b \in \mathbb{R}$ with a a positive odd integer, $0 < b < 1$ and $ab > 1 + \frac{3}{2}\pi$, then if we define $W : \mathbb{R} \to \mathbb{R}$ by

$$W(x) = \sum_{n=0}^{\infty} b^n \cos(\pi a^n x)$$

for all $x \in \mathbb{R}$, then the above series converges uniformly so W is continuous, but W is nowhere differentiable. The reason Corollary 2.5.2 does not apply here is that if $f_n(x) = b^n \cos(\pi a^n x)$ for all $x \in \mathbb{R}$ and $n \in \mathbb{N}$, then

$$f_n'(x) = -\pi(ab)^n \sin(\pi a^n x)$$

is Riemann integrable, but

$$f_n'\left(1\right) = -\pi(ab)^n \sin\left(\frac{\pi a^n}{2}\right) = (-1)^{\frac{a+1}{2}}(ab)^n$$

for all $n \in \mathbb{N}$ so

$$\sum_{n=0}^{\infty} f_n'\left(\frac{1}{2}\right)$$

diverges since $ab > 1$. In fact, Corollary 2.5.2 can be used to show that

$$\sum_{n=0}^{\infty} -\pi(ab)^n \sin(\pi a^n x)$$

does not converge uniformly on any closed interval of $\mathbb{R}$ with positive length.

Remark 2.5.4. Another natural question in regards to Theorem 2.5.1 (and thus Corollary 2.5.2) is that, if we knew the uniform limit f of the sequence of functions happened to be differentiable, does the limit of the sequence of derivatives converge to f'? Unfortunately, we obtain a quick answer of no based on the following example.

We claim that there exists a sequence $(f_n)_{n\geq 1}$ of real-valued continuously differentiable functions on a closed interval $[a, b]$ that converge uniformly to a continuously differentiable function $f : [a, b] \to \mathbb{R}$, but

$$f'(x) \neq \lim_{n\to\infty} f_n'(x)$$

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

To see this, for each $n \in \mathbb{N}$ let $f_n : [-\pi, \pi] \to \mathbb{R}$ be defined by

$$f_n(x) = \frac{1}{n} \sin(nx)$$

for all $x \in [-\pi, \pi]$. Since

$$|f_n(x)| \leq \frac{1}{n} |\sin(nx)| \leq \frac{1}{n}$$

for all $x \in [-\pi, \pi]$ and $n \in \mathbb{N}$, and since $\lim_{n\to\infty} \frac{1}{n} = 0$, we see that $(f_n)_{n\geq 1}$ converges uniformly to 0. However $f = 0$ is a continuously differentiable function with derivative 0 and

$$f_n'(x) = \cos(nx)$$

for all $x \in [-\pi, \pi]$ and $n \in \mathbb{N}$, so

$$\lim_{n\to\infty} f_n'(0) = \lim_{n\to\infty} \cos(0) = 1 \neq f'(0).$$

Hence the example is complete.

Of course the above example requires knowledge of the derivatives of cosine and sine, which we have yet to demonstrate with the definitions provided in Chapter 1. It is about time we rectify this delay.

2.6 Power Series

Using Corollary 2.5.2 we can now finally complete our construction of the exponential, cosine, and sine functions by describing their derivatives. As all three of these functions have a very specific form, it is useful for other applications to describe a larger collection of functions and derive their properties.

Definition 2.6.1. Given $c \in \mathbb{R}$, a *power series centred at* c is any series of the form
$$\sum_{n=0}^{\infty} a_n(x-c)^n$$
where x is a real variable and $(a_n)_{n\geq 0}$ is a sequence of real numbers.

Often we take $c = 0$ when discussing power series as any results that can be done at $c = 0$ can be translated to an arbitrary c. However, we will prove the following at an arbitrary c without the need to translate.

Theorem 2.6.2. *Let* $(a_n)_{n\geq 0}$ *be a sequence of real numbers and let* $c \in \mathbb{R}$*. Suppose* $x_0 \in \mathbb{R} \setminus \{c\}$ *is such that*
$$\sum_{n=0}^{\infty} a_n(x_0-c)^n$$
converges. Then for any $r \in \mathbb{R}$ *with* $0 < r < |x_0 - c|$*, the series of functions in* x
$$\sum_{n=0}^{\infty} a_n(x-c)^n \qquad \text{and} \qquad \sum_{n=1}^{\infty} na_n(x-c)^{n-1}$$
converge uniformly and absolutely on $[c-r, c+r]$*.*

Moreover, if $f : [c-r, c+r] \to \mathbb{R}$ *is defined by*
$$f(x) = \sum_{n=0}^{\infty} a_n(x-c)^n$$
for all $x \in [c-r, c+r]$*, then* f *is differentiable on* $[c-r, c+r]$ *with*
$$f'(x) = \sum_{n=1}^{\infty} na_n(x-c)^{n-1}$$
for all $x \in (c-r, c+r)$*.*

Proof. For each $n \in \mathbb{N} \cup \{0\}$, let $f_n : \mathbb{R} \to \mathbb{R}$ and $g_n : \mathbb{R} \to \mathbb{R}$ be defined by
$$f_n(x) = a_n(x-c)^n \qquad \text{and} \qquad g_n(x) = na_n(x-c)^{n-1}$$

for all $x \in \mathbb{R}$. Our goal is to use the Weierstrass M-Test (Theorem 2.2.15) to show that $\sum_{n=0}^{\infty} f_n$ and $\sum_{n=0}^{\infty} g_n$ converge absolutely, uniformly, and define continuous functions.

To begin, notice since

$$\sum_{n=0}^{\infty} a_n(x_0 - c)^n$$

converges, Corollary 1.2.6 implies that

$$\lim_{n\to\infty} a_n(x_0 - c)^n = 0.$$

Hence, by Corollary 1.1.6 implies there exists an $M > 0$ such that

$$|a_n(x_0 - c)^n| \leq M.$$

Therefore, we have for all $n \in \mathbb{N} \cup \{0\}$ and $x \in [c - r, c + r]$ that

$$|a_n(x - c)^n| = |a_n(x_0 - c)^n| \left|\frac{(x - c)^n}{(x - c_0)^n}\right| \leq M \left(\frac{r}{|x_0 - c|}\right)^n.$$

Therefore, since $0 \leq r < |x_0 - c|$, we know that the geometric series

$$\sum_{n=0}^{\infty} M \left(\frac{r}{|x_0 - c|}\right)^n$$

converges. Hence, since f_n is continuous for all $n \in \mathbb{N} \cup \{0\}$, the Weierstrass M-Test (Theorem 2.2.15) implies that $\sum_{n=0}^{\infty} f_n$ converges uniformly and absolutely to f on $[c - r, c + r]$ and f is continuous on $[c - r, c + r]$.

To obtain a similar result for $\sum_{n=0}^{\infty} g_n$, first we claim that if $r_0 = \frac{r}{|x_0 - c|}$ then

$$\sum_{n=0}^{\infty} Mnr_0^{n-1}$$

converges. To see this, if $b_n = Mnr_0^n$ for all $n \in \mathbb{N}$, then

$$\left|\frac{b_{n+1}}{b_n}\right| = \frac{M(n+1)|r_0|^n}{Mn|r_0|^{n-1}} = \frac{n+1}{n}|r_0|.$$

Hence

$$\lim_{n\to\infty} \left|\frac{b_{n+1}}{b_n}\right| = |r_0|.$$

Therefore, since $|r_0| < 1$ since $0 \leq r < |x_0 - c|$, we obtain that $\sum_{n=0}^{\infty} Mnr_0^n$ converges by the Ratio Test (Theorem 1.2.18).

Now notice for all $n \in \mathbb{N}$ and $x \in [c - r, c + r]$ that

$$|na_n(x - c)^{n-1}| = n|a_n(x_0 - c)^{n-1}| \left|\frac{(x - c)^{n-1}}{(x - c_0)^{n-1}}\right| \leq nM \left(\frac{r}{|x_0 - c|}\right)^{n-1}.$$

Since

$$\sum_{n=1}^{\infty} Mn\left(\frac{r}{|x_0 - c|}\right)^{n-1}$$

converges and since g_n is continuous for all $n \in \mathbb{N} \cup \{0\}$, the Weierstrass M-Test (Theorem 2.2.15) implies that $\sum_{n=0}^{\infty} g_n$ converges uniformly and absolutely to a continuous function g on $[c-r, c+r]$.

Finally, since $\sum_{n=0}^{\infty} f_n$ converges pointwise to f, since $f_n' = g_n$ is continuous (and thus Riemann integrable) for all $n \in \mathbb{N} \cup \{0\}$, and since $\sum_{n=0}^{\infty} f_n' = \sum_{n=1}^{\infty} g_n$ converges uniformly to a continuous function on $[c-r, c+r]$, Theorem 2.5.1 implies that f is differentiable on $[c-r, c+r]$ and

$$f'(x) = \sum_{n=0}^{\infty} f_n'(x) = \sum_{n=1}^{\infty} g_n(x) = \sum_{n=1}^{\infty} na_n(x-c)^{n-1}$$

for all $x \in (c-r, c+r)$ as desired. ∎

Remark 2.6.3. Given $c \in \mathbb{C}$ and a sequence $(a_n)_{n \geq 0}$ of complex numbers is not difficult to verify that the first part of the proof of Theorem 2.6.2 works for the complex power series

$$\sum_{n=0}^{\infty} a_n(z-c)^n \qquad \text{and} \qquad \sum_{n=1}^{\infty} a_n n(z-c)^{n-1}$$

where the interval $[c-r, c+r]$ is replaced with a closed disk of radius r centred at c. The second part of Theorem 2.6.2 concerning the second power series is the derivative of the first power series is true, but more complicated to prove since we would need to discuss the derivatives of functions on the complex plane and since we cannot use Theorem 2.5.1 in its present form as the proof relies heavily on the Fundamental Theorem of Calculus. The details of these results are best left for a complex analysis course (see MATH 3410).

With Theorem 2.6.2 in hand, we can immediately obtain some new convergent series.

Example 2.6.4. Recall for all $x \in (-1, 1)$, the geometric series $\sum_{n=0}^{\infty} x^n$ converges to the function

$$f(x) = \frac{1}{1-x}$$

Hence Theorem 2.6.2 implies that $f'(x) = \sum_{n=1}^{\infty} nx^{n-1}$. Hence

$$\sum_{n=1}^{\infty} nx^n = xf'(x) = \frac{x}{(1-x)^2}.$$

Therefore

$$\sum_{n=1}^{\infty} \frac{n}{2^n} = \frac{\frac{1}{2}}{\left(1-\frac{1}{2}\right)^2} = 2.$$

Of course, Theorem 2.6.2 finally allows us to completely prove the remaining properties of the exponential, cosine, and sine functions that we desired.

Corollary 2.6.5. *Consider the function* $f : \mathbb{R} \to (0, \infty)$ *defined by*

$$f(x) = e^x$$

for all $x \in \mathbb{R}$. *Then* f *is differentiable on its domain and* $f'(x) = f(x)$ *for all* $x \in \mathbb{R}$. *Moreover,* f *is strictly increasing, bijective function.*

The inverse of f *is the function* $\ln : (0, \infty) \to \mathbb{R}$ *and is called the* natural logarithm. *The natural logarithm is differentiable on its domain with* $\ln'(x) = \frac{1}{x}$ *for all* $x \in \mathbb{R}$.

Proof. Since

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

for all $x \in \mathbb{R}$, Theorem 2.6.2 implies that f is differentiable on any closed interval centred at 0 (and thus differentiable on $\mathbb{R}$) with

$$f'(x) = \sum_{n=1}^{\infty} n \frac{1}{n!} x^{n-1} = \sum_{n=1}^{\infty} \frac{1}{(n-1)!} x^{n-1} = e^x$$

as desired.

Notice that

$$f'(x) = e^x > 0$$

for all $x \in \mathbb{R}$ by Corollary 1.5.3. Therefore, by Corollary 1.5.3 and a result from MATH 2001, f is a strictly increasing continuous function from $\mathbb{R}$ to $(0, \infty)$ and thus has a differentiable inverse $\ln : (0, \infty) \to \mathbb{R}$. Moreover, by another result from MATH 2001, if $x \in (0, \infty)$ and $\ln(x) = y$, then $x = e^y$ and

$$\ln'(x) = \frac{1}{f'(y)} = \frac{1}{e^y} = \frac{1}{x}$$

as desired. ∎

Example 2.6.6. Since Corollary 1.5.7 proved that

$$\cos(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} \qquad \text{and} \qquad \sin(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1}$$

for all $x \in \mathbb{R}$, Theorem 2.6.2 implies that cos and sin are differentiable with

$$\sin'(x) = \sum_{n=0}^{\infty} (2n+1) \frac{(-1)^n}{(2n+1)!} x^{2n} = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} = \cos(x)$$

and

$$\begin{aligned}\cos'(x) &= \sum_{n=0}^{\infty} 2n\frac{(-1)^n}{(2n)!}x^{2n-1} \\ &= \sum_{n=1}^{\infty} \frac{(-1)^n}{(2n-1)!}x^{2n-1} \\ &= \sum_{m=0}^{\infty} \frac{(-1)^{m+1}}{(2m+1)!}x^{2m+1} \\ &= -\sum_{m=0}^{\infty} \frac{(-1)^m}{(2m+1)!}x^{2m+1} \\ &= -\sin(x)\end{aligned}$$

just as we knew to be true!

Remark 2.6.7. Using Example 2.6.6 together with Corollary 1.5.7, one can derive the known properties of cos and sin. In particular, since for $x \in (0,1)$ we know

$$\frac{1}{(4n+1)!}x^{4n+1} - \frac{1}{(4n+3)!}x^{4n+3} > 0$$

for all $n \geq 2$, we obtain that

$$\sin(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!}x^{2n+1} \geq x - \frac{1}{6}x^3 > 0.$$

Hence $\cos'(x) < 0$ for all $x \in (0,1)$, so cosine is a decreasing function on $(0,1)$. Moreover, since $\sin'(x) = \cos(x)$, which must be positive on some interval around 0, sine must be increasing on $(0,c)$ for some $c > 0$. Thus, the coupling $\cos'(x) = -\sin(x)$ and $\sin'(x) = \cos(x)$ imply that cosine must have a first root larger than 0, which is what we call $\frac{\pi}{2}$. From there, we can derive all the special angles for cosine and sine via Corollary 1.5.7, show the function $e^{i\theta}$ draws out a circle in the complex plane as θ moves from 0 to 2π, and that the area of the circle is π via integrals and trigonometric substitution. That is, our most basic understanding of trigonometry is a bi-product of series of functions!

When discussing power series, it is useful to keep track of the largest interval that Theorem 2.6.2 applies on in order to know where we can take the derivatives term-wise. We can encapsulate this information in the following.

Definition 2.6.8. Let $(a_n)_{n\geq 0}$ be a sequence of real numbers and let $c \in \mathbb{R}$. The *radius of convergence of the power series* $\sum_{n=0}^{\infty} a_n(x-c)^n$ is

$$R = \sup\left\{ |x_0 - c| \;\middle|\; x_0 \in \mathbb{R}, \sum_{n=0}^{\infty} a_n(x_0-c)^n \text{ converges} \right\}.$$

Remark 2.6.9. Let R be the radius of convergence of

$$\sum_{n=0}^{\infty} a_n(x-c)^n.$$

Clearly $R \in [0, \infty]$ by definition. Moreover, by Theorem 2.6.2, we know that $\sum_{n=0}^{\infty} a_n(x_0-c)^n$ converges, then so too must $\sum_{n=0}^{\infty} a_n(x-c)^n$ for all x such that $|x-c| < |x_0-c|$. Therefore, if $x \in (c-R, c+R)$, then $\sum_{n=0}^{\infty} a_n(x-c)^n$ converges by Theorem 2.6.2. Moreover if $x \notin [c-R, c+R]$, then $|x-c| > R$ so $\sum_{n=0}^{\infty} a_n(x-c)^n$ must diverge by the definition of the radius of convergences. However when $x = c-R$ or $x = c+R$, we do not have any information on whether $\sum_{n=0}^{\infty} a_n(x-c)^n$ converges as the following example shows.

Example 2.6.10. For $x \in \mathbb{R}$, consider the series

$$\sum_{n=1}^{\infty} \frac{1}{n} x^n.$$

Since $\lim_{n\to\infty} \frac{1}{n} x^n$ exists and is zero if and only if $x \in [-1, 1]$, the above series can converge only if $x \in [-1, 1]$ by Corollary 1.2.6. Moreover, since

$$\sum^{\infty} \frac{1}{n}(-1)^n$$

converges by Example 1.2.23, the radius of convergence of this power series around 0 is 1. Moreover this series converges when $x = -1$ whereas, when $x = 1$,

$$\sum_{n=1}^{\infty} \frac{1}{n} x^n = \sum_{n=1}^{\infty} \frac{1}{n}$$

does not converge by Corollary 1.2.15. Thus a power series may or may not converge at the boundary points of its radius of convergence.

2.7 Taylor's Theorem

By Theorem 2.6.2 we know that a power series is differentiable inside its radius of convergence. The following theorem extends this and informs us of the derivatives at specific points!

Theorem 2.7.1 (Taylor's Theorem). *Let $(a_n)_{n\geq 0}$ be a sequence of real numbers and let $c \in \mathbb{R}$. Suppose the power series*

$$f(x) = \sum_{n=0}^{\infty} a_n(x-c)^n$$

has radius of convergence $R > 0$. Then f is infinitely differentiable on $(c-R, c+R)$ and

$$a_n = \frac{f^{(n)}(c)}{n!}$$

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

Proof. We claim that for each $m \in \mathbb{N} \cup \{0\}$ that f is m-times differentiable on $(c - R, c + R)$ with

$$f^{(m)}(x) = \sum_{n=m}^{\infty} \frac{n!}{(n-m)!} a_n (x-c)^{n-m}$$

for all $x \in (c - R, c + R)$. To prove this claim, we will proceed by induction. Clearly the case $m = 0$ follows by the assumptions of the theorem. For the inductive step, if

$$f^{(m)}(x) = \sum_{n=m}^{\infty} \frac{n!}{(n-m)!} a_n (x-c)^{n-m}$$

for all $x \in (c-R, c+R)$, then since this power series has radius of convergence R, by considering every closed interval centred at c of radius less than R, Theorem 2.6.2 implies that $f^{(m)}$ is differentiable on $(c - R, c + R)$ with

$$\begin{aligned} f^{(m+1)}(x) &= \sum_{n=m}^{\infty} (n-m) \frac{n!}{(n-m)!} a_n (x-c)^{n-m-1} \\ &= \sum_{n=m+1}^{\infty} \frac{n!}{(n-m-1)!} a_n (x-c)^{n-m-1} \end{aligned}$$

for all $x \in (c - R, c + R)$. Hence the claim is complete.

Since $c \in (c - R, c + R)$, the above implies that

$$f^{(m)}(c) = \frac{m!}{(m-m)!} a_m = (m!) a_m$$

for all $m \in \mathbb{N}$ thereby completing the proof. ∎

Taylor's Theorem (Theorem 2.7.1) also enables us to show that any function that has a power series expansion at a point has a unique power series expansion!

Corollary 2.7.2. *Let $(a_n)_{n\geq 0}$ and $(b_n)_{n\geq 0}$ be sequences of real numbers and let $c \in \mathbb{R}$. If there exists an $R > 0$ such that*

$$\sum_{n=0}^{\infty} a_n (x-c)^n \qquad \text{and} \qquad \sum_{n=0}^{\infty} b_n (x-c)^n$$

converge and are equal on $(c - R, c + R)$, then $a_n = b_n$ for all $n \in \mathbb{N} \cup \{0\}$.

Proof. Define $f, g : (c - R, c + R) \to \mathbb{R}$ by

$$f(x) = \sum_{n=0}^{\infty} a_n (x-c)^n \qquad \text{and} \qquad g(x) = \sum_{n=0}^{\infty} b_n (x-c)^n$$

for all $x \in (c - R, c + R)$. Hence Taylor's Theorem (Theorem 2.7.1) implies that f and g are infinitely differentiable with

$$a_n = \frac{f^{(n)}(c)}{n!} \qquad \text{and} \qquad b_n = \frac{g^{(n)}(c)}{n!}$$

for all $n \in \mathbb{N} \cup \{0\}$. Therefore, since $f = g$ on $(c - R, c + R)$, this implies that

$$a_n = \frac{f^{(n)}(c)}{n!} = \frac{g^{(n)}(c)}{n!} = b_n$$

for all $n \in \mathbb{N} \cup \{0\}$ as desired. ∎

With the above results, we can finally turn our attention to attempting to approximate continuous functions with 'nicer' functions. Power series are one such approximation as if we can attempt to write any continuous function as an infinite series of polynomials and this can be quite useful in various situations and applications. Of course Taylor's Theorem (Theorem 2.7.1) tells us that only infinitely differentiable functions have power series. However, for infinitely differentiable functions, Taylor's Theorem (Theorem 2.7.1) tells us exactly what power series can approximate a function and Corollary 2.7.2 shows there is at most one power series that approximates a function. However, we still have yet to answer the question: "Given an infinitely differentiable function f and a point $c \in \mathbb{R}$, does the power series of

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

always have a non-zero radius of convergence?"

Unfortunately, no:

Example 2.7.3. Consider the function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} 0 & \text{if } x = 0 \\ e^{-\frac{1}{x^2}} & \text{if } x \neq 0 \end{cases}$$

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

$$\lim_{x \to 0} \left| \frac{e^{-\frac{1}{x^2}}}{x^n} \right| = \lim_{x \to \infty} \frac{e^{-x^2}}{\frac{1}{x^n}} = \lim_{x \to \infty} \frac{x^n}{e^{x^2}} = 0$$

by a few applications of L'Höpital's rule, it can be verified that f is infinitely differentiable at 0 with $f^{(n)}(0) = 0$ for all $n \in \mathbb{N}$. Since $f(x) \neq 0$ for all $x \in \mathbb{R} \setminus \{0\}$ by Corollary 1.5.3, we see that

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

only at $x = 0$.

2.8 Weierstrass Approximation Theorem

As the previous section shows, we appear to be out of luck in approximating continuous functions with their power series since it can only possibly work for infinitely differentiable functions and Example 2.7.3 shows that there are infinitely differentiable functions for which the power series only converges at a single point. However, given an infinitely differentiable function f, the power series of f at a point is only one specific possible approximation of f. Are there other approximations of continuous functions using polynomials and, if so, how can we compute them?

The main goal of this section is to show that every continuous function on a closed interval can be 'uniformly' approximated by polynomials. Hence, depending on the proof of this result, there may be hope of explicitly describing a polynomial approximations of any continuous function! The main theorem of this section, the Weierstrass Approximation Theorem (Theorem 2.8.6), can be demonstrated using some interesting ideas. The first requirement to prove the Weierstrass Approximation Theorem is the following stronger notion of continuity.

Definition 2.8.1. Let $I \subseteq \mathbb{R}$ be an interval. A function $f : I \to \mathbb{C}$ is said to be *uniformly continuous* on I if for all $\epsilon > 0$ there exists a $\delta > 0$ such that if $x, y \in I$ and $|x - y| < \delta$ then $|f(x) - f(y)| < \epsilon$.

Remark 2.8.2. Note the difference between Definitions 2.1.1 and 2.8.1 is that for a fixed ϵ, $\delta > 0$ need only work for a given point in Definition 2.1.1 whereas, for uniformly continuous functions, Definition 2.8.1 enforces that the same δ works for all points in the interval. That is, for uniform continuity, there is one δ to rule them all!

The above is more desirable than simple continuity in that having a δ that works for the whole interval seems far more powerful than at a single point. Unfortunately, not every continuous function is uniformly continuous as the following example shows.

Example 2.8.3. We claim that if $f : \mathbb{R} \to \mathbb{R}$ is defined by $f(x) = x^2$ for all $x \in \mathbb{R}$, then Definition 2.8.1 fails for $\epsilon = 2$ and thus f is not uniformly continuous. To see this, let $\delta > 0$ be arbitrary. Choose $n \in \mathbb{N}$ such that $\frac{1}{n} < \delta$ and let $x_n = n$ and $y_n = n + \frac{1}{n}$. Then $|x_n - y_n| < \frac{1}{n} < \delta$ yet

$$|f(x_n) - f(y_n)| = \left| n^2 - \left(n + \frac{1}{n} \right)^2 \right| = 2 + \frac{1}{n^2} \geq 2.$$

Hence f is not uniformly continuous on $\mathbb{R}$.

The above shows that x^2 is not uniformly continuous on all of $\mathbb{R}$ since x^2 grows too quickly as x tends to infinity. Consequently, one may ask, "Are things much nicer if we restrict to finite intervals?" For closed intervals, yes!

Theorem 2.8.4. *Let $a, b \in \mathbb{R}$ be such that $a < b$. If $f : [a, b] \to \mathbb{C}$ is continuous, then f is uniformly continuous.*

Proof. Let $f : [a, b] \to \mathbb{C}$ be continuous. Suppose to the contrary that f is not uniformly continuous. Hence there exists an $\epsilon > 0$ such that for all $\delta > 0$ there exists $x, y \in [a, b]$ such that $|x - y| < \delta$ and $|f(x) - f(y)| \geq \epsilon$. Therefore, for each $n \in \mathbb{N}$ there exist $x_n, y_n \in [a, b]$ with $|x_n - y_n| < \frac{1}{n}$ and $|f(x_n) - f(y_n)| \geq \epsilon$.

Since $[a, b]$ is closed and bounded, the Bolzano-Weierstrass Theorem implies there exists a subsequence $(x_{k_n})_{n\geq 1}$ of $(x_n)_{n\geq 1}$ that converges to some number $L \in [a, b]$. Since f is continuous, there exists an $N_1 \in \mathbb{N}$ such that $|f(x_{k_n}) - f(L)| < \frac{\epsilon}{2}$ for all $n \geq N_1$.

Consider the subsequence $(y_{k_n})_{n\geq 1}$ of $(y_n)_{n\geq 1}$. Notice for all $n \in \mathbb{N}$ that

$$|y_{k_n} - L| \leq |y_{k_n} - x_{k_n}| + |x_{k_n} - L| \leq \frac{1}{k_n} + |x_{k_n} - L| \leq \frac{1}{n} + |x_{k_n} - L|.$$

Therefore, since $\lim_{n\to\infty} |x_{k_n} - L| = 0$ and $\lim_{n\to\infty} \frac{1}{n} = 0$, we obtain that $\lim_{n\to\infty} y_{k_n} = L$. Since f is continuous this implies that there exists an $N_2 \in \mathbb{N}$ such that $|f(y_{k_n}) - f(L)| < \frac{\epsilon}{2}$ for all $n \geq N_2$.

Notice if $N = \max\{N_1, N_2\}$, then the above implies that

$$|f(x_{k_N}) - f(y_{k_N})| \leq |f(x_{k_N}) - f(L)| + |f(L) - f(y_{k_N})| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

thereby contradicting the fact that $|f(x_{k_N}) - f(y_{k_N})| \geq \epsilon$. Hence f is uniformly continuous on $[a, b]$. ■

The only other ingredient required before the proof of the Weierstrass Approximation Theorem (Theorem 2.8.6) is the following technical lemma.

Lemma 2.8.5. *If $x \in [-1, 1]$ and $n \in \mathbb{N}$, then*

$$(1 - x^2)^n \geq 1 - nx^2.$$

Proof. Clearly it suffices to consider $x \in [0, 1]$ as $(1 - (-x)^2)^n = (1 - x^2)^n$ and $1 - n(-x)^2 = 1 - nx^2$ for all $x \in [-1, 1]$.

Consider the functions $f, g : [0, 1] \to \mathbb{R}$ defined by

$$f(x) = (1 - x^2)^n \qquad \text{and} \qquad g(x) = 1 - nx^2$$

for all $x \in [0, 1]$. Clearly $f(0) = 1 = g(0)$. Furthermore, f and g are differentiable with

$$f'(x) = n(1 - x^2)^{n-1}(-2x) \qquad \text{and} \qquad g'(x) = -2nx.$$

Since $-2nx \leq 0$ and $0 \leq 1 - x^2 \leq 1$ for all $x \in [0, 1]$, we see that $f'(x) \geq g'(x)$ for all $x \in [0, 1]$. Hence it follows that $f(x) \geq g(x)$ for all $x \in [0, 1]$. ■

Onto the main attraction!

Theorem 2.8.6 (Weierstrass Approximation Theorem). *Let $a, b \in \mathbb{R}$ be such that $a < b$. If $f : [a, b] \to \mathbb{R}$ is continuous, then there exists a sequence $(p_n)_{n\geq 1}$ of polynomials that converge uniformly to f on $[a, b]$; that is, for all $\epsilon > 0$ there exists a polynomial p such that*

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

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

Proof. To begin, for simplicity in the more complicated arguments, we desire to reduce to the case that $a = 0$ and $b = 1$. Indeed suppose we have proved the result when $a = 0$ and $b = 1$. Let $f : [a, b] \to \mathbb{R}$ be an arbitrary continuous function and let $\epsilon > 0$ be arbitrary. Define the function $g : [0, 1] \to \mathbb{R}$ by

$$g(x) = f(a + (b - a)x)$$

for all $x \in [0, 1]$. Clearly g is continuous. Thus, by our assumptions, there exists a polynomial q such that

$$|q(x) - g(x)| < \epsilon$$

for all $x \in [0, 1]$. If we define

$$p(x) = q\left(\frac{x - a}{b - a}\right)$$

for all $x \in \mathbb{R}$, then p is also a polynomial. We claim that $|p(x) - f(x)| < \epsilon$ for all $x \in [a, b]$. To see this, notice if $x_0 \in [a, b]$, then $y_0 = \frac{x_0 - a}{b - a} \in [0, 1]$ and $x_0 = a + (b - a)y_0$ so

$$|p(x_0) - f(x_0)| = \left|q\left(\frac{x_0 - a}{b - a}\right) - f(a + (b - a)y_0)\right| = |q(y_0) - g(y_0)| < \epsilon.$$

Therefore, since $x \in [a, b]$ and $\epsilon > 0$ were arbitrary, the argument is complete. Hence it suffices to prove the theorem in the case $a = 0$ and $b = 1$.

Next, for simplicity in the more complicated arguments, we desire to further reduce to the case that $f(0) = f(1) = 0$. Indeed suppose we have proved the result for all continuous functions that vanish at 0 and at 1. Let $f : [0, 1] \to \mathbb{R}$ be an arbitrary continuous function and let $\epsilon > 0$ be arbitrary. Define the function $g : [0, 1] \to \mathbb{R}$ by

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

for all $x \in [0, 1]$. Clearly g is continuous since f is continuous. Moreover, it is elementary to verify that $g(0) = g(1) = 0$. Thus, by our assumptions, there exists a polynomial q such that

$$|q(x) - g(x)| < \epsilon$$

for all $x \in [0, 1]$. If we define

$$p(x) = q(x) + (f(0) + (f(1) - f(0))x)$$

for all $x \in \mathbb{R}$, then p is also a polynomial. Notice that

$$|p(x) - f(x)| = |q(x) + (f(0) + (f(1) - f(0))x) - f(x)| = |q(x) - g(x)| < \epsilon$$

for all $x \in [0, 1]$. Therefore, as $\epsilon > 0$ was arbitrary, the argument is complete. Hence it suffices to prove the theorem in the case $a = 0$, $b = 1$, and $f(0) = f(1) = 0$.

To obtain the desired result in the case that $a = 0$, $b = 1$, and $f(0) = f(1) = 0$, let us first discuss some motivation for the proof. What we will do is construct a sequence of non-negative polynomials with integrals 1 on $[-1, 1]$ that will have 'most of their weight' at 0. We will then average or 'convolve' f against these polynomials at various points to obtain new polynomials that will approximate the value of f at each point uniformly over $[-1, 1]$.

To begin the proof, let $\epsilon > 0$ be arbitrary. First note since f is continuous on $[0, 1]$ and $f(0) = f(1) = 0$ that we can extend f to a continuous function on $\mathbb{R}$ by defining $f(x) = 0$ for all $x \in (-\infty, 0) \cup (1, \infty)$. Since f is then continuous on $[-2, 2]$, f is uniformly continuous on $[-2, 2]$ by Theorem 2.8.4. Therefore, by decreasing the δ in the definition of uniform continuity if needed, exists a $0 < \delta < 1$ such that if $x \in [-1, 1]$ and $|t| < \delta$ then

$$|f(x + t) - f(x)| < \frac{1}{2}\epsilon.$$

With the above set-up, we embark on constructing the polynomials approximates of f. First, we need some specific polynomials that will aid in constructing the appropriate polynomials to approximate f. Notice for each $n \in \mathbb{N}$ that

$$\int_{-1}^{1} (1 - x^2)^n \, dx > 0$$

as $(1 - x^2)^n > 0$ for all $x \in (-1, 1)$. Hence for each $n \in \mathbb{N}$ there exists a $c_n > 0$ such that

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

Therefore, by Lemma 2.8.5,

$$\begin{aligned}
\frac{1}{c_n} &= \int_{-1}^{1} (1-x^2)^n \, dx \\
&= 2\int_{0}^{1} (1-x^2)^n \, dx \\
&\geq 2\int_{0}^{\frac{1}{\sqrt{n}}} (1-x^2)^n \, dx \\
&\geq 2\int_{0}^{\frac{1}{\sqrt{n}}} 1 - nx^2 \, dx \\
&= 2\left(x - \frac{n}{3}x^3\right)\Bigg|_{x=0}^{\frac{1}{\sqrt{n}}} \\
&= \frac{4}{3\sqrt{n}} \geq \frac{1}{\sqrt{n}}
\end{aligned}$$

and thus $0 < c_n \leq \sqrt{n}$ for all $n \in \mathbb{N}$.

For each $n \in \mathbb{N}$ define $q_n : \mathbb{R} \to \mathbb{R}$ by

$$q_n(x) = c_n(1-x^2)^n$$

for all $x \in \mathbb{R}$. Thus each q_n is a polynomial, $q_n(x) \geq 0$ for all $x \in [-1, 1]$, and

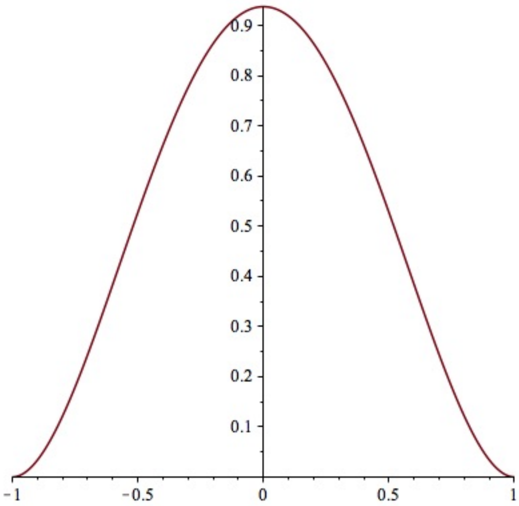
$$\int_{-1}^{1} q_n(x) \, dx = 1$$

by the definition of c_n. Moreover, notice by the definition of q_n that if $x \in [-1, -\delta] \cup [\delta, 1]$, then

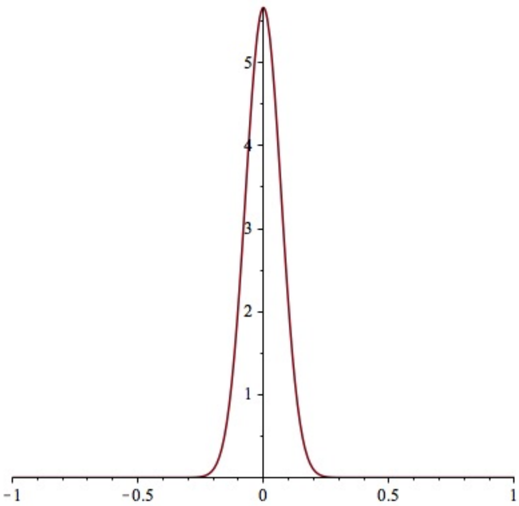
$$q_n(x) = c_n(1-x^2)^n \leq c_n(1-\delta^2)^n \leq \sqrt{n}(1-\delta^2)^n$$

since q_n is decreasing on $[\delta, 1]$ and increasing on $[-1, -\delta]$.

To add in the visual understanding of what an arbitrary q_n looks like, below is the graph of q_2:



and the graph of q_{100}.



Using the sequence of polynomials $(q_n)_{n\geq 1}$, we can now finally construct the polynomial approximates of f. For each $n \in \mathbb{N}$, define the "convolution of q_n and f" to be the function $q_n * f : [0, 1] \to \mathbb{R}$ defined by

$$(q_n * f)(x) = \int_{-1}^{1} q_n(t) f(x - t) \, dt$$

for all $x \in [0, 1]$. Clearly $q_n * f$ is a well-defined function since q_n is a polynomials, any translate of f is continuous, and the product of Riemann integrable functions is Riemann integrable. Moreover, since most of the weight of q_n is at $t = 0$, the value of $(q_n * f)(x)$ should be close to the value of $f(x)$. We will make this formal in a moment.

For now, we first claim $q_n * f$ are all polynomials. To see this, note

$$\begin{aligned}
(q_n * f)(x) &= \int_{-1}^{1} q_n(t) f(x-t)\,dt \\
&= \int_{x-1}^{x} q_n(t) f(x-t)\,dt && f \text{ is } 0 \text{ on } [0,1]^c \\
&= \int_{1}^{0} (-1) q_n(x-u) f(u)\,du && \text{substitute } u = x-t \\
&= \int_{0}^{1} q_n(x-u) f(u)\,du
\end{aligned}$$

However, note that $q_n(x-u)$ is a polynomial in x with coefficients being continuous functions in u and thus $q_n(x-u)f(u)$ is a polynomial in x with coefficients being continuous functions in u. Hence integrating $q_n(x-u)f(u)$ with respect to u can be performed by integrating the coefficients of the polynomial in x with respect to u thereby resulting in a polynomial in x. Hence $q_n * f$ is a polynomial on $[0,1]$.

Finally, we claim that if n is large enough then $|(q_n * f)(x) - f(x)| < \epsilon$ for all $x \in [0,1]$. To see this, first note since f is continuous on $[-2,2]$ that the Extreme Value Theorem implies there exists an $M > 0$ such that $|f(x)| \leq M$ for all $x \in [-2,2]$. Therefore, if $x \in [0,1]$, we see that

$$\begin{aligned}
&|(q_n * f)(x) - f(x)| \\
&= \left| \int_{-1}^{1} q_n(t) f(x-t)\,dt - f(x) \right| \\
&= \left| \int_{-1}^{1} q_n(t) f(x-t)\,dt - f(x) \int_{-1}^{1} q_n(t)\,dt \right| && \text{as } \int_{-1}^{1} q_n(x)\,dx = 1 \\
&= \left| \int_{-1}^{1} (f(x-t) - f(x)) q_n(t)\,dt \right| \\
&\leq \int_{-1}^{1} |f(x-t) - f(x)| q_n(t)\,dt && \text{as } q_n(x) \geq 0 \text{ on } [-1,1] \\
&= \int_{[-1,-\delta]\cup[\delta,1]} |f(x-t) - f(x)| q_n(t)\,dt + \int_{-\delta}^{\delta} |f(x-t) - f(x)| q_n(t)\,dt \\
&\leq \int_{[-1,-\delta]\cup[\delta,1]} 2M\sqrt{n}(1-\delta^2)^n\,dt + \int_{-\delta}^{\delta} \frac{\epsilon}{2} q_n(t)\,dt \\
&\leq 4\sqrt{n}M(1-\delta^2)^n(1-\delta) + \frac{\epsilon}{2}\int_{-1}^{1} q_n(t)\,dt \\
&= 4\sqrt{n}M(1-\delta^2)^n(1-\delta) + \frac{\epsilon}{2}.
\end{aligned}$$

Therefore, as $0 < 1 - \delta^2 < 1$ so

$$\lim_{n\to\infty} 4\sqrt{n}M(1-\delta^2)^n(1-\delta) = 0,$$

we see that for sufficiently large n that $\|(q_n * f) - f\|_\infty < \epsilon$. Hence, as $\epsilon > 0$ was arbitrary, the result follows. ∎

The Weierstrass Approximation Theorem (Theorem 2.8.6) is great in that if one can prove a result/property for polynomials that passes through uniform limits, one can extend the result to all continuous functions. Moreover, a careful analysis of the proof of the Weierstrass Approximation Theorem yields an explicit description of polynomial approximates to any continuous function f; namely $f * q_n$ (modulo the simplifications done at the start of the proof).

Unfortunately, the polynomials $f * q_n$ are not necessarily easy to compute. Indeed the first challenge is finding explicit descriptions of c_n for all $n \in \mathbb{N}$. Of course, we know that

$$\frac{1}{c_n} = \int_{-1}^{1} (1 - x^2)^n \, dx$$

and we can easily integrate any polynomial, but the formula for the precise value of c_n may not be nice (actually, it turns out that $c_n = \frac{(2n+1)!!}{2((2n)!!)}$ where $n!!$ is take the product of every other natural number starting at n and going down until 1 or 2). Then one needs to compute

$$(f * q_n)(x) = \int_{-1}^{1} f(x - t) q_n(t) \, dt$$

which means integrating f against a complicated polynomial. Generally to compute the values of

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

one would want to apply integration by parts several times (if f was n-times differentiable), which can be lengthy and potentially messy, and then one would need to take a linear combination of these polynomial integrands based on an expansion of q_n to get $f * q_n$.

So, although in theory when given a continuous function f on a closed interval we have explicit descriptions of polynomials that uniformly approximate f, the descriptions of these polynomials are nasty and the ability to compute these polynomials is not a simple task. Wouldn't it be far nicer if we could do such approximations with simple to compute and easily described polynomials?

Chapter 3

Series of Trigonometric Polynomials

It turns out that there is an approximation of continuous functions we can do that is simple to compute and easy to describe. The one caveat is that this approximation requires us to replace polynomials with series of 'trigonometric polynomials', known as Fourier series (pronounced fuor-ree-ay). The trigonometric polynomials are loosely the linear combinations of cosine and sine functions of integer frequencies. On the surface trigonometric functions may appear more difficult to work with, but there are numerous applications where working with trigonometric functions makes sense.

In fact, the original idea for approximating with trigonometric polynomials stems from physics and the solutions to specific differential equations. The simplest occurrence from physics in this direction is the idea of simple harmonic motion. In this system, a mass m is placed on a horizontal frictionless surface and is attached to a horizontal ideal spring which is fixed to a wall and the mass is set to oscillate. If $x(t)$ denotes the position of the mass at time t, then simple mechanics imply that

$$mx''(t) = -kx(t)$$

where k is Hooke's spring constant. It is not difficult to verify that $\cos(\omega t)$ and $\sin(\omega t)$ are solutions to this equation provided we choose

$$\omega = \sqrt{\frac{k}{m}}.$$

In fact, elementary dimension theory implies that every solution to this differential equation is a linear combination of $\cos(\omega t)$ and $\sin(\omega t)$. It is not difficult to verify that every such solution has the form $\cos(\omega t + t_0)$ and thus has a graph in a wave shape.

Much later in physics, it was the idea that particles have wave-like properties that led to wave-particle duality and the notion of the wave equation in Quantum Mechanics. In 1925, Schrödinger use these and other ideas from physics to derive the wave equation of a particle, work for which he was later awarded the Noble Prize in Physics in 1933. In its simplest one-spacial dimensional form, the wave function $\Psi(x,t)$ assigns a complex number to each point x at time t and satisfies the differential equation

$$-\frac{\hbar^2}{2m}\frac{\partial^2\Psi(x,t)}{\partial x^2} + V(x)\Psi(x,t) = i\hbar\frac{\partial\Psi(x,t)}{\partial t}$$

where $\hbar$ is Planck's reduced constant (Planck's constant divided by 2π), m the mass, and $V(x)$ a potential function at the point x. As noted above, both Ψ and this differential equations involved complex numbers, which was highly questionable in physics as one expects all observable values to be real numbers. Indeed, even Schrödinger stated, "What is unpleasant here, and indeed directly to be objected to, is the use of complex numbers."

However, due to Euler's formula, complex exponentials immediately allow us to combine cosine and sine functions of the same frequencies thereby allowing a unified approach to waves. Indeed, in the case that $V(x) = 0$, we claim that

$$\Psi(x,t) = Ae^{ikx-i\omega t}$$

is a solution to Schrödinger's equation where A is a (complex) constant. Indeed, (provided we can differentiate complex exponentials like real exponentials), we see that

$$\begin{aligned} -\frac{\hbar^2}{2m}\frac{\partial^2\Psi(x,t)}{\partial x^2} &= -\frac{\hbar^2}{2m}(ik)^2Ae^{ikx-i\omega t} = \frac{\hbar^2k^2}{2m}Ae^{ikx-i\omega t} \\ i\hbar\frac{\partial\Psi(x,t)}{\partial t} &= i\hbar(-i\omega)Ae^{ikx-i\omega t} = \hbar\omega Ae^{ikx-i\omega t}, \end{aligned}$$

so, provided

$$\omega = \frac{\hbar k^2}{2m},$$

we indeed have a solution to Schrödinger's equation. Moreover, dimension theory says these are the only solutions.

As solutions to more advanced versions of Schrödinger's equation can be obtained by approximating with linear combinations of these basic solutions, and as Schrödinger's equation is our basis for understanding much of chemistry and physics, the use of complex numbers in both the equation and solutions means nature works with complex numbers, not real ones! For us, the use of complex numbers is desirable in that mathematics, both pure and applied, becomes more elegant with their use. Thus we will use complex numbers throughout this chapter to obtain the desired approximations by trigonometric polynomials.

The motivation for why such approximations are expected in mathematics follows from linear algebra and orthogonal projections. Once the basic facts of these approximations are obtained, we will turn to the idea of convolution and the properties of q_n used in the proof of the Weierstrass Approximation Theorem (Theorem 2.8.6). If we can construct functions with similar properties to q_n using trigonometric functions and develop an analogue of the 'convolution' used in the Weierstrass Approximation Theorem, then hopefully we have convergences of these approximations. However, we need to pay particular attention to the form of convergence as lack of rigour in these arguments lead to some of the biggest plunders in mathematics!

3.1 Motivation for Fourier Series

To begin our study of Fourier series, we turn to the basic structures and underlying linear algebra to motivate our constructions. A reader that has yet to study abstract linear algebra should have the ability to comprehend this section without more complicated linear algebra, whereas the reader familiar with linear algebra can process the more difficult results near the end of the section with ease. We begin with some basic definitions that will be used throughout this chapter.

There are many different types and ways to view Fourier series. The version we will be working with (which is probably the simplest) involves the following set.

Definition 3.1.1. The *unit circle* (or 1-*torus*) is the set

$$\mathbb{T} = \{z \in \mathbb{C} \mid |z| = 1\} = \left\{e^{i\theta} \mid \theta \in (-\pi, \pi]\right\}.$$

Remark 3.1.2. The reason we want to consider $\mathbb{T}$ is that the functions we will want to be working with, namely $\cos(nx)$ and $\sin(nx)$ for $n \in \mathbb{N}$, are 2π-periodic (i.e. $f(x + 2\pi m) = f(x)$ for all $x \in \mathbb{R}$ and $m \in \mathbb{Z}$). Since given any 2π-periodic function f we know for any $x \in \mathbb{R}$ that f is constant on

$$\{x + 2\pi m \mid m \in \mathbb{Z}\},$$

we can view f as a function g on $\mathbb{T}$ by defining

$$g(e^{ix}) = f(x)$$

for all $x \in \mathbb{R}$. This identification creates a bijection between the 2π-periodic functions and the functions on $\mathbb{T}$. There are many benefits of working with $\mathbb{T}$ such as $\mathbb{T}$ is a closed and bounded (and thus compact) subset of $\mathbb{C}$.

It should be pointed out that we will often identify $(-\pi, \pi]$ with $\mathbb{T}$ via the map $x \mapsto e^{ix}$. Note when we do this, there are some particular behaviours with respect to convergence. Notice that although $\left(-\pi + \frac{1}{n}\right)_{n\geq 1}$ is a sequence in $(-\pi, \pi]$ which converges to $-\pi \notin (-\pi, \pi]$, working in $\mathbb{T}$ we see that

$$\lim_{n\to\infty} e^{i\left(-\pi+\frac{1}{n}\right)} = e^{-i\pi} = e^{i\pi}$$

and $\pi \in (-\pi, \pi]$. Thus, when using $(-\pi, \pi]$ for $\mathbb{T}$, we are actually equating $-\pi$ and π in terms of all possible convergences. Furthermore, given $x \in (-\pi, \pi]$ and $t \in \mathbb{R}$, it is possible that $x + t \notin (-\pi, \pi]$. However, we can think of $x + t$ as $e^{i(x+t)}$ which will then reduce to an element of $(-\pi, \pi]$, namely $x+t+2\pi m$ where m is the unique integer such that $x + t + 2\pi m \in (-\pi, \pi]$. Thus we can think of $\mathbb{T}$ as $(-\pi, \pi]$ with these particular topological behaviours.

Furthermore, we could repeat these same ideas using $[0, 2\pi)$ in place of $(-\pi, \pi]$. If we do this, then we would be working with the real numbers modulo 2π. The important part of $\mathbb{R}$ is this continuous loop of angles.

As we will be working with functions on $\mathbb{T}$, it is useful to introduce some notation about various sets of functions. As we will be working with integrals of complex-valued functions, we point the interested reader to Appendix B where such integrals are defined by taking the appropriate linear combination of their real and imaginary parts (and, by doing so, all main results for Riemann integrals extended to this setting). In particular, the most general functions we will be working with in this chapter are the following.

Notation 3.1.3. The *Riemann integrable functions on* $\mathbb{T}$ is the set

$$\mathcal{RI}(\mathbb{T}) = \{f : (-\pi, \pi] \to \mathbb{C} \mid f \text{ is bounded and Riemann integrable}\}.$$

Remark 3.1.4. Although the above definition defines elements of $\mathcal{RI}(\mathbb{T})$ as functions on $(-\pi, \pi]$, by the ideas of Remark 3.1.2 we can also view elements of $\mathcal{RI}(\mathbb{T})$ as functions on $\mathbb{T}$, or as 2π-periodic functions on $\mathbb{R}$. Thus, given $f \in \mathcal{RI}(\mathbb{T})$, we can make sense of $f(z)$ for all $z \in \mathbb{T}$ and $f(x)$ for all $x \in \mathbb{R}$. Typically we will use the latter.

Of course, 2π-periodic continuous functions our are main focus, which we can view via the following.

Notation 3.1.5. The *complex-valued continuous functions on* $\mathbb{T}$ are denoted by $\mathcal{C}(\mathbb{T})$.

Remark 3.1.6. Again, as per Remark 3.1.2, elements of $\mathcal{C}(\mathbb{T})$ can be viewed as 2π-periodic functions on $\mathbb{R}$ in which case $\mathcal{C}(\mathbb{T})$ becomes the set of all continuous 2π-periodic functions on $\mathbb{R}$. If we want to view $\mathbb{T} = (-\pi, \pi]$, then by the idea of identifying $-\pi$ and π in Remark 3.1.2, we see that

$$\mathcal{C}(\mathbb{T}) = \left\{ f : (-\pi, \pi] \to \mathbb{C} \;\middle|\; f \text{ continuous and } \lim_{x \searrow -\pi} f(x) = f(\pi) \right\}.$$

Note clearly $\mathcal{C}(\mathbb{T}) \subseteq \mathcal{RI}(\mathbb{T})$.

It is now time that we describe the functions we hope to approximate elements of $\mathcal{C}(\mathbb{T})$ (or possibly $\mathcal{RI}(\mathbb{T})$) by.

Definition 3.1.7. The *set of trigonometric polynomials on* $\mathbb{T}$ is the set

$$\mathcal{T}(\mathbb{T}) = \left\{ \sum_{k=-n}^{n} c_k e^{ikx} \,\middle|\, n \in \mathbb{N}, c_k \in \mathbb{C} \text{ for all } k \right\}$$

viewed as a subset of $\mathcal{C}(\mathbb{T})$.

If $n \in \mathbb{N}$, the *set of trigonometric polynomials on* $\mathbb{T}$ *of degree at most* n is the set

$$\mathcal{T}_n(\mathbb{T}) = \left\{ \sum_{k=-n}^{n} c_k e^{ikx} \,\middle|\, c_k \in \mathbb{C} \text{ for all } k \right\}$$

(i.e. a polynomial of degree at most n is a linear combination of powers of $(e^{ix})^k$ and $(e^{-ix})^k$ for $k \in \{0, 1, \ldots, n\}$).

A reader at this point may be confused as our initial motivation was to approximate elements of $\mathcal{C}(\mathbb{T})$ by linear combinations of $\cos(nx)$ and $\sin(nx)$ for $n \in \mathbb{N}$. The following shows that the trigonometric polynomials are indeed the functions we are looking for. In particular, note the complex number description of the trigonometric polynomials is more elegant than the following description.

Lemma 3.1.8. *For all* $n \in \mathbb{N}$

$$\mathcal{T}_n(\mathbb{T}) = \left\{ \sum_{k=0}^{n} a_k \cos(kx) + b_k \sin(kx) \,\middle|\, a_k, b_k \in \mathbb{C} \textit{ for all } k \right\}$$

as subsets of functions in $\mathcal{C}(\mathbb{T})$.

Proof. Let

$$T_n = \left\{ \sum_{k=0}^{n} a_k \cos(kx) + b_k \sin(kx) \,\middle|\, a_k, b_k \in \mathbb{C} \text{ for all } k \right\}.$$

To see that $\mathcal{T}_n(\mathbb{T}) \subseteq T_n$, notice for all $(c_k)_{k=-n}^{n} \subseteq \mathbb{C}$ that

$$\begin{aligned}
&\sum_{k=-n}^{n} c_k e^{ikx} \\
&= \sum_{k=-n}^{n} c_k \left(\cos(kx) + i\sin(kx)\right) \\
&= c_0 + \sum_{k=1}^{n} c_k \left(\cos(kx) + i\sin(kx)\right) + c_{-k}\left(\cos(-kx) + i\sin(-kx)\right) \\
&= c_0 \cos(0x) + \sum_{k=1}^{n} (c_k + c_{-k})\cos(kx) + i(c_k - c_{-k})\sin(kx) \in T_n.
\end{aligned}$$

Hence $\mathcal{T}_n(\mathbb{T}) \subseteq T_n$.

For the reverse inclusion, notice for all $(a_k)_{k=0}^n, (b_k)_{k=0}^n \subseteq \mathbb{C}$ that

$$\begin{aligned}
&\sum_{k=0}^n a_k \cos(kx) + b_k \sin(kx) \\
&= a_0 + \sum_{k=1}^n a_k \left(\frac{e^{ikx} + e^{-ikx}}{2} \right) + b_k \left(\frac{e^{ikx} - e^{-ikx}}{2} \right) \\
&= a_0 e^{i0x} + \sum_{k=1}^n \frac{a_k + b_k}{2} e^{ikx} + \frac{a_k - b_k}{2} e^{-ikx} \in \mathcal{T}_n(\mathbb{T}).
\end{aligned}$$

Hence $T_n \subseteq \mathcal{T}_n(\mathbb{T})$ thereby completing the proof. ∎

If we are going to approximate a function $f \in \mathcal{C}(\mathbb{T})$ using trigonometric polynomials, one idea for finding such approximations is to find the trigonometric polynomial that is 'closest' to f. Of course $\mathcal{T}(\mathbb{T})$ is an infinite dimensional vector space and thus it might be better to find the elements of $\mathcal{T}_n(\mathbb{T})$ that are 'closest' to f for each $n \in \mathbb{N}$. To do so, we need to define what we mean by 'closest'. This of course means we need some notion of distance. This is thus where our course intersects linear algebra in that we can import some notions of distance and, more important, orthogonality into this context. These notions will revolve around the following object.

Definition 3.1.9. The *inner product on* $\mathcal{RI}(\mathbb{T})$ is the function $\langle \cdot, \cdot \rangle : \mathcal{RI}(\mathbb{T}) \times \mathcal{RI}(\mathbb{T}) \to \mathbb{C}$ defined by

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

for all $f, g \in \mathcal{RI}(\mathbb{T})$.

Remark 3.1.10. As the complex conjugate and products of elements of $\mathcal{RI}(\mathbb{T})$ remain elements of $\mathcal{RI}(\mathbb{T})$, the inner product on $\mathcal{RI}(\mathbb{T})$ is well-defined. Moreover, the reason we added a $\frac{1}{2\pi}$ in front of the integral is so that if 1 represents the constant function that takes the value 1 at each point in $\mathbb{T}$, then

$$\langle 1, 1 \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} 1\, dx = 1.$$

Thus the $\frac{1}{2\pi}$ in front of the integral acts as a normalization. This will be explored further shortly.

However, the above inner product on $\mathcal{RI}(\mathbb{T})$ is not an inner product on $\mathcal{RI}(\mathbb{T})$ in the sense of linear algebra since if $f : \mathbb{T} \to \mathbb{C}$ is defined by

$$f(x) = \begin{cases} 1 & \text{if } x = 1 \\ 0 & \text{if } x \neq 1 \end{cases}$$

for all $x \in \mathbb{T}$, then $f \neq 0$ but $\langle f, f\rangle = 0$. Thus the reason we call this an inner product is that it shares all but this property with inner products and is indeed an inner product if we reduce to $\mathcal{C}(\mathbb{T})$. (Moreover, if one 'corrects' $\mathcal{RI}(\mathbb{T})$, then this will be indeed an inner product. However, that is a topic for MATH 4012).

Proposition 3.1.11. *The inner product on $\mathcal{RI}(\mathbb{T})$ has the following properties:*

a) $\langle 0, 0\rangle = 0$.

b) $\langle f, f\rangle \geq 0$ *for all* $f \in \mathcal{RI}(\mathbb{T})$.

c) If $f \in \mathcal{C}(\mathbb{T})$ *and* $\langle f, f\rangle = 0$, *then* $f = 0$.

d) $\langle \alpha_1 f_1 + \alpha_2 f_2, g\rangle = \alpha_1\langle f_1, g\rangle + \alpha_2\langle f_2, g\rangle$ *for all* $f_1, f_2, g \in \mathcal{RI}(\mathbb{T})$ *and* $\alpha_1, \alpha_2 \in \mathbb{C}$.

e) $\langle g, \alpha_1 f_1 + \alpha_2 f_2\rangle = \overline{\alpha_1}\langle g, f_1\rangle + \overline{\alpha_2}\langle g, f_2\rangle$ *for all* $f_1, f_2, g \in \mathcal{RI}(\mathbb{T})$ *and* $\alpha_1, \alpha_2 \in \mathbb{C}$.

f) $\overline{\langle f, g\rangle} = \langle g, f\rangle$ *for all* $f, g \in \mathcal{RI}(\mathbb{T})$.

Proof. Clearly a) is true. To see that b) is true, note for all $f \in \mathcal{RI}(\mathbb{T})$ that

$$\langle f, f\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|^2\, dx \geq 0$$

since $|f(x)|^2 \geq 0$ for all $x \in \mathbb{T}$.

To see that c) is true, let $f \in \mathcal{C}(\mathbb{T})$ be such that $\langle f, f\rangle = 0$. Then

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

Since f is continuous and thus $|f|^2$ is continuous, this implies by MATH 2001 that $|f(x)|^2 = 0$ for all $x \in \mathbb{T}$ and thus $f = 0$ as desired.

To see that d) is true, let $f_1, f_2, g \in \mathcal{RI}(\mathbb{T})$ and $\alpha_1, \alpha_2 \in \mathbb{C}$ be arbitrary. Then

$$\begin{aligned}
\langle \alpha_1 f_1 + \alpha_2 f_2, g\rangle &= \frac{1}{2\pi}\int_{-\pi}^{\pi} (\alpha_1 f_1(x) + \alpha_2 f_2(x))\overline{g(x)}\, dx \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} \alpha_1 f_1(x)\overline{g(x)} + \alpha_2 f_2(x)\overline{g(x)}\, dx \\
&= \alpha_1\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} f_1(x)\overline{g(x)}\, dx\right) + \alpha_2\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} f_2(x)\overline{g(x)}\, dx\right) \\
&= \alpha_1\langle f_1, g\rangle + \alpha_2\langle f_2, g\rangle
\end{aligned}$$

where the third equality comes from the fact that the integral is linear. Hence d) holds.

To see that e) is true, one can construct a proof using the same ideas as the proof of part d) or notice for all $f_1, f_2, g \in \mathcal{RI}(\mathbb{T})$ and $\alpha_1, \alpha_2 \in \mathbb{C}$ that parts d) and f) imply that

$$\begin{aligned}
\langle g, \alpha_1 f_1 + \alpha_2 f_2\rangle &= \overline{\langle \alpha_1 f_1 + \alpha_2 f_2, g\rangle} && \text{by f)}\\
&= \overline{\alpha_1\langle f_1, g\rangle + \alpha_2\langle f_2, g\rangle} && \text{by d)}\\
&= \overline{\alpha_1}\,\overline{\langle f_1, g\rangle} + \overline{\alpha_2}\,\overline{\langle f_2, g\rangle} && \text{by complex conjugates}\\
&= \overline{\alpha_1}\langle g, f_1\rangle + \overline{\alpha_2}\langle g, f_2\rangle && \text{by f).}
\end{aligned}$$

Thus it suffices to prove f).

To see that f) is true, notice for all $f, g \in \mathcal{RI}(\mathbb{T})$ that

$$\begin{aligned}
\overline{\langle f, g\rangle} &= \overline{\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\overline{g(x)}\,dx}\\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} \overline{f(x)\overline{g(x)}}\,dx\\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} g(x)\overline{f(x)}\,dx\\
&= \langle g, f\rangle
\end{aligned}$$

as desired. ∎

As remarked earlier, Proposition 3.1.11 shows that the inner product on $\mathcal{RI}(\mathbb{T})$ is indeed an inner product on $\mathcal{C}(\mathbb{T})$. Using the same idea from linear algebra, we obtain a notion of a length.

Definition 3.1.12. The *length of an element* $f \in \mathcal{RI}(\mathbb{T})$, denoted $\|f\|_2$, is defined to be

$$\|f\|_2 = \sqrt{\langle f, f\rangle} = \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|^2\,dx\right)^{\frac{1}{2}} \geq 0.$$

It is elementary to see that $\|f\|_2$ is well-defined and non-negative for all $f \in \mathcal{RI}(\mathbb{T})$. To show that this notion of a length has the properties one would expect of a length function (i.e. properties similar to the absolute value of complex numbers), we first need to prove the following.

Theorem 3.1.13 (Cauchy-Schwarz Inequality). *For all* $f, g \in \mathcal{RI}(\mathbb{T})$,

$$|\langle f, g\rangle| \leq \|f\|_2 \|g\|_2 .$$

Proof 1 (Linear Algebra Techniques with a Complication). In this first proof, we will present the classical proof of the Cauchy-Schwarz Inequality from linear algebra. However, there is a slight complication in that we are not quite working with an inner product on $\mathcal{RI}(\mathbb{T})$.

First, we claim that if $\|f\|_2 = 0$, then $\langle f, g\rangle = 0$ and thus the inequality hold. The best way to see this uses technology from MATH 4012. In particular, if

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

then it must be true that $f(x) = 0$ for "almost every $x \in \mathbb{T}$" so $f(x)\overline{g(x)} = 0$ "almost every $x \in \mathbb{T}$" and thus

$$|\langle f, g\rangle| = \left|\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\overline{g(x)}\,dx\right| = \left|\frac{1}{2\pi}\int_{-\pi}^{\pi} 0\,dx\right| = 0 \leq \|f\|_2\,\|g\|_2\,.$$

Of course, the above is easily seen to be true when $f \in \mathcal{C}(\mathbb{T})$. Moreover, if $f \in \mathcal{RI}(\mathbb{T})$, then f is "almost continuous" (see Lemma 3.3.8).

To prove this directly without further technology requires us to delve deep into Riemann sum computations that we outline here. Indeed if $\|f\|_2 = 0$, then for any $\epsilon > 0$ there exists a partition $\mathcal{P}$ of $[-\pi, \pi]$ such that $U(|f|^2, \mathcal{P}) < \epsilon$. This means for each interval in $\mathcal{P}$ that either the length of the interval 'is small' or $|f(x)|^2$ is 'small' on the interval. This can be used to show that $U(|f|, \mathcal{P}) < C(\epsilon + \sqrt{\epsilon})$ for some constant C depending only on f (i.e. for the intervals of 'small' length we use the fact that f is bounded to keep the upper Riemann sum 'small', and on the other intervals knowing $|f(x)|^2$ is 'small' on the interval implies $|f(x)|$ is 'small' on the interval). Therefore, since $g \in \mathcal{RI}(\mathbb{T})$, g is bounded so $U(|f\overline{g}|, \mathcal{P}) < CM(\epsilon + \sqrt{\epsilon})$ for some constant M depending on g. Therefore, since

$$\left|\int_{-\pi}^{\pi} f(x)\overline{g(x)}\,dx\right| \leq \int_{-\pi}^{\pi} |f(x)\overline{g(x)}|\,dx \leq U(|f\overline{g}|, \mathcal{P}) < CM(\epsilon + \sqrt{\epsilon}),$$

we obtain that $\langle f, g\rangle = 0$. The details are left as an exercise.

Therefore, if $\|f\|_2 = 0$, then $\langle f, g\rangle = 0$ and thus the inequality hold. By using similar arguments or Proposition 3.1.11 part f), if $\|g\|_2 = 0$, then $\langle f, g\rangle = 0$ and thus the inequality hold. Therefore, to complete the proof, we may assume that $\|f\|_2 \neq 0$ and $\|g\|_2 \neq 0$.

Choose $z \in \mathbb{C}$ with $|z| = 1$ such that

$$\langle zf, g\rangle = z\langle f, g\rangle = |\langle f, g\rangle|$$

(that is, if $\langle f, g\rangle = re^{i\theta}$ for $r \geq 0$ and $\theta \in [0, 2\pi)$, then take $z = e^{-i\theta}$). Notice by Proposition 3.1.11 that for all $t \in \mathbb{R}$

$$\begin{aligned} 0 &\leq \langle zf + tg, zf + tg\rangle \\ &= |z|^2\langle f, f\rangle + t\overline{z}\langle g, f\rangle + tz\langle f, g\rangle + t^2\langle g, g\rangle \\ &= \langle f, f\rangle + 2t|\langle f, g\rangle| + t^2\langle g, g\rangle. \end{aligned}$$

By substituting

$$t_0 = -\frac{|\langle f, g\rangle|}{\langle g, g\rangle}$$

which is well-defined as $\langle g, g\rangle = \|g\|_2^2 \neq 0$, we obtain that

$$0 \leq \langle f, f\rangle - 2\frac{|\langle f, g\rangle|^2}{\langle g, g\rangle} + \frac{|\langle f, g\rangle|^2}{\langle g, g\rangle} = \langle f, f\rangle - \frac{|\langle f, g\rangle|^2}{\langle g, g\rangle}.$$

By rearranging, we obtain that

$$|\langle f, g\rangle|^2 \leq \langle f, f\rangle\langle g, g\rangle,$$

which then implies the desired inequality by taking square roots. ∎

Proof 2 (Direct Proof). We will provide a direct proof of the Cauchy-Schwarz Inequality that is specific to $\mathcal{RI}(\mathbb{T})$.

First notice for all $z, w \in \mathbb{C}$ and $r \in \mathbb{R}$ with $r > 0$ that

$$0 \leq (r|z| - r^{-1}|w|)^2 \leq r^2|z|^2 - 2|z||w| + r^{-2}|w|^2$$

so

$$|z\overline{w}| = |z||\overline{w}| = |z||w| = |zw| \leq \frac{1}{2}\left(r^2|z|^2 + r^{-2}|w|^2\right).$$

Therefore, for all $f, g \in \mathcal{RI}(\mathbb{T})$ and for all $r \in \mathbb{R}$ with $r > 0$, we have that

$$\begin{aligned} |\langle f, g\rangle| &= \left|\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\overline{g(x)}\,dx\right| \\ &\leq \frac{1}{2\pi}\int_{-\pi}^{\pi} \left|f(x)\overline{g(x)}\right|\,dx \\ &\leq \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{1}{2}r^2|f(x)|^2 + \frac{1}{2}r^{-2}|g(x)|^2\,dx \\ &= \frac{1}{2}\left(r^2\|f\|_2^2 + r^{-2}\|g\|_2^2\right). \end{aligned}$$

To see that the above yields the desired inequality, we will consider a few cases. First, if $\|g\|_2 = 0$, then by sending r to 0 from above we obtain that

$$|\langle f, g\rangle| = 0 \leq \|f\|_2\|g\|_2$$

as desired. Similarly, if $\|f\|_2 = 0$, then by sending r to infinity we obtain that

$$|\langle f, g\rangle| = 0 \leq \|f\|_2\|g\|_2$$

as desired. Finally, if $\|f\|_2 \neq 0$ and $\|g\|_2 \neq 0$, then by setting

$$r = \sqrt{\frac{\|g\|_2}{\|f\|_2}} \in (0, \infty)$$

we obtain that

$$|\langle f, g\rangle| \leq \frac{1}{2}\left(\frac{\|g\|_2}{\|f\|_2}\|f\|^2 + \frac{\|f\|_2}{\|g\|_2}\|g\|^2\right) = \|f\|_2\|g\|_2$$

as desired. ∎

Using the Cauchy-Scwarz Inequality (Theorem 3.1.13), we can demonstrate this length function has the desired properties.

Proposition 3.1.14. *The length function on $\mathcal{RI}(\mathbb{T})$ has the following properties:*

a) $\|f\|_2 \geq 0$ *for all* $f \in \mathcal{RI}(\mathbb{T})$.

b) If $f \in \mathcal{C}(\mathbb{T})$ *and* $\|f\|_2 = 0$, *then* $f = 0$.

c) $\|\alpha f\|_2 = |\alpha| \|f\|_2$ *for all* $f \in \mathcal{RI}(\mathbb{T})$ *and* $\alpha \in \mathbb{C}$.

d) (Triangle Inequality) $\|f + g\|_2 \leq \|f\|_2 + \|g\|_2$ *for all* $f, g \in \mathcal{RI}(\mathbb{T})$.

Proof. Clearly a) and b) follow from Proposition 3.1.11 parts b) and c). To see that c) is true, notice for all $f \in \mathcal{RI}(\mathbb{T})$ and $\alpha \in \mathbb{C}$ that

$$\begin{aligned} \|\alpha f\|_2 &= \sqrt{\langle \alpha f, \alpha f\rangle} && \text{definition} \\ &= \sqrt{\alpha\overline{\alpha}\langle f, f\rangle} && \text{by Proposition 3.1.11 parts d) and e)} \\ &= \sqrt{|\alpha|^2\langle f, f\rangle} \\ &= |\alpha| \|f\|_2 \end{aligned}$$

as desired.

To see that d) is true, notice for all $f, g \in \mathcal{RI}(\mathbb{T})$ that

$$\begin{aligned} &\|f + g\|_2^2 \\ &= \langle f + g, f + g\rangle \\ &= \langle f, f\rangle + \langle f, g\rangle + \langle g, f\rangle + \langle g, g\rangle && \text{by Proposition 3.1.11 parts d) and e)} \\ &= \|f\|_2^2 + \langle f, g\rangle + \overline{\langle f, g\rangle} + \|g\|_2^2 && \text{by Proposition 3.1.11 part f)} \\ &= \|f\|_2^2 + 2\mathrm{Re}(\langle f, g\rangle) + \|g\|_2^2 \\ &\leq \|f\|_2^2 + 2|\langle f, g\rangle| + \|g\|_2^2 \\ &\leq \|f\|_2^2 + 2\|f\|_2^2 \|g\|_2^2 + \|g\|_2^2 && \text{by Cauchy-Schwarz (Theoremm 3.1.13)} \\ &= (\|f\|_2 + \|g\|_2)^2 . \end{aligned}$$

Therefore, by taking the square roots of both sides, the Triangle Inequality has been demonstrated. ∎

Now that we have a length function on $\mathcal{RI}(\mathbb{T})$, we can ask the following question, "Given $f \in \mathcal{RI}(\mathbb{T})$ and $n \in \mathbb{N}$, how can we find the element $p \in \mathcal{T}_n(\mathbb{T})$ such that $\|f - p\|_2$ is as small as possible?" Linear algebra dictates that this element p can be computed by taking the 'orthogonal projection' of f onto $\mathcal{T}_n(\mathbb{T})$. Thus we review this technology.

Definition 3.1.15. A non-empty set $S \subseteq \mathcal{RI}(\mathbb{T})$ is said to be an *orthonormal set* if for all $f, g \in S$ we have

$$\langle f, g\rangle = \begin{cases} 0 & \text{if } f \neq g \\ 1 & \text{if } f = g \end{cases}.$$

Luckily, there are natural orthonormal sets that span the trigonometric polynomials. Indeed consider the following.

Notation 3.1.16. For each $n \in \mathbb{Z}$, let $e_n : \mathbb{T} \to \mathbb{C}$ denote the function defined by

$$e_n(x) = e^{inx}$$

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

Theorem 3.1.17. *Let*

$$\mathcal{B} = \{e_n \mid n \in \mathbb{Z}\} \subseteq \mathcal{T}(\mathbb{T}).$$

Then $\mathcal{B}$ is an orthonormal set in $\mathcal{RI}(\mathbb{T})$.

Proof. First, notice for all $n \in \mathbb{Z}$ that

$$\begin{aligned}\langle e_n, e_n\rangle &= \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{inx}\overline{e^{inx}}\, dx \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{inx}e^{-inx}\, dx \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{inx-inx}\, dx \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{0}\, dx \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} 1\, dx = 1.\end{aligned}$$

Moreover, for all $n, m \in \mathbb{Z}$ with $n \neq m$, we see that

$$\begin{aligned}
\langle e_n, e_m \rangle &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{inx} \overline{e^{imx}} \, dx \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{inx} e^{-imx} \, dx \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(n-m)x} \, dx \\
&= \frac{1}{2\pi} \left(\frac{1}{i(n-m)} e^{i(n-m)x} \right) \Big|_{x=-\pi}^{\pi} \\
&= \frac{1}{2\pi} \frac{1}{i(n-m)} \left(e^{i(n-m)\pi} - e^{i(n-m)(-\pi)} \right) \\
&= \frac{1}{2\pi} \frac{1}{i(n-m)} \left(\cos((n-m)\pi) + i \sin((n-m)\pi) \right) \\
&\quad - \frac{1}{2\pi} \frac{1}{i(n-m)} \left(\cos(-(n-m)\pi) + i \sin(-(n-m)\pi) \right) \\
&= \frac{1}{2\pi} \frac{1}{i(n-m)} \left(\cos((n-m)\pi) + 0i - \cos((n-m)\pi) - 0i \right) \\
&= 0.
\end{aligned}$$

Hence $\mathcal{B}$ is an orthonormal set in $\mathcal{RI}(\mathbb{T})$. ∎

Once one has an orthogonal spanning set for a finite dimensional subspace, one can easily construct the orthogonal projection onto said subspace. In our context, this reduces to the following.

Theorem 3.1.18. *For each $n \in \mathbb{N}$ define $P_n : \mathcal{RI}(\mathbb{T}) \to \mathcal{T}_n(\mathbb{T})$ by*

$$P_n(f) = \sum_{k=-n}^{n} \langle f, e_k \rangle e_k$$

for all $f \in \mathcal{RI}(\mathbb{T})$. Then the following hold:

a) $P_n(\alpha_1 f_1 + \alpha_2 f_2) = \alpha_1 P_n(f_1) + \alpha_2 P_n(f_2)$ for all $f_1, f_2 \in \mathcal{RI}(\mathbb{T})$ and $\alpha_1, \alpha_2 \in \mathbb{C}$.

b) $P_n(f) = f$ for all $f \in \mathcal{T}_n(\mathbb{T})$.

c) $\langle f - P_n(f), p \rangle = 0$ for all $f \in \mathcal{RI}(\mathbb{T})$ and $p \in \mathcal{T}_n(\mathbb{T})$.

d) For all $f \in \mathcal{RI}(\mathbb{T})$,

$$\|f - P_n(f)\|_2 \leq \inf \{ \|f - p\|_2 \mid p \in \mathcal{T}_n(\mathbb{T}) \}.$$

Moreover, if $p \in \mathcal{T}_n(\mathbb{T})$ and $\|f - p\|_2 = \|f - P_n(f)\|_2$, then $p = P_n(f)$.

That is, P_n is the orthogonal projection of $\mathcal{RI}(\mathbb{T})$ onto $\mathcal{T}_n(\mathbb{T})$.

Proof. To see that a) is true, notice by Proposition 3.1.11 parts d) and e) that for all $f_1, f_2 \in \mathcal{RI}(\mathbb{T})$ and $\alpha_1, \alpha_2 \in \mathbb{C}$

$$\begin{aligned} P_n(\alpha_1 f_1 + \alpha_2 f_2) &= \sum_{k=-n}^{n} \langle \alpha_1 f_1 + \alpha_2 f_2, e_k \rangle \, e_k \\ &= \sum_{k=-n}^{n} \alpha_1 \langle f_1, e_k \rangle e_k + \alpha_2 \langle f_2, e_k \rangle e_k \\ &= \alpha_1 P_n(f_1) + \alpha_2 P_n(f_2) \end{aligned}$$

as desired.

To see that b) is true, let $f \in \mathcal{T}_n(\mathbb{T})$ be arbitrary. Thus there exists $(c_j)_{j=-n}^n \in \mathbb{C}$ such that

$$f = \sum_{j=-n}^{n} c_j e_j.$$

Therefore

$$\begin{aligned} P_n(f) &= \sum_{k=-n}^{n} \left\langle \sum_{j=-n}^{n} c_j e_j, e_k \right\rangle e_k \\ &= \sum_{k=-n}^{n} \sum_{j=-n}^{n} c_j \, \langle e_j, e_k \rangle \, e_k && \text{by Propositon 3.1.11 part d)} \\ &= \sum_{k=-n}^{n} c_k \, \langle e_k, e_k \rangle \, e_k && \text{by Theorem 3.1.17} \\ &= \sum_{k=-n}^{n} c_k e_k && \text{by Theorem 3.1.17} \\ &= f \end{aligned}$$

as desired.

To see that c) is true, fix $f \in \mathcal{RI}(\mathbb{T})$ and let $p \in \mathcal{T}_n(\mathbb{T})$ be arbitrary. Thus there exists $(c_j)_{j=-n}^n \in \mathbb{C}$ such that

$$p = \sum_{j=-n}^{n} c_j e_j.$$

Therefore, by similar arguments to those used in part b), we see that

$$\begin{aligned}\langle f - P_n(f), p\rangle &= \langle f, p\rangle - \langle P_n(f), p\rangle \\ &= \left\langle f, \sum_{j=-n}^{n} c_j e_j \right\rangle - \left\langle \sum_{k=-n}^{n} \langle f, e_k\rangle e_k, \sum_{j=-n}^{n} c_j e_j \right\rangle \\ &= \sum_{j=-n}^{n} \overline{c_j}\langle f, e_j\rangle - \sum_{j=-n}^{n}\sum_{k=-n}^{n} \langle f, e_k\rangle \overline{c_j}\langle e_k, e_j\rangle \\ &= \sum_{j=-n}^{n} \overline{c_j}\langle f, e_j\rangle - \sum_{j=-n}^{n} \langle f, e_j\rangle \overline{c_j} \\ &= 0\end{aligned}$$

as desired.

Finally, to see that d) is true, let $p \in \mathcal{T}_n(\mathbb{T})$ be arbitrary. Then

$$\begin{aligned}\|f-p\|_2^2 &= \|(f - P_n(f)) + (P_n(f) - p)\|_2^2 \\ &= \langle (f - P_n(f)) + (P_n(f) - p), (f - P_n(f)) + (P_n(f) - p)\rangle \\ &= \langle (f - P_n(f)), (f - P_n(f))\rangle + \langle (f - P_n(f)), (P_n(f) - p)\rangle \\ &\quad + \langle (P_n(f) - p), (f - P_n(f))\rangle + \langle (P_n(f) - p), (P_n(f) - p)\rangle \\ &= \|f - P_n(f)\|_2^2 + 0 + 0 + \|P_n(f) - p\|_2^2 \\ &\geq \|f - P_n(f)\|_2^2\end{aligned}$$

where the fourth equality comes from part c) and the fact that $P_n(f) - p \in \mathcal{T}_n(\mathbb{T})$. Hence the desired inequality holds. Moreover, if $\|f - p\|_2 = \|f - P_n(f)\|_2$, the above computation implies that $\|P_n(f) - p\|_2 = 0$ and thus $P_n(f) = p$ by Proposition 3.1.14 part b) since $P_n(f), p \in \mathcal{T}_n(\mathbb{T}) \subseteq \mathcal{C}(\mathbb{T})$. ∎

3.2 Basics of Fourier Series

Section 3.1 shows that given an $f \in \mathcal{RI}(\mathbb{T})$ and $n \in \mathbb{N}$, the element of $\mathcal{T}_n(\mathbb{T})$ that is closest to f is $P_n(f)$. Thus, if we desire to approximate f using trigonometric polynomials, it is natural to consider the sequence $(P_n(f))_{n\geq 1}$ for our approximations. The questions that remain are, "In what way do we desire to approximate f with $(P_n(f))_{n\geq 1}$ and will this work?"

In order to proceed with these questions, it is useful to analyze the specific form of $P_n(f)$ and note that the sequence $(P_n(f))_{n\geq 1}$ can be expressed in a nice way.

Definition 3.2.1. Let $f \in \mathcal{RI}(\mathbb{T})$ and let $n \in \mathbb{Z}$. The n^{th} *Fourier coefficient of* f, denoted $\widehat{f}(n)$, is

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

Definition 3.2.2. Let $f \in \mathcal{RI}(\mathbb{T})$ and let $n \in \mathbb{N}$. The n^{th} *partial Fourier series of* f, is the orthogonal projection of f onto $\mathcal{T}_n(\mathbb{T})$ from Theorem 3.1.18; that is,

$$P_n(f) = \sum_{k=-n}^{n} \langle f, e_k \rangle e_k = \sum_{k=-n}^{n} \widehat{f}(k) e_k = \sum_{k=-n}^{n} \widehat{f}(k) e^{ikx}.$$

As the above expressions show, $(P_n(f))_{n\geq 1}$ is quite a nice sequence being the sequence of partial sums of a series of functions. For terminology, we define the following.

Definition 3.2.3. Let $f \in \mathcal{RI}(\mathbb{T})$. The *Fourier series of* f, denoted $\mathcal{F}(f)$, is the series of functions

$$\mathcal{F}(f) = \sum_{k=-\infty}^{\infty} \widehat{f}(k) e^{ikx}.$$

Remark 3.2.4. With the above out of the way, we can turn our attention to our main question of this chapter, "Given an $f \in \mathcal{RI}(\mathbb{T})$, does $\mathcal{F}(f)$ converge to f and in which sense does it converge?"

It is useful to point our that if $f, g \in \mathcal{RI}(\mathbb{T})$ are such that f and g differ at exactly one point, then the above definitions shows

$$\widehat{f}(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx}\, dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} g(x) e^{-inx}\, dx = \widehat{g}(n)$$

for all $n \in bZ$. Hence $\mathcal{F}(f) = \mathcal{F}(g)$. Therefore, we cannot hope that $\mathcal{F}(f)(x) = f(x)$ for every $x \in \mathbb{T}$ since changing the value of f at a single point does not change $\mathcal{F}(f)$.

Luckily, if $f \in \mathcal{C}(\mathbb{T})$ we need not worry about the above example as we cannot change the value of f at a single point at remain in $\mathcal{C}(\mathbb{T})$. Moreover, since $P_n(f) \in \mathcal{T}_n(\mathbb{T}) \subseteq \mathcal{C}(\mathbb{T})$, we can indeed ask that $\mathcal{F}(f)$ is equal to f since $\mathcal{F}(f)$ will be a series of continuous functions and thus could equal f. However, as Example 2.2.3 shows, a pointwise convergent series of continuous functions need not be continuous. So we are left with the following questions:

Question 3.2.5. *If $f \in \mathcal{C}(\mathbb{T})$, does $\mathcal{F}(f)$ converge uniformly and, if so, does it converge uniformly to f?*

Question 3.2.6. *If $f \in \mathcal{C}(\mathbb{T})$, is $\mathcal{F}(f)(x) = f(x)$ for all $x \in \mathbb{T}$?*

Of course, a positive answer to Question 3.2.5 implies a positive answer to Question 3.2.6. In order to try and answer these questions, it is useful to develop the basic properties and look at some examples of Fourier series. We begin with the following.

Example 3.2.7. For $m \in \mathbb{N}$ and $(c_k)_{k=-m}^m \in \mathbb{C}$, consider the trigonometric polynomial

$$p(x) = \sum_{j=-m}^{m} c_j e^{ijx}$$

for all $x \in \mathbb{T}$. Notice for all $n \in \mathbb{N}$ that

$$\widehat{p}(n) = \left\langle \sum_{j=-m}^{m} c_j e_j, e_n \right\rangle = \sum_{j=-m}^{m} c_j \langle e_j, e_n \rangle = \begin{cases} c_n & \text{if } |n| \leq m \\ 0 & \text{otherwise} \end{cases}.$$

Hence for all $n \geq m$ we have that $P_n(p) = p$ (this can also be seen by Theorem 3.1.18 part b)) and thus $\mathcal{F}(p) = p$.

Example 3.2.8. For $m \in \mathbb{N}$, let $f, g : \mathbb{T} \to \mathbb{R}$ by defined by

$$f(x) = \cos(mx) \qquad \text{and} \qquad g(x) = \sin(mx)$$

for all $x \in \mathbb{T}$. Therefore, since $f = \frac{1}{2}(e_m + e_{-m})$ and $g = \frac{1}{2i}(e_m - e_{-m})$ are trigonometric polynomials, Example 3.2.7 implies that

$$\widehat{f}(n) = \begin{cases} \frac{1}{2} & \text{if } n = m \\ \frac{1}{2} & \text{if } n = -m \\ 0 & \text{otherwise} \end{cases} \qquad \text{and} \qquad \widehat{g}(n) = \begin{cases} \frac{1}{2i} & \text{if } n = m \\ -\frac{1}{2i} & \text{if } n = -m \\ 0 & \text{otherwise} \end{cases}.$$

Hence, for all $n \geq m$, we see that

$$P_n(f) = \frac{1}{2}e_m + \frac{1}{2}e_{-m} = f \qquad \text{and} \qquad P_n(g) = \frac{1}{2i}e_m - \frac{1}{2i}e_{-m} = g.$$

Hence $\mathcal{F}(f) = f$ and $\mathcal{F}(g) = g$.

Example 3.2.8 has wider reaching implications in that if $f : \mathbb{T} \to \mathbb{R}$, we would hope that we could express $\mathcal{F}(f)$ as a series of real functions instead of complex exponentials. This would be of particular interest to several applications in physics and applied mathematics. In order to do so, we desire some further properties of the Fourier coefficients. In particular, we desired to know how the Fourier coefficients behave under certain operations. In addition to the standard operations, the following will also be of use to us.

Definition 3.2.9. For each $f \in \mathcal{RI}(\mathbb{T})$, the *complex conjugate of* f is the function $\overline{f} : \mathbb{T} \to \mathbb{C}$ defined by $\overline{f}(x) = \overline{f(x)}$ for all $x \in \mathbb{T}$.

Definition 3.2.10. For each $f \in \mathcal{RI}(\mathbb{T})$ and $y \in \mathbb{R}$, the *translation of* f *by* y is the function $f_y : \mathbb{T} \to \mathbb{C}$ defined by $f_y(x) = f(x - y)$ for all $x \in \mathbb{T}$ (where, by $x - y$ we mean $x - y \mod 2\pi$).

Definition 3.2.11. For each $f \in \mathcal{RI}(\mathbb{T})$, the *inversion of* f is the function $\check{f} : \mathbb{T} \to \mathbb{C}$ defined by $\check{f}(x) = f(-x)$ for all $x \in \mathbb{T}$.

Proposition 3.2.12. *For all $f, g \in \mathcal{RI}(\mathbb{T})$, the following hold:*

a) $\widehat{(f+g)}(n) = \widehat{f}(n) + \widehat{g}(n)$ *for all* $n \in \mathbb{Z}$. *Hence* $\mathcal{F}(f+g) = \mathcal{F}(f) + \mathcal{F}(g)$.

b) $\widehat{(\alpha f)}(n) = \alpha \widehat{f}(n)$ *for all* $\alpha \in \mathbb{C}$ *and* $n \in \mathbb{Z}$. *Hence* $\mathcal{F}(\alpha f) = \alpha \mathcal{F}(f)$.

c) $\widehat{\overline{f}}(n) = \overline{\widehat{f}(-n)}$ *for all* $n \in \mathbb{Z}$.

d) $\widehat{f_y}(n) = e^{-iny}\widehat{f}(n)$ *for all* $y \in \mathbb{R}$ *and* $n \in \mathbb{Z}$.

e) $\widehat{\check{f}}(n) = \widehat{f}(-n)$ *for all* $n \in \mathbb{Z}$.

f) $|\widehat{f}(n)| \leq \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|\, dx$ *for all* $n \in \mathbb{Z}$.

Proof. Clearly a) and b) follow trivially by the linearity of the integral.

To see that c) is true, notice for all $f \in \mathcal{RI}(\mathbb{T})$ that

$$\begin{aligned}
\widehat{\overline{f}}(n) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \overline{f}(x) e^{-inx}\, dx \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} \overline{f(x)e^{inx}}\, dx \\
&= \overline{\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{inx}\, dx} \\
&= \overline{\widehat{f}(-n)}
\end{aligned}$$

as desired.

To see that d) is true, notice for all $f \in \mathcal{RI}(\mathbb{T})$ and $y \in \mathbb{R}$ that

$$\begin{aligned}
\widehat{f_y}(n) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f_y(x) e^{-inx}\, dx \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-y) e^{-inx}\, dx \\
&= \frac{1}{2\pi}\int_{-\pi-y}^{\pi-y} f(t) e^{-in(y+t)}\, dt && \text{substitute } t = x - y \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t) e^{-in(y+t)}\, dt && 2\pi\text{-periodic} \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t) e^{-iny} e^{-int}\, dt \\
&= \frac{1}{2\pi} e^{-iny}\int_{-\pi}^{\pi} f(t) e^{-int}\, dt \\
&= e^{-iny}\widehat{f}(n)
\end{aligned}$$

as desired.

To see that e) is true, notice for all $f \in \mathcal{RI}(\mathbb{T})$ that

$$\begin{aligned}\widehat{\check{f}}(n) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \check{f}(x)e^{-inx}\,dx \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(-x)e^{-inx}\,dx \\ &= \frac{1}{2\pi}\int_{\pi}^{-\pi} f(t)e^{-in(-t)}(-1)\,dt \quad \text{substitute } t=-x \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-i(-n)t}\,dt \\ &= \widehat{f}(-n)\end{aligned}$$

as desired.

Finally, to see that f) is true, notice that

$$|\widehat{f}(n)| = \left|\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}\,dx\right| \leq \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|f(x)e^{-inx}\right|\,dx = \frac{1}{2\pi}\int_{-\pi}^{\pi}|f(x)|\,dx$$

as desired. ■

Remark 3.2.13. Note that nothing has been said about the Fourier coefficients of a product of two elements in $\mathcal{RI}(\mathbb{T})$. In particular, "if $f, g \in \mathcal{RI}(\mathbb{T})$, is it true that

$$(\widehat{fg})(n) = \widehat{f}(n)\widehat{g}(n)$$

for all $n \in \mathbb{Z}$?" It turns out this answer is no! After all, integrals often do not behave well with respect to products. For example, if $f(x) = x$ for all $x \in (-\pi, \pi]$, then

$$(\widehat{f}(0))^2 = \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\,dx\right)^2 = 0^2 = 0$$

whereas

$$(\widehat{f^2})(0) = \frac{1}{2\pi}\int_{-\pi}^{\pi} x^2\,dx = \frac{1}{2\pi}\left(\frac{1}{3}x^3\right)\Big|_{x=-\pi}^{\pi} = \frac{\pi^2}{3} \neq (\widehat{f}(0))^2.$$

First, using Proposition 3.2.12, we immediately obtain the real-valued version of Fourier series often used in physics and applied mathematics.

Theorem 3.2.14. *Let $f \in \mathcal{RI}(\mathbb{T})$ be real-valued. For each $n \in \mathbb{N} \cup \{0\}$, let*

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos(nx)\,dx \qquad \text{and} \qquad b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin(nx)\,dx.$$

Then the following hold:

a) $a_n, b_n \in \mathbb{R}$ for all $n \in \mathbb{N} \cup \{0\}$ and $b_0 = 0$, *b)* $P_n(f)(x) = \frac{1}{2}a_0 + \sum_{k=1}^{N} a_k \cos(kx) + b_k \sin(kx)$ *for all* $n \in \mathbb{N}$ *and* $x \in \mathbb{T}$,

c) $a_n = 2\mathrm{Re}(\widehat{f}(n))$ *for all* $n \in \mathbb{N} \cup \{0\}$, *and*

d) $b_n = -2\mathrm{Im}(\widehat{f}(n))$ *for all* $n \in \mathbb{N}$.

In particular,

$$\mathcal{F}(f) = \frac{1}{2}a_0 + \sum_{k=1}^{\infty} a_k \cos(kx) + b_k \sin(kx).$$

Proof. It is elementary to see that $a_k, b_k \in \mathbb{R}$ for all $k \in \mathbb{N} \cup \{0\}$ and that $b_k = 0$ (since $\sin(0x) = 0$ for all $x \in \mathbb{T}$. To see the desired formula for $P_n(f)$, note by part c) of Proposition 3.2.12 and the fact that f is real-valued (so $\overline{f} = f$) that

$$\widehat{f}(k) = \widehat{\overline{f}}(k) = \overline{\widehat{f}(-k)}$$

for all $k \in \mathbb{Z}$. Hence $\widehat{f}(0) = \overline{\widehat{f}(0)}$, so $\widehat{f}(0) \in \mathbb{R}$. Moreover, notice for all $n \in \mathbb{N}$ and $x \in \mathbb{T}$ that

$$\begin{aligned}
P_n(f)(x) &= \sum_{k=-n}^{n} \widehat{f}(k)e^{ikx} \\
&= \widehat{f}(0)e^0 + \sum_{k=1}^{n} \widehat{f}(k)e^{ikx} + \overline{\widehat{f}(k)}e^{-ikx} \\
&= \widehat{f}(0)e^0 + \sum_{k=1}^{n} \mathrm{Re}(\widehat{f}(k))\left(e^{ikx} + e^{-ikx}\right) + i\mathrm{Im}(\widehat{f}(k))\left(e^{ikx} - e^{-ikx}\right) \\
&= \mathrm{Re}\left(\widehat{f}(0)\right)1 + \sum_{k=1}^{n} 2\mathrm{Re}(\widehat{f}(k))\cos(kx) - 2\mathrm{Im}(\widehat{f}(k))\sin(kx).
\end{aligned}$$

However, since f is real-valued, we have for all $k \in \mathbb{N} \cup \{0\}$ that

$$\begin{aligned}
\mathrm{Re}(\widehat{f}(k)) &= \mathrm{Re}\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-ikx}\,dx\right) \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{Re}\left(f(x)e^{-ikx}\right)\,dx \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\mathrm{Re}\left(e^{-ikx}\right)\,dx \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\cos(kx)\,dx = \frac{a_k}{2}
\end{aligned}$$

and

$$\begin{aligned}\operatorname{Im}(\widehat{f}(k)) &= \operatorname{Im}\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-ikx}\,dx\right)\\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \operatorname{Im}\left(f(x)e^{-ikx}\right)\,dx\\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\operatorname{Im}\left(e^{-ikx}\right)\,dx\\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)(-\sin(kx))\,dx = -\frac{b_k}{2}.\end{aligned}$$

Hence

$$P_n(f)(x) = \frac{1}{2}a_0 + \sum_{k=1}^{N} a_k\cos(kx) + b_k\sin(kx)$$

for all $n \in \mathbb{N}$ and $x \in \mathbb{T}$ as desired. Moreover, the above computations show the formulae for a_n and b_n hold. ∎

Remark 3.2.15. Recall from Example 2.3.4 that if $a, b \in \mathbb{R}$ with a a positive integer, $0 < b < 1$, and $4ab > 1 + \frac{3}{2}\pi$, then the Weierstrass function

$$W(x) = \sum_{n=0}^{\infty} b^n \cos(\pi a^n x)$$

converged uniformly, was continuous, but was nowhere differentiable. As the series definition for W is exactly the Fourier series of W by Theorem 3.2.14 (well, with πx instead of x – a simple change of variables rectifies this), Fourier series can uniformly approximate even awful functions!

Using Proposition 3.2.12, it is about time we compute some Fourier series of some functions that are not trigonometric polynomials. We start with the simplest function.

Example 3.2.16. Define $f : \mathbb{T} \to \mathbb{C}$ by $f(x) = x$ for all $x \in (-\pi, \pi]$. Notice that

$$\widehat{f}(0) = \frac{1}{2\pi}\int_{-\pi}^{\pi} x\,dx = 0.$$

Moreover, if $n \in \mathbb{Z} \setminus \{0\}$, then since

$$\int_{-\pi}^{\pi} e^{-inx}\,dx = \langle 1, e_n\rangle = 0,$$

we obtain by using integration by parts that

$$\begin{aligned}
2\pi\widehat{f}(n) &= \int_{-\pi}^{\pi} xe^{-inx}\,dx \\
&= \left(-\frac{1}{in}xe^{-inx}\right)\Big|_{x=-\pi}^{\pi} - \int_{-\pi}^{\pi} -\frac{1}{in}e^{-inx}\,dx \\
&= \left(-\frac{1}{in}xe^{-inx}\right)\Big|_{x=-\pi}^{\pi} + 0 \\
&= -\frac{1}{in}\pi e^{-in\pi} + \frac{1}{in}(-\pi)e^{in\pi} \\
&= -\frac{\pi}{in}\left(e^{-in\pi} + e^{in\pi}\right) \\
&= \frac{\pi i}{n}(2\cos(n\pi)) \\
&= \frac{2\pi}{n}(-1)^n i
\end{aligned}$$

since $n \in \mathbb{Z}$. Hence $\widehat{f}(n) = \frac{1}{n}(-1)^n i$ for all $n \in \mathbb{Z} \setminus \{0\}$. Thus

$$\mathcal{F}(f)(x) = \sum_{\substack{n=-\infty \\ n\neq 0}}^{\infty} \frac{1}{n}(-1)^n i e^{inx}.$$

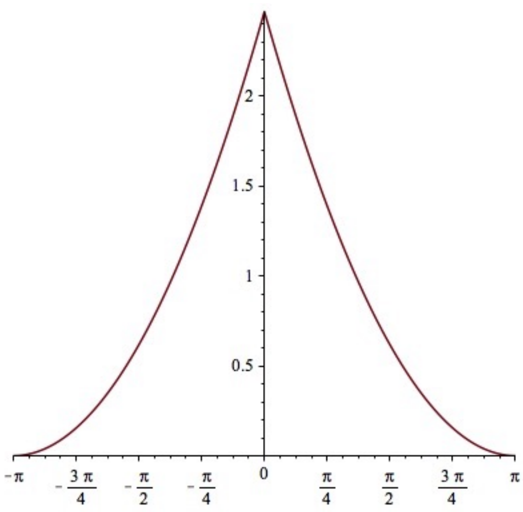
However, since f is real-valued, by using Theorem 3.2.14, we see if

$$a_n = 2\mathrm{Re}(\widehat{f}(n)) = 0 \quad \text{and} \quad b_n = -2\mathrm{Im}(\widehat{f}(n)) = \frac{2}{n}(-1)^{n+1}$$

for all $n \in \mathbb{Z}$, then

$$\mathcal{F}(f)(x) = \sum_{n=1}^{\infty} \frac{2}{n}(-1)^{n+1}\sin(nx).$$

Example 3.2.17. Define $f : \mathbb{T} \to \mathbb{C}$ by $f(x) = \frac{1}{4}(x-\pi)^2$ for all $x \in [0, 2\pi)$; that is, as a function on $[-\pi, \pi]$, f is the following:



Thus $f \in \mathcal{C}(\mathbb{T})$.

Notice that

$$\begin{aligned}\widehat{f}(0) &= \frac{1}{2\pi}\int_0^{2\pi} \frac{1}{4}(x-\pi)^2\,dx \\ &= \frac{1}{2\pi}\left(\frac{1}{12}(x-\pi)^3\right)\Big|_{x=0}^{2\pi} \\ &= \frac{1}{2\pi}\left(\frac{1}{12}\pi^3 - \frac{1}{12}(-\pi)^3\right) = \frac{1}{2\pi}\frac{\pi^3}{6} = \frac{\pi^2}{12}.\end{aligned}$$

Moreover, if $n \in \mathbb{Z}\setminus\{0\}$, then, by using integration by parts, we obtain that

$$\begin{aligned}2\pi\widehat{f}(n) &= \int_0^{2\pi} \frac{1}{4}(x-\pi)^2 e^{-inx}\,dx \\ &= \left(\frac{1}{4}(x-\pi)^2\frac{1}{-in}e^{-inx}\right)\Big|_{x=0}^{2\pi} - \int_0^{2\pi}\frac{1}{2}(x-\pi)\frac{1}{-in}e^{-inx}\,dx \\ &= \left(\frac{1}{4}\pi^2\frac{1}{-in}1 - \frac{1}{4}(-\pi)^2\frac{1}{-in}1\right) + \int_0^{2\pi}\frac{1}{2ni}(x-\pi)e^{-inx}\,dx \\ &= 0 + \left(\frac{1}{2ni}(x-\pi)\frac{1}{-in}e^{-inx}\right)\Big|_{x=0}^{2\pi} - \int_0^{2\pi}\frac{1}{2ni}\frac{1}{-in}e^{-inx}\,dx \\ &= \left(\frac{1}{2n^2}(x-\pi)e^{-inx}\right)\Big|_{x=0}^{2\pi} - \frac{1}{2n^2}\int_0^{2\pi}e^{-inx}\,dx \\ &= \left(\frac{1}{2n^2}\pi 1 - \frac{1}{2n^2}(-\pi)1\right) + 0 \\ &= \frac{\pi}{n^2}.\end{aligned}$$

Hence $\widehat{f}(n) = \frac{1}{2n^2}$ for all $n \in \mathbb{Z} \setminus \{0\}$. Thus

$$\mathcal{F}(f)(x) = \frac{\pi^2}{12} + \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \frac{1}{2n^2} e^{inx}.$$

However, since f is real-valued, by using Theorem 3.2.14, we see if

$$a_n = 2\mathrm{Re}(\widehat{f}(n)) = \frac{1}{n^2} \quad \text{and} \quad b_n = -2\mathrm{Im}(\widehat{f}(n)) = 0$$

for all $n \in \mathbb{Z}$, then

$$\mathcal{F}(f)(x) = \frac{\pi^2}{12} + \sum_{n=1}^{\infty} \frac{1}{n^2} \cos(nx).$$

Remark 3.2.18. For a moment, suppose if f was as in Example 3.2.17 then $f(0) = \mathcal{F}(f)(0)$. Thus we would have that

$$\frac{1}{4}(0 - \pi)^2 = \frac{\pi^2}{12} + \sum_{n=1}^{\infty} \frac{1}{n^2} \cos(n0).$$

So

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

Therefore, if the answer to Question 3.2.6 is yes, we have an answer to one case of Question 1.2.17.

To conclude this section, we note we can construct functions that have pre-described Fourier series provided the desired Fourier coefficients are 'nice' enough. To do this, we first need the following.

Remark 3.2.19. It is necessary for us to discuss infinite series summed over the integers. Note that this discussion was avoided in the above discussion of Fourier series since we were always taking the limits of the projections onto $\mathcal{T}_n(\mathbb{T})$ and thus pairing $-n$ with n for all $n \in \mathbb{N}$. In the case of absolute convergences, this causes absolutely no issues.

To begin this discussion, suppose $(a_n)_{n \in \mathbb{Z}}$ is a sequence of non-negative real numbers. To define

$$\sum_{n=-\infty}^{\infty} a_n,$$

consider the bijection $\sigma_0 : \mathbb{N} \to \mathbb{Z}$ defined by

$$\sigma_0(n) = \begin{cases} 0 & \text{if } n = 1 \\ -\frac{n}{2} & \text{if } n \text{ is even} \\ -\frac{n-1}{2} & \text{if } n \text{ is odd and } n \neq 1 \end{cases}.$$

We then define

$$L = \sum_{n=1}^\infty a_{\sigma_0(n)}$$

provided the series converges. However, since this series converges absolutely, Theorem 1.3.2 implies that we could replace σ_0 with any bijection $\sigma : \mathbb{N} \to \mathbb{Z}$ and the resulting series would still converge to L. Hence we have a well-defined notice of a convergent series of non-negative real numbers summed over the integers. Moreover, it is also useful to note that

$$L = \lim_{N\to\infty} \sum_{n=1}^{2N+1} a_{\sigma_0(n)} = \lim_{N\to\infty} \sum_{n=-N}^{N} a_n$$

and Theorem 1.2.9 implies that $\sum_{n=-\infty}^\infty a_n$ converges if and only if

$$\sup\left(\left\{ \sum_{n\in F} a_n \,\middle|\, F \subseteq \mathbb{Z} \text{ finite}\right\}\right) < \infty.$$

Therefore, if $\Omega \subseteq \mathbb{C}$ and $(f_n)_{n\in\mathbb{Z}}$ are continuous functions on Ω such that if

$$0 \leq M_n = \sup(\{|f_n(x)| \mid x \in \Omega\}) < \infty$$

for all $n \in \mathbb{Z}$ then $\sum_{n=-\infty}^\infty M_n < \infty$, then the Weierstrass M-Test (Theorem 2.2.15) together with (using the real and imaginary parts and) the above definitions imply that $f : \Omega \to \mathbb{C}$ defined by

$$f(x) = \lim_{N\to\infty} \sum_{n=1}^{2N+1} f_{\sigma_0(n)}(x)$$

for all $x \in \Omega$ is a well-defined continuous function on Ω. Moreover, the Weierstrass M-Test implies the series converges uniformly on Ω. Finally, the above argument also shows that we could replace σ_0 with any bijection $\sigma : \mathbb{N} \to \mathbb{Z}$ and that

$$f(x) = \lim_{N\to\infty} \sum_{n=1}^{2N+1} f_{\sigma_0(n)}(x) = \lim_{N\to\infty} \sum_{n=-N}^{N} f_n(x)$$

for all $x \in \Omega$.

Proposition 3.2.20. *Let $(z_n)_{n\in\mathbb{Z}}$ be complex numbers such that*

$$\sum_{n=-\infty}^\infty |z_n| < \infty.$$

Then the function $f : \mathbb{T} \to \mathbb{C}$ defined by

$$f(x) = \sum_{n=-\infty}^\infty z_n e^{inx}$$

for all $x \in \mathbb{T}$ is a well-defined element of $\mathcal{C}(\mathbb{T})$ such that

$$\widehat{f}(n) = z_n$$

for all $n \in \mathbb{Z}$. Moreover, the series description of f converges uniformly.

Proof. Since

$$\sup(\{|z_n e^{-inx}| \mid x \in \Omega\}) = |z_n|$$

for all $n \in \mathbb{Z}$ and since $\sum_{n=-\infty}^{\infty} |z_n| < \infty$, Remark 3.2.19 implies that if we define

$$f(x) = \sum_{n=-\infty}^{\infty} z_n e^{inx}$$

for all $x \in \mathbb{T}$, then f is a well-defined element of $\mathcal{C}(\mathbb{T})$. Moreover, since the above series converges uniformly by Remark 3.2.19, we obtain for all $k \in \mathbb{Z}$ that

$$\begin{aligned}
\widehat{f}(k) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-ikx}\,dx \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} \sum_{n=-\infty}^{\infty} z_n e^{inx}e^{-ikx}\,dx \\
&= \sum_{n=-\infty}^{\infty} z_n \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{inx}e^{-ikx}\,dx\right) && \text{by Corollary 2.4.5} \\
&= z_n && \text{by Theorem 3.1.17}
\end{aligned}$$

■

3.3 Convolutions

In the hopes of answering Questions 3.2.5 and 3.2.6, it is useful to revisit the proof of the Weierstrass Approximation Theorem (Theorem 2.8.6). There, a convolution (i.e. complicated multiplication) against certain fixed polynomials was introduced in order to approximate our given function. The goal of this section is to introduce the same idea in the context of Fourier series. In fact, we will see that this convolution is the correct multiplication to use so that the Fourier coefficients of a 'product' is the product of the Fourier coefficients.

The convolution we will use throughout the chapter is the following.

Definition 3.3.1. Given $f, g \in \mathcal{RI}(\mathbb{T})$, the *convolution of f and g* is the function $f * g : \mathbb{T} \to \mathbb{C}$ defined by

$$(f * g)(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(y)g(x-y)\,dy$$

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

It is important to note that $f * g$ is well-defined since the translation and product of elements of $\mathcal{RI}(\mathbb{T})$ are elements of $\mathcal{RI}(\mathbb{T})$. The most elementary and useful example to see that this is the convolution we should be considering is the following.

Example 3.3.2. Let $f \in \mathcal{RI}(\mathbb{T})$ and $n \in \mathbb{Z}$ be arbitrary. Then for all $x \in \mathbb{T}$ we see that

$$\begin{aligned}(f * e_n)(x) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(y)e^{in(x-y)}\,dy \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(y)e^{inx}e^{-iny}\,dy \\ &= e^{inx}\frac{1}{2\pi}\int_{-\pi}^{\pi} f(y)e^{-iny}\,dy \\ &= \widehat{f}(n)e^{inx}.\end{aligned}$$

Hence $f * e_n = \widehat{f}(n)e_n$.

Thus, in order to realize the partial Fourier series via convolutions, it remains only to show that convolutions are linear in the second entry. Consequently, let us examine the elementary properties of convolutions. In order to facilitate the proofs of e) and f) in the following, we require Fubini's Theorem (Theorem C.1.1). At this time we can only prove Fubini's Theorem for continuous functions. Those that take MATH 4012 can upgrade parts e) and f) later in their academic careers.

Proposition 3.3.3. *For all $f, g, h \in \mathcal{RI}(\mathbb{T})$, the following properties hold:*

a) $f * g = g * f$.

b) $f * (g + h) = (f * g) + (f * h)$.

c) $(g + h) * f = (g * f) + (h * f)$

d) $(cf) * g = c(f * g) = f * (cg)$ *for all* $c \in \mathbb{C}$.

e) $(f * g) * h = f * (g * h)$ *provided* $f, g, h \in \mathcal{C}(\mathbb{T})$.

f) $\widehat{(f * g)}(n) = \widehat{f}(n)\widehat{g}(n)$ *provided* $f, g \in \mathcal{C}(\mathbb{T})$.

Proof. To see that a) is true, notice for all $f, g \in \mathcal{RI}(\mathbb{T})$ and $x \in \mathbb{T}$ that

$$\begin{aligned}
(f * g)(x) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y)g(x-y)\,dy \\
&= \frac{1}{2\pi} \int_{x+\pi}^{x-\pi} f(x-t)g(t)(-1)\,dt && \text{substitute } t = x-y \\
&= \frac{1}{2\pi} \int_{x-\pi}^{x+\pi} f(x-t)g(t)\,dt \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x-t)g(t)\,dt && 2\pi\text{-perodicity} \\
&= (g * f)(x).
\end{aligned}$$

Therefore, since $x \in \mathbb{T}$ was arbitrary, $f * g = g * f$.

To see that b) and c) are true, notice that c) will follow from b) due to a). To see that b) is true, notice for all $f, g, h \in \mathcal{RI}(\mathbb{T})$ and $x \in \mathbb{T}$ that

$$\begin{aligned}
(f * (g+h))(x) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y)(g+h)(x-y)\,dy \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y)(g(x-y)+h(x-y))\,dy \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y)g(x-y)+f(y)h(x-y)\,dy \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y)g(x-y)\,dy + \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y)h(x-y)\,dy \\
&= (f * g)(x) + (f * h)(x).
\end{aligned}$$

Therefore, since $x \in \mathbb{T}$ was arbitrary, $f * (g+h) = (f * g) + (f * h)$.

To see that d) is true, notice by a) that it suffices to prove $(cf) * g = c(f * g)$. To see this, notice for all $f, g \in \mathcal{RI}(\mathbb{T})$, $c \in \mathbb{C}$, and $x \in \mathbb{T}$ that

$$\begin{aligned}
((cf) * g)(x) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} (cf)(y)g(x-y)\,dy \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} cf(y)g(x-y)\,dy \\
&= c\frac{1}{2\pi} \int_{-\pi}^{\pi} f(y)g(x-y)\,dy \\
&= c((f * g)(x)).
\end{aligned}$$

Therefore, since $x \in \mathbb{T}$ was arbitrary, $(cf) * g = c(f * g)$.

To see that e) is true, notice for all $f, g, h \in \mathcal{C}(\mathbb{T})$ and $x \in \mathbb{T}$ that

$$
\begin{aligned}
&((f * g) * h)(x) \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} (f * g)(y) h(x - y)\, dy \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} f(q) g(y - q)\, dq \right) h(x - y)\, dy \\
&= \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} f(q) g(y - q) h(x - y)\, dq\, dy \\
&= \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} f(q) g(y - q) h(x - y)\, dy\, dq && \text{by Fubini's Theorem} \\
&= \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi-q}^{\pi-q} f(q) g(t) h(x - (t + q))\, dt\, dq && \text{substitute } t = y - q \\
&= \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} f(q) g(t) h((x - q) - t))\, dt\, dq && 2\pi\text{-perodicity} \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(q) \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} g(t) h((x - q) - t))\, dt \right) dq \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(q) (g * h)(x - q)\, dq \\
&= (f * (g * h))(x).
\end{aligned}
$$

Therefore, since $x \in \mathbb{T}$ was arbitrary, $(f * g) * h = f * (g * h)$.

Finally, to see that f) is true, notice for all $f, g \in \mathcal{C}(\mathbb{T})$ and $n \in \mathbb{Z}$ that

$$
\begin{aligned}
\widehat{(f * g)}(n) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} (f * g)(x) e^{-inx}\, dx \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} f(y) g(x - y)\, dy \right) e^{-inx}\, dx \\
&= \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} f(y) g(x - y) e^{-inx}\, dy\, dx \\
&= \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} f(y) g(x - y) e^{-inx}\, dx\, dy && \text{by Fubini's Theorem} \\
&= \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi-y}^{\pi-y} f(y) g(t) e^{-in(t+y)}\, dt\, dy && \text{substitute } t = x - y \\
&= \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} f(y) e^{-iny} g(t) e^{-int}\, dt\, dy && 2\pi\text{-perodicity} \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y) e^{-iny} \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} g(t) e^{-int}\, dt \right) dy \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y) e^{-iny} \widehat{g}(n)\, dy \\
&= \widehat{f}(n) \widehat{g}(n)
\end{aligned}
$$

as desired. ∎

Proposition 3.3.3 easily enables us to describe the convolution of any element of $\mathcal{RI}(\mathbb{T})$ against a trigonometric polynomial.

Example 3.3.4. Let $f \in \mathcal{RI}(\mathbb{T})$ and let $p \in Trig(\mathbb{T})$. Thus there exists an $n \in \mathbb{N}$ and $(c_k)_{k=-n}^{n} \subseteq \mathbb{C}$ such that

$$p(x) = \sum_{k=-n}^{n} c_k e^{ikx}$$

for all $x \in \mathbb{T}$. Therefore, by Example 3.3.2 and Proposition 3.3.3, we have that

$$(f * p)(x) = \sum_{k=-n}^{n} c_k \widehat{f}(k) e^{ikx}$$

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

Clearly in order to get the n^{th} partial Fourier series, we need to consider the following trigonometric polynomials.

Definition 3.3.5. For $n \in \mathbb{N}$, the n^{th} *Dirichlet kernel* (pronounced duh-ri-klet), denote D_n, is the trigonometric polynomial defined by

$$D_n(x) = \sum_{k=-n}^{n} e^{ikx}$$

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

Corollary 3.3.6. *For all $f \in \mathcal{RI}(\mathbb{T})$ and $n \in \mathbb{N}$, $f * D_n = P_n(f)$.*

Proof. This immediately follows from Example 3.3.4 along with the definitions of D_n and $P_n(f)$. ∎

Now that we have the Dirichlet kernels, it would be useful to better understand these trigonometric polynomials. In particular, the following lemma easily enables us to graph these functions.

Lemma 3.3.7. *For all $n \in \mathbb{N}$*

$$D_n(x) = \frac{\sin\left(\left(n + \frac{1}{2}\right)x\right)}{\sin\left(\frac{1}{2}x\right)}$$

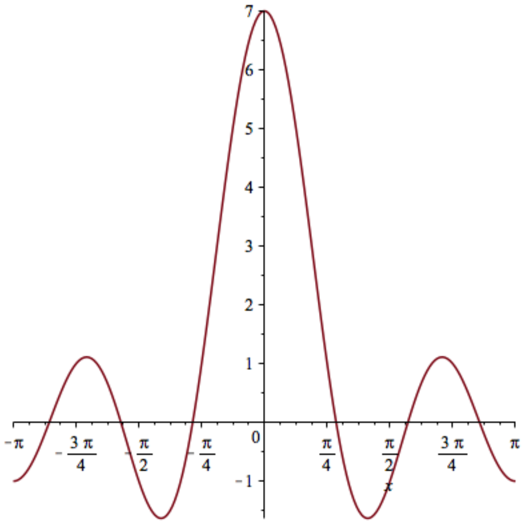
as a continuous function on $\mathbb{T}$ (i.e. the formula for $D_n(0)$ should be interpreted as the limit as x tends to 0).

Proof. Notice for all $x \in \mathbb{T}$ that

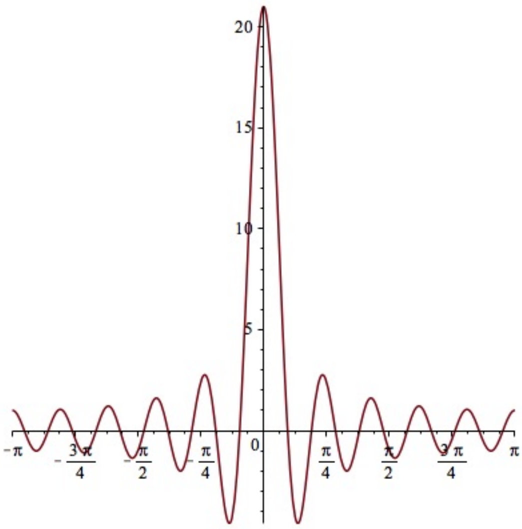
$$\begin{aligned} D_n(x) &= e^{-inx} \sum_{k=0}^{2n} e^{ikx} \\ &= e^{-inx} \sum_{k=0}^{2n} (e^{ix})^k \\ &= e^{-inx} \frac{e^{i(2n+1)x} - 1}{e^{ix} - 1} \\ &= e^{-inx} \frac{e^{i(2n+1)x} - 1}{e^{i\frac{1}{2}x}(e^{i\frac{1}{2}x} - e^{-i\frac{1}{2}x})} \\ &= e^{-i\left(n+\frac{1}{2}\right)x} \frac{e^{i(2n+1)x} - 1}{e^{i\frac{1}{2}x} - e^{-i\frac{1}{2}x}} \\ &= \frac{e^{i\left(n+\frac{1}{2}\right)x} - e^{-i\left(n+\frac{1}{2}\right)x}}{e^{i\frac{1}{2}x} - e^{-i\frac{1}{2}x}} \\ &= \frac{2i \sin\left(\left(n+\frac{1}{2}\right)x\right)}{2i \sin\left(\frac{1}{2}x\right)} \\ &= \frac{\sin\left(\left(n+\frac{1}{2}\right)x\right)}{\sin\left(\frac{1}{2}x\right)} \end{aligned}$$

as desired. ∎

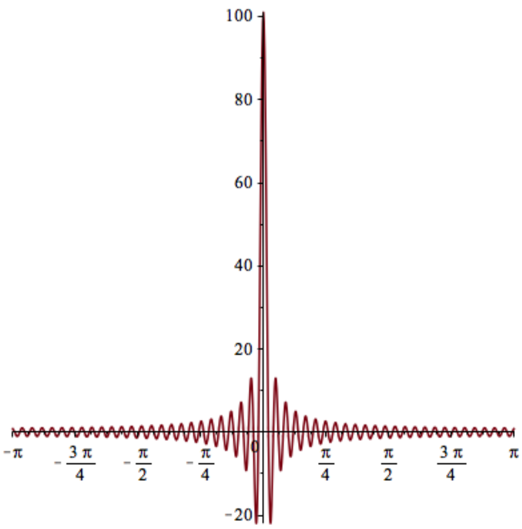
Using Lemma 3.3.7, it is easy to graph D_n for all $n \in \mathbb{N}$. In particular, the graph of D_3 is as follows:



The graph of D_{10} is as follows:



Finally the graph of D_{50} is as follows:



As these graphs do not look too different than those of the functions used in the proof of the Weierstrass Approximation Theorem (Theorem 2.8.6), perhaps there is hope in solving Questions 3.2.5 and 3.2.6 in the same way!

Before we head in that direction, it is useful to obtain one more piece of information about convolutions. In particular, we desire to prove that $f * g$ is actually continuous for all $f, g \in \mathcal{RI}(\mathbb{T})$. This can be useful for us when we want to consider the convolution of functions that do not include the Dirichlet kernel. In order to obtain this fact, we need to be able to 'approximate' elements of $\mathcal{RI}(\mathbb{T})$ with elements from $\mathcal{C}(\mathbb{T})$ via the following.

Lemma 3.3.8. *Let $f \in \mathcal{RI}(\mathbb{T})$ and let $M > 0$ be such that $|f(x)| \leq M$ for all $x \in \mathbb{T}$. There exists a sequence $(g_n)_{n\geq 1}$ of elements of $\mathcal{C}(\mathbb{T})$ such that*

$$\sup(\{|g_n(x)| \mid x \in \mathbb{T}\}) \leq 2M$$

for all $n \in \mathbb{N}$ and

$$\lim_{n\to\infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x) - g_n(x)| \, dx = 0.$$

Proof. To begin, we first claim that if $f \in \mathcal{RI}(\mathbb{T})$ is real-valued, then for any $\epsilon > 0$ there exists a $g \in \mathcal{C}(\mathbb{T})$ such that

$$\sup(\{|g(x)| \mid x \in \mathbb{T}\}) \leq M$$

and

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x) - g(x)| \, dx < \epsilon.$$

This will be done by first approximating f with a step function, and then correcting the step function to a continuous function by using very step line segments.

To see the above claim is true, let $\epsilon > 0$ be arbitrary. Since $f \in \mathcal{RI}(\mathbb{T})$ is real-valued, there exists a partition $\mathcal{P}$ of $\mathbb{T}$ such that

$$U(f, \mathcal{P}) - L(f, \mathcal{P}) < \epsilon\pi.$$

Since $\mathcal{P}$ is a partition of $\mathbb{T}$, we can write $\mathcal{P} = \{t_k\}_{k=0}^{n}$ where

$$-\pi = t_0 < t_1 < \cdots < t_{n-1} < t_n = \pi.$$

Therefore, if

$$M_k = \sup(\{f(x) \mid x \in [t_{k-1}t_k]\}) \quad \text{and} \quad m_k = \inf(\{f(x) \mid x \in [t_{k-1}t_k]\}),$$

then

$$U(f, \mathcal{P}) - L(f, \mathcal{P}) = \sum_{k=1}^{n} (M_k - m_k)(t_k - t_{k-1}).$$

Define the step function $h : \mathbb{T} \to \mathbb{R}$ by

$$h(x) = \begin{cases} M_k & \text{if } x \in [t_{k-1}, t_k) \text{ and } k \neq n \\ M_n & \text{if } x \in [t_{n-1}t_n] \end{cases}$$

for all $x \in \mathbb{T}$. Then

$$\sup(\{|h(x)| \mid x \in \mathbb{T}\}) = \sup(\{M_k \mid k \in \{1, \ldots, n\}\}) \leq M$$

by construction. Moreover

$$\begin{aligned}
\frac{1}{2\pi}\int_{-\pi}^{\pi}|f(x)-h(x)|\,dx &= \frac{1}{2\pi}\sum_{k=1}^{n}\int_{t_{k-1}}^{t_k}|f(x)-h(x)|\,dx \\
&= \frac{1}{2\pi}\sum_{k=1}^{n}\int_{t_{k-1}}^{t_k} M_k - f(x)\,dx \\
&\leq \frac{1}{2\pi}\sum_{k=1}^{n}\int_{t_{k-1}}^{t_k} M_k - m_k\,dx \\
&= \frac{1}{2\pi}\sum_{k=1}^{n}(M_k-m_k)(t_k-t_{k-1}) \\
&= \frac{1}{2\pi}(U(f,\mathcal{P})-L(f,\mathcal{P})) \\
&< \frac{1}{2\pi}(\epsilon\pi) = \frac{\epsilon}{2}.
\end{aligned}$$

It remains only to correct h using steep line segments to obtain a continuous function.

Let

$$\delta = \min\left(\left\{\frac{\pi\epsilon}{4(M+1)}\right\}\bigcup\left\{\frac{t_k-t_{k-1}}{2}\;\middle|\; k\in\{1,\ldots,n\}\right\}\right) > 0.$$

Notice by our choice of δ that no two intervals of the form $(t_k-\delta, t_k+\delta)$ overlap. Equating $t_n=\pi$ with $t_0=-\pi$ modulo 2π, define $g:\mathbb{T}\to\mathbb{R}$ by

$$g(x) = \begin{cases} h(x) & \text{if } x\in[t_{k-1}+\delta, t_k-\delta] \\ \frac{(h(t_k+\delta)-h(t_k-\delta))(x-(t_k-\delta))}{(t_k+\delta)-(t_k-\delta)} + h(t_k-\delta) & \text{if } x\in(t_k-\delta, t_k+\delta) \end{cases}$$

for all $x\in\mathbb{T}$. By construction $g\in\mathcal{C}(\mathbb{T})$ and

$$\sup(\{|g(x)| \mid x\in\mathbb{T}\}) \leq \sup(\{|h(x)| \mid x\in\mathbb{T}\}) \leq M.$$

Moreover

$$\begin{aligned}
\frac{1}{2\pi}\int_{-\pi}^{\pi}|h(x)-g(x)|\,dx &= \frac{1}{2\pi}\sum_{k=1}^{n}\int_{t_k-\delta}^{t_k+\delta}|h(x)-g(x)|\,dx \\
&\leq \frac{1}{2\pi}\sum_{k=1}^{n}\int_{t_k-\delta}^{t_k+\delta}2M\,dx \\
&= \frac{1}{2\pi}4\delta M \\
&< \frac{1}{2\pi}4\left(\frac{\pi\epsilon}{4(M+1)}\right)M < \frac{\epsilon}{2}.
\end{aligned}$$

Therefore

$$\begin{aligned}\frac{1}{2\pi}\int_{-\pi}^{\pi}|f(x)-g(x)|\,dx &\leq \frac{1}{2\pi}\int_{-\pi}^{\pi}|f(x)-h(x)|+|h(x)-g(x)|\,dx\\ &< \frac{\epsilon}{2}+\frac{\epsilon}{2}=\epsilon\end{aligned}$$

thereby completing the proof of the claim.

To prove the statement in the lemma, let $f \in \mathcal{RI}(\mathbb{T})$ be arbitrary. By letting $f_1 = \mathrm{Re}(f) \in \mathcal{RI}(\mathbb{T})$ and $f_2 = \mathrm{Im}(f) \in \mathcal{RI}(\mathbb{T})$, the above claim implies that for every $n \in \mathbb{N}$ there exists $g_{n,1}, g_{n,2} \in \mathcal{C}(\mathbb{T})$ such that

$$\sup(\{|g_{n,j}(x)| \mid x \in \mathbb{T}\}) \leq \sup(\{|f_j(x)| \mid x \in \mathbb{T}\}) \leq M$$

for all $n \in \mathbb{N}$ and $j \in \{1, 2\}$, and

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}|f(x)-g_{n,j}(x)|\,dx < \frac{1}{2n}$$

for $n \in \mathbb{N}$ and $j \in \{1, 2\}$. Therefore, if for each $n \in \mathbb{N}$ we define

$$g_n = g_{n,1} + ig_{n,2},$$

then $g_n \in \mathcal{C}(\mathbb{T})$ for all $n \in \mathbb{N}$,

$$\sup(\{|g_n(x)| \mid x \in \mathbb{T}\}) \leq \sup(\{|g_{n,1}(x)| + |g_{n,2}(x)| \mid x \in \mathbb{T}\}) \leq 2M$$

for all $n \in \mathbb{N}$, and

$$\begin{aligned}\frac{1}{2\pi}\int_{-\pi}^{\pi}|f(x)-g_n(x)|\,dx &\leq \frac{1}{2\pi}\int_{-\pi}^{\pi}|f_1(x)-g_{n,1}(x)|+|f_2(x)-g_{n,2}(x)|\,dx\\ &< \frac{1}{2n}+\frac{1}{2n}=\frac{1}{n}\end{aligned}$$

so

$$\lim_{n\to\infty}\frac{1}{2\pi}\int_{-\pi}^{\pi}|f(x)-g_n(x)|\,dx = 0$$

as desired. ∎

Theorem 3.3.9. *For all $f, g \in \mathcal{RI}(\mathbb{T})$, $f * g \in \mathcal{C}(\mathbb{T})$.*

Proof. To prove this result, we will first deal with the case that $g \in \mathcal{C}(\mathbb{T})$ and then use Lemma 3.3.8 to improve the result to an arbitrary $g \in \mathcal{RI}(\mathbb{T})$.

To begin, suppose $f \in \mathcal{RI}(\mathbb{T})$ and $g \in \mathcal{C}(\mathbb{T})$. To see that $f * g$ is continuous, we will show that $f * g$ is uniformly continuous on $\mathbb{T}$. Thus let $\epsilon > 0$ be arbitrary. Since $f \in \mathcal{RI}(\mathbb{T})$,

$$M = \frac{1}{2\pi}\int_{-\pi}^{\pi}|f(y)|\,dy < \infty.$$

Since $g \in \mathcal{C}(\mathbb{T})$, g is uniformly continuous on $\mathbb{T}$ by Theorem 2.8.4. Hence there exists a $\delta > 0$ such that if $t_1, t_2 \in \mathbb{T}$ and $|t_1 - t_2| < \delta$ (view this distance as modulo 2π), then $|g(t_1) - g(t_1)| < \frac{\epsilon}{M+1}$. Therefore, if $x_1, x_2 \in \mathbb{T}$ and $|x_1 - x_2| < \delta$ (modulo 2π), then for all $y \in \mathbb{T}$ we have $|(x_1 - y) - (x_2 - y)| < \delta$ so

$$\begin{aligned}
&|(f * g)(x_1) - (f * g)(x_2)| \\
&= \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y) g(x_1 - y)\, dy - \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y) g(x_2 - y)\, dy \right| \\
&= \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y)(g(x_1 - y) - g(x_2 - y))\, dy \right| \\
&\leq \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(y)||g(x_1 - y) - g(x_2 - y)|\, dy \\
&\leq \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(y)| \frac{\epsilon}{M+1}\, dy \\
&= \frac{\epsilon}{M+1} \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(y)|\, dy \\
&= \frac{\epsilon}{M+1} M < \epsilon.
\end{aligned}$$

Therefore, as $\epsilon > 0$ was arbitrary, $f * g$ is uniformly continuous on $\mathbb{T}$ in the case that $g \in \mathcal{C}(\mathbb{T})$.

To prove the arbitrary case, let $f, g \in \mathcal{RI}(\mathbb{T})$ be arbitrary. To see that $f * g$ is continuous on $\mathbb{T}$, we will show that $f * g$ is uniformly a uniformly limit of continuous functions on $\mathbb{T}$. To construct this sequence, note by Lemma 3.3.8 that exists a sequence $(g_n)_{n \geq 1}$ of elements of $\mathcal{C}(\mathbb{T})$ such that

$$\lim_{n \to \infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} |g(x) - g_n(x)|\, dx = 0.$$

Since $g_n \in \mathcal{C}(\mathbb{T})$, by the first case of this proof we know that $(f * g_n)_{n \geq 1}$ is a sequence of continuous functions on $\mathbb{T}$. Hence it suffices by Theorem 2.2.9 to prove that $(f * g_n)_{n \geq 1}$ converges uniformly to $f * g$ on $\mathbb{T}$.

To see that $(f * g_n)_{n \geq 1}$ converges uniformly to $f * g$ on $\mathbb{T}$, let $\epsilon > 0$ be arbitrary. Since $f \in \mathcal{RI}(\mathbb{T})$, there exists an $M \in \mathbb{R}$ such that $|f(x)| \leq M$ for all $x \in \mathbb{T}$. Moreover, since

$$\lim_{n \to \infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} |g(x) - g_n(x)|\, dx = 0$$

there exists an $N \in \mathbb{N}$ such that

$$\lim_{n \to \infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} |g(x) - g_n(x)|\, dx < \frac{\epsilon}{M+1}$$

for all $n \geq N$. Hence for all $n \geq N$ and $x \in \mathbb{T}$, we have that

$$\begin{aligned}
&|(f * g)(x) - (f * g_n)(x)| \\
&= \left|\frac{1}{2\pi}\int_{-\pi}^{\pi} f(y)g(x-y)\,dy - \frac{1}{2\pi}\int_{-\pi}^{\pi} f(y)g_n(x-y)\,dy\right| \\
&= \left|\frac{1}{2\pi}\int_{-\pi}^{\pi} f(y)(g(x-y) - g_n(x-y))\,dy\right| \\
&\leq \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(y)||g(x-y) - g_n(x-y|)\,dy \\
&\leq \frac{1}{2\pi}\int_{-\pi}^{\pi} M|g(x-y) - g_n(x-y)|\,dy && |f(y)| \leq M \\
&= \frac{M}{2\pi}\int_{-\pi}^{\pi} |g(x-y) - g_n(x-y)|\,dy \\
&= \frac{M}{2\pi}\int_{x+\pi}^{x-\pi} |g(t) - g_n(t)|(-1)\,dt && \text{substitute } t = x-y \\
&= \frac{M}{2\pi}\int_{x-\pi}^{x+\pi} |g(t) - g_n(t)|\,dt \\
&= \frac{M}{2\pi}\int_{-\pi}^{\pi} |g(t) - g_n(t)|\,dt && 2\pi\text{-periodic} \\
&= M\frac{\epsilon}{M+1} < \epsilon.
\end{aligned}$$

Therefore, as $\epsilon > 0$ was arbitrary, $(f * g_n)_{n\geq 1}$ converges uniformly to $f * g$ on $\mathbb{T}$ thereby completing the proof. ∎

Remark 3.3.10. Using the same ideas as in Theorem 3.3.9, it is possible to extend parts e) and f) of Proposition 3.3.3. Indeed, the proof of Theorem 3.3.9 shows that we can replace any element of $\mathcal{RI}(\mathbb{T})$ in a convolution with an element of $\mathcal{C}(\mathbb{T})$ from Lemma 3.3.8 and be uniformly close to the original convolution based on a bound for the other function (which always exist since we are dealing with $\mathcal{RI}(\mathbb{T})$). Thus part e) of Proposition 3.3.3 can be shown (with some technical details) to hold when $\mathcal{RI}(\mathbb{T})$ is replaced with $\mathcal{C}(\mathbb{T})$ and part f) of Proposition 3.3.3 will also extend once we prove the following. We will not provide a proof of these extensions of parts e) and f) of Proposition 3.3.3 as we are focusing on $\mathcal{C}(\mathbb{T})$ and as we will not be using them later in the course.

Lemma 3.3.11. *Let $f \in \mathcal{RI}(\mathbb{T})$ and let $(f_n)_{n\geq 1}$ be a sequence of elements of $\mathcal{RI}(\mathbb{T})$ such that*

$$\lim_{n\to\infty} \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x) - f_n(x)|\,dx = 0.$$

Then

$$\lim_{n\to\infty} \sup_{k\in\mathbb{Z}} \left|\widehat{f}(k) - \widehat{f_n}(k)\right| = 0.$$

Proof. Notice for all $n, k \in \mathbb{N}$ that

$$\begin{aligned}\left|\widehat{f}(k) - \widehat{f_n}(k)\right| &= \left|\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-ikx}\,dx - \frac{1}{2\pi}\int_{-\pi}^{\pi} f_n(x)e^{-ikx}\,dx\right| \\ &= \left|\frac{1}{2\pi}\int_{-\pi}^{\pi} (f(x) - f_n(x))e^{-ikx}\,dx\right| \\ &\leq \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x) - f_n(x)||e^{-ikx}|\,dx \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x) - f_n(x)|\,dx.\end{aligned}$$

Hence for all $n \in \mathbb{N}$,

$$0 \leq \sup_{k\in\mathbb{Z}} \left|\widehat{f}(k) - \widehat{f_n}(k)\right| \leq \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x) - f_n(x)|\,dx.$$

Therefore, since

$$\lim_{n\to\infty} \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x) - f_n(x)|\,dx = 0,$$

the result follows. ∎

3.4 Summability Kernels

Now that we have a convolution description of the partial Fourier series using the Dirichlet kernels, we turn our attention back to the Weierstrass Approximation Theorem (Theorem 2.8.6). The proof of the Weierstrass Approximation Theorem showed that the convolutions of a continuous function f against specific polynomials converged uniformly to f. Thus, if we desire the same thing in the context of Fourier series, it should be enlightening to see what properties these polynomials had.

Example 3.4.1. Recall the construction used in the Weierstrass Approximation Theorem (Theorem 2.8.6) where for each $n \in \mathbb{N}$ we defined $q_n : [-1, 1] \to \mathbb{R}$ by

$$q_n(x) = c_n(1 - x^2)^n$$

for all $x \in [-1, 1]$ where

$$\frac{1}{c_n} = \int_{-1}^{1} (1 - x^2)^n\,dx.$$

By looking at the final computation in the proof of the Weierstrass Approximation Theorem, the key properties of q_n needed to show that $q_n * f$ converged uniformly to f were

(I) $\int_{-1}^{1} q_n(t)\,dt = 1$,

(II) $\int_{-1}^{1} |q_n(t)|\, dt$ was bounded (in this case by 1 as $q_n(t) \geq 0$), and

(III) $\lim_{n\to\infty} \int_{[-1,-\delta]\cup[\delta,1]} |q_n(t)|\, dt = 0$ for all $\delta > 0$.

Note that (III) was accomplished since for any δ we could make $|q_n(t)|$ arbitrary small on $[-1, -\delta] \cup [\delta, 1]$ by letting n be sufficiently large.

Thus, if we are to repeat the proof of the Weierstrass Approximation Theorem (Theorem 2.8.6) in this context, we desire similar properties for functions on $\mathbb{T}$. To describe such collections of functions, we define the following.

Definition 3.4.2. A *summability kernel* is a sequence $(k_n)_{n\geq 1}$ of elements of $\mathcal{C}(\mathbb{T})$ such that

(I) $\frac{1}{2\pi}\int_{-\pi}^{\pi} k_n(y)\, dy = 1$ for all $n \in \mathbb{N}$,

(II) there exists an $M > 0$ such that $\frac{1}{2\pi}\int_{-\pi}^{\pi} |k_n(y)|\, dy \leq M$ for all $n \in \mathbb{N}$, and

(III) $\lim_{n\to\infty} \frac{1}{2\pi}\int_{\delta\leq|y|\leq\pi} |k_n(y)|\, dy = 0$ for all $\delta \in (0, \pi)$.

Repeating part of the proof of Weierstrass Approximation Theorem (Theorem 2.8.6), we have the following.

Theorem 3.4.3. *Let $f \in \mathcal{RI}(\mathbb{T})$ and let $(k_n)_{n\geq 1}$ be a summability kernel. If $x \in \mathbb{T}$ is a point of continuity of f, then*

$$\lim_{n\to\infty} (f * k_n)(x) = f(x).$$

Moreover, if I is a closed interval in $\mathbb{T}$ (e.g. $I = \mathbb{T}$) and f is continuous at each point in I, then $(f * k_n)_{n\geq 1}$ converges uniformly to f on I.

Proof. To begin, fix $f \in \mathcal{RI}(\mathbb{T})$. Since $f \in \mathcal{RI}(\mathbb{T})$, there exists a $K > 0$ such that

$$|f(x)| \leq K$$

for all $x \in \mathbb{T}$. Moreover, since $(k_n)_{n\geq 1}$ is a summability kernel, property (II) of a summability kernel implies there exists an $M > 0$ such that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |k_n(x)|\, dx \leq M$$

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

To prove the first part of the theorem, suppose $x \in \mathbb{T}$ is a point of continuity of f. To see that

$$\lim_{n\to\infty} (f * k_n)(x) = f(x),$$

let $\epsilon > 0$ be arbitrary. Since f is continuous at x, there exists a $\delta > 0$ such that if $y \in (-\delta, +\delta)$, then

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

Moreover, since $(k_n)_{n\geq 1}$ is a summability kernel, property (III) of a summability kernel implies there exists an $N \in \mathbb{N}$ such that if $n \geq N$ then

$$\frac{1}{2\pi}\int_{\delta\leq|y|\leq\pi} |k_n(y)|\, dy < \frac{\epsilon}{4K}.$$

Hence, for all $n \geq N$, we have that

$$\begin{aligned}
&|(f * k_n)(x) - f(x)| \\
&= |(k_n * f)(x) - f(x)| && \text{by Proposition 3.3.3} \\
&= \left|\frac{1}{2\pi}\int_{-\pi}^{\pi} k_n(y) f(x-y)\, dy - f(x)\right| \\
&= \left|\frac{1}{2\pi}\int_{-\pi}^{\pi} k_n(y) f(x-y)\, dy - f(x)\frac{1}{2\pi}\int_{-\pi}^{\pi} k_n(y)\, dy\right| && \text{by (I)} \\
&= \left|\frac{1}{2\pi}\int_{-\pi}^{\pi} k_n(y)(f(x-y) - f(x))\, dy\right| \\
&\leq \frac{1}{2\pi}\int_{-\pi}^{\pi} |k_n(y)||f(x-y) - f(x)|\, dy \\
&= \frac{1}{2\pi}\int_{-\delta}^{\delta} |k_n(y)||f(x-y) - f(x)|\, dy \\
&\quad + \frac{1}{2\pi}\int_{\delta\leq|y|\leq\pi} |k_n(y)||f(x-y) - f(x)|\, dy \\
&\leq \frac{1}{2\pi}\int_{-\delta}^{\delta} |k_n(y)|\frac{\epsilon}{2M}\, dy + \frac{1}{2\pi}\int_{\delta\leq|y|\leq\pi} |k_n(y)|(2K)\, dy \\
&\leq \frac{\epsilon}{2M}\frac{1}{2\pi}\int_{-\pi}^{\pi} |k_n(y)|\, dy + (2K)\frac{1}{2\pi}\int_{\delta\leq|y|\leq\pi} |k_n(y)|\, dy \\
&\leq \frac{\epsilon}{2M}M + (2K)\frac{\epsilon}{4K} = \epsilon.
\end{aligned}$$

Therefore, as $\epsilon > 0$ was arbitrary, the proof of the first part of the theorem is complete.

To see the second part of the theorem is true, we simply need to adapt the above proof to show that given an $\epsilon > 0$, the same N works for all $x \in I$. Thus suppose f is continuous at each point of a closed interval I and fix an $\epsilon > 0$. Note Theorem 2.8.4 implies that f is uniformly continuous on I. Therefore, if $I = \mathbb{T}$ then the δ chosen above for a single x can be chosen to work for all $x \in \mathbb{T}$ and the same computation show that $|(f*k_n)(x) - f(x)| < \epsilon$ for all $n \geq N$ and $x \in \mathbb{T}$ thereby completing the proof. If $I \neq \mathbb{T}$, there is a slight technicality we need to deal with.

Since f is uniformly continuous on I, there a $\delta_0 > 0$ such that if $y \in (-\delta_0, \delta_0)$ AND $x - y \in I$, then

$$|f(x) - f(x-y)| < \frac{\epsilon}{4M}.$$

Let $x_1, x_2 \in I$ be two endpoints of I. Since f is continuous at x_1 and x_2, there exists $\delta_1, \delta_2 > 0$ such that for $j \in \{1, 2\}$ if $y \in (-\delta_j, \delta_j)$, then

$$|f(x_j) - f(x_j - y)| < \frac{\epsilon}{4M}.$$

Let $\delta = \min(\{\delta_0, \delta_1, \delta_2\}) > 0$. Therefore, if $x \in I$ and $y \in (-\delta, \delta)$, then either $x - y \in I$ so

$$|f(x) - f(x-y)| < \frac{\epsilon}{4M}$$

or $x - y \notin I$ so $x - y$ must be with δ of x_j for some $j \in \{1, 2\}$ so x is within δ of the same x_j so

$$|f(x) - f(x-y)| \leq |f(x) - f(x_j)| + |f(x_j) - f(x-y)| < \frac{\epsilon}{4M} + \frac{\epsilon}{4M} = \frac{\epsilon}{2M}.$$

Hence the same computation used above shows that $|(f * k_n)(x) - f(x)| < \epsilon$ for all $n \geq N$ and $x \in \mathbb{T}$ thereby completing the proof. ∎

With Theorem 3.4.3 complete, the only thing that remains in order to confirm a positive answer to Question 3.2.5 is to show that the Dirichlet kernel is a summability kernel. Right?

Proposition 3.4.4. *For all $n \in \mathbb{N}$,*

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} D_n(x)\, dx = 1.$$

However

$$\lim_{n\to\infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} |D_n(x)|\, dx = \infty.$$

Proof. To see the first claim, notice by Theorem 3.1.17 that

$$\begin{aligned} \frac{1}{2\pi} \int_{-\pi}^{\pi} D_n(x)\, dx &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{k=-n}^{n} e^{ikx}\, dx \\ &= \frac{1}{2\pi} \sum_{k=-n}^{n} \int_{-\pi}^{\pi} e^{ikx}\, dx = 1. \end{aligned}$$

To see the second claim, notice for all $n \in \mathbb{N}$ that

$$
\begin{aligned}
&\frac{1}{2\pi}\int_{-\pi}^{\pi} |D_n(x)|\,dx \\
&= \frac{1}{2\pi}\int_{0}^{2\pi} |D_n(x)|\,dx && 2\pi\text{-periodicity} \\
&= \frac{1}{2\pi}\int_{0}^{2\pi} \left|\frac{\sin\left(\left(n+\frac{1}{2}\right)x\right)}{\sin\left(\frac{1}{2}x\right)}\right|\,dx && \text{by Lemma 3.3.7} \\
&\geq \frac{1}{2\pi}\int_{0}^{2\pi} \left|\frac{\sin\left(\left(n+\frac{1}{2}\right)x\right)}{\frac{1}{2}x}\right|\,dx && \text{as } |\sin(x)| \leq |x| \\
&= \frac{1}{2\pi}\int_{0}^{(2n+1)\pi} \left|\frac{\sin(y)}{\frac{1}{2}\frac{1}{n+\frac{1}{2}}y}\right|\left(\frac{1}{n+\frac{1}{2}}\right)dy && \text{let } y = \left(n+\frac{1}{2}\right)x \\
&= \frac{1}{\pi}\int_{0}^{(2n+1)\pi} \left|\frac{\sin(y)}{y}\right|\,dy \\
&= \frac{1}{\pi}\sum_{k=0}^{2n}\int_{\pi k}^{\pi(k+1)} \frac{|\sin(y)|}{y}\,dy \\
&\geq \frac{1}{\pi}\sum_{k=0}^{2n}\int_{\pi k}^{\pi(k+1)} \frac{|\sin(y)|}{\pi(k+1)}\,dy \\
&= \frac{1}{\pi}\sum_{k=0}^{2n}\frac{2}{\pi(k+1)} = \frac{2}{\pi^2}\sum_{k=0}^{2n}\frac{1}{k+1}.
\end{aligned}
$$

Therefore, since $\sum_{k=1}^{\infty}\frac{1}{k}$ diverges, the result follows. ∎

As Proposition 3.4.4 shows the Dirichlet kernel is not a summability kernel, we are out of luck when it comes to applying Theorem 3.4.3 to show that the Fourier series converges uniformly for any continuous function. Of course, this does not show that Question 3.2.5 has a negative answer; just that our approach does not work.

3.5 Fejér's Kernel

Although we are out of luck in showing a positive answer to Question 3.2.5 using summability kernels, we can use summability kernels to obtain important information about Fourier series. The idea on how we can proceed is the following result that says it is much easier for the averages of a sequence to converge than it is for the original sequence to converge.

Lemma 3.5.1. *Let $(z_n)_{n\geq 0}$ be a sequence of complex numbers that converges to $L \in \mathbb{C}$. For each $n \in \mathbb{N}$, let*

$$\sigma_n = \frac{z_0 + z_1 + z_2 + \cdots + z_n}{n+1}.$$

Then $(\sigma_n)_{n\geq 1}$ converges to L.

Proof. To see that $(\sigma_n)_{n\geq 1}$ converges to L, let $\epsilon > 0$ be arbitrary. Since $(z_n)_{n\geq 0}$ converges to L, $(z_n)_{n\geq 0}$ is bounded. Therefore there exists an $M \in \mathbb{R}$ such that $|z_n| \leq M$ for all $n \in \mathbb{N}$. Moreover, since $(z_n)_{n\geq 0}$ converges to L, there exists an $N_1 \in \mathbb{N}$ such that

$$|z_n - L| < \frac{\epsilon}{2}$$

for all $n \geq N_1$. Since

$$\lim_{n\to\infty} \frac{(|L| + M)N_1}{n+1} = 0,$$

there exists an $N > N_1$ such that

$$\frac{(|L| + M)N_1}{n+1} < \frac{\epsilon}{2}$$

for all $n \geq N$. Hence for all $n \geq N$ we have that

$$\begin{aligned}
|L - \sigma_n| &= \left| L - \frac{1}{n+1}\sum_{k=0}^{n} z_k \right| \\
&= \left| \frac{1}{n+1}\sum_{k=0}^{n} L - z_k \right| \\
&\leq \frac{1}{n+1}\sum_{k=0}^{n} |L - z_k| \\
&= \frac{1}{n+1}\sum_{k=N_1}^{n} |L - z_k| + \frac{1}{n+1}\sum_{k=0}^{N_1-1} |L - z_k| \\
&\leq \frac{1}{n+1}\sum_{k=N_1}^{n} \frac{\epsilon}{2} + \frac{1}{n+1}\sum_{k=0}^{N_1-1} (|L| + M) \\
&\leq \frac{n - N_1}{n+1}\frac{\epsilon}{2} + \frac{(|L| + M)N_1}{n+1} \\
&\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.
\end{aligned}$$

Therefore, since $\epsilon > 0$ was arbitrary, $(\sigma_n)_{n\geq 1}$ converges to L as desired. ■

Remark 3.5.2. In light of Lemma 3.5.1, let us examine taking pointwise averages of Fourier series in the hopes of obtaining more information about convergence.

To begin, let $f \in \mathcal{RI}(\mathbb{T})$ be arbitrary. If $(P_n(f))_{n\geq 1}$ is going to converge pointwise to f, then by Lemma 3.5.1 we must have that f is the pointwise limit of

$$\frac{1}{n+1}\sum_{m=0}^{n} P_m(f) = \frac{1}{n+1}\sum_{m=0}^{n} f * D_m = f * \left(\frac{1}{n+1}\sum_{m=0}^{n} D_m\right).$$

Notice for all $x \in \mathbb{T}$ that

$$\begin{aligned}\frac{1}{n+1}\sum_{m=0}^{n} D_m(x) &= \frac{1}{n+1}\sum_{m=0}^{n}\sum_{k=-m}^{m} e^{ikx} \\ &= \frac{1}{n+1}\sum_{k=-n}^{n}\sum_{m=|k|}^{n} e^{ikx} \\ &= \frac{1}{n+1}\sum_{k=-n}^{n} (n-|k|+1)e^{ikx} \\ &= \sum_{k=-n}^{n}\left(1-\frac{|k|}{n+1}\right)e^{ikx}.\end{aligned}$$

Thus we make the following definitions.

Definition 3.5.3. For $n \in \mathbb{N}$, the n^{th} *Fejér kernel* (pronounced fay-yer), denote F_n, is the trigonometric polynomial defined by

$$F_n(x) = \frac{1}{n+1}\sum_{k=0}^{n} D_k(x) = \sum_{k=-n}^{n}\left(1-\frac{|k|}{n+1}\right)e^{ikx}$$

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

Definition 3.5.4. Let $f \in \mathcal{RI}(\mathbb{T})$ and let $n \in \mathbb{N}$. The n^{th} *Cesàro sum of f* (pronounced suh-zaa-row), denoted $\sigma_n(f)$, is

$$\sigma_n(f) = \frac{1}{n+1}\sum_{k=0}^{n} P_k(f) = f * \left(\frac{1}{n+1}\sum_{k=0}^{n} D_k\right) = f * F_n.$$

Thus for all $x \in \mathbb{T}$,

$$\sigma_n(f)(x) = \sum_{k=-n}^{n}\left(1-\frac{|k|}{n+1}\right)\widehat{f}(k)e^{ikx}.$$

As illustrated via Lemma 3.5.1, the reason to examine the Cesàro sums is their relation to pointwise convergence of the Fourier series. In fact, we have the following.

Lemma 3.5.5. *Let $f \in \mathcal{RI}(\mathbb{T})$. If $x_0 \in \mathbb{T}$ and $\lim_{n\to\infty} P_n(f)(x_0)$ exists, then $\lim_{n\to\infty} \sigma_n(f)(x_0)$ exists and*

$$\lim_{n\to\infty} \sigma_n(f)(x_0) = \lim_{n\to\infty} P_n(f)(x_0).$$

Moreover, if I is a closed interval of $\mathbb{T}$ (e.g. $I = \mathbb{T}$), g is a continuous function on I, and $(P_n(f))_{n\geq 1}$ converges uniformly to g on I, then $(\sigma_n(f))_{n\geq 1}$ converges uniformly to g on I.

Proof. Note the first part of the lemma follows immediately from the second part of the lemma by letting $I = \{x_0\}$. Thus we will prove the second part of the lemma. The proof will be achieved by modifying the proof of Lemma 3.5.1.

To see the second part of the lemma, suppose $(P_n(f))_{n\geq 1}$ converges uniformly to $g \in \mathcal{C}(\mathbb{T})$ on a closed interval $I \subseteq \mathbb{T}$. Thus Proposition 2.2.11 implies there exists an $M \in \mathbb{R}$ such that

$$|P_n(f)(x)| \leq M$$

for all $x \in I$ and $n \in \mathbb{N}$.

Since $(P_n(f))_{n\geq 1}$ converges uniformly to g on I, there exists an $N_1 \in \mathbb{N}$ such that

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

for all $n \geq N_1$ and $x \in I$. Since

$$\lim_{n\to\infty} \frac{(M_0 + M)N_1}{n+1} = 0,$$

there exists an $N > N_1$ such that

$$\frac{(M_0 + M)N_1}{n+1} < \frac{\epsilon}{2}$$

for all $n \geq N$. Hence for all $n \geq N$ and $x \in I$ we have that

$$\begin{aligned}
&|g(x) - \sigma_n(f)(x)| \\
&= \left| g(x) - \frac{1}{n+1} \sum_{k=0}^{n} P_k(f)(x) \right| \\
&= \left| \frac{1}{n+1} \sum_{k=0}^{n} g(x) - P_k(f)(x) \right| \\
&\leq \frac{1}{n+1} \sum_{k=0}^{n} |g(x) - P_k(f)(x)| \\
&= \frac{1}{n+1} \sum_{k=N_1}^{n} |g(x) - P_k(f)(x)| + \frac{1}{n+1} \sum_{k=0}^{N_1-1} |g(x) - P_k(f)(x)| \\
&\leq \frac{1}{n+1} \sum_{k=N_1}^{n} \frac{\epsilon}{2} + \frac{1}{n+1} \sum_{k=0}^{N_1-1} (M_0 + M) \\
&\leq \frac{n - N_1}{n+1} \frac{\epsilon}{2} + \frac{(M_0 + M)N_1}{n+1} \\
&\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.
\end{aligned}$$

Therefore, since $\epsilon > 0$ was arbitrary, $(\sigma_n(f))_{n\geq 1}$ converges uniformly to g on I as desired. ∎

Due to Lemma 3.5.5, perhaps we have a method for obtaining a negative answer to Questions 3.2.5 and 3.2.6. Indeed if we can show that the Cesàro sum do not converge in the appropriate sense, then the Fourier series do not converge in the corresponding sense. This will not be the case as the Fejér kernels are actually a summability kernel. To see this, we can repeat the idea of Lemma 3.3.7 to obtain a formula for the Fejér kernels.

Lemma 3.5.6. *For all $n \in \mathbb{N}$*

$$F_n(x) = \frac{1}{n+1} \left(\frac{\sin\left(\frac{n+1}{2}x\right)}{\sin\left(\frac{x}{2}\right)} \right)^2$$

as a continuous function on $\mathbb{T}$ (i.e. the formula for $F_n(0)$ should be interpreted as the limit as x tends to 0).

Proof. To being, first notice for all $x \in \mathbb{T}$ that

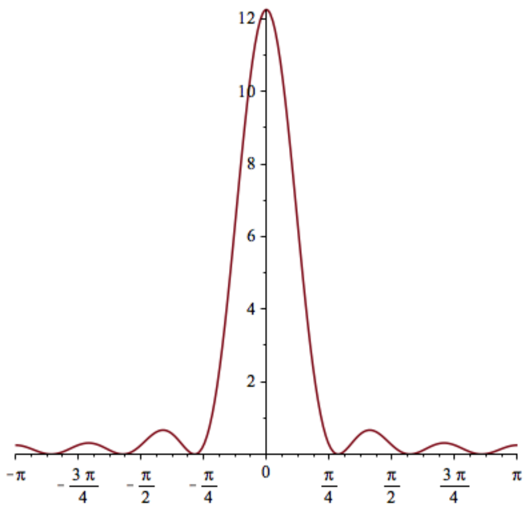
$$\sin^2\left(\frac{x}{2}\right) = \left(\frac{e^{i\frac{x}{2}} - e^{-i\frac{x}{2}}}{2i} \right)^2 = \frac{1}{2} - \frac{1}{4}e^{ix} - \frac{1}{4}e^{-ix}.$$

Therefore

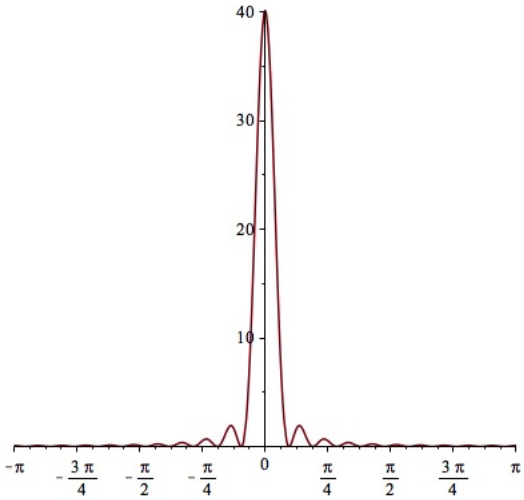
$$
\begin{aligned}
&\sin^2\left(\frac{x}{2}\right) F_n(x) \\
&= \left(\frac{1}{2} - \frac{1}{4}e^{ix} - \frac{1}{4}e^{-ix}\right) \sum_{k=-n}^{n} \left(1 - \frac{|k|}{n+1}\right) e^{ikx} \\
&= \left(\frac{1}{2} - \frac{1}{4}\left(1 - \frac{1}{n+1}\right) - \frac{1}{4}\left(1 - \frac{1}{n+1}\right)\right) e^0 \\
&\qquad - \frac{1}{4}\left(1 - \frac{n}{n+1}\right) e^{i(n+1)x} - \frac{1}{4}\left(1 - \frac{n}{n+1}\right) e^{-i(n+1)x} \\
&\qquad + \left(\frac{1}{2}\left(1 - \frac{n}{n+1}\right) - \frac{1}{4}\left(1 - \frac{n-1}{n}\right)\right) e^{inx} \\
&\qquad + \left(\frac{1}{2}\left(1 - \frac{n}{n+1}\right) - \frac{1}{4}\left(1 - \frac{n-1}{n}\right)\right) e^{-inx} \\
&\qquad + \sum_{k=1}^{n-1} \left(\frac{1}{2}\left(1 - \frac{k}{n+1}\right) - \frac{1}{4}\left(1 - \frac{k-1}{n+1}\right) - \frac{1}{4}\left(1 - \frac{k+1}{n+1}\right)\right) e^{ikx} \\
&\qquad + \sum_{k=1}^{n-1} \left(\frac{1}{2}\left(1 - \frac{k}{n+1}\right) - \frac{1}{4}\left(1 - \frac{k-1}{n+1}\right) - \frac{1}{4}\left(1 - \frac{k+1}{n+1}\right)\right) e^{-ikx} \\
&= \frac{1}{n+1}\left(\frac{1}{2} - \frac{1}{4}e^{i(n+1)x} - \frac{1}{4}e^{-i(n+1)x} + 0 + 0\right) \\
&= \frac{1}{n+1}\sin^2\left(\frac{n+1}{2}x\right)
\end{aligned}
$$

as desired. ∎

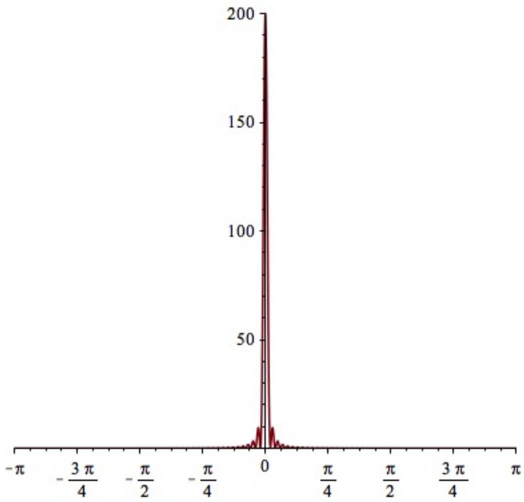
Using Lemma 3.5.6, it is easy to graph F_n for all $n \in \mathbb{N}$. In particular, the graph of F_3 is as follows:



The graph of F_{10} is as follows:



Finally, the graph of F_{50} is as follows:



In particular, these graphs look far more like the graphs of the functions used in the proof of the Weierstrass Approximation Theorem (Theorem 2.8.6). This is due to the fact we can prove the following.

Theorem 3.5.7. *The Fejèr's kernel has the following properties:*

a) $F_n(x) \geq 0$ *for all* $x \in \mathbb{T}$.

b) $F_n(-x) = F_n(x)$ *for all* $x \in \mathbb{T}$.

c) $\frac{1}{2\pi} \int_{-\pi}^{\pi} F_n(x)\, dx = 1$.

d) For all $\delta \in (0, \pi)$, $\lim_{n\to\infty} \sup\left(\{|F_n(x)| \mid \delta \leq |x| \leq \pi\}\right) = 0$.

Hence the Fejèr's kernel is a summability kernel.

Proof. Clearly a) follows from Lemma 3.5.6. Moreover, by Lemma 3.5.6,

$$F_n(-x) = \frac{1}{n+1}\left(\frac{\sin\left(\frac{n+1}{2}(-x)\right)}{\sin\left(\frac{(-x)}{2}\right)}\right)^2 = \frac{1}{n+1}\left(\frac{-\sin\left(\frac{n+1}{2}x\right)}{-\sin\left(\frac{x}{2}\right)}\right)^2 = F_n(x)$$

for all $x \in \mathbb{T}$ so b) follows.

To see that c) is true, notice since for all $k \in \mathbb{Z}$ we know that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{ikx} = \begin{cases} 1 & \text{if } k = 0 \\ 0 & \text{otherwise} \end{cases}$$

that

$$\begin{aligned} \frac{1}{2\pi}\int_{-\pi}^{\pi} F_n(x)\,dx &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{k=-n}^{n}\left(1 - \frac{|k|}{n+1}\right)e^{ikx}\,dx \\ &= \frac{1}{2\pi}\sum_{k=-n}^{n}\left(1 - \frac{|k|}{n+1}\right)\int_{-\pi}^{\pi} e^{ikx}\,dx = 1. \end{aligned}$$

To see that d) is true, let $\delta \in (0, \pi)$ be arbitrary. By Lemma 3.5.6, we know that

$$0 \leq F_n(x) \leq \frac{1}{n+1}\frac{1}{\sin\left(\frac{x}{2}\right)^2}$$

for all $x \in \mathbb{T} \setminus \{0\}$. Therefore, since $x \mapsto \sin(\frac{x}{2})$ is increasing on (δ, π), if $x \in \mathbb{T}$ and $\delta \leq |x| \leq \pi$, then

$$0 \leq F_n(x) \leq \frac{1}{n+1}\frac{1}{\sin\left(\frac{\delta}{2}\right)^2}.$$

Hence

$$0 \leq \int_{\delta\leq|x|\leq\pi} F_n(x)\,dx \leq \int_{\delta\leq|x|\leq\pi}\frac{1}{n+1}\frac{1}{\sin\left(\frac{\delta}{2}\right)^2}\,dx \leq \frac{2(\pi-\delta)}{n+1}\frac{1}{\sin\left(\frac{\delta}{2}\right)^2}.$$

Therefore, since

$$\lim_{n\to\infty}\frac{2(\pi-\delta)}{n+1}\frac{1}{\sin\left(\frac{\delta}{2}\right)^2} = 0$$

as δ is fixed, d) holds.

Finally, to see that $(F_n)_{n\geq1}$ is a summability kernel, note that property (I) of Definition 3.4.2 follows from part c), property (II) of Definition 3.4.2 follows from parts a) and c), and property (iii) of Definition 3.4.2 follows from part d). Hence $(F_n)_{n\geq1}$ is a summability kernel. ∎

Our knowledge of summability kernels immediately gives us some positive results towards answering Questions 3.2.5 and 3.2.6 in the affirmative.

Theorem 3.5.8 (Fejér's Theorem). *Let $f \in \mathcal{RI}(\mathbb{T})$. If $x \in \mathbb{T}$ is a point of continuity of f, then*

$$\lim_{n\to\infty} \sigma_n(f)(x) = f(x).$$

Moreover, if I is a closed interval in $\mathbb{T}$ (e.g. $I = \mathbb{T}$) and f is continuous on I, then $(\sigma_n(f))_{n\geq 1}$ converges uniformly to f on I.

Proof. The result immediately follows from Theorem 3.4.3 and Theorem 3.5.7. ∎

Of course Fejér's Theorem (Theorem 3.5.8 does not answer Questions 3.2.5 and 3.2.6 as we do not have the appropriate direction in Lemma 3.5.5. However, Fejér's Theorem can be used to obtain some additional knowledge about Fourier coefficients.

Corollary 3.5.9. *Let $f \in \mathcal{RI}(\mathbb{T})$. The following are true:*

a) If $x \in \mathbb{T}$ is a point of continuity of f and $\widehat{f}(n) = 0$ for all $n \in \mathbb{Z}$, then $f(x) = 0$.

b) If $f \in \mathcal{C}(\mathbb{T})$ and $\widehat{f}(n) = 0$ for all $n \in \mathbb{Z}$, then $f = 0$.

c) If $f, g \in \mathcal{C}(\mathbb{T})$ and $\widehat{f}(n) = \widehat{g}(n)$ for all $n \in \mathbb{Z}$, then $f = g$.

Proof. To see that a) is true, let $x \in \mathbb{T}$ be a point of continuity of f and suppose $\widehat{f}(n) = 0$ for all $n \in \mathbb{Z}$. Then $\sigma_n(f) = 0$ for all $n \in \mathbb{N}$ by definition. Therefore, since x is a point of continuity of f, Fejér's Theorem (Theorem 3.5.8) implies that

$$f(x) = \lim_{n\to\infty} \sigma_n(f)(x) = 0$$

as desired.

Next, note that b) follows immediately from part a). Finally, to see that c) is true, let $f, g \in \mathcal{C}(\mathbb{T})$ be such that $\widehat{f}(n) = \widehat{g}(n)$ for all $n \in \mathbb{Z}$. Therefore, $h = f - g \in \mathcal{C}(\mathbb{T})$ is such that

$$\widehat{h}(n) = (\widehat{f - g})(n) = \widehat{f}(n) - \widehat{g}(n) = 0$$

for all $n \in \mathbb{Z}$. Hence part b) implies that $h = 0$ so $f = g$ as desired. ∎

In addition, Fejér's Theorem (Theorem 3.5.8 allows us to prove that we can get an affirmative answer to Question 3.2.5 if we replace the partial Fourier series of a function with a different sequence of trigonometric polynomials.

Theorem 3.5.10. *Every element of $\mathcal{C}(\mathbb{T})$ can be uniformly approximated by trigonometric polynomials.*

Proof. Let $f \in \mathcal{C}(\mathbb{T})$ be arbitrary. Then $(\sigma_n(f))_{n\geq 1}$ is a sequence of trigonometric polynomials that converge uniformly to f by Fejér's Theorem (Theorem 3.5.8). Hence the result follows. ∎

Theorem 3.5.10 also lets us prove a fact that we would expect if we were to have the convergence of the Fourier series of functions.

Theorem 3.5.11 (The Riemann-Lebesgue Lemma). *If $f \in \mathcal{RI}(\mathbb{T})$, then*

$$\lim_{n\to\infty} \widehat{f}(n) = 0 \qquad \text{and} \qquad \lim_{n\to-\infty} \widehat{f}(n) = 0.$$

Proof. To see that the result is true, fix $f \in \mathcal{RI}(\mathbb{T})$ and let $\epsilon > 0$ be arbitrary. By Lemma 3.3.8, there exists a $g \in \mathcal{C}(\mathbb{T})$ such that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x) - g(x)|\, dx < \frac{\epsilon}{2}.$$

Therefore Proposition 3.2.12 implies that

$$|\widehat{f}(n) - \widehat{g}(n)| = |\widehat{(f-g)}(n)| \leq \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x) - g(x)|\, dx < \frac{\epsilon}{2}$$

for all $n \in \mathbb{Z}$.

By Theorem 3.5.10 there exists a $p \in \mathcal{T}(\mathbb{T})$ such that

$$|g(x) - p(x)| < \frac{\epsilon}{2}$$

for all $x \in \mathbb{T}$. Therefore, again by Proposition 3.2.12 implies that

$$|\widehat{g}(n) - \widehat{p}(n)| = |\widehat{(g-p)}(n)| \leq \frac{1}{2\pi}\int_{-\pi}^{\pi} |g(x) - p(x)|\, dx \leq \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\epsilon}{2}\, dx < \frac{\epsilon}{2}$$

for all $n \in \mathbb{Z}$. However, $p \in \mathcal{T}_N(\mathbb{T})$ for some $N \in \mathbb{N}$ so $\widehat{p}(n) = 0$ for all $n \in \mathbb{Z}$ such that $|n| \geq N$ by Example 3.2.7. Hence for all $n \in \mathbb{Z}$ with $|n| \geq N$, we have that

$$|\widehat{f}(n)| \leq |\widehat{f}(n) - \widehat{g}(n)| + |\widehat{g}(n) - \widehat{p}(n)| + |\widehat{p}(n)| < \frac{\epsilon}{2} + \frac{\epsilon}{2} + 0 = \epsilon.$$

Therefore, since $\epsilon > 0$ was arbitrary,

$$\lim_{n\to\infty} \widehat{f}(n) = 0 \qquad \text{and} \qquad \lim_{n\to-\infty} \widehat{f}(n) = 0$$

as desired. ∎

3.6 The Poisson Kernel

While we are on the topic of summability kernels, there is another kernel that is quite important in complex analysis and certain applications in applied mathematics.

Definition 3.6.1. For $r \in [0, 1)$, the r^{th} *Poisson kernel* (pronounced pwan-ssawn), denoted P_r, is the series of trigonometric polynomials defined by

$$P_r(x) = \sum_{n=-\infty}^{\infty} r^{|n|} e^{inx}$$

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

Remark 3.6.2. Note for all $r \in [0, 1)$ that since

$$\sum_{n=-\infty}^{\infty} r^{|n|} < \infty$$

by Example 1.2.3, Remark 3.2.19 implies that P_r is a well-defined element of $\mathcal{C}(\mathbb{T})$ with the above series converging uniformly and absolutely on $\mathbb{T}$ and

$$\widehat{P_r}(n) = r^{|n|}$$

for all $n \in \mathbb{Z}$.

Like with Fejér kernel, there is an alternative formula for the Poisson kernel.

Lemma 3.6.3. *For all* $r \in [0, 1)$

$$P_r(x) = \frac{1 - r^2}{1 - 2r\cos(x) + r^2} > 0$$

as a continuous function on $\mathbb{T}$.

Proof. Fix $r \in [0, 1)$ and let $x \in \mathbb{T}$ be arbitrary. Note when $r = 0$ the result is trivial. Otherwise when $r \neq 0$ notice that

$$1-2r\cos(x)+r^2 = 1-r(e^{ix}+e^{-ix})+r^2 = (1-re^{ix})(1-re^{-ix}) = |1-re^{ix}|^2.$$

Hence, since $r \in (0, 1)$ so

$$|re^{ix}| = r < 1,$$

we see that $|1 - re^{ix}|^2 \geq (1-r)^2 > 0$ for all $x \in \mathbb{T}$. Moreover, by Remark 3.6.2, we have that

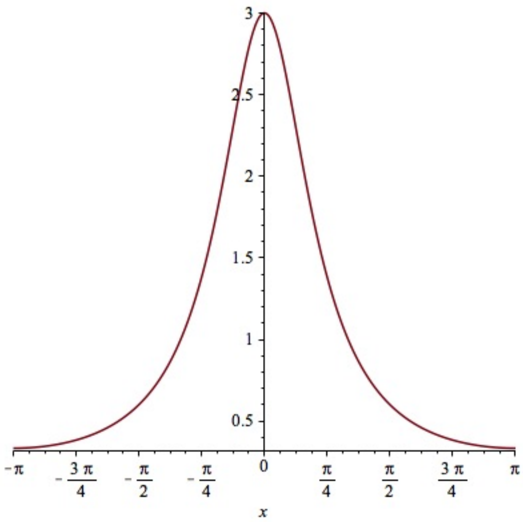
$$\begin{aligned}
&(1 - 2r\cos(x) + r^2)P_r(x) \\
&= \lim_{N\to\infty} (1 - r(e^{ix} + e^{ix}) + r^2) \sum_{n=-N}^{N} r^{|n|} e^{inx} \\
&= \lim_{N\to\infty} \left((1+r^2) - r^2 - r^2\right) e^0 - r^{N+1}e^{i(N+1)x} - r^{N+1}e^{-i(N+1)x} \\
&\quad + \left((1+r^2)r^N - rr^{N-1}\right) e^{iNx} + \left((1+r^2)r^N - rr^{N-1}\right) e^{-iNx} \\
&\quad + \sum_{n=1}^{N-1} \left((1+r^2)r^n - r(r^{n+1}) - r(r^{n-1})\right) e^{inx} \\
&\quad + \sum_{n=1}^{N-1} \left((1+r^2)r^n - r(r^{n+1}) - r(r^{n-1})\right) e^{-inx} \\
&= \lim_{N\to\infty} (1-r^2) - r^{N+1}e^{i(N+1)x} - r^{N+1}e^{-i(N+1)x} + r^{N+2}e^{iNx} + r^{N+2}e^{-iNx} + 0 + 0 \\
&= (1-r^2)
\end{aligned}$$

as $r \in (0,1)$. Hence

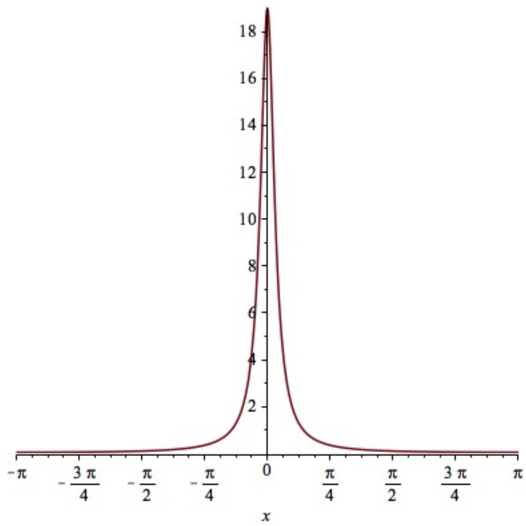
$$P_r(x) = \frac{1-r^2}{1 - 2r\cos(x) + r^2}$$

for all $x \in \mathbb{T}$. Moreover, as the both the numerator and denominator are positive, $P_r(x) > 0$ for all $x \in \mathbb{T}$ as desired. ∎

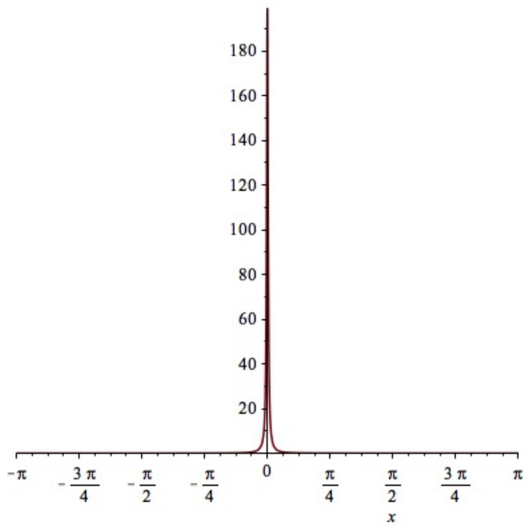
Using Lemma 3.6.3, it is easy to graph P_r for all $r \in [0,1)$. In particular, the graph of $P_{0.5}$ is as follows:



The graph of $P_{0.9}$ is as follows:



Finally, the graph of $P_{0.99}$ is as follows:



It turns out that the Poisson kernel is a summability kernel in the sense that if in the definition of a summability kernel (Definition 3.4.2) we replace $n \in \mathbb{N}$ with $n \to \infty$ with $[0, 1)$ and $r \to 1$, the same properties hold!

Lemma 3.6.4. *The Poisson kernel is a summability kernel.*

Proof. To see that property (I) of Definition 3.4.2 holds, note since the Poisson kernel converges uniformly by Remark 3.6.2 that its real and imaginary parts converge uniformly. Hence Corollary 2.4.5 (together with considering the real and imaginary parts) implies for all $r \in [0, 1)$ that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} P_r(x)\,dx = \sum_{n=-\infty}^{\infty} \frac{1}{2\pi}\int_{-\pi}^{\pi} r^{|n|}e^{inx}\,dx = \sum_{n=-\infty}^{\infty} \frac{1}{2\pi} r^{|n|}\int_{-\pi}^{\pi} e^{inx}\,dx = 1.$$

Hence property (I) of Definition 3.4.2 holds. Moreover, since $P_r(x) \geq 0$ for $x \in \mathbb{T}$ and $r \in [0,1)$, property (II) of Definition 3.4.2 immediately holds.

To see that property (III) of Definition 3.4.2 holds, let $\delta \in (0,\pi)$ be arbitrary. Since $x \mapsto \cos(x)$ is decreasing on $(\delta,\pi]$, we obtain from Lemma 3.6.3 that

$$0 \leq P_r(x) = \frac{1-r^2}{1-2r\cos(x)+r^2} \leq \frac{1-r^2}{1-2r\cos(\delta)+r^2}$$

for all $\delta \leq |x| \leq \pi$ and $r \in [0,1)$. Hence

$$\begin{aligned}\frac{1}{2\pi}\int_{\delta\leq|x|\leq\pi} |P_r(x)|\,dx &\leq \frac{1}{2\pi}\int_{\delta\leq|x|\leq\pi} \frac{1-r^2}{1-2r\cos(\delta)+r^2}\,dx \\ &= \frac{2(\pi-\delta)}{2\pi}\left(\frac{1-r^2}{1-2r\cos(\delta)+r^2}\right).\end{aligned}$$

However, since

$$\lim_{r\nearrow 1} 1-2r\cos(\delta)+r^2 = 2-2\cos(\delta) > 0$$

and $\lim_{r\nearrow 1} 1-r^2 = 0$, we obtain that

$$\lim_{r\nearrow 1}\frac{1}{2\pi}\int_{\delta\leq|x|\leq\pi} |P_r(x)|\,dx = 0.$$

Therefore, as $\delta \in (0,\pi)$ was arbitrary, the Poisson kernel is a summability kernel as desired. ∎

In order to apply the fact that the Poisson kernel is a summability kernel to obtain convergent series, we simply need to define the following.

Definition 3.6.5. Let $f \in \mathcal{RI}(\mathbb{T})$ and let $r \in [0,1)$. The r^{th} *Abel sum of f*, denoted $A_r(f)$, is

$$A_r(f) = f * P_r.$$

A description of the Abel sums without the need for convolution is easy to obtain.

Lemma 3.6.6. *Let $f \in \mathcal{RI}(\mathbb{T})$ and let $r \in [0,1)$. Then for all $x \in \mathbb{T}$,*

$$A_r(f)(x) = \sum_{n=-\infty}^{\infty} r^{|n|}\widehat{f}(n)e^{inx}.$$

Moreover

$$\widehat{A_r(f)}(n) = r^{|n|}\widehat{f}(n)$$

for all $n \in \mathbb{Z}$.

Proof. Recall the Poisson kernel converges uniformly by Remark 3.6.2. Therefore

$$\sum_{n=-\infty}^{\infty} r^{|n|} f(y) e^{in(x-y)}$$

converges uniformly over $y \in \mathbb{T}$ for any $x \in \mathbb{T}$ since $f \in \mathcal{R}(\mathbb{T})$ is bounded. Hence Corollary 2.4.5 (together with considering the real and imaginary parts) implies for all $r \in [0, 1)$ and $x \in \mathbb{T}$ that

$$\begin{aligned} A_r(f)(x) &= (f * P_r)(x) \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y) P_r(x-y)\, dy \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{n=-\infty}^{\infty} r^{|n|} f(y) e^{in(x-y)}\, dy \\ &= \sum_{n=-\infty}^{\infty} r^{|n|} \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} f(y) e^{in(x-y)}\, dy \right) \\ &= \sum_{n=-\infty}^{\infty} r^{|n|} \widehat{f}(n) e^{inx} && \text{by Example 3.3.2} \end{aligned}$$

thereby completing the first statement of the lemma.

To see that the second statement of the lemma is true, recall from the Riemann-Lebesgue Lemma (Theorem 3.5.11) that

$$\lim_{n\to\infty} \widehat{f}(n) = 0 \quad \text{and} \quad \lim_{n\to-\infty} \widehat{f}(n) = 0.$$

Therefore, $(\widehat{f}(n))_{n\in\mathbb{Z}}$ is bounded so there exists an $M \in \mathbb{R}$ such that

$$\left| \widehat{f}(n) \right| \leq M$$

for all $n \in \mathbb{Z}$. Hence for all $r \in [0, 1)$

$$\sum_{n=-\infty}^{\infty} \left| r^{|n|} \widehat{f}(n) e^{in\theta} \right| \leq \sum_{n=-\infty}^{\infty} M r^{|n|}$$

converges by Example 1.2.3. Therefore Proposition 3.2.20 implies that

$$\widehat{A_r(f)}(n) = r^{|n|} \widehat{f}(n)$$

for all $n \in \mathbb{Z}$ as desired. ∎

Of course, because the Poisson kernel is a summability kernel, we have the following.

Theorem 3.6.7. *Let $f \in \mathcal{RI}(\mathbb{T})$. If $x \in \mathbb{T}$ is a point of continuity of f, then*

$$\lim_{r \nearrow 1} A_f(f)(x) = f(x).$$

Moreover, if I is a closed interval in $\mathbb{T}$ (e.g. $I = \mathbb{T}$) and f is continuous on I, then $(A_r(f))_{r\in[0,1)}$ converges uniformly to f on I as r tends to 1.

Proof. The result immediately follows from Theorem 3.4.3 and Lemma 3.6.4. ∎

3.7 Failure of Fourier Series Convergence

With the above discussions complete, we arrive at the following which shows our attempts to show a positive answer to Questions 3.2.5 and 3.2.6 were folly. In fact, one of the biggest in mathematics was a mistaken proof that answers to Questions 3.2.5 and 3.2.6 was yes!

Theorem 3.7.1. *There exists a $f \in \mathcal{C}(\mathbb{T})$ such that $\mathcal{F}(f)$ diverges at a point. Hence the Fourier series of f does not converge pointwise nor uniformly to f on $\mathbb{T}$.*

Proof. To construct the desired function f, we will first look at a specific element g of $\mathcal{RI}(\mathbb{T})$ whose Fourier series is not absolutely summable. We will then use a series of translated partial Fourier of g to obtain f. This will be done so that when we take a certain sequence of partial Fourier series of f, each breaks into a bounded term plus a term involving the negative integer terms of a partial Fourier of g which will be unbounded. This idea of having only the negative integer terms of a partial Fourier series was derived to breaking the inherit symmetry of having both n and $-n$ in the definition of partial Fourier series.

Consider the function $g : \mathbb{T} \to \mathbb{C}$ defined by

$$g(x) = \begin{cases} i(\pi - x) & \text{if } 0 < x \leq \pi \\ i(-\pi - x) & \text{if } -\pi < x < 0 \end{cases}$$

for all $x \in \mathbb{T}$ (often g is called the sawtooth function due to its graph on $\mathbb{R}$). Notice if $h : \mathbb{T} \to \mathbb{C}$ is defined by $h(x) = x$ for all $x \in (-\pi, \pi]$, then

$$g(x) = -ih(x - \pi)$$

for all $x \in \mathbb{T}$. Hence Example 3.2.16 together with Proposition 3.2.12 implies that

$$\widehat{g}(n) = -ie^{-in\pi}\widehat{h}(n) = -i(-1)^n \left(\frac{1}{n}(-1)^n i\right) = \frac{1}{n}$$

for all $n \in \mathbb{Z}$ and $\widehat{g}(0) = 0$. Thus

$$\mathcal{F}(g) = \sum_{n \in \mathbb{Z} \setminus \{0\}} \frac{1}{n} e^{inx}$$

which clearly is not absolutely summable since $\sum_{n=1}^{\infty} \frac{1}{n}$ diverges.

For each $N \in \mathbb{N}$, consider the N^{th} partial Fourier series of g,

$$P_N(g)(x) = \sum_{1 \leq |n| \leq N} \frac{1}{n} e^{inx}.$$

One piece of information we will need for later use is the following.

Lemma 3.7.2. *There exists a constant $M \in \mathbb{R}$ such that*

$$|P_N(g)(x)| \leq M$$

for all $N \in \mathbb{N}$ and $x \in \mathbb{T}$ (that is, $(P_n(g))_{n \geq 1}$ is uniformly bounded).

Proof. We will use the Abel sums of g to obtain a bound for $P_N(g)$. For all $N \in \mathbb{N}$, let $r_N = 1 - \frac{1}{N} \in (0, 1)$. Then, for all $x \in \mathbb{T}$, we have that

$$\begin{aligned}
&|P_N(g)(x) - A_{r_N}(g)(x)| \\
&= \left| \sum_{1 \leq |n| \leq N} \frac{1}{n} e^{inx} - \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \frac{1}{n} r_N^{|n|} e^{inx} \right| \\
&\leq \left| \sum_{1 \leq |n| \leq N} \left(\frac{1}{n} - r_N^{|n|} \frac{1}{n} \right) e^{inx} \right| + \left| \sum_{|n| \geq N+1} \frac{1}{n} r_N^{|n|} e^{inx} \right| \\
&\leq 2 \sum_{n=1}^{N} \frac{1}{n} (1 - r_N^n) + 2 \sum_{n=N+1}^{\infty} \frac{1}{n} r_N^n \\
&\leq 2 \sum_{n=1}^{N} (1 - r_N^n) + 2 \sum_{n=N+1}^{\infty} \frac{1}{N} r_N^n \\
&\leq 2 \sum_{n=1}^{N} (1 - r_N) + \frac{2}{N} \sum_{n=N+1}^{\infty} r_N^n \\
&= 2N(1 - r_N) + \frac{2}{N} \frac{r_N^{N+1}}{1 - r_N} \\
&\leq 2N(1 - r_N) + \frac{2}{N} \frac{1}{1 - r_N} \\
&= 2N \left(1 - \left(1 - \frac{1}{N} \right) \right) + \frac{2}{N} \frac{1}{1 - \left(1 - \frac{1}{N} \right)} = 4.
\end{aligned}$$

However, since $|g(x)| \leq \pi$ for all $x \in \mathbb{T}$, we see for all $x \in \mathbb{T}$ and $N \in \mathbb{N}$ that

$$\begin{aligned}
|A_{r_N}(g)(x)| &= \left|\frac{1}{2\pi}\int_{-\pi}^{\pi} g(y)P_{r_N}(x-y)\,dy\right| \\
&\leq \frac{1}{2\pi}\int_{-\pi}^{\pi} |g(y)|P_{r_N}(x-y)\,dy && \text{as } P_{r_N}(x-y) > 0 \\
&\leq \frac{1}{2\pi}\int_{-\pi}^{\pi} \pi P_{r_N}(x-y)\,dy \\
&\leq \frac{1}{2\pi}\int_{-\pi}^{\pi} \pi P_{r_N}(y)\,dy && \text{translation and inversion invariance} \\
&= \pi && \text{by summability kernel properties.}
\end{aligned}$$

Hence

$$|P_N(g)(x)| \leq 4 + \pi$$

for all $x \in \mathbb{T}$ and $N \in \mathbb{N}$ as desired. ∎

For technical purposes in order to compute the partial sums of our desired function f, for each $N \in \mathbb{N}$ let $\widetilde{g}_N \in \mathcal{T}_n(\mathbb{T})$ be defined by

$$\widetilde{g}_N(x) = \sum_{n=-N}^{-1} \frac{1}{n} e^{inx}$$

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

One piece of information we will need for later use is the following.

Lemma 3.7.3. *For all $N \in \mathbb{N}$,*

$$|\widetilde{g}_N(0)| \geq \ln(N).$$

Proof. Notice for all $N \in \mathbb{N}$ that

$$\begin{aligned}
|\widetilde{g}_N(0)| &= \sum_{n=1}^{N} \frac{1}{n} \\
&\geq \sum_{n=1}^{N-1} \int_n^{n+1} \frac{1}{x}\,dx \\
&= \int_1^N \frac{1}{x}\,dx = \ln(N)
\end{aligned}$$

as desired. ∎

To obtain construct f, we first need some translate of the functions we have constructed above. For all $N \in \mathbb{N}$, define $h_N : \mathbb{T} \to \mathbb{C}$ and $\widetilde{h}_N : \mathbb{T} \to \mathbb{C}$ by

$$h_N(x) = e^{i(2N)x} P_N(g)(x) \qquad \text{and} \qquad \widetilde{h}_N(x) = e^{i(2N)x}\widetilde{g}_N(x)$$

for all $x \in \mathbb{T}$. In particular, we see that

$$h_N(x) = \sum_{1 \leq |n| \leq N} \frac{1}{n} e^{i(2N+n)x} \in \mathcal{T}_{3N}(\mathbb{T})$$

$$\widetilde{h}_N(x) = \sum_{n=-N}^{-1} \frac{1}{n} e^{i(2N+n)x} \in \mathcal{T}_{2N-1}(\mathbb{T}).$$

Note if $k \in \mathbb{N}$ then e^{-ikx} does not appear in the descriptions of h_N and $\widetilde{h}_N$. Moreover, by analyzing these formulae for h_N and $\widetilde{h}_N$, we immediately obtain the following piece of information we will need for later use.

Lemma 3.7.4. *For all $N \in \mathbb{N}$, if $K \in \mathbb{N}$ then*

$$P_K(h_N) = \begin{cases} 0 & \text{if } K < N \\ \widetilde{h}_N & \text{if } K = 2N \\ h_N & \text{if } K \geq 3N \end{cases}.$$

For our final construction before defining f, for each $k \in \mathbb{N}$ let

$$N_k = 3^{2^k}.$$

Thus

$$N_{k+1} > 3N_k$$

for all $k \in \mathbb{N}$.

Define $f : \mathbb{T} \to \mathbb{C}$ by

$$f(x) = \sum_{k=1}^{\infty} \frac{1}{k^2} h_{N_k}(x)$$

for all $x \in \mathbb{T}$. Since h_N is a trigonometric polynomial for all $N \in \mathbb{N}$ and thus an element of $\mathcal{C}(\mathbb{T})$, since

$$\left| \frac{1}{k^2} h_{N_k}(x) \right| = \left| \frac{1}{k^2} e^{i(2N_k)x} P_{N_k}(g)(x) \right| \leq \frac{1}{k^2} \left| e^{i(2N_k)x} \right| |P_{N_k}(g)(x)| \leq \frac{1}{k^2} M$$

for all $k \in \mathbb{N}$ and $x \in \mathbb{T}$, and since

$$\sum_{k=1}^{\infty} \frac{M}{k^2}$$

converges, the Weierstrass M-Test (Theorem 2.2.15) (applied to the real and imaginary parts) implies that f is well-define, the series defining f converges uniformly and absolutely, and $f \in \mathcal{C}(\mathbb{T})$. Therefore, since

$$\sum_{k=1}^{\infty} \frac{1}{k^2} h_{N_k}(y) D_n(x-y)$$

converges uniformly over $y \in \mathbb{T}$ for any $x \in \mathbb{T}$ since D_n is continuous and thus bounded, by Corollary 2.4.5 (applied to the real and imaginary parts) for all $n \in \mathbb{N}$ we see that

$$\begin{aligned}
P_n(f)(x) &= (f * D_n)(x) \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(\sum_{k=1}^{\infty} \frac{1}{k^2} h_{N_k}(y) \right) D_n(x-y)\, dy \\
&= \sum_{k=1}^{\infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(\frac{1}{k^2} h_{N_k}(y) \right) D_n(x-y)\, dy \\
&= \sum_{k=1}^{\infty} \frac{1}{k^2} P_n(h_{N_k})(x).
\end{aligned}$$

for all $x \in \mathbb{T}$. However, since $N_{k+1} > 3N_k$ for all $k \in \mathbb{N}$, notice for all $m \in \mathbb{N}$ that

$$\begin{aligned}
|P_{2N_m}(f)(0)| &= \left| \sum_{k=1}^{\infty} \frac{1}{k^2} P_{2N_m}(h_{N_k})(0) \right| \\
&= \left| \sum_{k=1}^{m} \frac{1}{k^2} P_{2N_m}(h_{N_k})(0) \right| \\
&= \left| \frac{1}{m^2} \widetilde{h}_{N_m}(0) + \sum_{k=1}^{m-1} \frac{1}{k^2} h_{N_k}(0) \right| \\
&\geq \frac{1}{m^2} |\widetilde{h}_{N_m}(0)| - \sum_{k=1}^{m-1} \frac{1}{k^2} |h_{N_k}(0)| \\
&\geq \frac{1}{m^2} \ln(N_m) - \sum_{k=1}^{m-1} \frac{1}{k^2} M \\
&\geq \frac{1}{m^2} \ln\left(3^{2^m}\right) - \sum_{k=1}^{\infty} \frac{1}{k^2} M \\
&\geq \frac{\ln(3) 2^m}{m^2} - \sum_{k=1}^{\infty} \frac{1}{k^2} M.
\end{aligned}$$

Therefore, since

$$\sum_{k=1}^{\infty} \frac{1}{k^2} M$$

converges and

$$\lim_{m \to \infty} \frac{\ln(3) 2^m}{m^2} = \infty,$$

we obtain that $(P_N(f)(0))_{N \geq 1}$ is not bounded and thus cannot possibly converge to $f(0)$. ∎

3.8 Instances of Uniform Convergence

As we have obtained negative answers to Questions 3.2.5 and 3.2.6, the best we can hope for is to show specific but very general cases where the Fourier series do converge uniformly or pointwise. In this section, we will look at some subsets of $\mathcal{C}(\mathbb{T})$ on which the Fourier series converge uniformly. The first such result follows from Proposition 3.2.20.

Corollary 3.8.1. *If $f \in \mathcal{C}(\mathbb{T})$ is such that*

$$\sum_{n=-\infty}^{\infty} |\widehat{f}(n)| < \infty,$$

then $(P_n(f))_{n\geq 1}$ converges uniformly to f on $\mathbb{T}$. Hence

$$f(x) = \mathcal{F}(f)(x) = \sum_{n=-\infty}^{\infty} \widehat{f}(n)e^{inx}$$

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

Proof. Note by Proposition 3.2.20 that if we define $g : \mathbb{T} \to \mathbb{C}$ by

$$g(x) = \sum_{n=-\infty}^{\infty} \widehat{f}(n)e^{inx}$$

for all $x \in \mathbb{T}$, then g is a well-defined element of $\mathcal{C}(\mathbb{T})$ such that

$$\widehat{g}(n) = \widehat{f}(n)$$

for all $n \in \mathbb{Z}$. Hence the uniqueness of the Fourier coefficients (Corollary 3.5.9) implies that $g = f$. Moreover, since Proposition 3.2.20 implies the series description of g converges uniformly to g on $\mathbb{T}$, $(P_n(f))_{n\geq 1}$ converges uniformly to f on $\mathbb{T}$ as desired. ∎

Using the above, we have an answer to specific case of Question 1.2.17.

Corollary 3.8.2. $\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$.

Proof. Recall by Example 3.2.17 that if we define $f : \mathbb{T} \to \mathbb{C}$ by $f(x) = \frac{1}{4}(x - \pi)^2$ for all $x \in [0, 2\pi)$, then

$$\mathcal{F}(f)(x) = \frac{\pi^2}{12} + \sum_{\substack{n=-\infty \\ n\neq 0}}^{\infty} \frac{1}{2n^2}e^{inx} = \frac{\pi^2}{12} + \sum_{n=1}^{\infty} \frac{1}{n^2}\cos(nx).$$

Since $f \in \mathcal{C}(\mathbb{T})$ and since

$$\sum_{n=-\infty}^{\infty} |\widehat{f}(n)| < \infty,$$

Corollary 3.8.1 implies that

$$f(0) = \frac{\pi^2}{12} + \sum_{n=1}^{\infty} \frac{1}{n^2} \cos(n(0)) = \frac{\pi^2}{12} + \sum_{n=1}^{\infty} \frac{1}{n^2}.$$

Therefore, since $f(0) = \frac{\pi^2}{4}$, we obtain that

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{4} - \frac{\pi^2}{12} = \frac{\pi^2}{6}$$

as desired. ∎

Our other main way to obtain continuous functions whose Fourier series converges uniformly on $\mathbb{T}$ is to consider the Cesàro sums. Indeed the following theorem implies that given a function with sufficiently nice Fourier coefficients, the Fourier series converges uniformly on $\mathbb{T}$ if and only if the Cesàro sum converge uniformly on $\mathbb{T}$. Since we know Cesàro sums converges uniformly on $\mathbb{T}$ for any element of $\mathcal{C}(\mathbb{T})$ by Fejér's Theorem (Theorem 3.5.8), we obtain the uniform convergence of Fourier series in some settings.

Theorem 3.8.3 (Hardy's Theorem). *Let $f \in \mathcal{RI}(\mathbb{T})$ be such that*

$$\sup\left\{\left|n\widehat{f}(n)\right| \mid n \in \mathbb{Z}\right\} < \infty.$$

If $x_0 \in \mathbb{T}$, then $\lim_{n\to\infty} P_n(f)(x_0)$ exists if and only if $\lim_{n\to\infty} \sigma_n(f)(x_0)$, in which case

$$\lim_{n\to\infty} P_n(f)(x_0) = \lim_{n\to\infty} \sigma_n(f)(x_0).$$

Furthermore, if I is a closed interval of $\mathbb{T}$ and g is a continuous function on I, then $(P_n(f))_{n\geq 1}$ converges uniformly to g on I if and only if $(\sigma_n(f))_{n\geq 1}$ converges uniformly to g on I.

Proof. As before, note the first part of the lemma follows immediately from the second part of the lemma by letting $I = \{x_0\}$. To see the second part of the statement is true, note by Lemma 3.5.5 that if $(P_n(f))_{n\geq 1}$ converges uniformly to g on I then $(\sigma_n(f))_{n\geq 1}$ converges uniformly to g on I.

For the converse statement, suppose $(\sigma_n(f))_{n\geq 1}$ converges uniformly to g on I. To see that $(P_n(f))_{n\geq 1}$ converges uniformly to g on I, let $\epsilon > 0$ be arbitrary. Our goal will be to decompose $P_n(f)$ into three terms, two which involve $\sigma_{n'}(f)$ for some $n' \in \mathbb{N}$, and one which can be made arbitrarily small using the assumptions of the theorem. We begin with the required estimates to manage this arbitrarily small term.

Since

$$\sup\left\{\left|n\widehat{f}(n)\right| \mid n \in \mathbb{Z}\right\} < \infty,$$

there exists an $M_1 \in \mathbb{R}$ such that

$$\left|n\widehat{f}(n)\right| \leq M_1$$

for all $n \in \mathbb{Z}$. For each $x \in \mathbb{R}$, let $\lfloor x \rfloor$ denoting the greatest integer less than or equal to x (i.e. the floor of x). Moreover, let

$$\alpha = 1 + \frac{1}{6M_1 + 1}\epsilon > 1.$$

Then for all $n \in \mathbb{N}$ we have that that

$$\begin{aligned}
\sum_{n<|k|\leq \lfloor \alpha n \rfloor} \left| \widehat{f}(k) \right| &\leq 2 \sum_{n<k\leq \lfloor \alpha n \rfloor} \frac{M_1}{k} \\
&\leq 2\frac{M_1}{n}\left(\lfloor \alpha n \rfloor - n\right) \\
&\leq 2\frac{M_1}{n}(\alpha - 1)n \\
&= 2M_1(\alpha - 1) < \frac{\epsilon}{3}.
\end{aligned}$$

Moreover, since $\alpha > 1$, there exists an $N_0 \in \mathbb{N}$ such that $\lfloor \alpha n \rfloor - n > 0$ for all $n \geq N_0$.

To return to the proof at hand, notice for all $n \geq N_0$ and $x \in I$ that

$$\begin{aligned}
&P_n(f)(x) \\
&= \sum_{0\leq |k| \leq n} \widehat{f}(k)e^{ikx} \\
&= \frac{1}{\lfloor \alpha n \rfloor - n} \sum_{0\leq |k| \leq n} \left(\lfloor \alpha n \rfloor - n\right) \widehat{f}(k)e^{ikx} \\
&= \frac{1}{\lfloor \alpha n \rfloor - n} \sum_{0\leq |k| \leq n} \left(\left(\lfloor \alpha n \rfloor + 1 - |k|\right) - (n + 1 - |k|)\right) \widehat{f}(k)e^{ikx} \\
&= \frac{\lfloor \alpha n \rfloor + 1}{\lfloor \alpha n \rfloor - n} \sum_{0\leq |k| \leq n} \left(1 - \frac{|k|}{\lfloor \alpha n \rfloor + 1}\right) e^{ikx} \\
&\qquad - \frac{n+1}{\lfloor \alpha n \rfloor - n} \sum_{0\leq |k| \leq n} \left(1 - \frac{|k|}{n+1}\right) \widehat{f}(k)e^{ikx} \\
&= \frac{\lfloor \alpha n \rfloor + 1}{\lfloor \alpha n \rfloor - n} \sum_{0\leq |k| \leq n} \left(1 - \frac{|k|}{\lfloor \alpha n \rfloor + 1}\right) \widehat{f}(k)e^{ikx} - \frac{n+1}{\lfloor \alpha n \rfloor - n}\sigma_n(f)(x) \\
&= \frac{\lfloor \alpha n \rfloor + 1}{\lfloor \alpha n \rfloor - n} \sum_{0\leq |k| \leq \lfloor \alpha n \rfloor} \left(1 - \frac{|k|}{\lfloor \alpha n \rfloor + 1}\right) \widehat{f}(k)e^{ikx} \\
&\qquad - \frac{\lfloor \alpha n \rfloor + 1}{\lfloor \alpha n \rfloor - n} \sum_{n< |k| \leq \lfloor \alpha n \rfloor} \left(1 - \frac{|k|}{\lfloor \alpha n \rfloor + 1}\right) \widehat{f}(k)e^{ikx} - \frac{n+1}{\lfloor \alpha n \rfloor - n}\sigma_n(f)(x) \\
&= \frac{\lfloor \alpha n \rfloor + 1}{\lfloor \alpha n \rfloor - n}\sigma_{\lfloor \alpha n \rfloor}(f)(x) - \frac{\lfloor \alpha n \rfloor + 1}{\lfloor \alpha n \rfloor - n} \sum_{n< |k| \leq \lfloor \alpha n \rfloor} \left(1 - \frac{|k|}{\lfloor \alpha n \rfloor + 1}\right) \widehat{f}(k)e^{ikx} \\
&\qquad - \frac{n+1}{\lfloor \alpha n \rfloor - n}\sigma_n(f)(x).
\end{aligned}$$

Since $\alpha > 1$ and
$$\frac{\alpha}{\alpha-1} - \frac{1}{\alpha-1} = 1,$$
we have for all $n \geq N_0$ and $x \in I$ that
$$\begin{aligned}|g(x) - P_n(f)(x)| \leq & \left|\frac{\alpha}{\alpha-1}g(x) - \frac{\lfloor \alpha n\rfloor + 1}{\lfloor \alpha n\rfloor - n}\sigma_{\lfloor \alpha n\rfloor}(f)(x)\right| \\ & + \left|\frac{\lfloor \alpha n\rfloor + 1}{\lfloor \alpha n\rfloor - n}\sum_{n<|k|\leq\lfloor \alpha n\rfloor}\left(1 - \frac{|k|}{\lfloor \alpha n\rfloor + 1}\right)\widehat{f}(k)e^{ikx}\right| \\ & + \left|\frac{1}{\alpha-1}g(x) - \frac{n+1}{\lfloor \alpha n\rfloor - n}\sigma_n(f)(x)\right|.\end{aligned}$$

We will now show that each of these three terms can be made uniformly small on I provided n is sufficiently large. Notice by our choice of α that for all $n \geq N_0$
$$\begin{aligned}&\left|\frac{\lfloor \alpha n\rfloor + 1}{\lfloor \alpha n\rfloor - n}\sum_{n<|k|\leq\lfloor \alpha n\rfloor}\left(1 - \frac{|k|}{\lfloor \alpha n\rfloor + 1}\right)\widehat{f}(k)e^{ikx}\right| \\ &\leq \frac{\lfloor \alpha n\rfloor + 1}{\lfloor \alpha n\rfloor - n}\sum_{n<|k|\leq\lfloor \alpha n\rfloor}\left(1 - \frac{|k|}{\lfloor \alpha n\rfloor + 1}\right)\left|\widehat{f}(k)\right| \\ &\leq \sum_{n<|k|\leq\lfloor \alpha n\rfloor}\left(\frac{\lfloor \alpha n\rfloor + 1 - |k|}{\lfloor \alpha n\rfloor - n}\right)\left|\widehat{f}(k)\right| \\ &\leq \sum_{n<|k|\leq\lfloor \alpha n\rfloor}\left|\widehat{f}(k)\right| \\ &\leq \frac{\epsilon}{3}.\end{aligned}$$

To see the other two terms can be made uniformly small on I, note since
$$\lim_{n\to\infty}\frac{\lfloor \alpha n\rfloor}{n} = \alpha,$$
that
$$\lim_{n\to\infty}\frac{\lfloor \alpha n\rfloor + 1}{\lfloor \alpha n\rfloor - n} = \frac{\alpha}{\alpha-1} \qquad \text{and} \qquad \lim_{n\to\infty}\frac{n+1}{\lfloor \alpha n\rfloor - n} = \frac{1}{\alpha-1}.$$
Therefore, since $(\sigma_n(f))_{n\geq 1}$ converges uniformly to g on I,
$$\left(\frac{\lfloor \alpha n\rfloor + 1}{\lfloor \alpha n\rfloor - n}\sigma_{\lfloor \alpha n\rfloor}(f)\right)_{n\geq 1} \qquad \text{and} \qquad \left(\frac{n+1}{\lfloor \alpha n\rfloor - n}\sigma_n(f)\right)_{n\geq 1}$$
converge uniformly to
$$\frac{\alpha}{\alpha-1}g \qquad \text{and} \qquad \frac{1}{\alpha-1}g$$

respectively on I by Proposition 2.2.12. Thus there exists an $N \in \mathbb{N}$ such that

$$\left|\frac{\alpha}{\alpha-1}g(x) - \frac{\lfloor \alpha n\rfloor + 1}{\lfloor \alpha n\rfloor - n}\sigma_{\lfloor \alpha n\rfloor}(f)(x)\right| \leq \frac{\epsilon}{3} \quad \text{and}$$
$$\left|\frac{1}{\alpha-1}g(x) - \frac{n+1}{\lfloor \alpha n\rfloor - n}\sigma_{n}(f)(x)\right| \leq \frac{\epsilon}{3}$$

for all $n \geq N$ and $x \in I$. Hence for all $n \geq N$ and $x \in I$, we have that

$$|g(x) - P_n(f)(x)| < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon.$$

Therefore, since $\epsilon > 0$ was arbitrary, $(P_n(f))_{n\geq 1}$ converges uniformly to g on I as desired. ∎

One important example that can be obtained from Hardy's Theorem (Theorem 3.8.3) is the following.

Corollary 3.8.4. *Let $f \in \mathcal{C}(\mathbb{T})$ be continuously differentiable on $\mathbb{T}$. Then $(P_n(f))_{n\geq 1}$ converges uniformly to f on $\mathbb{T}$. Hence*

$$f(x) = \mathcal{F}(f)(x) = \sum_{n=-\infty}^{\infty} \widehat{f}(n)e^{inx}$$

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

Proof. Let $f \in \mathcal{C}(\mathbb{T})$ be continuously differentiable on $\mathbb{T}$. Since $(\sigma_n(f))_{n\geq 1}$ converges uniformly to f on $\mathbb{T}$ by Fejér's Theorem (Theorem 3.5.8), it suffices by Hardy's Theorem (Theorem 3.8.3) to show that

$$\sup\left\{\left|n\widehat{f}(n)\right| \mid n \in \mathbb{Z}\right\} < \infty.$$

Notice that $f' \in \mathcal{C}(\mathbb{T})$ (i.e. the derivatives of 2π-periodic functions are 2π-periodic). Since f' is continuous on $\mathbb{T}$, the Extreme Value Theorem implies there exists an $M \in \mathbb{R}$ such that

$$|f'(x)| \leq M$$

for all $x \in \mathbb{T}$. Therefore, if $n \in \mathbb{N} \setminus \{0\}$ then, by using integration by parts, we see that

$$\begin{aligned}
&\left|2\pi\widehat{f}(n)\right| \\
&= \left|\int_{-\pi}^{\pi} f(x)e^{-inx}\,dx\right| \\
&= \left|\left(f(x)\frac{1}{-in}e^{-inx}\right)\Big|_{x=-\pi}^{\pi} - \int_{-\pi}^{\pi} f'(x)\frac{1}{-in}e^{-inx}\,dx\right| \\
&= \left|\frac{1}{in}\int_{-\pi}^{\pi} f'(x)e^{-inx}\,dx\right| && 2\pi\text{-periodic} \\
&\leq \frac{1}{n}\int_{-\pi}^{\pi} |f'(x)e^{-inx}|\,dx \\
&\leq \frac{1}{n}\int_{-\pi}^{\pi} M\,dx \\
&= \frac{1}{n}2\pi M.
\end{aligned}$$

Hence

$$\left|\widehat{f}(n)\right| \leq M\frac{1}{n}$$

for all $n \in \mathbb{N} \setminus \{0\}$ thereby completing the proof. ∎

3.9 Instances of Pointwise Convergence

By Theorem 3.7.1 we know there exists continuous functions for which the Fourier series does not converge uniformly. Often for applications it is sufficient that the Fourier series converges pointwise. Thus in this section, we will analyze when the Fourier series converges pointwise. Instead of starting with Fourier series, we will look at the Cesàro sums. This is due to the fact that Hardy's Theorem (Theorem 3.8.3) implies if the Cesàro sums converge pointwise and the Fourier coefficients are sufficiently nice, then the Fourier series converges pointwise.

The following is a very general result that shows the Cesàro sums converge at a point provided a specific condition is met, which we will see happens in a lot of cases.

Theorem 3.9.1 (Lebesgue's Theorem). *Let $f \in \mathcal{RI}(\mathbb{T})$, let $x_0 \in \mathbb{T}$, and let $L \in \mathbb{C}$. Suppose*

$$\lim_{h \searrow 0} \frac{1}{h}\int_0^h \left|\frac{f(x_0+x)+f(x_0-x)}{2} - L\right| dx = 0.$$

Then $\lim_{n\to\infty} \sigma_n(f)(x_0) = L$.

Proof. To prove this result, we will analyze $|\sigma_n(f)(x_0) - L|$ using integrals and the Fejér kernels, and divide the integral into three parts each of which we can show is small.

To begin, we will introduce some notation. Define $\varphi : \mathbb{T} \to \mathbb{R}$ by

$$\varphi(x) = f(x_0 + x) + f(x_0 - x) - 2L$$

for all $x \in \mathbb{T}$, and define $\Phi : [0, \pi] \to [0, \infty)$ by

$$\Phi(h) = \int_0^h |\varphi(x)|\, dx$$

for all $h \in [0, \pi]$. Note $\varphi \in \mathcal{RI}(\mathbb{T})$ so Φ is well-defined. Moreover, by the Fundamental Theorem of Calculus, Φ is differentiable on $(0, \pi)$ with

$$\Phi'(h) = |\varphi(h)|$$

for all $h \in (0, \pi)$.

To proceed with the proof, let $\epsilon > 0$ be arbitrary. Note that instead of inputting the correct multiple of ϵ at each step in the proof, we will end with a constant multiple of ϵ in our final bound on $|\sigma_n(f)(x_0) - L|$, which is sufficient.

By the assumptions of the theorem, we know that

$$\lim_{h \searrow 0} \frac{1}{h}\Phi(h) = 0.$$

Moreover, by Lemma 3.5.6, we see for all $n \in \mathbb{N}$ and $x \in \mathbb{T}$ that

$$\begin{aligned} F_n(x) &= \frac{1}{n+1}\left(\frac{\sin\left(\left(\frac{n+1}{2}\right)x\right)}{\sin\left(\frac{x}{2}\right)}\right)^2 \\ &\leq \frac{1}{n+1}\left(\frac{\left(\frac{n+1}{2}\right)x}{\sin\left(\frac{x}{2}\right)}\right)^2 \\ &= (n+1)\left(\frac{\frac{x}{2}}{\sin\left(\frac{x}{2}\right)}\right)^2. \end{aligned}$$

Therefore, since

$$\lim_{x \to 0} \frac{\frac{x}{2}}{\sin\left(\frac{x}{2}\right)} = 1$$

by the Fundamental Trigonometric Limit, there exists a $\delta > 0$ such that

$$\begin{aligned} &\frac{1}{h}\Phi(h) < \epsilon && \text{for all } h \in (0, \delta] \text{ and} \\ &F_n(x) \leq 2(n+1) && \text{for all } x \in (-\delta, \delta) \text{ and } n \in \mathbb{N}. \end{aligned}$$

By Theorem 3.5.7, there exists an $N \in \mathbb{N}$ such that if $n \geq N$ then $\frac{1}{n} < \delta$ and

$$0 \leq F_n(x) < \epsilon$$

for all $\delta \leq |x| \leq \pi$. Hence, by using properties from Theorem 3.5.7, for all $n \geq N$ we have that

$$\begin{aligned}
&|\sigma_n(f)(x_0) - L| \\
&= |(F_n * f)(x_0) - L| \\
&= \left|\frac{1}{2\pi}\int_{-\pi}^{\pi} F_n(y) f(x_0 - y)\, dy - L\right| \\
&= \left|\frac{1}{2\pi}\int_{0}^{\pi} F_n(y) f(x_0 - y)\, dy + \frac{1}{2\pi}\int_{-\pi}^{0} F_n(y) f(x_0 - y)\, dy - L\right| \\
&= \left|\frac{1}{2\pi}\int_{0}^{\pi} F_n(y) f(x_0 - y)\, dy + \frac{1}{2\pi}\int_{\pi}^{0} F_n(-t) f(x_0 + t)(-1)\, dt - L\right| \\
&= \left|\frac{1}{2\pi}\int_{0}^{\pi} F_n(y) f(x_0 - y)\, dy + \frac{1}{2\pi}\int_{0}^{\pi} F_n(t) f(x_0 + t)\, dt - L\right| \\
&= \left|\frac{1}{2\pi}\int_{0}^{\pi} F_n(y)(f(x_0 - y) + f(x_0 + y))\, dy - L\right| \\
&= \left|\frac{1}{2\pi}\int_{0}^{\pi} F_n(y)(f(x_0 - y) + f(x_0 + y))\, dy - L\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} F_n(y)\, dy\right)\right| \\
&= \left|\frac{1}{2\pi}\int_{0}^{\pi} F_n(y)(f(x_0 - y) + f(x_0 + y))\, dy - 2L\left(\frac{1}{2\pi}\int_{0}^{\pi} F_n(y)\, dy\right)\right| \\
&= \left|\frac{1}{2\pi}\int_{0}^{\pi} F_n(y)(f(x_0 - y) + f(x_0 + y) - 2L)\, dy\right| \\
&= \left|\frac{1}{2\pi}\int_{0}^{\pi} F_n(y)\varphi(y)\, dy\right| \\
&\leq \frac{1}{2\pi}\int_{0}^{\pi} F_n(y)|\varphi(y)|\, dy \\
&= \frac{1}{2\pi}\int_{0}^{\frac{1}{n}} F_n(y)|\varphi(y)|\, dy + \frac{1}{2\pi}\int_{\frac{1}{n}}^{\delta} F_n(y)|\varphi(y)|\, dy + \frac{1}{2\pi}\int_{\delta}^{\pi} F_n(y)|\varphi(y)|\, dy.
\end{aligned}$$

To complete the proof it suffices to show that each of these three terms are small for all $n \geq N$. For the first term, notice since $\frac{1}{n} < \delta$ for all $n \geq N$ that

$$\begin{aligned}
\frac{1}{2\pi}\int_{0}^{\frac{1}{n}} F_n(y)|\varphi(y)|\, dy &\leq \frac{1}{2\pi}\int_{0}^{\frac{1}{n}} 2(n+1)|\varphi(y)|\, dy \\
&= \frac{n+1}{\pi}\Phi\left(\frac{1}{n}\right) \\
&\leq \frac{2n}{\pi}\Phi\left(\frac{1}{n}\right) \\
&\leq \frac{2}{\pi}\left(\frac{1}{\frac{1}{n}}\Phi\left(\frac{1}{n}\right)\right) \\
&\leq \frac{2}{\pi}\epsilon,
\end{aligned}$$

which is just a constant multiple of ϵ. For the third term, notice that

$$\begin{aligned}\frac{1}{2\pi}\int_\delta^\pi F_n(y)|\varphi(y)|\,dy &\leq \frac{1}{2\pi}\int_\delta^\pi \epsilon|\varphi(y)|\,dy\\ &\leq \frac{\epsilon}{2\pi}\int_0^\pi |\varphi(y)|\,dy\\ &= \frac{\epsilon}{2\pi}\Phi(\pi)\end{aligned}$$

which is just a constant multiple of ϵ since Φ depends only on f and L which are fixed. Finally, for the second term, since

$$\sin\left(\frac{x}{2}\right) \geq \frac{x}{\pi}$$

for all $x \in (0, \pi)$ by elementary calculus, Lemma 3.5.6 implies that

$$0 \leq F_n(x) \leq \frac{1}{n+1}\frac{1}{\sin^2\left(\frac{x}{2}\right)} \leq \frac{\pi^2}{(n+1)x^2}$$

for all $x \in (0, \pi)$ so

$$\begin{aligned}&\frac{1}{2\pi}\int_{\frac{1}{n}}^\delta F_n(y)|\varphi(y)|\,dy\\ &\leq \frac{1}{2\pi}\int_{\frac{1}{n}}^\delta \frac{\pi^2}{(n+1)y^2}|\varphi(y)|\,dy\\ &= \int_{\frac{1}{n}}^\delta \frac{\pi}{2(n+1)y^2}\Phi'(y)\,dy\\ &= \left(\frac{\pi}{2(n+1)y^2}\Phi(y)\right)\Big|_{y=\frac{1}{n}}^\delta - \int_{\frac{1}{n}}^\delta -\frac{\pi}{4(n+1)y^3}\Phi(y)\,dy\\ &= \left(\frac{\pi}{2(n+1)\delta^2}\Phi(\delta) - \frac{\pi}{2(n+1)\left(\frac{1}{n}\right)^2}\Phi\left(\frac{1}{n}\right)\right) + \int_{\frac{1}{n}}^\delta \frac{\pi}{(n+1)y^3}\Phi(y)\,dy\\ &\leq \left(\frac{\pi}{2(n+1)\delta}\epsilon - 0\right) + \int_{\frac{1}{n}}^\delta \frac{\pi}{(n+1)y^2}\epsilon\,dy\\ &= \frac{\pi}{2(n+1)\delta}\epsilon + \left(-\frac{\pi}{(n+1)y}\epsilon\right)\Big|_{y=\frac{1}{n}}^\delta\\ &= \frac{\pi}{2(n+1)\delta}\epsilon + \left(-\frac{\pi}{(n+1)\delta}\epsilon + \frac{\pi}{(n+1)\frac{1}{n}}\epsilon\right)\\ &= \frac{\pi}{1+\frac{1}{n}}\epsilon - \frac{\pi}{2(n+1)\delta}\epsilon\\ &\leq \frac{\pi}{1+\frac{1}{n}}\epsilon\\ &\leq \pi\epsilon.\end{aligned}$$

Therefore, combining these three estimates, we obtain that

$$\begin{aligned}|\sigma_n(f)(x_0) - L| &\leq \frac{2}{\pi}\epsilon + \pi\epsilon + \frac{\epsilon}{2\pi}\Phi(\pi) \\ &= \left(\frac{2}{\pi} + \pi + \frac{1}{2\pi}\Phi(\pi)\right)\epsilon.\end{aligned}$$

Therefore, since $\frac{2}{\pi} + \pi + \frac{1}{2\pi}\Phi(\pi)$ is a constant that does not depend on ϵ and $\epsilon > 0$ was arbitrary,

$$\lim_{n\to\infty} \sigma_n(f)(x_0) = L$$

as desired. ∎

To see a case where Lebesgue's Theorem (Theorem 3.9.1) can be utilized, we prove the following which allows us to handle jump discontinuities.

Corollary 3.9.2 (Fejér's Theorem by Zygmund). *Let $f \in \mathcal{RI}(\mathbb{T})$. If $x_0 \in \mathbb{T}$, and*

$$\omega_f(x_0) = \frac{1}{2}\lim_{h\searrow 0} f(x_0 + h) + \frac{1}{2}\lim_{h\searrow 0} f(x_0 - h)$$

exists, then

$$\lim_{n\to\infty} \sigma_n(f)(x_0) = \omega_f(x_0).$$

Proof. To see that this result is true, we claim that

$$\lim_{h\searrow 0} \frac{1}{h}\int_0^h \left|\frac{f(x_0 + x) + f(x_0 - x)}{2} - \omega_f(x_0)\right| dx = 0.$$

Once this is established, the result then follows from Lebesgue's Theorem (Theorem 3.9.1) with $L = \omega_f(x_0)$.

To see that the above claim is true, let $\epsilon > 0$. Since

$$\omega_f(x_0) = \frac{1}{2}\lim_{h\searrow 0} f(x_0 + h) + \frac{1}{2}\lim_{h\searrow 0} f(x_0 - h)$$

there exists an $h_0 \in (0, \infty)$ such that if $0 < h \leq h_0$, then

$$\left|\frac{f(x_0 + h) + f(x_0 - h)}{2} - \omega_f(x_0)\right| \leq \epsilon.$$

Thus for all $0 < h \leq h_0$ we have that

$$0 \leq \frac{1}{h}\int_0^h \left|\frac{f(x_0 + x) + f(x_0 - x)}{2} - \omega_f(x_0)\right| dx \leq \frac{1}{h}\int_0^h \epsilon\, dx = \frac{1}{h}(h\epsilon) = \epsilon.$$

Therefore, as $\epsilon > 0$ was arbitrary, the claim and thus proof are complete. ∎

In order to use Lebesgue's Theorem (Theorem 3.9.1) to show that the Cesàro sums of f converge to $f(x_0)$ at various points x_0, we want x_0 to satisfy the following condition.

Definition 3.9.3. Let $f \in \mathcal{RI}(\mathbb{T})$. A point $x_0 \in \mathbb{T}$ is said to be a *Lebesgue point of f* if

$$\lim_{h \searrow 0} \frac{1}{h} \int_0^h \left| \frac{f(x_0 + x) + f(x_0 - x)}{2} - f(x_0) \right| dx = 0.$$

It turns out that given $f \in \mathcal{RI}(\mathbb{T})$ 'most' points are Lebesgue points.

Theorem 3.9.4. *If $f \in \mathcal{RI}(\mathbb{T})$, then "almost every point" in $\mathbb{T}$ is a Lebesgue point of f. Consequently, Lebesgue's Theorem (Theorem 3.9.1) implies that*

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

for "almost every" $x_0 \in \mathbb{T}$.

Proof. The proof of this theorem requires material from MATH 4012. It may or may not be covered based on who is teaching MATH 4012. ■

Now let us return and see what we can prove about the pointwise convergence of Fourier series without the need of Cesàro sums and Hardy's Theorem (Theorem 3.8.3). We start with the following.

Lemma 3.9.5. *If $f \in \mathcal{RI}(\mathbb{T})$ is such that*

$$\int_{-\pi}^{\pi} \left| \frac{f(x)}{x} \right| dx < \infty,$$

then $\lim_{n \to \infty} P_n(f)(0) = 0$.

Proof. Let $f \in \mathcal{RI}(\mathbb{T})$ be arbitrary. Using Lemma 3.3.7, notice for all $n \in \mathbb{N}$

that

$$\begin{aligned}
P_n(f)(0) &= (f * D_n)(0) \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(y)D_n(-y)\,dy \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(y)\frac{\sin\left(\left(n+\frac{1}{2}\right)(-y)\right)}{\sin\left(\frac{(-y)}{2}\right)}\,dy \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(y)\frac{-\sin\left(\left(n+\frac{1}{2}\right)y\right)}{-\sin\left(\frac{y}{2}\right)}\,dy \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(y)\frac{\sin\left(\left(n+\frac{1}{2}\right)y\right)}{\sin\left(\frac{y}{2}\right)}\,dy \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(y)\frac{\sin(ny)\cos\left(\frac{y}{2}\right)+\cos(ny)\sin\left(\frac{y}{2}\right)}{\sin\left(\frac{y}{2}\right)}\,dy \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(y)\frac{\sin(ny)\cos\left(\frac{y}{2}\right)}{\sin\left(\frac{y}{2}\right)}\,dy + \frac{1}{2\pi}\int_{-\pi}^{\pi} f(y)\cos(ny)\;dy
\end{aligned}$$

(where the integral can be split since both integrands are the product of a Riemann integrable function against a continuous function and thus Riemann integrable). We will show that

$$\lim_{n\to\infty}\frac{1}{2\pi}\int_{-\pi}^{\pi} f(y)\cos(ny)\;dy = 0 \quad \text{and} \tag{3.1}$$

$$\lim_{n\to\infty}\frac{1}{2\pi}\int_{-\pi}^{\pi} f(y)\frac{\sin(ny)\cos\left(\frac{y}{2}\right)}{\sin\left(\frac{y}{2}\right)}\,dy = 0 \tag{3.2}$$

thereby completing the proof.

To see that (3.1) is true, let $f_1 = \mathrm{Re}(f)$ and $f_2 = \mathrm{Im}(f)$. Therefore, for all $n \in \mathbb{N}$, we see that

$$\begin{aligned}
&\frac{1}{2\pi}\int_{-\pi}^{\pi} f(y)\cos(ny)\;dy \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} (f_1(y)+if_2(y))\cos(ny)\;dy \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} f_1(y)\cos(ny)\;dy + \frac{i}{2\pi}\int_{-\pi}^{\pi} f_2(y)\cos(ny)\;dy \\
&= 2\mathrm{Re}\left(\widehat{f_1}(n)\right) + 2i\mathrm{Re}\left(\widehat{f_2}(n)\right)
\end{aligned}$$

by Theorem 3.2.14 since f_1 and f_2 are real-valued. Therefore, since the Riemann-Lebesgue Lemma (Theorem 3.5.11) implies that

$$\lim_{n\to\infty}\widehat{f_1}(n) = 0 \quad \text{and} \quad \lim_{n\to\infty}\widehat{f_2}(n) = 0,$$

it follows that (3.1) is true.

To see that (3.2) is true, define $g, g_1, g_2 : \mathbb{T} \to \mathbb{C}$ by

$$g(x) = \frac{f(x)\cos\left(\frac{x}{2}\right)}{\sin\left(\frac{x}{2}\right)}, \quad g_1(x) = \frac{f_1(x)\cos\left(\frac{x}{2}\right)}{\sin\left(\frac{x}{2}\right)}, \quad \text{and} \quad g_2(x) = \frac{f_2(x)\cos\left(\frac{x}{2}\right)}{\sin\left(\frac{x}{2}\right)}$$

for all $x \in \mathbb{T}$. Clearly $g_1 = \mathrm{Re}(g)$, $g_2 = \mathrm{Im}(g)$, and

$$\begin{aligned}
&\frac{1}{2\pi}\int_{-\pi}^{\pi} f(y)\frac{\sin(ny)\cos\left(\frac{y}{2}\right)}{\sin\left(\frac{y}{2}\right)}\,dy \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} g(y)\sin(ny)\,dy \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} (g_1(y) + ig_2(y))\sin(ny)\;dy \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} g_1(y)\sin(ny)\;dy + \frac{i}{2\pi}\int_{-\pi}^{\pi} g_2(y)\sin(ny)\;dy \\
&= -2\mathrm{Im}\left(\widehat{g_1}(n)\right) - 2i\mathrm{Im}\left(\widehat{g_2}(n)\right)
\end{aligned}$$

by Theorem 3.2.14 since g_1 and g_2 are real-valued. Therefore, since the Riemann-Lebesgue Lemma (Theorem 3.5.11) implies that

$$\lim_{n\to\infty} \widehat{g_1}(n) = 0 \quad \text{and} \quad \lim_{n\to\infty} \widehat{g_2}(n) = 0,$$

it follows that (3.2) is true... modulo one caveat... $g \notin \mathcal{RI}(\mathbb{T})$ since g may not be bounded at zero! So does the above actually work?

To see the above is valid, first note since

$$\left|\tan\left(\frac{x}{2}\right)\right| \geq |x|$$

for all $x \in (-\pi, \pi)$ that

$$\int_{-\pi}^{\pi} |g(y)|\,dy = \int_{-\pi}^{\pi} \frac{|f(y)|}{\left|\tan\left(\frac{y}{2}\right)\right|}\,dy \leq \int_{-\pi}^{\pi} \frac{2|f(y)|}{|y|}\,dy < \infty \tag{3.3}$$

by the assumptions of the theorem. It is this fact that we can use to correct the proof.

First, due to (3.3), it follows that the conclusions of Proposition 3.2.12 hold whenever we use a function that is a linear combination of g and elements of $\mathcal{RI}(\mathbb{T})$. Therefore Theorem 3.2.14 is still valid as the proof of Theorem 3.2.14 relies only on Proposition 3.2.12. Furthermore, since Proposition 3.2.12 holds whenever we use a function that is a linear combination of g and elements of $\mathcal{RI}(\mathbb{T})$, to prove the Riemann-Lebesgue Lemma (Theorem 3.5.11) for g it remains only to show that for all $\epsilon > 0$ there exists an $h \in \mathcal{C}(\mathbb{T})$ so that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |g(y) - h(y)|\,dy < \epsilon.$$

The proof of this fact can be obtained by looking at the proof of Lemma 3.3.8. Indeed ignoring the upper bounds on elements of $\mathcal{C}(\mathbb{T})$, it suffices to prove there is a step function h_0 on $\mathbb{T}$ such that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |g(y) - h_0(y)|\, dy < \frac{\epsilon}{2}.$$

By (3.3), it follows that there exists a $\delta > 0$ such that

$$\int_{-\delta}^{\delta} |g(y)|\, dy < \frac{\epsilon}{4}$$

so by constructing a step function that is 0 on $[-\delta, \delta]$ and constructed on $\mathbb{T} \setminus [-\delta, \delta]$ via the same ideas as the proof of Lemma 3.3.8, the desired h_0 will be constructed.

(The reason we did not work with such functions throughout is that we only needed these facts in this one result and doing so would make other results more difficult to prove.)

(Alternatively, one can revisit this after taking MATH 4012. Indeed, in pretty much every result in this chapter, $\mathcal{RI}(\mathbb{T})$ can be replaced with the Lebesgue integrable functions $L_1(\mathbb{T})$ and (3.3) shows $g \in L_1(\mathbb{T})$.) ∎

Theorem 3.9.6 (Dirichlet-Dini's Test). *Let $f \in \mathcal{RI}(\mathbb{T})$ and let $x_0 \in \mathbb{T}$. Suppose $L \in \mathbb{C}$ is such that*

$$\int_0^{\pi} \frac{1}{x}\left|\frac{f(x_0 + x) + f(x_0 - x)}{2} - L\right| dx < \infty.$$

Then $\lim_{n\to\infty} P_n(f)(x_0) = L$.

Proof. Fix $f \in \mathcal{RI}(\mathbb{T})$, $x_0 \in \mathbb{T}$, and $L \in \mathbb{C}$ satisfying the assumptions of the theorem. Define $g : \mathbb{T} \to \mathbb{C}$ by

$$g(x) = \frac{f(x_0 + x) + f(x_0 - x)}{2} - L$$

for all $x \in \mathbb{T}$. Then $g \in \mathcal{RI}(\mathbb{T})$ and, since $g(-x) = g(x)$ for all $x \in \mathbb{T}$,

$$\int_{-\pi}^{\pi} \left|\frac{g(x)}{x}\right| dx = 2\int_0^{\pi} \left|\frac{g(x)}{x}\right| dx < \infty,$$

by assumption. Hence Lemma 3.9.5 implies that

$$\lim_{n\to\infty} P_n(g)(0) = 0.$$

However, notice by Proposition 3.3.3 and Proposition 3.4.4, we have for all $n \in \mathbb{N}$ that

$$\begin{aligned}
&P_n(g)(0) \\
&= (g * D_n)(0) \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} g(y) D_n(-y)\, dy \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(\frac{f(x_0 + y) + f(x_0 - y)}{2} - L \right) D_n(-y)\, dy \\
&= \frac{1}{4\pi} \int_{-\pi}^{\pi} f(x_0 + y) D_n(-y)\, dy + \frac{1}{4\pi} \int_{-\pi}^{\pi} f(x_0 - y) D_n(-y)\, dy \\
&\quad - \frac{1}{2\pi} \int_{-\pi}^{\pi} L D_n(-y)\, dy \\
&= \frac{1}{4\pi} \int_{\pi}^{-\pi} f(x_0 - r) D_n(r)(-1)\, dr + \frac{1}{4\pi} \int_{-\pi}^{\pi} f(x_0 - y) D_n(y)\, dy - L \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x_0 - y) D_n(y)\, dy - L \\
&= (D_n * f)(x_0) - L \\
&= (f * D_n)(x_0) - L \\
&= P_n(f)(x_0) - L.
\end{aligned}$$

Thus the result follows. ∎

One example where Dirichlet-Dini's Test (Theorem 3.9.6) applies is the following.

Corollary 3.9.7. *Let $f \in \mathcal{RI}(\mathbb{T})$ and suppose f is differentiable at a point $x_0 \in \mathbb{T}$. Then*

$$f(x_0) = \lim_{n \to \infty} P_n(f)(x_0).$$

Proof. Let $f \in \mathcal{RI}(\mathbb{T})$ be such that f is differentiable at a point $x_0 \in \mathbb{T}$. Therefore

$$\lim_{x \searrow 0} \frac{1}{x} \left| \frac{f(x_0 + x) + f(x_0 - x)}{2} - f(x_0) \right| = |f'(x_0)|$$

so

$$x \mapsto \frac{1}{x} \left| \frac{f(x_0 + x) + f(x_0 - x)}{2} - f(x_0) \right|$$

is a bounded function on $(0, \pi]$. Thus

$$\int_0^{\pi} \frac{1}{x} \left| \frac{f(x_0 + x) + f(x_0 - x)}{2} - f(x_0) \right| dx < \infty$$

so the Dirichlet-Dini's Test (Theorem 3.9.6) implies that

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

as desired. ∎

3.10 Mean-Square Convergence

To conclude this chapter, we return to our motivation for constructing Fourier series, namely orthogonal projections. Using this linear algebra concepts, we can describe another form of convergence and obtain more information about the Fourier coefficients. In particular, part a) of the following looks like convergence where length has replaced the absolute value. This idea will be further explored in MATH 4011.

Theorem 3.10.1 (Mean-Square Convergence). *For all $f \in \mathcal{RI}(\mathbb{T})$,*

a) $\lim_{n\to\infty} \|P_n(f) - f\|_2 = 0$.

b) $\|f\|_2 = \sqrt{\sum_{n=-\infty}^{\infty} |\widehat{f}(n)|^2}$.

Proof. To see that a) is true, we will use the fact that elements of $\mathcal{RI}(\mathbb{T})$ can be approximated with elements of $\mathcal{C}(\mathbb{T})$ 'sufficiently well' and elements of $\mathcal{C}(\mathbb{T})$ can be approximated with elements of $\mathcal{T}(\mathbb{T})$ 'sufficiently well'.

Fix $f \in \mathcal{RI}(\mathbb{T})$ and let $\epsilon > 0$ be arbitrary. Since $f \in \mathcal{RI}(\mathbb{T})$, there exists an $M \in \mathbb{R}$ such that $|f(x)| \leq M$ for all $x \in \mathbb{T}$. By Lemma 3.3.8, there exists an $g \in \mathcal{C}(\mathbb{T})$ such that

$$|g(x)| \leq 2M$$

for all $x \in \mathbb{T}$ and

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x) - g(x)|\, dx < \frac{\epsilon^2}{4(3M+1)}.$$

Moreover, by Theorem 3.5.10, there exists $p \in \mathcal{T}(\mathbb{T})$ such that

$$|g(x) - p(x)| < \frac{\epsilon}{2}$$

for all $x \in \mathbb{T}$. Hence

$$\begin{aligned}
\|f - p\|_2 &\leq \|f - g\|_2 + \|g - p\|_2 \\
&= \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x) - g(x)|^2\, dx\right)^{\frac{1}{2}} + \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} |g(x) - p(x)|^2\, dx\right)^{\frac{1}{2}} \\
&\leq \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x) - g(x)|(|f(x)| + |g(x)|)\, dx\right)^{\frac{1}{2}} + \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\epsilon^2}{4}\, dx\right)^{\frac{1}{2}} \\
&\leq \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x) - g(x)|(3M)\, dx\right)^{\frac{1}{2}} + \frac{\epsilon}{2} \\
&\leq \left((3M)\frac{\epsilon^2}{4(3M+1)}\right)^{\frac{1}{2}} + \frac{\epsilon}{2} \\
&\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.
\end{aligned}$$

Since $p \in \mathcal{T}(\mathbb{T})$, there exists an $N \in \mathbb{N}$ such that $p \in \mathcal{T}_N(\mathbb{T})$. Therefore, if $n \geq N$ then $p \in \mathcal{T}_n(\mathbb{T})$ and thus part d) of Theorem 3.1.18 implies that

$$\|f - P_n(f)\|_2 \leq \|f - p\|_2 < \epsilon.$$

Therefore since $\epsilon > 0$ was arbitrary, the proof of a) is complete.

To see that b) is true, notice for all $n \in \mathbb{N}$ that

$$\|f\|_2 = \|(f - P_n(f)) + P_n(f)\|_2 \leq \|f - P_n(f)\|_2 + \|P_n(f)\|_2$$

so

$$\|f\|_2 - \|P_n(f)\|_2 \leq \|f - P_n(f)\|_2 .$$

Similarly

$$\begin{aligned} \|P_n(f)\|_2 &= \|(P_n(f) - f) + f\|_2 \\ &\leq \|P_n(f) - f\|_2 + \|f\|_2 \\ &= |-1| \, \|f - P_n(f)\|_2 + \|f\|_2 \\ &= \|f - P_n(f)\|_2 + \|f\|_2 \end{aligned}$$

so

$$\|P_n(f)\|_2 - \|f\|_2 \leq \|f - P_n(f)\|_2 .$$

Hence

$$|\|f\|_2 - \|P_n(f)\|_2| \leq \|f - P_n(f)\|_2 .$$

Therefore, part a) implies that

$$\lim_{n\to\infty} |\|f\|_2 - \|P_n(f)\|_2| = 0$$

so

$$\begin{aligned}
\|f\|_2 &= \lim_{n\to\infty} \|P_n(f)\|_2 \\
&= \lim_{n\to\infty} \sqrt{\langle P_n(f), P_n(f)\rangle} \\
&= \lim_{n\to\infty} \sqrt{\left\langle \sum_{k=-n}^{n} \widehat{f}(k)e_k, \sum_{j=-n}^{n} \widehat{f}(j)e_j \right\rangle} \\
&= \lim_{n\to\infty} \sqrt{\sum_{k=-n}^{n}\sum_{j=-n}^{n} \widehat{f}(k)\overline{\widehat{f}(j)}\,\langle e_k, e_j\rangle} \\
&= \lim_{n\to\infty} \sqrt{\sum_{k=-n}^{n} \widehat{f}(k)\overline{\widehat{f}(k)}\,\langle e_k, e_k\rangle} && \substack{\langle e_j,e_k\rangle=0 \\ \text{if } j\neq k} \\
&= \lim_{n\to\infty} \sqrt{\sum_{k=-n}^{n} \widehat{f}(k)\overline{\widehat{f}(k)}} && \langle e_k,e_k\rangle=1 \\
&= \lim_{n\to\infty} \sqrt{\sum_{k=-n}^{n} |\widehat{f}(k)|^2} \\
&= \sqrt{\sum_{k=-\infty}^{\infty} |\widehat{f}(k)|^2} && \substack{\text{taking the square root} \\ \text{is continuous}}
\end{aligned}$$

as desired. ∎

One immediate corollary of the Mean-Square Convergence (Theorem 3.10.1) is that we can compute the value of the following sum.

Corollary 3.10.2. $\sum_{n=1}^{\infty} \frac{1}{n^4} = \frac{\pi^4}{90}$.

Proof. Recall from Example 3.2.17 that if $f : \mathbb{T} \to \mathbb{C}$ is defined by $f(x) = \frac{1}{4}(x-\pi)^2$ for all $x \in [0, 2\pi)$, then

$$\widehat{f}(n) = \frac{1}{2n^2}$$

for all $n \in \mathbb{Z}$ with $n \neq 0$ and $\widehat{f}(0) = \frac{\pi^2}{12}$. Therefore

$$\sum_{n=-\infty}^{\infty} \left|\widehat{f}(n)\right|^2 = \frac{\pi^4}{144} + \sum_{\substack{n=-\infty \\ n\neq 0}}^{\infty} \frac{1}{|2n^2|^2} = \frac{\pi^4}{144} + \frac{1}{2}\sum_{n=1}^{\infty}\frac{1}{n^4}.$$

Thus the Mean-Square Convergence (Theorem 3.10.1) implies that

$$\begin{aligned}
\sum_{n=1}^{\infty} \frac{1}{n^4} &= -\frac{\pi^4}{72} + 2\left\|f\right\|_2^2 \\
&= -\frac{\pi^4}{72} + \frac{1}{\pi}\int_0^{2\pi} |f(x)|^2\,dx \\
&= -\frac{\pi^4}{72} + \frac{1}{\pi}\int_0^{2\pi} \frac{1}{16}(x-\pi)^4\,dx \\
&= -\frac{\pi^4}{72} + \left(\frac{1}{16\pi}\frac{1}{5}(x-\pi)^5\right)\Big|_{x=0}^{2\pi} \\
&= -\frac{\pi^4}{72} + \left(\frac{1}{80\pi}(\pi)^5 - \frac{1}{80\pi}(-\pi)^5\right) \\
&= -\frac{\pi^4}{72} + \frac{\pi^4}{40} \\
&= \frac{\pi^4}{90}
\end{aligned}$$

as desired. ∎

Another immediate corollary of the Mean-Square Convergence (Theorem 3.10.1) is the ability to compute the integral of a product of functions via a series of Fourier coefficients.

Corollary 3.10.3 (Parseval's Identity). *For all $f, g \in \mathcal{RI}(\mathbb{T})$,*

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\overline{g(x)}\,dx = \sum_{n=-\infty}^{\infty} \widehat{f}(n)\overline{\widehat{g}(n)}.$$

Proof. It is elementary to verify for any $z, w \in \mathbb{C}$ that

$$z\overline{w} = \frac{1}{4}\left(|z+w|^2 - |z-w|^2 + i|z+iw|^2 - i|z-iw|^2\right).$$

Hence the Mean-Square Convergence (Theorem 3.10.1) implies that

$$\begin{aligned}
&\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\overline{g(x)}\,dx \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{1}{4}\left(|f(x)+g(x)|^2 - |f(x)-g(x)|^2 + i|f(x)+ig(x)|^2 - i|f(x)-ig(x)|^2\right)dx \\
&= \frac{1}{4}\left(\|f+g\|^2 - \|f-g\|^2 + i\,\|f+ig\|^2 - i\,\|f-ig\|^2\right) \\
&= \frac{1}{4}\left(\sum_{n=-\infty}^{\infty} |\widehat{f}(n)+\widehat{g}(n)|^2 - |\widehat{f}(n)-\widehat{g}(n)|^2 + i|\widehat{f}(n)+i\widehat{g}(n)|^2 - i|\widehat{f}(n)-i\widehat{g}(n)|^2\right) \\
&= \sum_{n=-\infty}^{\infty} \widehat{f}(n)\overline{\widehat{g}(n)}
\end{aligned}$$

as desired. ∎

Our final mentioned (at this time) corollary of the Mean-Square Convergence (Theorem 3.10.1) is that if "$P_n(f)(x)$ stayed away from $f(x)$ at too many points", then the integral

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x) - P_n(f)|^2\, dx$$

would stay large thereby yielding a contradiction. This is formally seen as follows.

Corollary 3.10.4. *If $f \in \mathcal{RI}(\mathbb{T})$, then $\lim_{n\to\infty} P_n(f)(x) = f(x)$ for 'almost every' $x \in \mathbb{T}$.*

Proof. The proof of this result is beyond the ability of this course as it requires technology from MATH 4012. ∎

One can actually improve upon the Mean-Square Convergence (Theorem 3.10.1):

Theorem 3.10.5. *If $f \in \mathcal{RI}(\mathbb{T})$ and $p \in (1,\infty)$, then*

$$\lim_{n\to\infty} \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x) - P_n(f)(x)|^p\, dx = 0.$$

Proof. The proof of this result is beyond the ability of this course as it requires technology from MATH 4011. ∎

Chapter 4

Applications of Series of Functions

In this final chapter, we will look at some applications of the theory developed throughout the course, mainly the results in Chapter 3. This will demonstrate the power of series, and more specifically Fourier series, in analysis.

Of course the applications studied in this chapter are just a sampling of what can be done. One example we wish we had time for would be a complete proof of the Central Limit Theorem in probability. Said proof can be accomplished by taking a suitable transformation of the distributions of averages of random variables, taking a limit of the transformations, realizing the limit is the transform of a Gaussian distribution, and showing this implies the average of random variables tends to the Gaussian distribution in a certain sense. This is very reminiscent of Fourier series, but requires a fair bit of technology that would take all the time we have for applications. Consequently, we will focus on these multiple applications instead.

4.1 Isoperimetric Inequality

For our first application of Fourier series, we will look at the question in geometry of maximizing the area contained in a region of $\mathbb{R}^2$ by a closed string of a fixed length. Of course our intuition says the optimal shape should be a circle. However, we have seen weird things in this course such as a function that is continuous but nowhere differentiable, so perhaps our intuition is incorrect. Modulo some assumptions (which are made in the following definition), it turns out Fourier series can be used to show our intuition is correct.

Definition 4.1.1. A *parametrized curve in* $\mathbb{R}^2$ is a set $\Gamma \subseteq \mathbb{R}^2$ such that there exists a continuous function $\gamma : [a, b] \to \mathbb{R}^2$ such that $\Gamma = \mathrm{Range}(\gamma)$. The function γ is called a *parametrization of* Γ.

It is said that Γ is *closed* if $\gamma(a) = \gamma(b)$.

It is said that Γ is *simple* if Γ does not intersect itself except possibly at the endpoints; that is if $t_1, t_2 \in [a, b]$ are such that $t_1 \neq t_2$, then $\gamma(t_1) \neq \gamma(t_2)$ provided $\{x_1, x_2\} \neq \{a, b\}$.

Finally, it is said that a simple closed curve Γ is *smooth* if γ is continuously differentiable on $[a, b]$ (with the one-sided derivatives at the endpoints agreeing) and $\gamma'(t) \neq \vec{0}$ for all $t \in [a, b]$.

Remark 4.1.2. It should be pointed out that a curve being 'smooth' can mean different things in different contexts by various authors. For us, we will just need the above definition in our main result and thus opt for this definition of 'smooth' in order to simplify discussions.

Remark 4.1.3. It is not difficult to see that if $\gamma : [a, b] \to \mathbb{R}^2$ is the parametrization of a simple closed curve, then the x- and y-terms of γ are continuous periodic functions with period $b - a$. We will make use of this fact in order to apply Fourier series provided we can reduce to $\mathbb{T}$. Indeed this can often be done via the following.

Definition 4.1.4. Let $\gamma : [a, b] \to \mathbb{R}^2$ be the parametrization of a simple, closed, smooth curve Γ. A *re-parametrization* of $\gamma : [a, b] \to \mathbb{R}^2$ is a function $\varphi : [c, d] \to \mathbb{R}^2$ such that there exists a continuously differentiable bijective function $\psi : [c, d] \to [a, b]$ with non-vanishing derivative such that

$$\varphi(s) = \gamma(\psi(s))$$

for all $s \in [c, d]$.

Note a re-parametrization yields the same subset of $\mathbb{R}^2$ and we can always re-parametrized so that we are working on $\mathbb{T}$. Of course, the hope in differential geometry is that the parametrization does not affect the quantities one desires to study. The first quantity we desire to study is the following.

Definition 4.1.5. Given a simple, closed, smooth curve Γ parametrized by $\gamma : [a, b] \to \mathbb{R}^2$ where $\gamma(t) = (x(t), y(t))$ for $x, y : [a, b] \to \mathbb{R}$, the *length of* Γ is defined to be

$$\ell(\Gamma) = \int_a^b \left\|\gamma'(t)\right\|_2 \, dt = \int_a^b \sqrt{|x'(t)|^2 + |y'(t)|^2} \, dt.$$

Remark 4.1.6. It is not difficult to see that the above is the correct formula for length based on elementary calculus arguments (i.e. distance travelled is the integral of speed).

Remark 4.1.7. It is necessary to show that the length of a simple, closed, smooth curve Γ does not depend on the parametrization. Indeed suppose Γ is parametrized by $\gamma : [a, b] \to \mathbb{R}^2$, $\varphi : [c, d] \to \mathbb{R}^2$, and $\psi : [c, d] \to [a, b]$ is a continuously differentiable bijective function with non-vanishing derivative such that

$$\varphi(s) = \gamma(\psi(s))$$

for all $s \in [c, d]$. Notice this implies

$$\varphi'(s) = \gamma'(\psi(s))\psi'(s)$$

for all $s \in [c, d]$. Moreover either $\psi'(s) > 0$ for all $s \in (c, d)$ or $\psi'(s) < 0$ for all $s \in (c, d)$. The former is the case that γ and φ draw Γ in the same orientation whereas the latter is the case that γ and φ draw Γ in different orientations. Therefore we obtain that

$$\begin{aligned} \int_c^d \left\|\varphi'(s)\right\|_2 \, ds &= \int_c^d \left\|\gamma'(\psi(s))\right\|_2 |\psi'(s)| \, ds \\ &= \int_a^b \left\|\gamma'(t)\right\|_2 \, dt \end{aligned}$$

via a change of variables. (Well, there is actually a lot to prove here. First, if $\psi(c) = a$ and $\psi(d) = b$, then $\psi'(s) > 0$ and the above holds via by the Change Rule. If $\psi(c) = b$ and $\psi(d) = a$, then $\psi'(s) < 0$ so we would introduce a negative sign when implementing the Change Rule, which would be corrected as we would need to change the order of the bounds of the integral. Finally, if we re-parametrized Γ with a different parametrization where we do not start and end at $\gamma(a) = \gamma(b)$, the same proof works by using periodicity.) Hence the length of Γ is well-defined.

Remark 4.1.8. Given a simple, closed, smooth curve Γ, we can always re-parametrize Γ to ensure Γ is parametrized via a continuously differentiable map $\gamma : [0, 2\pi] \to \mathbb{R}^2$ with non-vanishing derivative. Moreover, it is possible to assume that γ has *constant speed*; that is, there exists a $c \in \mathbb{R}$ such that $|\gamma'(t)| = c$ for all $t \in [0, 2\pi]$ (specifically, we would expect $c = \frac{\ell(\Gamma)}{2\pi}$). Indeed, consider the map $f : [0, 2\pi] \to [0, 2\pi]$ by

$$f(s) = \frac{2\pi}{\ell(\Gamma)} \int_0^s \left\|\gamma'(t)\right\|_2 \, dt$$

for all $s \in [0, 2\pi]$. By the definition of the length of Γ and since γ' does not vanish, f is a continuously differentiable bijective function with

$$f'(s) = \frac{2\pi}{\ell(\Gamma)} \left\|\gamma'(s)\right\|_2 \neq 0$$

for all $s \in [0, 2\pi]$. By elementary calculus, $\psi = f^{-1} : [0, 2\pi] \to [0, 2\pi]$ is a continuously differentiable bijective function with non-vanishing derivative such that

$$\psi'(s) = \frac{1}{f'(\psi(s))} = \frac{1}{\frac{2\pi}{\ell(\Gamma)} \left\|\gamma'(\psi(s))\right\|_2}$$

for all $s \in [0, 2\pi]$. Therefore if $\varphi : [0, 2\pi] \to \mathbb{R}^2$ is defined by

$$\varphi(s) = \gamma(\psi(s))$$

for all $s \in [0, 2\pi]$, then

$$\|\varphi'(s)\|_2 = \|\gamma'(\psi(s))\psi'(s)\|_2 = \|\gamma'(\psi(s))\|_2 \frac{1}{\frac{2\pi}{\ell(\Gamma)} \|\gamma'(\psi(s))\|_2} = \frac{\ell(\Gamma)}{2\pi}$$

for all $s \in [0, 2\pi]$ as desired.

By our Fourier series arguments, we can now assume that Γ is continuously differentiable map $\gamma : [-\pi, \pi] \to \mathbb{R}^2$ with constant speed. Thus we are in Fourier series territory now to study γ.

Of course, the other geometric quantity we require is the following.

Definition 4.1.9. Given a simple closed curve Γ parametrized by $\gamma : [a, b] \to \mathbb{R}^2$ where $\gamma(t) = (x(t), y(t))$ where $x, y : [a, b] \to \mathbb{R}$, the *area enclosed by* Γ is

$$A(\Gamma) = \frac{1}{2}\left| \int_a^b x(t)y'(t) - y(t)x'(t)\, dt \right|.$$

Remark 4.1.10. Green's Theorem from multivariate calculus implies the above is the correct formula for the area enclosed by a simple, closed, smooth curve. As Green's Theorem is the stable of any multivariate calculus course, we omit the proof.

Remark 4.1.11. Again It is necessary to show that the area enclosed by a simple, closed, smooth curve Γ does not depend on the parametrization. Indeed suppose Γ is parametrized by $\gamma : [a, b] \to \mathbb{R}^2$, $\varphi : [c, d] \to \mathbb{R}^2$, and $\psi : [c, d] \to [a, b]$ is a continuously differentiable bijective function with non-vanishing derivative such that

$$\varphi(s) = \gamma(\psi(s))$$

for all $s \in [c, d]$. Notice this implies

$$\varphi'(s) = \gamma'(\psi(s))\psi'(s)$$

for all $s \in [c, d]$. Thus, if $x, y : [a, b] \to \mathbb{R}$ and $x_0, y_0 : [c, d] \to \mathbb{R}$ are such that

$$\gamma(t) = (x(t), y(t)) \quad \text{and} \quad \varphi(s) = (x_0(s), y_0(s))$$

for all $t \in [a, b]$ and $s \in [c, d]$, then

$$\begin{aligned} x_0(s) &= x(\psi(x)) & y_0(s) &= y(\psi(s)) \\ x_0'(s) &= x'(\psi(s))\psi'(s) \text{ and} & y_0'(s) &= y'(\psi(s))\psi'(s) \end{aligned}$$

for all $s \in [c, d]$. Thus, by the same technical arguments as in Remark 4.1.7, we have that

$$\begin{aligned}
&\frac{1}{2}\left|\int_c^d x_0(s)y_0'(s) - y_0(s)x_0'(s)\,ds\right| \\
&= \frac{1}{2}\left|\int_c^d x(\psi(s))y'(\psi(s))\psi'(s) - y(\psi(s))x'(\psi(s))\psi'(s)\,ds\right| \\
&= \frac{1}{2}\left|\int_a^b x(t)y'(t) - y(t)x'(t)\,dt\right|
\end{aligned}$$

Hence the area enclosed by a simple, closed, smooth curve is well-defined.

With the above technicalities out of the way, we can discuss our main result.

Theorem 4.1.12 (Isoperimetric Inequality). *If Γ is a simple, closed, smooth curve in $\mathbb{R}^2$, then*

$$A(\Gamma) \leq \frac{\ell(\Gamma)^2}{4\pi}.$$

Moreover, this inequality is an equality if and only if Γ is a circle.

Proof. Note if Γ is a circle of radius r, then we know that $\ell(\Gamma) = 2\pi r$ and $A(\Gamma) = \pi r^2$ so equality easily holds. To proceed with the remainder of the proof, we will first show that the inequality is true, and then shows based on the proof that the inequality holds exactly when Γ is a circle.

To see that the inequality holds, let $\gamma : [-\pi, \pi] \to \mathbb{R}^2$ be a parametrization of Γ of constant speed. Thus if $x, y : [-\pi, \pi] \to \mathbb{R}^2$ are such that $\gamma(t) = (x(t), y(t))$ for all $t \in [-\pi, \pi]$, then there exists a $c \in \mathbb{R}$ such that $c > 0$ and

$$|x'(t)|^2 + |y'(t)|^2 = c^2$$

for all $t \in [-\pi, \pi]$. Therefore

$$\ell(\Gamma) = \int_{-\pi}^{\pi} \sqrt{|x'(t)|^2 + |y'(t)|^2}\,dt = \int_{-\pi}^{\pi} c\,dt = 2\pi c$$

Thus

$$c = \frac{\ell(\Gamma)}{2\pi}.$$

By assumptions, $x', y' \in \mathcal{C}(\mathbb{T})$. Moreover, notice for all $n \in \mathbb{Z}$ that

$$\begin{aligned}
\widehat{x'}(n) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} x'(t)e^{-int}\,dt \\
&= \left(\frac{1}{2\pi}x(t)e^{-int}\right)\Big|_{t=-\pi}^{\pi} - \frac{1}{2\pi}\int_{-\pi}^{\pi} x(t)(-in)e^{-int}\,dt \\
&= 0 + \frac{in}{2\pi}\int_{-\pi}^{\pi} x(t)e^{-int}\,dt \\
&= in\widehat{x}(n)
\end{aligned}$$

and similarly $\widehat{y'}(n) = in\widehat{y}(n)$. Therefore, the Mean Square Convergence (Theorem 3.10.1), we obtain that

$$\begin{aligned}
&\sum_{n=-\infty}^{\infty} n^2 \left(|\widehat{x}(n)|^2 + |\widehat{y}(n)|^2\right) \\
&= \sum_{n=-\infty}^{\infty} \left|\widehat{x'}(n)\right|^2 + \left|\widehat{y'}(n)\right|^2 \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} |x'(t)|^2\,dx + \frac{1}{2\pi}\int_{-\pi}^{\pi} |y'(t)|^2\,dx \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} |x'(t)|^2 + |y'(t)|^2\,dt \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} c^2\,dt \\
&= c^2. \qquad (4.1)
\end{aligned}$$

Furthermore, Parseval's Identity (Corollary 3.10.3) implies that

$$\begin{aligned}
A(\Gamma) &= \frac{1}{2}\left|\int_{-\pi}^{\pi} x(t)y'(t) - x'(t)y(t)\,dt\right| \\
&= \pi\left|\sum_{n=-\infty}^{\infty} \widehat{x}(n)\overline{\widehat{y'}(n)} - \overline{\widehat{x'}(n)}\widehat{y}(n)\right| \\
&= \pi\left|\sum_{n=-\infty}^{\infty} \widehat{x}(n)\overline{in\widehat{y}(n)} - \overline{in\widehat{x}(n)}\widehat{y}(n)\right| \\
&= \pi\left|\sum_{n=-\infty}^{\infty} -in(\widehat{x}(n)\overline{\widehat{y}(n)} - \overline{\widehat{x}(n)}\widehat{y}(n))\right| \\
&\leq \pi\sum_{n=-\infty}^{\infty} n|\widehat{x}(n)\overline{\widehat{y}(n)} - \overline{\widehat{x}(n)}\widehat{y}(n)| \qquad (4.2) \\
&\leq \pi\sum_{n=-\infty}^{\infty} 2n|\widehat{x}(n)||\widehat{y}(n)| \qquad (4.3) \\
&\leq \pi\sum_{n=-\infty}^{\infty} n(|\widehat{x}(n)|^2 + |\widehat{y}(n)|^2) \qquad (4.4) \\
&\leq \pi\sum_{n=-\infty}^{\infty} n^2(|\widehat{x}(n)|^2 + |\widehat{y}(n)|^2) \qquad n^2 \geq n \qquad (4.5) \\
&= \pi c^2 \\
&= \pi\left(\frac{\ell(\Gamma)}{2\pi}\right)^2 \\
&= \frac{\ell(\Gamma)^2}{4\pi}
\end{aligned}$$

as desired.

To prove that if $A(\Gamma) = \frac{\ell(\Gamma)^2}{4\pi}$ then Γ is a circle, note that equality in the Isoperimetric Inequality implies that (4.2), (4.3), (4.4), and (4.5) are equalities by the above proof. Note equality in (4.5) implies that $|\widehat{x}(n)|^2 + |\widehat{y}(n)|^2 = 0$ for all n such that $|n| \geq 2$. Thus $a_n = b_n = 0$ for all $n \in \mathbb{Z}$ with $|n| \geq 2$. Thus x and y must be trigonometric polynomials of degree at most 1 so we may write

$$x(t) = a_1 e^{it} + a_0 + a_{-1}e^{-it} \quad \text{and} \quad y(t) = b_1 e^{it} + b_0 + b_{-1}e^{-it}$$

for some $a_1, a_0, a_{-1}, b_1, b_0, b_{-1} \in \mathbb{C}$.

Since x and y are real-valued, this implies $a_0, b_0 \in \mathbb{R}$, $a_{-1} = \overline{a_1}$, and $b_{-1} = \overline{b_1}$ (see Theorem 3.2.14). Moreover, equality in (4.4) implies

$$2|a_1||b_1| = |a_1|^2 + |b_1|^2$$

and thus $|a_1| = |b_1|$. Furthermore, notice equation (4.1) implies that

$$c^2 = \sum_{n=-1}^{1} n^2(|a_n|^2 + |b_n|^2) = 2(|a_1|^2 + |b_1|^2) = 4|a_1|^2$$

Therefore

$$|a_1| = |b_1| = \frac{c}{2}$$

so we can write

$$a_1 = \frac{c}{2}e^{i\theta_a} \quad \text{and} \quad b_1 = \frac{c}{2}e^{i\theta_b}$$

for some $\theta_a, \theta_b \in [-\pi, \pi]$.

By examining the inequalities from (4.2) to (4.3) to (4.4), equality implies that

$$\begin{aligned} 2|a_1\overline{b_1} - \overline{a_1}b_1| &= \sum_{n=-1}^{1} n|\widehat{x}(n)\overline{\widehat{y}(n)} - \overline{\widehat{x}(n)}\widehat{y}(n)| \\ &= \sum_{n=-1}^{1} n^2(|a_n|^2 + |b_n|^2) = c^2. \end{aligned}$$

Thus, subbing in the expressions for a_1 and b_1, we obtain that

$$\begin{aligned} c^2 &= 2\left|\left(\frac{c}{2}e^{i\theta_a}\right)\left(\frac{c}{2}e^{-i\theta_b}\right) - \left(\frac{c}{2}e^{-i\theta_a}\right)\left(\frac{c}{2}e^{i\theta_a}\right)\right| \\ &= \frac{c^2}{2}\left|e^{i(\theta_a-\theta_b)} - e^{-i(\theta_a-\theta_b)}\right| \\ &= c^2|\sin(\theta_a - \theta_b)|. \end{aligned}$$

Hence $|\sin(\theta_a - \theta_b)| = 1$ so $\theta_a - \theta_b = \frac{k\pi}{2}$ for some $k \in \mathbb{Z}$. Therefore

$$\begin{aligned} x(t) &= a_1 e^{it} + a_0 + a_{-1} e^{-it} \\ &= a_1 e^{it} + a_0 + \overline{a_1} e^{-it} \\ &= a_0 + \frac{c}{2}\left(e^{i(t-\theta_a)} + e^{-i(t-\theta_a)}\right) \\ &= a_0 + \frac{c}{2}\cos(t - \theta_a) \end{aligned}$$

and

$$\begin{aligned} y(t) &= b_1 e^{it} + b_0 + b_{-1} e^{-it} \\ &= b_1 e^{it} + a_0 + \overline{b_1} e^{-it} \\ &= b_0 + \frac{c}{2}\left(e^{i(t-\theta_b)} + e^{-i(t-\theta_b)}\right) \\ &= b_0 + \frac{c}{2}\cos(t - \theta_b) \\ &= b_0 + \frac{c}{2}\cos\left(t - \theta_a + \frac{k\pi}{2}\right) \\ &= b_0 \pm \frac{c}{2}\sin(t - \theta_a) \end{aligned}$$

(where $+$ is used when $\frac{k-1}{2}$ is even and $-$ is used when $\frac{k-1}{2}$ is odd). Thus $\gamma(t) = (x(t), y(t))$ is the parametrization of a circle, so Γ is a circle as desired. ∎

4.2 Weyl's Equidistribution Theorem

Our next application of Fourier series is rooted in *ergodic theory*; the branch of mathematics that studies properties of dynamical systems. To understand the basic concepts of ergodic theory, consider the interval $[0, 1)$ where we work modulo 1 and a bijective function $f : [0, 1) \to [0, 1)$. The hope is to understand the behaviour of f^n on points of $[0, 1)$ as n varies. For example, for a fixed number $\gamma \in (0, 1)$, consider f defined by $f(x) = x + \gamma$. For a fixed $x_0 \in [0, 1)$, what can be said about $(f^n(x_0))_{n \geq 1}$? In particular, if we were to plot these points, what does their distribution look like? Does it clump up or spread out? In particular, does it have the following property?

Definition 4.2.1. A sequence $(x_n)_{n \geq 1}$ in $[0, 1)$ is said to be *equidistributed* if for all $(a, b) \subseteq [0, 1)$,

$$\lim_{N\to\infty} \frac{|\{1 \leq n \leq N \mid x_n \in (a, b)\}|}{N} = b - a$$

where, given $X \subseteq \mathbb{N}$, $|X|$ denotes the number of elements in X.

Remark 4.2.2. Equidistributed clearly means 'equally distributed'; that is, if a sequence $(x_n)_{n\geq 1}$ is equidistributed, then the probability a term in $(x_n)_{n\geq 1}$ lies in an open interval of $[0,1)$ is approximately the length of the interval as we add more and more points. This means that asymptotically $(x_n)_{n\geq 1}$ is distributed with respect to the uniform distribution. Thus $(x_n)_{n\geq 1}$ will be as 'spread out as possible' in $[0,1)$. Moreover $(x_n)_{n\geq 1}$ must intersect every interval; a property in topology known as being *dense*. Thus being able to show sequences are equidistributed is desirable.

It is not difficult to show that certain sequences are not equidistributed.

Example 4.2.3. For all $n \in \mathbb{N}$, let $x_n = \frac{1}{2^n}$. Then $(x_n)_{n\geq 1}$ is not equidistributed since

$$\frac{\left|\left\{1 \leq n \leq N \;\middle|\; x_n \in \left(\frac{1}{2}, 1\right)\right\}\right|}{N} = 0 \neq \frac{1}{2}$$

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

Example 4.2.4. Consider the sequence

$$0, \frac{1}{2}, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, \frac{1}{2}, \frac{1}{8}, \frac{1}{2}, \frac{3}{8}, \frac{1}{2}, \frac{5}{8}, \frac{1}{2}, \frac{7}{8}, \frac{1}{2}, \frac{1}{16}, \ldots.$$

It is not difficult to see that this sequence intersects every interval, but is not equidistributed since any interval (a,b) around $\frac{1}{2}$ will yield

$$\frac{|\{1 \leq n \leq N \mid x_n \in (a,b)\}|}{N} \geq \frac{1}{2} - \frac{1}{N},$$

and there are intervals of length less than $\frac{1}{2}$ centred at $\frac{1}{2}$.

For more examples, we note one useful way to construct sequences in $[0,1)$ is to construct a sequence in $\mathbb{R}$ and use the following 'modulo 1' function.

Definition 4.2.5. Given $x \in \mathbb{R}$, also called the *integer part of* x (the *floor of* x), denoted $\lfloor x \rfloor$ is the greatest integer less than or equal to x.

Given $x \in \mathbb{R}$, the *fractional part of* x, denoted $[x]$, is defined to be

$$[x] = x - \lfloor x \rfloor \in [0,1).$$

By working 'modulo 1', we note if $x, y \in [0,1)$, then $[x+y] \in [0,1)$. Similarly other operations behave well with respect to taking the fractional part.

Example 4.2.6. For another example of a non-equidistributed sequences, fix $q \in \mathbb{Q} \cap [0,1)$ and consider the sequence $([nq])_{n\geq 1}$. If $q = \frac{a}{b}$ with $\gcd(a,b) = 1$ then

$$[nq] \in \left\{0, \frac{1}{b}, \frac{2}{b}, \ldots, \frac{b-1}{b}\right\}$$

for all $n \in \mathbb{N}$. Thus clearly $([nq])_{n\geq 1}$ is not equidistributed as it does not intersect $\left(0, \frac{1}{b}\right)$.

However, if we take an irrational number, then the exact opposite occurs:

Theorem 4.2.7 (Weyl's Equidistributed Theorem). *If $\gamma \in \mathbb{R}$ is irrational, then the sequence $([n\gamma])_{n\geq 1}$ is equidistributed in $[0,1)$.*

The way we will demonstrate the Weyl's Equidistributed Theorem (Theorem 4.2.7) is by demonstrating the following criterion for a sequence to be equidistributed.

Theorem 4.2.8 (Weyl's Criterion). *A sequence $(x_n)_{n\geq 1}$ in $[0,1)$ is equidistributed in $[0,1)$ if and only if*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} e^{2\pi k x_n i} = 0$$

for all $k \in \mathbb{Z} \setminus \{0\}$.

Indeed, Weyl's Equidistributed Theorem (Theorem 4.2.7) can be proved using Weyl's Criterion (Theorem 4.2.8) as follows.

Proof of Weyl's Equidistributed Theorem via Weyl's Criterion. Let $\gamma \in \mathbb{R}$ be irrational. To see that $([n\gamma])_{n\geq 1}$ is equidistributed in $[0,1)$, it suffices by Weyl's Criterion (Theorem 4.2.8) to show that

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} e^{2\pi k[n\gamma] i} = 0$$

for all $k \in \mathbb{Z} \setminus \{0\}$.

Fix $k \in \mathbb{Z} \setminus \{0\}$ and notice that

$$e^{2\pi k\gamma i} = \cos(2\pi k\gamma) + i\sin(2\pi k\gamma) \neq 1$$

as γ is irrational. Therefore, for all $N \in \mathbb{N}$

$$\begin{aligned}
\frac{1}{N}\sum_{n=1}^{N} e^{2\pi k[n\gamma]i} &= \frac{1}{N}\sum_{n=1}^{N} e^{2\pi k n\gamma i} && 2\pi\text{-periodicity} \\
&= \frac{1}{N}\sum_{n=1}^{N} \left(e^{2\pi k\gamma i}\right)^n \\
&= \frac{1}{N} e^{2\pi k\gamma i}\frac{(e^{2\pi k\gamma i})^{N+1} - 1}{e^{2\pi k\gamma i} - 1} && \text{geometric series.}
\end{aligned}$$

Since for all $N \in \mathbb{N}$

$$\left|\frac{1}{N}\sum_{n=1}^{N} e^{2\pi k[n\gamma]i}\right| \leq \frac{1}{N}\frac{2}{|e^{2\pi k\gamma i} - 1|},$$

we obtain that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} e^{2\pi k [n\gamma] i} = 0$$

Thus, as $k \in \mathbb{Z} \setminus \{0\}$ was arbitrary, Weyl's Criterion (Theorem 4.2.8) implies that $([n\gamma])_{n\geq 1}$ is equidistributed in $[0, 1)$. ∎

Remark 4.2.9. To be able to prove Weyl's Criterion (Theorem 4.2.8), we desire to connect the definition of equidistribution in $[0, 1)$ to notions of functions and integrals we have seen in this course. To do this, let $(x_n)_{n\geq 1}$ be a sequence in $[0, 1)$, let $A \subseteq [0, 1)$, and define the *characteristic function of A* to be the function $\chi_A : [0, 1) \to \mathbb{R}$ defined by

$$\chi_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}.$$

Notice for any $(a, b) \subseteq [0, 1)$ that

$$\frac{1}{N} \sum_{n=1}^{N} \chi_{(a,b)}(x_n) = \frac{|\{1 \leq n \leq N \mid x_n \in (a, b)\}|}{N}$$

and

$$\int_0^1 \chi_{(a,b)}(x)\, dx = b - a.$$

Therefore, by the definition of equidistributed, $(x_n)_{n\geq 1}$ is equidistributed in $[0, 1)$ if and only if

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \chi_{(a,b)}(x_n) = \int_0^1 \chi_{(a,b)}(x)\, dx$$

for all $(a, b) \subseteq [0, 1)$. Consequently, it is easy to see how the following result may be of use in the proof of Weyl's Criterion (Theorem 4.2.8).

Lemma 4.2.10. *Let $(x_n)_{n\geq 1}$ be a sequence in $[0, 1)$. Then*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} e^{2\pi k x_n i} = 0$$

for all $k \in \mathbb{Z} \setminus \{0\}$ if and only if

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(x_n) = \int_0^1 f(x)\, dx$$

for all continuous periodic functions $f : \mathbb{R} \to \mathbb{R}$ with period 1.

Proof. To begin, suppose

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} f(x_n) = \int_0^1 f(x)\,dx.$$

for all continuous periodic functions $f : \mathbb{R} \to \mathbb{R}$ with period 1. By using the linearity of the integral, the above limit holds for all continuous periodic functions $f : \mathbb{R} \to \mathbb{C}$ with period 1.

Fix $k \in \mathbb{Z} \setminus \{0\}$ and let $f : \mathbb{R} \to \mathbb{C}$ be defined by

$$f(x) = e^{2\pi kxi}$$

for all $x \in \mathbb{R}$. Thus f is a continuous periodic function with period 1 so the above implies that

$$\begin{aligned} 0 &= \int_0^1 e^{2\pi kxi}\,dx \\ &= \int_0^1 f(x)\,dx \\ &= \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} f(x_n) \\ &= \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} e^{2\pi kx_n i} \end{aligned}$$

as desired.

For the converse direction, suppose

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} e^{2\pi kx_n i} = 0$$

for all $k \in \mathbb{Z} \setminus \{0\}$. The idea of the proof is that this assumption implies the conclusions for trigonometric polynomials (using $e^{2\pi kxi}$ instead of e^{ikx}) and thus will hold since we can uniformly approximate periodic functions with trigonometric polynomials.

Let $\epsilon > 0$ be arbitrary. Consider the function $g : \mathbb{R} \to \mathbb{R}$ defined by

$$g(x) = f\left(\frac{x}{2\pi}\right)$$

for all $x \in \mathbb{R}$. Since f is continuous and periodic with period 1, $g \in \mathcal{C}(\mathbb{T})$. Therefore, Theorem 3.5.10 implies there exists a $p \in \mathcal{T}(\mathbb{T})$ such that

$$|g(x) - p(x)| < \frac{\epsilon}{3}$$

for all $x \in [0, 2\pi)$. Since $p \in \mathcal{T}(\mathbb{T})$, there exists an $m \in \mathbb{N}$ and $\{a_k\}_{k=-m}^{m} \in \mathbb{C}$ such that

$$p(x) = \sum_{k=-m}^{m} a_k e^{ikx}$$

for all $x \in [0, 2\pi)$. Therefore, if $q : \mathbb{R} \to \mathbb{R}$ is defined by

$$q(x) = p(2\pi x) = \sum_{k=-m}^{m} a_k e^{2\pi kxi}$$

for all $x \in \mathbb{R}$, then for all $x \in [0, 1)$ we have that

$$|f(x) - q(x)| = |g(2\pi x) - p(2\pi x)| < \frac{\epsilon}{3}.$$

Notice that

$$\int_0^1 q(x)\,dx = \int_0^1 \sum_{k=-N}^{N} a_k e^{2\pi kxi}\,dx = \sum_{k=-m}^{m} a_k \int_0^1 e^{2\pi kxi}\,dx = a_0$$

and thus

$$\begin{aligned}
&\left|\frac{1}{N}\sum_{n=1}^{N} q(x_n) - \int_0^1 q(x)\,dx\right| \\
&= \left|\frac{1}{N}\sum_{n=1}^{N} q(x_n) - a_0\right| \\
&= \left|\frac{1}{N}\sum_{n=1}^{N}\sum_{k=-m}^{m} a_k e^{2\pi kx_n i} - a_0\right| \\
&= \left|\frac{1}{N}\sum_{\substack{k=-m \\ k\neq 0}}^{m}\sum_{n=1}^{N} a_k e^{2\pi kx_n i}\right| \qquad \text{since } \frac{1}{N}Na_0e^0 = a_0 \\
&\leq \sum_{\substack{k=-m \\ k\neq 0}}^{m} |a_k| \left|\frac{1}{N}\sum_{n=1}^{N} e^{2\pi kx_n i}\right|.
\end{aligned}$$

Therefore, since

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} e^{2\pi kx_n i} = 0$$

for all $k \in \mathbb{Z} \setminus \{0\}$, there exists an $N_0 \in \mathbb{N}$ such that

$$\left|\frac{1}{N}\sum_{n=1}^{N} q(x_n) - \int_0^1 q(x)\,dx\right| < \frac{\epsilon}{3}$$

for all $N \geq N_0$.

Thus, for all $N \geq N_0$ we have that

$$\begin{aligned}
&\left|\frac{1}{N}\sum_{n=1}^{N} f(x_n) - \int_0^1 f(x)\,dx\right| \\
&\leq \left|\frac{1}{N}\sum_{n=1}^{N} f(x_n) - \frac{1}{N}\sum_{n=1}^{N} q(x_n)\right| + \left|\frac{1}{N}\sum_{n=1}^{N} q(x_n) - \int_0^1 q(x)\,dx\right| \\
&\quad + \left|\int_0^1 q(x)\,dx - \int_0^1 f(x)\,dx\right| \\
&\leq \frac{1}{N}\left(\sum_{n=1}^{N} |f(x_n) - q(x_n)|\right) + \frac{\epsilon}{3} + \int_0^1 |q(x) - f(x)|\,dx \\
&\leq \frac{1}{N}\left(\sum_{n=1}^{N} \frac{\epsilon}{3}\right) + \frac{\epsilon}{3} + \int_0^1 \frac{\epsilon}{3}\,dx \\
&= \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon.
\end{aligned}$$

Therefore, as $\epsilon > 0$ was arbitrary, the result is complete. ∎

With that lemma out of the way, it is now possible to prove our desired result.

Proof of Weyl's Criterion (Theorem 4.2.8). Let $(x_n)_{n\geq 1}$ be a sequence in $[0, 1)$. Suppose that

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} e^{2\pi k x_n i} = 0$$

for all $k \in \mathbb{Z} \setminus \{0\}$. To see that $(x_n)_{n\geq 1}$ is equidistributed in $[0, 1)$, by Remark 4.2.9 we simply need to upgrade Lemma 4.2.10 to include characteristic functions of intervals.

Let $(a, b) \subseteq [0, 1]$ be such that $a < b$ and let $\epsilon > 0$ be arbitrary. Consider the 1-periodic functions $f_+ : \mathbb{R} \to \mathbb{R}$ and $f_- : \mathbb{R} \to \mathbb{R}$ defined by

$$f_+(x) = \begin{cases} 1 & \text{if } x \in [a, b] \\ \frac{1}{\epsilon}(x - a + \epsilon) & \text{if } x \in [a - \epsilon, a] \setminus [a, b] \\ \frac{1}{\epsilon}(b + \epsilon - x) & \text{if } x \in [b, b + \epsilon] \setminus [a, b] \\ 0 & \text{otherwise} \end{cases}$$

$$f_-(x) = \begin{cases} 1 & \text{if } x \in [a + \epsilon, b - \epsilon] \\ \frac{1}{\epsilon}(x - a) & \text{if } x \in [a, a + \epsilon] \\ \frac{1}{\epsilon}(b - x) & \text{if } x \in [b - \epsilon, b] \\ 0 & \text{otherwise} \end{cases}.$$

Thus f_+ and f_- are continuous functions such that

$$\int_a^b f_-(x)\,dx = (b-a) - 2\epsilon, \qquad \int_a^b f_+(x)\,dx \leq (b-a) + 2\epsilon$$

and

$$f_-(x) \leq \chi_{(a,b)}(x) \leq f_+(x)$$

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

Recall Lemma 4.2.10 implies that

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^N f_+(x_n) = \int_0^1 f_+(x)\,dx$$
$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^N f_-(x_n) = \int_0^1 f_-(x)\,dx.$$

Therefore, since for all $N \in \mathbb{N}$,

$$\frac{1}{N}\sum_{n=1}^N f_-(x_n) \leq \frac{1}{N}\sum_{n=1}^N \chi_{(a,b)}(x_n) \leq \frac{1}{N}\sum_{n=1}^N f_+(x_n),$$

we obtain that

$$\begin{aligned}
(b-a) - 2\epsilon &= \int_0^1 f_-(x)\,dx. \\
&\leq \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^N f_-(x_n) \\
&\leq \liminf_{N\to\infty} \frac{1}{N}\sum_{n=1}^N \chi_{(a,b)}(x_n) \\
&\leq \limsup_{N\to\infty} \frac{1}{N}\sum_{n=1}^N \chi_{(a,b)}(x_n) \\
&\leq \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^N f_+(x_n) \\
&= \int_0^1 f_+(x)\,dx \\
&\leq (b-a) + 2\epsilon.
\end{aligned}$$

Thus, as $\epsilon > 0$ was arbitrary, we obtain that

$$b - a \leq \liminf_{N\to\infty} \frac{1}{N}\sum_{n=1}^N \chi_{(a,b)}(x_n) \leq \limsup_{N\to\infty} \frac{1}{N}\sum_{n=1}^N \chi_{(a,b)}(x_n) \leq b - a.$$

Hence

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^N \chi_{(a,b)}(x_n) = b - a = \int_0^1 \chi_{(a,b)}(x)\,dx.$$

Therefore, since $(a, b) \subseteq [0, 1)$ was arbitrary, $(x_n)_{n\geq 1}$ is equidistributed in $[0, 1)$ by Remark 4.2.9.

Conversely, suppose that $(x_n)_{n\geq 1}$ is equidistributed in $[0, 1)$. To see that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} e^{2\pi k x_n i} = 0$$

for all $k \in \mathbb{Z} \setminus \{0\}$, it suffices by Lemma 4.2.10 to show that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(x_n) = \int_0^1 f(x)\, dx$$

for all continuous periodic functions $f : \mathbb{R} \to \mathbb{R}$ with period 1. The idea is to use the fact $(x_n)_{n\geq 1}$ is equidistributed in $[0, 1)$ implies the desired limit holds for characteristic functions of intervals and extend this to linear combinations of characteristic functions of intervals and then to all continuous functions via Riemann sums.

First we generalize Remark 4.2.9 to more intervals. Since $(x_n)_{n\geq 1}$ is equidistributed in $[0, 1)$, Remark 4.2.9 implies that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{n} \chi_{(a,b)}(x_n) = b - a.$$

Thus

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{n} \chi_{(0,1)}(x_n) = 1.$$

so

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{n} \chi_{[0,0]}(x_n) = 0.$$

Since for any $b \leq 1$ we have that

$$\chi_{[0,b]}(x) = 1 - \chi_{(b,1)}(x)$$

for all $x \in [0, 1)$ we obtain that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{n} \chi_{[0,b]}(x_n) = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{n} 1 - \chi_{(b,1)}(x_n) = 1 - (1 - b) = b.$$

Similarly, for any $0 < a < b \leq 1$, we have that

$$\chi_{[a,b]}(x) = 1 - \chi_{[0,0]}(x) - \chi_{(0,a)}(x) - \chi_{(b,1)}(x)$$

for all $x \in [0, 1)$ so

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{n} \chi_{[a,b]}(x_n) = 1 - 0 - a - (1 - b) = b - a.$$

Returning to the desired direction, fix an arbitrary continuous periodic function $f : \mathbb{R} \to \mathbb{R}$ with period 1 and let $\epsilon > 0$ be arbitrary. Since f is Riemann integrable, there exists a partition $\mathcal{P} = \{t_k\}_{k=0}^{\ell}$ of $[0,1]$ where

$$0 = t_0 < t_1 < \cdots < t_\ell = 1$$

such that if

$$\begin{aligned} M_k &= \sup(\{f(x) \mid x \in [t_{k-1}, t_k]\}) \quad \text{and} \\ m_k &= \inf(\{f(x) \mid x \in [t_{k-1}, t_k]\}) \end{aligned}$$

for all $k \in \{1, 2, \ldots, n\}$ and

$$\begin{aligned} U(f, \mathcal{P}) &= \sum_{k=1}^{\ell} M_k(t_k - t_{k-1}) \quad \text{and} \\ L(f, \mathcal{P}) &= \sum_{k=1}^{\ell} m_k(t_k - t_{k-1}), \end{aligned}$$

then

$$L(f, \mathcal{P}) \leq \int_0^1 f(x)\,dx \leq U(f, \mathcal{P}) \leq L(f, \mathcal{P}) + \epsilon.$$

Define $f_-, f_+ : [0,1] \to \mathbb{R}$ by

$$\begin{aligned} f_+(x) &= \sum_{k=1}^{\ell} M_k \chi_{[t_{k-1}, t_k]}(x) \quad \text{and} \\ f_-(x) &= \sum_{k=1}^{\ell} m_k \chi_{(t_{k-1}, t_k)}(x) \end{aligned}$$

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

$$f_-(x) \leq f(x) \leq f_+(x)$$

for all $x \in [0,1]$ so

$$\frac{1}{N} \sum_{n=1}^{N} f_-(x_n) \leq \frac{1}{N} \sum_{n=1}^{N} f(x_n) \leq \frac{1}{N} \sum_{n=1}^{N} f_+(x_n)$$

for all $N \in \mathbb{N}$. Since

$$\begin{aligned} \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f_+(x_n) &= \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \sum_{k=1}^{\ell} M_k \chi_{[t_{k-1}, t_k]}(x_n) \\ &= \sum_{k=1}^{\ell} M_k \left(\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \chi_{[t_{k-1}, t_k]}(x_n) \right) \\ &= \sum_{k=1}^{\ell} M_k(t_k - t_{k-1}) \\ &= U(f, \mathcal{P}). \end{aligned}$$

and similarly

$$\begin{aligned}
\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} f_-(x_n) &= \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N}\sum_{k=1}^{\ell} m_k \chi_{(t_{k-1},t_k)}(x_n) \\
&= \sum_{k=1}^{\ell} m_k \left(\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \chi_{(t_{k-1},t_k)}(x_n)\right) \\
&= \sum_{k=1}^{\ell} m_k (t_k - t_{k-1}) \\
&= L(f,\mathcal{P}).
\end{aligned}$$

we obtain that

$$\begin{aligned}
\int_0^1 f(x)\,dx - \epsilon &\leq L(f,\mathcal{P}) \\
&\leq \liminf_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} f(x_n) \\
&\leq \limsup_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} f(x_n) \\
&\leq U(f,\mathcal{P}) \\
&\leq L(f,\mathcal{P}) + \epsilon \\
&\leq \int_0^1 f(x)\,dx + \epsilon.
\end{aligned}$$

Therefore, since $\epsilon > 0$ was arbitrary, we obtain that

$$\int_0^1 f(x)\,dx \leq \liminf_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} f(x_n) \leq \limsup_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} f(x_n) \leq \int_0^1 f(x)\,dx$$

so

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} f(x_n) = \int_0^1 f(x)\,dx.$$

Hence, since f was arbitrary, the proof is complete. ∎

4.3 Solution to the Heat Equation

Fourier series can also be used to solve problems in physics and many other areas. One example of this is the so-called heat equation. Here we have a fixed position object and are looking at its temperature $u(x, y, t)$ at point (x, y) at time t. Using physics, it is possible to show that u satisfies the differential equation

$$\frac{\sigma}{\kappa}\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}$$

where σ constant called specific heat of the material and κ conductivity of the material.

The goal of this section is to solve the steady-state heat equation; that is, when $\frac{\partial u}{\partial t} = 0$. This represents determining the temperature distribution $u(x, y)$ of an object at equilibrium. In particular, our solution will focus on the case of a circle (or narrow disk). Since every point in the open disk

$$D = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < 1\}$$

can be though of in polar coordinates with radius at most 1, we can represent D as $[0, 1) \times \mathbb{T}$. The solutions to the steady-state heat equation then follow from considering the Poisson kernel.

Theorem 4.3.1. *Let $f \in \mathcal{RI}(\mathbb{T})$ and define $u : [0, 1) \times \mathbb{T} \to \mathbb{C}$ by*

$$u(r, \theta) = (f * P_r)(\theta)$$

for all $r \in (0, 1)$ and $\theta \in \mathbb{T}$. Then the following are true:

a) If $\theta \in \mathbb{T}$ is a point of continuity of f, then

$$\lim_{r \nearrow 1} u(r, \theta) = f(\theta).$$

Moreover, if $f \in \mathcal{C}(\mathbb{T})$, then the limit is uniform over $\theta \in \mathbb{T}$.

b) u is twice continuously derivatives in the open unit disc and

$$\Delta(u) = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0.$$

c) If $f \in \mathcal{C}(\mathbb{T})$, then u is the unique solution to the steady-state heat equation in the open unit disc with boundary values f; that is, u is the unique twice continuously differentiable function on the open unit disc such that $\Delta u = 0$ and

$$\lim_{r \nearrow 1} u(r, \theta) = f(\theta)$$

uniformly over all $\theta \in \mathbb{T}$.

Proof. Clearly part a) follows immediately from Theorem 3.6.7. To see that part b) is true, we will be working with polar coordinates.

First recall from the Riemann-Lebesgue Lemma (Theorem 3.5.11) that

$$\lim_{n \to \infty} \widehat{f}(n) = 0 \quad \text{and} \quad \lim_{n \to -\infty} \widehat{f}(n) = 0.$$

Therefore, $(\widehat{f}(n))_{n \in \mathbb{Z}}$ is bounded so there exists an $M \in \mathbb{R}$ such that

$$\left|\widehat{f}(n)\right| \leq M$$

for all $n \in \mathbb{Z}$. Note for all $r \in [0, 1)$ that

$$\sum_{n=-\infty}^{\infty} \left| nr^{|n|-1} \widehat{f}(n) e^{in\theta} \right| \leq \sum_{n=-\infty}^{\infty} Mnr^{|n|-1}$$

converges by a simple application of the Ratio Test (Theorem 1.2.18). Therefore, by using the Weierstrass M-Test (Theorem 2.2.15) and the real and imaginary parts of u, we may to conclude that

$$\sum_{n=-\infty}^{\infty} nr^{|n|-1} \widehat{f}(n) e^{in\theta}$$

converges uniformly over $r \in [0, 1)$, and by using Corollary 2.5.2, we obtain that u is differentiable with respect to r and

$$\frac{\partial u}{\partial r}(r, \theta) = \sum_{n=-\infty}^{\infty} |n| r^{|n|-1} \widehat{f}(n) e^{in\theta}.$$

(note this technically only gives us u is differentiable with respect to r on $(0, 1)$, but this can be extended to $r = 0$ by some simple computations.)

Similarly arguments show that u is twice (actually infinitely) continuously differentiable with

$$\begin{aligned}
\frac{\partial u}{\partial \theta}(r, \theta) &= \sum_{n=-\infty}^{\infty} inr^{|n|} \widehat{f}(n) e^{in\theta} \\
\frac{\partial^2 u}{\partial r^2}(r, \theta) &= \sum_{n=-\infty}^{\infty} |n|(|n| - 1) r^{|n|-2} \widehat{f}(n) e^{in\theta} \qquad \text{and} \\
\frac{\partial^2 u}{\partial \theta^2}(r, \theta) &= \sum_{n=-\infty}^{\infty} -n^2 r^{|n|} \widehat{f}(n) e^{in\theta}.
\end{aligned}$$

Recall by Appendix C.3 that in polar coordinates

$$\Delta u = \frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial u^2}{\partial \theta^2}.$$

Therefore, we obtain that

$$\begin{aligned}
\Delta u &= \sum_{n=-\infty}^{\infty} \left(|n|(|n|-1) r^{|n|-2} + \frac{1}{r}\left(|n| r^{|n|-1} \right) + \frac{1}{r^2}\left(-n^2 r^{|n|} \right) \right) \widehat{f}(n) e^{in\theta} \\
&= \sum_{n=-\infty}^{\infty} \left((|n|^2 - |n|) r^{|n|-2} + |n| r^{|n|-2} - n^2 r^{|n|-2} \right) \widehat{f}(n) e^{in\theta} \\
&= 0
\end{aligned}$$

as desired.

To see that c) is true, suppose $v : [0,1) \times \mathbb{T} \to \mathbb{C}$ is twice continuously differentiable function on the open unit disc such that $\Delta v = 0$ and

$$\lim_{r \nearrow 1} v(r,\theta) = f(\theta)$$

uniformly over $\theta \in \mathbb{T}$. For all $r \in (0,1)$ and $n \in \mathbb{Z}$, let

$$a_n(r) = \frac{1}{2\pi} \int_{-\pi}^{\pi} v(r,\theta) e^{-in\theta}\, d\theta$$

(i.e. the n^{th} Fourier coefficient of $\theta \mapsto v(r,\theta)$). Our goal is to show that

$$a_n(r) = r^{|n|} \widehat{f}(n)$$

for all $n \in \mathbb{Z}$ and $r \in (0,1)$. Indeed since for all $r \in (0,1)$ we know that

$$r^{|n|} \widehat{f}(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} u(r,\theta) e^{-in\theta}\, d\theta$$

by Lemma 3.6.6, the uniqueness of the Fourier coefficients for continuous functions from Corollary 3.5.9 implies that

$$v(r,\theta) = u(r,\theta)$$

for all $\theta \in \mathbb{T}$ and $r \in (0,1)$, which implies $v = u$ by continuity thereby completing the proof.

To prove the desired formula for $a_n(r)$, we will construct and solve second order differential equations using $\Delta v = 0$. Note by the Leibniz Integration Rule (Theorem C.2.1) that a_n is twice continuously differentiable on $(0,1)$. Moreover, since for all $n \in \mathbb{Z}$, $r \in (0,1)$, and $\theta \in \mathbb{T}$ we know that

$$0 = e^{-in\theta} \Delta v(r,\theta) = e^{-in\theta} \left(\frac{\partial^2 v}{\partial r^2} + \frac{1}{r}\frac{\partial v}{\partial r} + \frac{1}{r^2}\frac{\partial v^2}{\partial \theta^2} \right)(r,\theta),$$

by integrating with respect to θ and applying the Leibniz Integration Rule

(Theorem C.2.1) we obtain that

$$\begin{aligned}
0 &= \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-in\theta}\left(\frac{\partial^2 v}{\partial r^2} + \frac{1}{r}\frac{\partial v}{\partial r} + \frac{1}{r^2}\frac{\partial v^2}{\partial \theta^2}\right)(r,\theta)\, d\theta \\
&= \frac{\partial^2}{\partial r^2}\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} v(r,\theta)e^{-in\theta}\, d\theta\right) + \frac{1}{r}\frac{\partial}{\partial r}\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} v(r,\theta)e^{-in\theta}\, d\theta\right) \\
&\quad + \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{1}{r^2}\frac{\partial v^2}{\partial \theta^2}(r,\theta)e^{-in\theta}\, d\theta \\
&= a_n''(r) + \frac{1}{r}a_n'(r) + \frac{1}{2\pi r^2}\int_{-\pi}^{\pi} \frac{\partial v^2}{\partial \theta^2}(r,\theta)e^{-in\theta}\, d\theta \\
&= a_n''(r) + \frac{1}{r}a_n'(r) + \frac{1}{2\pi r^2}\left(0 - \int_{-\pi}^{\pi} \frac{\partial v}{\partial \theta}(r,\theta)(-in)e^{-in\theta}\, d\theta\right) \\
&= a_n''(r) + \frac{1}{r}a_n'(r) - \frac{1}{2\pi r^2}\left(\int_{-\pi}^{\pi} \frac{\partial v}{\partial \theta}(r,\theta)(-in)e^{-in\theta}\, d\theta\right) \\
&= a_n''(r) + \frac{1}{r}a_n'(r) - \frac{1}{2\pi r^2}\left(0 - \int_{-\pi}^{\pi} v(r,\theta)(-n^2)e^{-in\theta}\, d\theta\right) \\
&= a_n''(r) + \frac{1}{r}a_n'(r) - \frac{n^2}{2\pi r^2}\int_{-\pi}^{\pi} v(r,\theta)e^{-in\theta}\, d\theta \\
&= a_n''(r) + \frac{1}{r}a_n'(r) - \frac{n^2}{r^2}a_n(r).
\end{aligned}$$

To solve this second-order differential equation for $a_n(r)$, fix $n \in \mathbb{Z}$ and let $b_n(r) = \frac{a_n(r)}{r^n}$ for all $r \in (0,1)$. Thus b_n is twice continuously differentiable on $(0,1)$ with

$$\begin{aligned}
a_n'(r) &= nr^{n-1}b_n(r) + r^n b_n'(r) \quad \text{and} \\
a_n''(r) &= n(n-1)r^{n-2}b_n(r) + 2nr^{n-1}b_n'(r) + r^n b_n''(r).
\end{aligned}$$

Therefore, we obtain that

$$\begin{aligned}
0 &= \left(n(n-1)r^{n-2}b_n(r) + 2nr^{n-1}b_n'(r) + r^n b_n''(r)\right) \\
&\quad + \frac{1}{r}\left(nr^{n-1}b_n(r) + r^n b_n'(r)\right) - \frac{n^2}{r^2}r^n b_n(r) \\
&= r^n b_n''(r) + (2n+1)r^{n-1}b_n'(r).
\end{aligned}$$

Thus

$$0 = rb_n''(r) + (2n+1)b_n'(r) = (rb_n'(r) + 2nb_n(r))'.$$

Therefore, there must exists a constant $c_n \in \mathbb{C}$ such that

$$rb_n'(r) + 2nb_n(r) = c_n$$

for all $r \in (0,1)$. To solve this differential equation for $b_n(r)$, by breaking the discussion via the real and imaginary parts of b_n, we may assume that b_n is real-valued in the following computations.

In the case $n = 0$, we see that $rb_0'(r) = c_0$ so

$$a_0(r) = b_0(r) = \int b_0'(r)\,dr = \int \frac{c}{r}\,dr = d_0 + c_0\ln(r)$$

for some constant d_0. Otherwise, when $n \neq 0$, let

$$h_n(r) = b_n(r) - \frac{c_n}{2n}.$$

Therefore

$$rh_n'(r) + 2nh_n(r) = 0$$

so

$$\frac{h_n'(r)}{h_n(r)} = -\frac{2n}{r}$$

and thus

$$\ln(h_n(r)) = \int \frac{h_n'(r)}{h_n(r)}\,dt = \int \frac{2n}{r}\,dt = -2n\ln(r) + \ln(d_n) = \ln(d_n r^{-2n})$$

for some constant d_n. Therefore

$$h_n(r) = d_n r^{-2n}$$

so

$$b_n(r) = d_n r^{-2n} + \frac{c_n}{2n}$$

and thus

$$a_n(r) = d_n r^{-n} + \frac{c_n}{2n} r^n.$$

Since v is continuous on the open unit disk, v must be bounded near zero. Therefore, since

$$a_n(r) = \frac{1}{2\pi}\int_{-\pi}^{\pi} v(r,\theta)e^{-in\theta}\,d\theta,$$

we must have that $a_n(r)$ is bounded as r tends to 0 from above. Therefore, by examining our formulae for $a_n(r)$, we must have that $c_0 = 0$ and $d_n = 0$ for all $n \in \mathbb{Z}\setminus\{0\}$ so that $a_0(r) = s_0$ and

$$a_n(r) = s_n r^{|n|}$$

for all $n \in \mathbb{Z}\setminus\{0\}$ where s_n are constants. However, since

$$\lim_{r\nearrow 1} v(r,\theta) = f(\theta)$$

uniformly over $\theta \in \mathbb{T}$, we obtain for all $n \in \mathbb{Z}$ that

$$\begin{aligned}
s_n &= \lim_{r \nearrow 1} s_n r^{|n|} \\
&= \lim_{r \nearrow 1} a_n(r) \\
&= \lim_{r \nearrow 1} \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} v(r,\theta) e^{-in\theta} \, d\theta \right) \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(\lim_{r \nearrow 1} v(r,\theta) \right) e^{-in\theta} \, d\theta & \text{by Theorem 2.4.4} \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\theta) e^{-in\theta} \, d\theta \\
&= \widehat{f}(n).
\end{aligned}$$

Hence

$$a_n(r) = r^{|n|} \widehat{f}(n)$$

for all $n \in \mathbb{Z}$ thereby completing the proof by the above discussions. ∎

To conclude this section, we point out that Theorem 4.3.1 also is of importance in complex analysis. Indeed u represents the unique harmonic function on the closed unit disk with boundary values f, and the study of harmonic functions is vital to complex analysis.

4.4 The Riemann Zeta Function

To conclude our applications chapter, we look at one instance of series that is not related to Fourier series, but one of the most important series in not only number theory, but pure mathematics:

Definition 4.4.1. The *Riemann zeta function (on* $(1,\infty)$*)* is the function $\zeta : (1,\infty) \to \mathbb{R}$ defined by

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$$

for all $s \in (1,\infty)$.

Remark 4.4.2. It is elementary to see that $\zeta(s)$ is well-defined for all $s \in (1,\infty)$ as the p-test (Corollary 1.2.15) implies the series $\sum_{n=1}^{\infty} \frac{1}{n^s}$ converges absolutely for all $s \in (1,\infty)$.

In fact, ζ can be extended to a large portion of the complex plane. Indeed it is not difficult to extend ζ to all $s \in \mathbb{C}$ with $\mathrm{Re}(s) > 1$ via the same series definition. To see this, note for all $s = a + bi$ with $a, b \in \mathbb{R}$ and $a > 1$ that

$$\frac{1}{n^s} = \frac{1}{n^a n^{bi}} = \frac{1}{n^a} e^{-\ln(n^{bi})} = \frac{1}{n^a} e^{-ib\ln(n)}$$

(modulo understanding complex logarithms), so

$$\sum_{n=1}^{\infty} \left| \frac{1}{n^s} \right| = \sum_{n=1}^{\infty} \frac{1}{n^a}$$

converges absolutely and thus $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$ converges for all $s \in \mathbb{C}$ with $\mathrm{Re}(s) > 1$.

For our next extension, notice for all $s \in \mathbb{C}$ with $\mathrm{Re}(s) > 1$ that

$$\begin{aligned}
(1 - 2^{1-s})\zeta(s) &= \sum_{n=1}^{\infty} \frac{1}{n^s} - \frac{2}{(2n)^s} \\
&= \sum_{n=1}^{\infty} \frac{1}{n^s} - \sum_{n=1}^{\infty} \frac{2}{(2n)^s} \\
&= \left(\sum_{n=1}^{\infty} \frac{1}{(2n-1)^s} + \sum_{n=1}^{\infty} \frac{1}{(2n)^s} \right) - \sum_{n=1}^{\infty} \frac{2}{(2n)^s} \\
&= \sum_{n=1}^{\infty} \frac{1}{(2n-1)^s} - \sum_{n=1}^{\infty} \frac{1}{(2n)^s} \\
&= \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^s}
\end{aligned}$$

(since all series involved converge absolutely by the same arguments as above). Therefore

$$\zeta(s) = \frac{1}{1 - 2^{1-s}} \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^s}$$

for all $s \in \mathbb{C}$ with $\mathrm{Re}(s) > 1$. It is not difficult to see that this definition of ζ extends to all $s \in \mathbb{R}$ with $0 < s < 1$ by the Alternating Series Test (Theorem 1.2.22).

It is more complicated to extend ζ to all $s \in \mathbb{C}$ with $0 < \mathrm{Re}(s) < 1$ as the above series will not be 'alternating'. As this requires knowledge of analytic extensions from complex analysis, we will not pursue this. However, with some knowledge of Fourier series on $\mathbb{R}$ (another topic we do not have time for), it is possible to show *Riemann's functional equation*

$$\zeta(s) = 2^s \pi^{s-1} \sin\left(\frac{\pi s}{2}\right) \Gamma(1-s)\zeta(1-s)$$

where Γ is the *Gamma function* defined by

$$\Gamma(s) = \int_0^{\infty} x^{s-1} e^{-x} \, dx.$$

(In fact, $\Gamma(n) = (n-1)!$). Modulo complex analysis, this lets one extend ζ to all $s \in \mathbb{C}$ with $\mathrm{Re}(s) < 0$.

Of course, our focus on the Riemann zeta function will not lie in these complex analysis results, but the connection between the Riemann zeta function and prime numbers. Unfortunately, we will not go so far as to study the Riemann Hypothesis (that is, the question of whether or not all zeros of ζ in the region $0 < \mathrm{Re}(s) < 1$ occur on the line $\mathrm{Re}(s) = \frac{1}{2}$, which has implications for quantum computing and factorization of primes), but to establish some elementary facts. To do this, we must first discuss infinite products.

Definition 4.4.3. Let $(a_n)_{n\geq 1}$ be a sequence of positive real numbers and for each $N \in \mathbb{N}$ let

$$P_N = a_1 a_2 \cdots a_N = \prod_{k=1}^{N} a_k.$$

The infinite product $\prod_{n=1}^{\infty} a_n$ is said to *converge to* L if $(P_N)_{N\geq 1}$ converges to L.

The most basic connection between the Riemann zeta function and prime numbers can be seen via the following result.

Theorem 4.4.4 (Euler's Product Formula). *For all* $s \in (1, \infty)$*,*

$$\zeta(s) = \prod_{p \text{ a prime}} \frac{1}{1 - \frac{1}{p^s}}.$$

Proof. Fix $s \in (1, \infty)$. Recall by the Fundamental Theorem of Arithmetic that every $n \in \mathbb{N}$ has a unique factorization into a product of powers of prime numbers. Therefore for all $N, M \in \mathbb{N}$, we have that

$$\prod_{\substack{p \text{ a prime} \\ p \leq N}} \left(\sum_{k=0}^{M} \left(\frac{1}{p^k} \right)^s \right) \leq \sum_{n=1}^{\infty} \frac{1}{n^s}.$$

Since $0 \leq \frac{1}{p^s} < 1$ for all primes p, we have via geometric series that

$$\frac{1}{1 - \frac{1}{p^s}} = \sum_{k=0}^{\infty} \left(\frac{1}{p^k} \right)^s$$

and every term in the series is positive. Therefore

$$\prod_{\substack{p \text{ a prime} \\ p \leq N}} \frac{1}{1 - \frac{1}{p^s}} \leq \sum_{n=1}^{\infty} \frac{1}{n^s}$$

for all $N \in \mathbb{N}$. However, notice that if $n \in \mathbb{N}$ and $n \leq N$, then n is a product of powers of primes that are all at most N. Therefore

$$\begin{aligned}\sum_{n=1}^{N} \frac{1}{n^s} &\leq \prod_{\substack{p \text{ a prime} \\ p \leq N}} \sum_{k=0}^{\infty} \left(\frac{1}{p^k}\right)^s \\ &= \prod_{\substack{p \text{ a prime} \\ p \leq N}} \frac{1}{1 - \frac{1}{p^s}} \leq \sum_{n=1}^{\infty} \frac{1}{n^s}.\end{aligned}$$

Therefore, by taking the limit as N tends to infinity, we obtain that

$$\prod_{p \text{ a prime}} \frac{1}{1 - \frac{1}{p^s}}$$

converges and is equal to $\zeta(s)$ as desired. ∎

To finish off the course, we note the following series diverges, which implies there are a lot of primes.

Corollary 4.4.5. *The series*

$$\sum_{p \text{ a prime}} \frac{1}{p}$$

diverges.

Proof. We will show for any $N \in \mathbb{N}$ that

$$\sum_{\substack{p \text{ a prime} \\ p \leq N}} \frac{1}{p} > \ln(\ln(N)) - \frac{1}{2}$$

from which the result clearly follows.

To begin, notice that

$$
\begin{aligned}
\ln\left(\prod_{\substack{p \text{ a prime}\\ p\leq N}} \frac{1}{1-\frac{1}{p}}\right) &= \sum_{\substack{p \text{ a prime}\\ p\leq N}} -\ln\left(1-\frac{1}{p}\right) \\
&= \sum_{\substack{p \text{ a prime}\\ p\leq N}} \sum_{k=1}^{\infty} \frac{1}{k}\left(\frac{1}{p}\right)^k && \text{by Remark 2.4.6} \\
&= \sum_{\substack{p \text{ a prime}\\ p\leq N}} \frac{1}{p} + \sum_{k=2}^{\infty} \frac{1}{k}\left(\frac{1}{p}\right)^k \\
&< \sum_{\substack{p \text{ a prime}\\ p\leq N}} \frac{1}{p} + \frac{1}{2}\sum_{k=2}^{\infty} \frac{1}{p^k} \\
&= \sum_{\substack{p \text{ a prime}\\ p\leq N}} \frac{1}{p} + \frac{1}{2}\frac{\frac{1}{p^2}}{1-\frac{1}{p}} \\
&= \left(\sum_{\substack{p \text{ a prime}\\ p\leq N}} \frac{1}{p}\right) + \frac{1}{2}\sum_{\substack{p \text{ a prime}\\ p\leq N}} \frac{1}{p(p-1)} \\
&< \left(\sum_{\substack{p \text{ a prime}\\ p\leq N}} \frac{1}{p}\right) + \frac{1}{2}\sum_{n=2}^{N} \frac{1}{n(n-1)} \\
&= \left(\sum_{\substack{p \text{ a prime}\\ p\leq N}} \frac{1}{p}\right) + \frac{1}{2}\sum_{n=2}^{N} \frac{1}{n-1} - \frac{1}{n} \\
&= \left(\sum_{\substack{p \text{ a prime}\\ p\leq N}} \frac{1}{p}\right) + \frac{1}{2}\left(1-\frac{1}{N}\right) \\
&< \left(\sum_{\substack{p \text{ a prime}\\ p\leq N}} \frac{1}{p}\right) + \frac{1}{2}.
\end{aligned}
$$

Moreover, by the same ideas as used in the proof of Theorem 4.4.4, we have

that

$$\begin{aligned}
\prod_{\substack{p \text{ a prime} \\ p \leq N}} \frac{1}{1-\frac{1}{p}} &= \prod_{\substack{p \text{ a prime} \\ p \leq N}} \sum_{k=0}^{\infty} \left(\frac{1}{p}\right)^k && \text{geometric series} \\
&\geq \sum_{k=1}^{N} \frac{1}{k} && \text{Fundamental Theorem of Arithmetic} \\
&> \sum_{k=1}^{N} \int_k^{k+1} \frac{1}{x}\, dx \\
&= \int_1^{N+1} \frac{1}{x}\, dx \\
&> \int_1^{N} \frac{1}{x}\, dx \\
&= \ln(N).
\end{aligned}$$

Combining these two, we have due to the fact that the natural logarithm is increasing on $(0, \infty)$ that

$$\sum_{\substack{p \text{ a prime} \\ p \leq N}} \frac{1}{p} > \ln(\ln(N)) - \frac{1}{2}$$

as desired. ∎

Appendix A

Complex Numbers

In this appendix chapter, we will quickly review the basics of complex numbers.

Definition A.0.1. The *complex numbers*, denoted $\mathbb{C}$, are the set

$$\mathbb{C} = \{a + bi \mid a, b \in \mathbb{R}\}$$

where i denotes a fixed symbol.

Given a complex number $z = a + bi$ where $a, b \in \mathbb{R}$, the number a is called *real part of* z and is denoted by$\operatorname{Re}(z)$, and the number b is called the *imaginary part of* z and is denoted by $\operatorname{Im}(z)$.

The symbol i in a complex number is meant to denote "$\sqrt{-1}$". To be specific, we will equip $\mathbb{C}$ with binary operations of addition and multiplication so that $(0+1i)(0+1i) = -1$ so that indeed i is a complex solution to $x^2 = -1$.

Definition A.0.2. The binary operations $+ : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$ and $\cdot : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$ defined by

$$(a+bi)+(c+di) = (a+c)+(b+d)i \quad \text{and} \quad (a+bi)(c+di) = (ac-bd)+(ad+bc)i$$

for all $a, b, c, d \in \mathbb{R}$ are called *complex addition* and *complex multiplication* respectively.

Example A.0.3. It is not difficult to see that

$$(1 + 2i) + (3 + 4i) = 4 + 6i \qquad \text{and} \qquad (1 + 2i)(3 + 4i) = -5 + 10i.$$

Moreover, since

$$i^2 = (0 + 1i)(0 + 1i) = -1 + 0i,$$

we do indeed have that i is a complex solution to $x^2 = -1$. In addition, it is not difficult to see that $-i$ is also a complex solution to $x^2 = -1$.

In order for $\mathbb{C}$ to be as nice to work with as $\mathbb{R}$, we require complex addition and multiplication to have specific properties. To be specific, we want the following.

Theorem A.0.4. *The set of complex numbers $\mathbb{C}$ together with complex addition and multiplication is a* field*; that is,*

(i) (Commutativity of Addition) $z + w = w + z$ for all $z, w \in \mathbb{C}$.

(ii) (Associativity of Addition) $z + (w + u) = (z + w) + u$ for all $z, w, u \in \mathbb{C}$.

(iii) (Additive Unit) There exists a $0 \in \mathbb{C}$ such that $z + 0 = z$ for all $z \in \mathbb{C}$ (i.e. $0 = 0 + 0i$).

(iv) (Additive Inverses) For all $z \in \mathbb{C}$ there exists a $-z \in \mathbb{C}$ such that $z + (-z) = 0$ (i.e. $-(a + bi) = (-a) + (-b)i$).

(v) (Commutativity of Multiplication) $zw = wz$ for all $z, w \in \mathbb{C}$.

(vi) (Associativity of Multiplication) $z(wu) = (zw)u$ for all $z, w, u \in \mathbb{C}$.

(vii) (Multiplicative Unit) There exists a $1 \in \mathbb{C}$ such that $1z = z$ for all $z \in \mathbb{C}$ (i.e. $1 = 1 + 0i$).

(viii) (Multiplicative Inverses) For all $z \in \mathbb{C} \setminus \{0\}$ there exists a $z^{-1} \in \mathbb{C}$ such that $z^{-1}z = 1$.

(ix) (Distributivity) $z(w + u) = (zw) + (zu)$ for all $z, w, u \in \mathbb{C}$.

Proof. Let $z, w, u \in \mathbb{C}$ be arbitrary. Hence there exists $a, b, c, d, x, y \in \mathbb{R}$ such that

$$z = a + bi, \quad w = c + di, \quad \text{and} \quad u = x + yi.$$

We will now examine each of the above nine properties for these arbitrary elements of $\mathbb{C}$ and demonstrate the property holds using the analogous property for real numbers.

(i) Commutativity of Addition: Notice that

$$z + w = (a + c) + (b + d)i = (c + a) + (d + b)i = w + z$$

due to the commutativity of addition of real numbers. Thus commutativity of addition of complex numbers has been demonstrated.

(ii) Associativity of Addition: Notice that

$$\begin{aligned} z + (w + u) &= (a + bi) + ((c + x) + (d + y)i) \\ &= (a + (c + x)) + (b + (d + y))i \\ &= ((a + c) + x) + ((b + d) + y)i \\ &= ((a + c) + (b + d)i) + (x + yi) \\ &= (z + w) + u \end{aligned}$$

where the third equality holds due to the associativity of addition of real numbers. Thus associativity of addition of complex numbers has been demonstrated.

(iii) Additive Unit: Notice with $0 = 0 + 0i$ that

$$z + 0 = (a + 0) + (b + 0)i = a + bi = z$$

due to the property of the zero element of $\mathbb{R}$. Thus the complex numbers have an additive unit.

(iv) Additive Inverses: Let $-z = (-a) + (-b)i$ where $-a$ and $-b$ are the additive inverses of a and b in $\mathbb{R}$. then

$$z + (-z) = (a + (-a)) + (b + (-b))i = 0 + 0i = 0$$

as desired. Thus the complex numbers have additive inverses.

(v) Commutativity of Multiplication: Notice that

$$zw = (ac - bd) + (ad + bc)i = (ca - db) + (cb + da)i = wz$$

due to the commutativity of addition and multiplication of real numbers. Thus commutativity of multiplication of complex numbers has been demonstrated.

(vi) Associativity of Multiplication: Notice that

$$\begin{aligned}
z(wu) &= (a + bi)((c + di)(x + yi)) \\
&= (a + bi)((cx - dy) + (cy + dx)i) \\
&= (a(cx - dy) - b(cy + dx)) + (a(cy + dx) + b(cx - dy))i \\
&= (acx - ady - bcy - bdx) + (acy + adx + bcx - bdy)i \\
&= (acx - bdx - ady - bcy) + (acy - bdy + adc + bcx)i \\
&= ((ac - bd)x - (ad + bc)y) + ((ac - bd)y + (ad + bc)x)i \\
&= ((ac - bd) + (ad + bc)i)(x + yi) \\
&= ((a + bi)(c + di))(x + yi) \\
&= (zw)u
\end{aligned}$$

where commutativity and associativity of addition (fifth equality) and distributivity of real numbers (fourth and sixth equalities) have been used. Thus associativity of multiplication of complex numbers has been demonstrated.

(vii) Multiplicative Unit: Notice with $1 = 1 + 0i$ that

$$1z = (1a - 0b) + (1b + 0a)i = a + bi = z$$

due to the property of the one element of $\mathbb{R}$. Thus the complex numbers have a multiplicative unit.

(viii) Multiplicative Inverses: Assume $z \neq 0$. Thus $a \neq 0$ or $b \neq 0$ so $a^2 + b^2 > 0$. Define

$$z^{-1} = \frac{a}{a^2+b^2} + \frac{-b}{a^2+b^2} i,$$

which makes sense since $\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2} \in \mathbb{R}$ since non-zero real numbers have multiplicative inverses. Notice that

$$\begin{aligned} z^{-1}z &= \left(\left(\frac{a}{a^2+b^2}\right)a - \left(\frac{-b}{a^2+b^2}\right)b\right) + \left(\left(\frac{a}{a^2+b^2}\right)b + \left(\frac{-b}{a^2+b^2}\right)a\right)i \\ &= \left(\frac{a^2+b^2}{a^2+b^2}\right) + \left(\frac{ab-ba}{a^2+b^2}\right)i \\ &= 1 + 0i = 1 \end{aligned}$$

due to the commutative of multiplication and properties of addition of real numbers. Hence z^{-1} is indeed the multiplicative inverses of z. Thus the existence of multiplicative inverses for non-zero complex numbers has been demonstrated.

(ix) Distributivity: Notice that

$$\begin{aligned} z(w+u) &= (a+bi)((c+di)+(x+yi)) \\ &= (a+bi)((c+x)+(d+y)i) \\ &= (a(c+x)-b(d+y)) + (a(d+y)+b(c+x))i \\ &= (ac+ax-bd-dy) + (ad+ay+bc+bx)i \\ &= ((ac-bd)+(ax-dy)) + ((ad+bc)+(ay+bx))i \\ &= ((ac-bd)+(ad+bc)i) + ((ax-dy)+(ay+bx)i) \\ &= ((a+bi)(c+di)) + ((a+bi)(x+yi)) \\ &= (zw) + (zu) \end{aligned}$$

where commutativity and associativity of addition (fourth equality) and distributivity of real numbers (third equality) have been used. Thus distributivity of complex numbers has been demonstrated.

As all nine properties have now been demonstrated for arbitrary complex numbers, the set of complex numbers together with addition and multiplication are a field. ∎

Remark A.0.5. Embedded in the proof of Theorem A.0.4 is the formula for the inverse of a non-zero complex number. Indeed if $z = a + bi$ where $a, b \in \mathbb{R}$ is such that $z \neq 0$, then

$$z^{-1} = \frac{a}{a^2+b^2} + \frac{-b}{a^2+b^2} i.$$

For example,

$$(3+4i)^{-1} = \frac{3}{25} - \frac{4}{25} i.$$

Note we can also write

$$z^{-1} = \frac{a - bi}{a^2 + b^2}.$$

When dealing with complex numbers, both the numerator and denominator are important common quantities that are worthy of developing further.

Definition A.0.6. Given a complex number $z = a + bi$ where $a, b \in \mathbb{R}$, the *absolute value* (or *length* or *modulus*) of z, denoted $|z|$, is the quantity

$$|z| = \sqrt{a^2 + b^2}.$$

Definition A.0.7. Given a complex number $z = a + bi$ where $a, b \in \mathbb{R}$, the *complex conjugate of* z, denoted $\overline{z}$, is the quantity

$$\overline{z} = a + (-b)i.$$

Using our knowledge of the inverse of a non-zero complex number, we have the following.

Corollary A.0.8. *If $z \in \mathbb{C} \setminus \{0\}$, then $z^{-1} = \frac{1}{|z|^2}\overline{z}$.*

There are many other properties of the absolute value and complex conjugates that are worthy of recording.

Proposition A.0.9. *For all $z, w \in \mathbb{C}$, the following are true:*

a) $\mathrm{Re}(z) = \frac{z+\overline{z}}{2}$.

b) $\mathrm{Im}(z) = \frac{z-\overline{z}}{2i}$.

c) $\overline{z + w} = \overline{z} + \overline{w}$.

d) $\overline{zw} = \overline{z}\,\overline{w}$.

e) $|z| = \sqrt{\overline{z}z}$.

f) $|\mathrm{Re}(z)| \leq |z|$.

g) $|\mathrm{Im}(z)| \leq |z|$.

h) $|\overline{z}| = |z|$.

i) $|zw| = |z||w|$.

j) *If $z \neq 0$, then* $|z^{-1}| = |z|^{-1}$.

k) $|z + w| \leq |z| + |w|$.

l) $||z| - |w|| \leq |z - w|$.

Proof. Let $z, w \in \mathbb{C}$ be arbitrary. Thus there exists $a, b, c, d \in \mathbb{R}$ such that $z = a + bi$ and $w = c + di$.

To see that a) is true, notice

$$\frac{z + \overline{z}}{2} = \frac{(a + bi) + (a - bi)}{2} = \frac{2a}{2} = a = \mathrm{Re}(z)$$

as desired.

To see that b) is true, notice

$$\frac{z - \overline{z}}{2i} = \frac{(a + bi) - (a - bi)}{2i} = \frac{2bi}{2i} = b = \mathrm{Im}(z)$$

as desired.

To see that c) is true, notice

$$\begin{aligned}\overline{z + w} &= \overline{(a + c) + (b + d)i} \\ &= (a + c) + (-(b + d))i \\ &= (a + (-b)i) + (c + (-d)i) \\ &= \overline{z} + \overline{w}.\end{aligned}$$

as desired.

To see that d) is true, notice that

$$\begin{aligned}\overline{zw} &= \overline{(ac - bd) + (ad + bc)i} \\ &= (ac - bd) + (-(ad + bc))i \\ &= (ac - (-b)(-d)) + (a(-d) + (-b)c)i \\ &= (a + (-b)i)(c + (-d)i) \\ &= \overline{z}\,\overline{w}\end{aligned}$$

as desired.

To see that e) is true, notice that

$$\begin{aligned}\sqrt{\overline{z}z} &= \sqrt{(a + (-b)i)(a + bi)} \\ &= \sqrt{(a^2 - (-b)b) + (ab + a(-b))i} \\ &= \sqrt{a^2 + b^2} = |z|\end{aligned}$$

as desired.

To see that f) and g) are true, note that

$$a^2 \leq a^2 + b^2 \qquad \text{and} \qquad b^2 \leq a^2 + b^2,$$

so

$$|\mathrm{Re}(z)| = |a| = \sqrt{a^2} \leq \sqrt{a^2 + b^2} = |z|$$

and

$$|\text{Im}(z)| = |b| = \sqrt{b^2} \leq \sqrt{a^2 + b^2} = |z|$$

as desired.

To see that h) is true, note that

$$|\overline{z}| = |a + (-b)i| = \sqrt{a^2 + (-b)^2} = \sqrt{a^2 + b^2} = |z|$$

as desired.

To see that i) is true, notice that

$$zw = (ac - bd) + (ad + bc)i$$

so that

$$\begin{aligned}
|zw| &= \sqrt{(ac - bd)^2 + (ad + bc)^2} \\
&= \sqrt{(a^2c^2 - 2abcd + b^2d^2) + (a^2d^2 + 2abcd + b^2c^2)} \\
&= \sqrt{a^2c^2 + b^2d^2 + a^2d^2 + b^2c^2} \\
&= \sqrt{(a^2 + b^2)(c^2 + d^2)} \\
&= \sqrt{(a^2 + b^2)}\sqrt{(c^2 + d^2)} \\
&= |z||w|
\end{aligned}$$

as desired.

To see that j) is true, notice by i) that

$$|z^{-1}||z| = |z^{-1}z| = |1| = 1,$$

so $|z^{-1}| = |z|^{-1}$ as desired.

To see that k) is true, notice that

$$\begin{aligned}
|z + w|^2 &= (\overline{z + w})\,(z + w) && \text{by e)} \\
&= (\overline{z} + \overline{w})\,(z + w) && \text{by c)} \\
&= \overline{z}z + \overline{z}w + \overline{w}z + \overline{w}w && \\
&= |z|^2 + \overline{z}w + \overline{\overline{z}w} + |w|^2 && \text{by d) and e)} \\
&= |z|^2 + 2\text{Re}(\overline{z}w) + |w|^2 && \text{by a)} \\
&\leq |z|^2 + 2|\overline{z}w| + |w|^2 && \text{by f)} \\
&= |z|^2 + 2|\overline{z}||w| + |w|^2 && \text{by i)} \\
&= |z|^2 + 2|z||w| + |w|^2 && \text{by h)} \\
&= (|z| + |w|)^2.
\end{aligned}$$

Therefore, by taking the square root of both sides of the inequality, the desired result is obtained.

Finally, to see that l) is true, first notice by k) that

$$|z| = |(z - w) + w| \leq |z - w| + |w|.$$

Therefore

$$|z| - |w| \leq |z - w|.$$

Similarly, notice by k) and i) that

$$\begin{aligned} |w| &= |(w - z) + z| \\ &\leq |w - z| + |z| \\ &= |(-1)(z - w)| + |z| \\ &= |-1||z - w| + |z| = |z - w| + |z|. \end{aligned}$$

Hence

$$|w| - |z| \leq |z - w|.$$

Therefore, by combining $|z| - |w| \leq |z - w|$ and $|w| - |z| \leq |z - w|$, we obtain that

$$||z| - |w|| \leq |z - w|$$

as desired. ∎

Appendix B

Differentiation and Integration in $\mathbb{C}$

In this section, we will extend the notions of derivatives and integrals of real-valued functions on closed intervals to complex-valued functions. In particular, it will be demonstrated that every elementary notion for derivative and integrals seen in previous courses extend to these complex-valued functions with ease.

B.1 Complex Differentiation

As is natural with calculus, we begin with differentiation. To simplify notation, we introduce the following.

Notation B.1.1. For $a, b \in \mathbb{R}$ with $a < b$, the vector space of all complex-valued functions on $[a, b]$ is denoted by $\mathcal{F}([a, b], \mathbb{C})$. Recall if $f \in \mathcal{F}([a, b], \mathbb{C})$, then the real and imaginary parts of f are the functions $\text{Re}(f), \text{Im}(f) : [a, b] \to \mathbb{R}$ defined by

$$\text{Re}(f)(x) = \text{Re}(f(x)) \qquad \text{and} \qquad \text{Im}(f)(x) = \text{Im}(f(x))$$

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

Definition B.1.2. Let $f \in \mathcal{F}([a, b], \mathbb{C})$ and let $x_0 \in (a, b)$. It is said that f is *differentiable at* x_0 if

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

exists in $\mathbb{C}$. When f is differentiable at x_0, the *derivative of* f *at* x_0, denoted $f'(x_0)$, is

$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} \in \mathbb{C}.$$

Finally, it is said that f is *differentiable on* $[a, b]$ if f is differentiable at each point in (a, b) and f is continuous on $[a, b]$. The *derivative of* f *on* (a, b) is the function $f' : (a, b) \to \mathbb{C}$ whose value at $x_0 \in (a, b)$ is $f'(x_0)$ as defined above.

Of course, it is possible to develop the properties of derivatives of complex-valued functions on intervals by mirroring the corresponding real-valued results. However, as limits of complex numbers converge if and only if their real and imaginary parts converge, the theory of derivatives of complex-valued functions on intervals reduces down to the theory of derivatives of real-valued functions on intervals.

Theorem B.1.3. *Let* $f \in \mathcal{F}([a, b], \mathbb{C})$*, let* $f_1 = \mathrm{Re}(f)$*, let* $f_2 = \mathrm{Im}(f)$*, and let* $x_0 \in (a, b)$*. Then* f *is differentiable at* x_0 *if and only if* f_1 *and* f_2 *are differentiable at* x_0*. When* f *is differentiable at* x_0*,*

$$f'(x_0) = f_1'(x_0) + i f_2'(x_0).$$

Finally, f *is differentiable on* $[a, b]$ *if and only if* f_1 *and* f_2 *are differentiable on* $[a, b]$*, in which case* $f' = f_1' + i f_2'$*.*

Proof. Since for all $x_0 \in (a, b)$ and $h \in \mathbb{R}$ with $x_0 + h \in [a, b]$ we have that

$$\frac{f(x_0 + h) - f(x_0)}{h} = \frac{f_1(x_0 + h) - f_1(x_0)}{h} + i\frac{f_2(x_0 + h) - f_2(x_0)}{h},$$

it follows that

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

exists if and only if

$$\lim_{h \to 0} \frac{f_1(x_0 + h) - f_1(x_0)}{h} \quad \text{and} \quad \lim_{h \to 0} \frac{f_2(x_0 + h) - f_2(x_0)}{h}$$

exist. Moreover, this implies that

$$f'(x_0) = f_1'(x_0) + i f_2'(x_0).$$

Finally, since f is continuous if and only if f_1 and f_2 are continuous, the result follows. ∎

It is worthwhile to see that one specific function behaves incredibly well with respect to this definition of differentiation thereby further supporting why this definition of differentiation for complex-valued functions on intervals is desirable.

Example B.1.4. Let $a, b \in \mathbb{R}$, let $\alpha = a + bi \in \mathbb{C}$, and define $f : \mathbb{R} \to \mathbb{C}$ by

$$f(x) = e^{\alpha x} = e^{(a+bi)x} = e^{ax}\cos(bx) + ie^{ax}\sin(bx)$$

for all $x \in \mathbb{R}$. Then f is differentiable at each point in $\mathbb{R}$ with

$$\begin{aligned} f'(x) &= (e^{ax}\cos(bx))' + i\,(e^{ax}\sin(bx))' \\ &= (ae^{ax}\cos(bx) - be^{ax}\sin(bx)) + i\,(ae^{ax}\sin(bx) + be^{ax}\cos(bx)) \\ &= (a+bi)e^{ax}\cos(bx) + (-b+ia)e^{ax}\sin(bx) \\ &= (a+bi)e^{ax}\cos(bx) + i(a+bi)e^{ax}\sin(bx) \\ &= (a+bi)\left(e^{ax}\cos(bx) + ie^{ax}\sin(bx)\right) \\ &= \alpha f(x). \end{aligned}$$

Hence

$$(e^{\alpha x})' = \alpha e^{\alpha x}$$

for all $\alpha \in \mathbb{C}$.

Of course, some results, such as the following, immediately import from the theory of derivatives of real-valued functions.

Corollary B.1.5. *Let $f \in \mathcal{F}([a, b], \mathbb{C})$ and let $x_0 \in (a, b)$. If f is differentiable at x_0, then f is continuous at x_0.*

Proof. Since f is differentiable at x_0, $\mathrm{Re}(f)$ and $\mathrm{Im}(f)$ are differentiable at x_0. Therefore $\mathrm{Re}(f)$ and $\mathrm{Im}(f)$ are continuous at x_0 so $f = \mathrm{Re}(f) + i\mathrm{Im}(f)$ is continuous at x_0. ∎

Again using the results for derivatives of real-valued functions, certain operations behave well with respect to differentiation.

Corollary B.1.6. *Let $f, g \in \mathcal{F}([a, b], \mathbb{C})$ and let $x_0 \in (a, b)$. If f and g are differentiable at x_0, then $f + g$ is differentiable at x_0 with*

$$(f + g)'(x_0) = f'(x_0) + g'(x_0).$$

Proof. Let $f_1 = \mathrm{Re}(f)$, $f_2 = \mathrm{Im}(f)$, $g_1 = \mathrm{Re}(g)$, and $g_2 = \mathrm{Im}(g)$. Since

$$f + g = (f_1 + if_2) + (g_1 + ig_2) = (f_1 + g_1) + i(f_2 + g_2)$$

we have that

$$\mathrm{Re}(f + g) = f_1 + g_1 \qquad \text{and} \qquad \mathrm{Im}(f + g) = f_2 + g_2.$$

Therefore, since f and g are differential at x_0, we know that f_1, f_2, g_1, and g_2 are differentiable at x_0 and thus $\mathrm{Re}(f + g)$ and $\mathrm{Im}(f + g)$ are differentiable at x_0 by results from the real case with

$$\begin{aligned}(f+g)'(x_0) &= (f_1+g_1)'(x_0) + i(f_2+g_2)'(x_0)\\ &= (f_1'(x_0)+g_1'(x_0)) + i(f_2'(x_0)+g_2'(x_0))\\ &= (f_1'(x_0)+if_2'(x_0)) + (g_1'(x_0)+ig_2'(x_0))\\ &= f'(x_0)+g'(x_0)\end{aligned}$$

as desired. ∎

To see that scalar multiplication behaves well with respect to differentiation, it is easier to generalize the product rule first.

Theorem B.1.7 (Product Rule). *Let $f, g \in \mathcal{F}([a,b],\mathbb{C})$ and let $x_0 \in (a,b)$. If f and g are differentiable at x_0, then fg is differentiable at x_0 with*

$$(fg)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0).$$

Proof. Let $f_1 = \mathrm{Re}(f)$, $f_2 = \mathrm{Im}(f)$, $g_1 = \mathrm{Re}(g)$, and $g_2 = \mathrm{Im}(g)$. Since

$$fg = (f_1+if_2)(g_1+ig_2) = (f_1g_1 - f_2g_2) + i(f_1g_2+f_2g_1)$$

we have that

$$\mathrm{Re}(fg) = f_1g_1 - f_2g_2 \qquad \text{and} \qquad \mathrm{Im}(fg) = f_1g_2+f_2g_1.$$

Therefore, since f and g are differential at x_0, we know that f_1, f_2, g_1, and g_2 are differentiable at x_0 and thus $\mathrm{Re}(fg)$ and $\mathrm{Im}(fg)$ are differentiable at x_0 by the real-valued product rule with

$$\begin{aligned}(fg)'(x_0) &= (f_1g_1-f_2g_2)'(x_0) + i(f_1g_2+f_2g_1)'(x_0)\\ &= ((f_1g_1)'(x_0) - (f_2g_2)'(x_0)) + i\left((f_1g_2)'(x_0) + (f_2g_1)'(x_0)\right)\\ &= ((f_1'(x_0)g_1(x_0) + f_1(x_0)g_1'(x_0)) - (f_2'(x_0)g_2(x_0) + f_2(x_0)g_2'(x_0))\\ &\qquad + i\left((f_1'(x_0)g_2(x_0) + f_1(x_0)g_2'(x_0)) + (f_2'(x_0)g_1(x_0)+f_2(x_0)g_1'(x_0)\right)\\ &= f_1'(x_0)\left(g_1(x_0)+ig_2(x_0)\right) + f_2'(x_0)\left(-g_2(x_0)+ig_1(x_0)\right)\\ &\qquad + \left(f_1(x_0)+if_2(x_0)\right)g_1'(x_0) + \left(-f_2(x_0)+if_1(x_0)\right)g_2'(x_0)\\ &= f_1'(x_0)g(x_0) + if_2'(x_0)g(x_0) + f(x_0)g_1'(x_0) + if(x_0)g_2'(x_0)\\ &= (f_1'(x_0)+if_2'(x_0))g(x_0) + f(x_0)(g_1'(x_0)+ig_2'(x_0))\\ &= f'(x_0)g(x_0)+f(x_0)g'(x_0)\end{aligned}$$

as desired. ∎

Corollary B.1.8. *Let $f \in \mathcal{F}([a,b],\mathbb{C})$ and let $x_0 \in (a,b)$. If f is differentiable at x_0 and $\alpha \in \mathbb{C}$, then αf is differentiable at x_0 with*

$$(\alpha f)'(x_0) = \alpha f'(x_0).$$

Proof. Define $g \in \mathcal{F}([a,b],\mathbb{C})$ by $g(x) = \alpha$ for all $x \in [a,b]$. It is elementary to see based on the definition of the derivative that g is differentiable with $g'(x) = 0$ for all $x \in (a,b)$. Hence the Product Rule implies $fg = \alpha f$ is differentiable at x_0 with

$$(\alpha f)'(x_0)(fg)'(x_0) = f'(x_0)g(x_0) + f(x_0g'(x_0) = \alpha f'(x_0) + 0 = \alpha f'(x_0)$$

as desired. ∎

Of course, the quotient rule also generalizes.

Theorem B.1.9 (Quotient Rule). *Let $f, g \in \mathcal{F}([a,b],\mathbb{C})$ and let $x_0 \in (a,b)$. If g is differentiable at x_0 and $g(x_0) \neq 0$, then $\frac{1}{g}$ is differentiable at x_0 with*

$$\left(\frac{1}{g}\right)'(x_0) = -\frac{g'(x_0)}{(g(x_0))^2}.$$

Therefore, if in addition f is differentiable at x_0, then $\frac{f}{g}$ is differentiable at x_0 with

$$\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{(g(x_0))^2}.$$

Proof. Instead of appealing to the real and imaginary parts, it is much easier to return to the definition of the derivative. Suppose g is differentiable at x_0 and $g(x_0) \neq 0$. Since g is differentiable at x_0, g is continuous at x_0 and therefore, since $g(x_0) \neq 0$, there exists a $\delta > 0$ such that if $x \in \mathbb{R}$ and $|x - x_0| < \delta$, then $x \in (a,b)$ and $g(x) \neq 0$. Thus for all $h \in \mathbb{R}$ with $0 < |h| < \delta$ we have that

$$\frac{\frac{1}{g(x_0+h)} - \frac{1}{g(x_0)}}{h} = \frac{\frac{g(x_0)-g(x_0+h)}{g(x_0)g(x_0+h)}}{h} = \frac{g(x_0) - g(x_0+h)}{g(x_0)g(x_0+h)h}.$$

Since g is continuous at x_0, we know that

$$\lim_{h\to 0} \frac{1}{g(x_0+h)} = \frac{1}{g(x_0)}.$$

Moreover, since g is differentiable at x_0, we know that

$$\lim_{h\to 0} \frac{g(x_0) - g(x_0+h)}{h} = -g'(x_0).$$

Hence

$$\lim_{h\to 0} \frac{\frac{1}{g(x_0+h)} - \frac{1}{g(x_0)}}{h} = -\frac{g'(x_0)}{(g(x_0))^2}.$$

Thus $\frac{1}{g}$ is differentiable at x_0 with

$$\left(\frac{1}{g}\right)'(x_0) = -\frac{g'(x_0)}{(g(x_0))^2}.$$

Finally, if in addition f is differentiable at x_0, then $\frac{f}{g} = f\frac{1}{g}$ is differentiable at x_0 by the product rule with

$$\begin{aligned}\left(\frac{f}{g}\right)'(x_0) &= f'(x_0)\frac{1}{g(x_0)} + f(x_0)\left(\frac{1}{g}\right)'(x_0)\\ &= \frac{f'(x_0)}{g(x_0)} - f(x_0)\frac{g'(x_0)}{(g(x_0))^2}\\ &= \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{(g(x_0))^2}\end{aligned}$$

as desired. ∎

Finally, we have a version of the chain rule for the composition of a real-valued function with a complex-valued function.

Theorem B.1.10 (Chain Rule). *Let I and J be open intervals, let $g : I \to \mathbb{C}$, and let $f : J \to \mathbb{R}$ be such that $f(J) \subseteq I$. Suppose that $a \in J$, f is differentiable at a, and g is differentiable at $f(a)$. Then $g \circ f : I \to \mathbb{C}$ is differentiable at a and*

$$(g \circ f)'(a) = g'(f(a))f'(a).$$

Proof. Let $g_1 = \mathrm{Re}(g)$ and $g_2 = \mathrm{Im}(g)$. Then

$$(g \circ f)(x) = g_1(f(x)) + ig_2(f(x)) = (g_1 \circ f)(x) + i(g_2 \circ f)(x)$$

for all $x \in J$. Therefore, by the chain rule for real-valued functions, $g \circ f$ is differentiable at a with

$$\begin{aligned}(g \circ f)'(a) &= (g_1 \circ f)'(a) + i(g_2 \circ f)'(a)\\ &= g_1'(f(a))f'(a) + ig_2'(f(a))f'(a)\\ &= g'(f(a))f'(a)\end{aligned}$$

as desired. ∎

B.2 Complex Integration

With the basics of differentiation of a complex-valued function on an interval complete, we turn our attention to integration. As differentiation can be done via the real and imaginary parts, the following definition should be no surprise.

Definition B.2.1. Given $f \in \mathcal{F}([a,b],\mathbb{C})$, it is said that f is *Riemann integrable* if $\mathrm{Re}(f)$ and $\mathrm{Im}(f)$ are Riemann integrable. When f is Riemann integrable, the *complex-valued Riemann integral of f on* $[a,b]$ is defined to be the complex number

$$\int_a^b f(x)\,dx = \int_a^b \mathrm{Re}(f)(x)\,dx + i\int_a^b \mathrm{Im}(f)(x)\,dx.$$

Using the above definition and by exploiting results in the real-valued setting, we automatically generalize the Fundamental Theorems of Calculus.

Theorem B.2.2 (Fundamental Theorem of Calculus, Part I). *Let $f : [a, b] \to \mathbb{C}$ be continuous and define $F : [a, b] \to \mathbb{C}$ by*

$$F(x) = \int_a^x f(t)\, dt$$

for all $x \in [a, b]$. Then F is continuous on $[a, b]$, differentiable on (a, b), and $F'(x) = f(x)$ for all $x \in (a, b)$.

Proof. Let $f_1 = \mathrm{Re}(f)$, $f_2 = \mathrm{Im}(f)$, $F_1 = \mathrm{Re}(F)$, and $F_2 = \mathrm{Im}(F)$. By the definition of the complex-valued Riemann integral, we have that

$$F_1(x) = \int_a^x f_1(t)\, dt \quad \text{and} \quad F_2(x) = \int_a^x f_2(t)\, dt$$

for all $x \in [a, b]$. Since f is continuous, f_1 and f_2 are continuous. Therefore, by the real-valued version of the Fundamental Theorem of Calculus, we obtain that F_1 and F_2 are continuous on $[a, b]$, differentiable on (a, b), and $F_1'(x) = f_1(x)$ and $F_2'(x) = f_2(x)$ for all $x \in (a, b)$. Thus it follows from Theorem B.1.3 that F is continuous on $[a, b]$, differentiable on (a, b), and $F'(x) = f(x)$ for all $x \in (a, b)$. ∎

Theorem B.2.3 (Fundamental Theorem of Calculus, II). *Let $f, F : [a, b] \to \mathbb{C}$ be such that f is Riemann integrable on $[a, b]$, F is continuous on $[a, b]$, F is differentiable on (a, b), and $F'(x) = f(x)$ for all $x \in (a, b)$. Then*

$$\int_a^b f(t)\, dt = F(b) - F(a).$$

Proof. Let $f_1 = \mathrm{Re}(f)$, $f_2 = \mathrm{Im}(f)$, $F_1 = \mathrm{Re}(F)$, and $F_2 = \mathrm{Im}(F)$. Therefore f_1 and f_2 are Riemann integrable on $[a, b]$, F_1 and F_2 are continuous on $[a, b]$, F_1 and F_2 are differentiable on (a, b), and $F_1'(x) = f_1(x)$ and $F_2'(x) = f_2(x)$ for all $x \in (a, b)$ by Theorem B.1.3. Therefore, by the real-valued version of the Fundamental Theorem of Calculus, we obtain that

$$\int_a^b f_1(t)\, dt = F_1(b) - F_1(a) \quad \text{and} \quad \int_a^b f_2(t)\, dt = F_2(b) - F_2(a).$$

Thus

$$\begin{aligned} \int_a^b f(t)\, dt &= \int_a^b f_1(t)\, dt + i \int_a^b f_2(t)\, dt \\ &= (F_1(b) - F_1(a)) + i(F_2(b) - F_2(a)) \\ &= (F_1(b) + iF_2(b)) - (F_1(a) + iF_2(a)) = F(b) - F(a) \end{aligned}$$

as desired. ∎

As with real-valued functions, the Fundamental Theorems of Calculus immediately enables us to integrate any function we know the anti-derivative of. In particular, the exponentials in Example B.1.4 are particularly easy yet useful in this course.

Example B.2.4. Let $a, b \in \mathbb{R}$, let $\alpha = a + bi \in \mathbb{C} \setminus \{0\}$, and define $f : \mathbb{R} \to \mathbb{C}$ by

$$f(x) = e^{\alpha x} = e^{(a+bi)x} = e^{ax}\cos(bx) + ie^{ax}\sin(bx)$$

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

$$f'(x) = \alpha e^{\alpha x}$$

for all $x \in \mathbb{R}$, we see that

$$\left(\frac{1}{\alpha}f\right)'(x) = e^{\alpha x}$$

for all $x \in \mathbb{R}$. Thus the Fundamental Theorem of Calculus implies for all $c, d \in \mathbb{R}$ with $c < d$ that

$$\int_c^d e^{\alpha x}\, dx = \frac{1}{\alpha}e^{\alpha d} - \frac{1}{\alpha}e^{\alpha c}.$$

To conclude this discussion, it is particularly useful to know what operations on Riemann integrable functions produce Riemann integrable functions and how the values of the integrals relate. We begin with the following.

Proposition B.2.5. *Let $f, g \in \mathcal{F}([a, b], \mathbb{C})$ be Riemann integrable. Then the following are true:*

a) $f + g$ is Riemann integrable.

b) $\overline{f}$ is Riemann integrable.

c) fg is Riemann integrable.

d) $|f|$ is Riemann integrable.

Proof. Let $f_1 = \mathrm{Re}(f)$, $f_2 = \mathrm{Im}(f)$, $g_1 = \mathrm{Re}(g)$, and $g_2 = \mathrm{Im}(g)$. Since f and g are Riemann integrable, f_1, f_2, g_1, and g_2 are Riemann integrable. Therefore, since

$$\begin{aligned}
\mathrm{Re}(f+g) &= f_1 + g_1, \\
\mathrm{Im}(f+g) &= f_2 + g_2, \\
\mathrm{Re}(\overline{f}) &= f_1, \\
\mathrm{Im}(\overline{f}) &= -f_2, \\
\mathrm{Re}(fg) &= f_1 g_1 - f_2 g_2, \\
\mathrm{Im}(fg) &= f_1 g_2 + f_2 g_1, \text{ and} \\
|f| &= \sqrt{f_1^2 + f_2^2},
\end{aligned}$$

we see that all of these functions are Riemann integrable and thus $f+g$, $\overline{f}$, fg, and $|f|$ are Riemann integrable. ∎

Of course, it is unsurprising that the complex integral is linear.

Proposition B.2.6. *Let $f, g \in \mathcal{F}([a,b],\mathbb{C})$ be Riemann integrable. Then*

$$\int_a^b f(x) + g(x)\, dx = \int_a^b f(x)\, dx + \int_a^b g(x)\, dx.$$

Moreover, if $\alpha \in \mathbb{C}$, then

$$\int_a^b \alpha f(x)\, dx = \alpha \int_a^b f(x)\, dx.$$

Proof. Let $f_1 = \mathrm{Re}(f)$, $f_2 = \mathrm{Im}(f)$, $g_1 = \mathrm{Re}(g)$, and $g_2 = \mathrm{Im}(g)$. Since

$$\mathrm{Re}(f+g) = f_1 + g_1 \quad \text{and} \quad \mathrm{Im}(f+g) = f_2 + g_2,$$

we obtain by the definition of the complex integral that

$$\begin{aligned}
&\int_a^b f(x) + g(x)\, dx \\
&= \left(\int_a^b f_1(x) + g_1(x)\, dx \right) + i \left(\int_a^b f_2(x) + g_2(x)\, dx \right) \\
&= \left(\int_a^b f_1(x)\, dx + \int_a^b g_1(x)\, dx \right) + i \left(\int_a^b f_2(x)\, dx + \int_a^b g_2(x)\, dx \right) \\
&= \left(\int_a^b f_1(x)\, dx + i \int_a^b f_2(x)\, dx \right) + \left(\int_a^b g_1(x)\, dx + i \int_a^b g_2(x)\, dx \right) \\
&= \int_a^b f(x)\, dx + \int_a^b g(x)\, dx
\end{aligned}$$

as desired.

For the second part of the proof, write $\alpha = c + di$ where $c, d \in \mathbb{R}$. Since

$$\mathrm{Re}(\alpha f) = cf_1 - df_2 \quad \text{and} \quad \mathrm{Im}(\alpha f) = cf_2 + df_1,$$

we obtain by the definition of the complex integral that

$$\begin{aligned}
&\int_a^b \alpha f \, dx \\
&= \left(\int_a^b cf_1(x) - df_2(x)\, dx\right) + i\left(\int_a^b cf_2(x) + df_1(x)\, dx\right) \\
&= \left(c\int_a^b f_1(x)\, dx - d\int_a^b f_2(x)\, dx\right) + i\left(c\int_a^b f_2(x)\, dx + d\int_a^b f_1(x)\, dx\right) \\
&= (c+di)\int_a^b f_1(x)\, dx + (-d+ic)\int_a^b f_2(x)\, dx \\
&= \alpha\int_a^b f_1(x)\, dx + \alpha i\int_a^b f_2(x)\, dx \\
&= \alpha\left(\int_a^b f_1(x)\, dx + i\int_a^b f_2(x)\, dx\right) = \alpha\int_a^b f(x)\, dx
\end{aligned}$$

as desired. ∎

In addition, integrating a complex conjugation is almost trivial.

Proposition B.2.7. *Let $f \in \mathcal{F}([a,b],\mathbb{C})$ be Riemann integrable. Then*

$$\int_a^b \overline{f(x)}\, dx = \overline{\int_a^b f(x)\, dx}.$$

Proof. Let $f_1 = \mathrm{Re}(f)$ and $f_2 = \mathrm{Im}(f)$. Since

$$\mathrm{Re}(\overline{f}) = f_1 \quad \text{and} \quad \mathrm{Im}(\overline{f}) = -f_2,$$

we obtain by the definition of the complex integral that

$$\begin{aligned}
&\int_a^b \overline{f(x)}\, dx \\
&= \int_a^b f_1(x)\, dx + i\int_a^b -f_2(x)\, dx \\
&= \int_a^b f_1(x)\, dx - i\int_a^b f_2(x)\, dx \\
&= \overline{\int_a^b f_1(x)\, dx + i\int_a^b f_2(x)\, dx} \\
&= \overline{\int_a^b f(x)\, dx}
\end{aligned}$$

as desired. ∎

We recall from integrating real-valued functions that integrals of products need not be the product of the integrals. However, integration by parts still works.

Theorem B.2.8 (Integration by Parts). *Let $h_1, h_2 \in \mathcal{F}([a,b],\mathbb{C})$ be continuously differentiable functions. Then*

$$\int_a^b h_1'(x)h_2(x)\,dx = (h_1(b)h_2(b) - h_1(a)h_2(a)) - \int_a^b h_1(x)h_2'(x)\,dx.$$

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

$$f(x) = h_1(x)h_2(x)$$

for all $x \in [a,b]$. Since h_1 and h_2 are continuous on $[a,b]$, we obtain that f is continuous on $[a,b]$. Moreover, since h_1 and h_2 are differentiable on (a,b), the product rule implies that

$$f'(x) = h_1'(x)h_2(x) + h_1(x)h_2'(x)$$

for all $x \in (a,b)$. Furthermore, since h_1 and h_2 are continuously differentiable, we see that $f' = h_1'h_2 + h_1h_2'$ is Riemann integrable on $[a,b]$. Therefore, the Fundamental Theorem of Calculus II implies that

$$\begin{aligned} h_1(b)h_2(b) - h_1(a)h_2(a) &= f(b) - f(a) \\ &= \int_a^b f'(x)\,dx \\ &= \int_a^b h_1'(x)h_2(x) + h_1(x)h_2'(x)\,dx \\ &= \int_a^b h_1'(x)h_2(x)\,dx + \int_a^b h_1(x)h_2'(x)\,dx. \end{aligned}$$

Thus, by rearranging this equation, the result follows. ∎

Finally, the following relates the integrals of a function and its absolute value. However, as the proof does not follow from the real-valued results, we need to be slightly clever.

Theorem B.2.9. *Let $f \in \mathcal{F}([a,b],\mathbb{C})$ be Riemann integrable. Then*

$$\left|\int_a^b f(x)\,dx\right| \leq \int_a^b |f(x)|\,dx.$$

Proof. Since

$$\int_a^b f(x)\,dx \in \mathbb{C},$$

there exists a $z \in \mathbb{C}$ such that $|z| = 1$ and

$$z\int_a^b f(x)\,dx = \left|\int_a^b f(x)\,dx\right|$$

(i.e. if $\int_a^b f(x)\,dx = re^{i\theta}$ for some $r \geq 0$ and $\theta \in [0, 2\pi)$, then take $z = e^{-i\theta}$). Therefore we have that

$$\begin{aligned} 0 &\leq \left| \int_a^b f(x)\,dx \right| \\ &= z \int_a^b f(x)\,dx \\ &= \int_a^b zf(x)\,dx \\ &= \int_a^b \mathrm{Re}(zf(x))\,dx + i \int_a^b \mathrm{Im}(zf(x))\,dx. \end{aligned}$$

However, the above inequality implies it must be true that

$$\int_a^b \mathrm{Im}(zf(x))\,dx = 0$$

as the only way a real number and a complex number are equal is if the imaginary part of the complex number is zero. Therefore we have that

$$\begin{aligned} \left| \int_a^b f(x)\,dx \right| &= \int_a^b \mathrm{Re}(zf(x))\,dx \\ &\leq \int_a^b |zf(x)|\,dx \qquad\qquad \text{as } \mathrm{Re}(zf(x)) \leq |zf(x)| \\ &= \int_a^b |z||f(x)|\,dx \\ &= \int_a^b |f(x)|\,dx \end{aligned}$$

as desired. ∎

Appendix C

Multivariate Calculus Theorems

In this appendix chapter, we will provide proofs to the multivariate calculus results that are used in this course.

C.1 Fubini's Theorem

We begin with the fundamental result of changing the order of integral signs.

Theorem C.1.1 (Fubini's Theorem). *If $f : [a, b] \times [c, d] \to \mathbb{R}$ is continuous, then*

$$y \mapsto \int_a^b f(x, y)\, dx \quad \text{and} \quad x \mapsto \int_c^d f(x, y)\, dy$$

are continuous on $[c, d]$ and $[a, b]$ respectively, and

$$\iint_{[a,b]\times[c,d]} f(x, y)\, dA = \int_a^b \int_c^d f(x, y)\, dy\, dx = \int_c^d \int_a^b f(x, y)\, dx\, dy.$$

Before we get to the proof of Fubini's Theorem (Theorem C.1.1), we first desire to present a proof that if $f : [a, b] \times [c, d] \to \mathbb{R}$ is continuous, then f is Riemann integrable over $[a, b] \times [c, d]$. To do this, we re-imagine the notion of uniform continuity presented in Definition 2.8.1.

Definition C.1.2. A function $f : [a, b] \times [c, d] \to \mathbb{R}$ is said to be *uniformly continuous* on I if for all $\epsilon > 0$ there exists a $\delta > 0$ such that if $(x_1, y_1), (x_2, y_2) \in [a, b] \times [c, d]$, $|x_1 - x_2| < \delta$, and $|y_1 - y_2| < \delta$, then $|f(x_1, y_1) - f(x_2, y_2)| < \epsilon$.

Unsurprisingly, we can generalize the proof of Theorem 2.8.4 to the following.

Theorem C.1.3. *If $f : [a, b] \times [c, d] \to \mathbb{R}$ is continuous, then f is uniformly continuous.*

Proof. Let $f : [a, b] \times [c, d] \to \mathbb{R}$ be continuous. Suppose to the contrary that f is not uniformly continuous. Hence there exists an $\epsilon > 0$ such that for all $\delta > 0$ there exists $(x_1, y_1), (x_2, y_2) \in [a, b] \times [c, d]$ such that $|x_1 - x_2| < \delta$, $|y_1 - y_2| < \delta$, and $|f(x_1, y_1) - f(x_2, y_2)| \geq \epsilon$. Therefore, for each $n \in \mathbb{N}$ there exist $(x_n, y_n), (x'_n, y'_n) \in [a, b] \times [c, d]$ with $|x_n - x'_n| < \frac{1}{n}$, $|y_n - y'_n| < \delta$, and $|f(x_n, y_n) - f(x'_n, y'_n)| \geq \epsilon$.

Since $[a, b]$ is closed and bounded, the Bolzano-Weierstrass Theorem implies there exists a subsequence $(x_{k_n})_{n\geq1}$ of $(x_n)_{n\geq1}$ that converges to some number $L \in [a, b]$. Subsequently, since $[c, d]$ is closed and bounded, the Bolzano-Weierstrass Theorem implies there exists a subsequence $(y_{m_{k_n}})_{n\geq1}$ of $(y_{k_n})_{n\geq1}$ that converges to some number $K \in [c, d]$. Note $(x_{m_{k_n}})_{n\geq1}$ still converges to L since $(x_{k_n})_{n\geq1}$ does.

Since f is continuous on $[a, b] \times [c, d]$, there exists an $N_1 \in \mathbb{N}$ such that $|f(x_{m_{k_n}}, y_{m_{k_n}}) - f(L, K)| < \frac{\epsilon}{2}$ for all $n \geq N_1$.

Consider the subsequence $(x'_{m_{k_n}})_{n\geq1}$ of $(x'_n)_{n\geq1}$. Notice for all $n \in \mathbb{N}$ that

$$\begin{aligned}
|x'_{m_{k_n}} - L| &\leq |x'_{m_{k_n}} - x_{m_{k_n}}| + |x_{m_{k_n}} - L| \\
&\leq \frac{1}{m_{k_n}} + |x_{m_{k_n}} - L| \\
&\leq \frac{1}{n} + |x_{m_{k_n}} - L|.
\end{aligned}$$

Therefore, since $\lim_{n\to\infty} |x_{m_{k_n}} - L| = 0$ and $\lim_{n\to\infty} \frac{1}{n} = 0$, we obtain that $\lim_{n\to\infty} x'_{m_{k_n}} = L$. Similarly, $\lim_{n\to\infty} y'_{m_{k_n}} = K$. Since f is continuous this implies that there exists an $N_2 \in \mathbb{N}$ such that $|f(x'_{m_{k_n}}, y'_{m_{k_n}}) - f(L, K)| < \frac{\epsilon}{2}$ for all $n \geq N_2$.

Notice if $N = \max\{N_1, N_2\}$, then the above implies that

$$\begin{aligned}
&|f(x_{m_{k_n}}, y_{m_{k_n}}) - f(x'_{m_{k_n}}, y'_{m_{k_n}})| \\
&\leq |f(x_{m_{k_n}}, y_{m_{k_n}}) - f(L)| + |f(L) - f(x'_{m_{k_n}}, y'_{m_{k_n}})| \\
&< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon
\end{aligned}$$

thereby contradicting the fact that $|f(x_{m_{k_n}}, y_{m_{k_n}}) - f(x'_{m_{k_n}}, y'_{m_{k_n}})| \geq \epsilon$. Hence f is uniformly continuous on $[a, b] \times [c, d]$. ■

It is the property of uniform continuity that implies continuous functions are Riemann integrable. The following demonstrates the 2-variable version, whereas the 1-variable version (which should have been proved in MATH 2001) follows by similar arguments.

Corollary C.1.4. *If $f : [a, b] \times [c, d] \to \mathbb{R}$ is continuous, then f is Riemann integrable on $[a, b] \times [c, d]$.*

Proof. Let $\epsilon > 0$. Since f is continuous on $[a,b]\times[c,d]$, f is uniformly continuous on $[a,b]\times[c,d]$. Thus there exists a $\delta > 0$ such that if $(x_1,y_1),(x_2,y_2)\in[a,b]\times[c,d]$ are such that $|x_1-x_2|<\delta$ and $|y_1-y_2|<\delta$, then

$$|f(x_1,y_1)-f(x_2,y_2)| < \frac{\epsilon}{(b-a)(d-c)}.$$

Let $\mathcal{P}$ be any partition of $[a,b]\times[c,d]$ with interval lengths at most δ (which can always be constructed). Write $\mathcal{P}=\{\{(x_i,y_j)\}_{i=0}^n\}_{j=0}^m$ where

$$\begin{aligned} a &= x_0 < x_1 < \cdots < x_n = b \\ c &= y_0 < y_1 < \cdots < y_m = d. \end{aligned}$$

For all $(i,j)\in\{1,\ldots,n\}\times\{1,\ldots,m\}$, let

$$\begin{aligned} M_{i,j} &= \sup(\{f(x,y) \mid (x,y)\in[x_{i-1},x_i]\times[y_{j-1},y_j]\}) \\ m_{i,j} &= \inf(\{f(x,y) \mid (x,y)\in[x_{i-1},x_i]\times[y_{j-1},y_j]\}. \end{aligned}$$

Then

$$\begin{aligned} U(f,\mathcal{P}) &= \sum_{i=1}^n\sum_{j=1}^m M_{i,j}(x_i-x_{i-1})(y_j-y_{j-1}) \\ L(f,\mathcal{P}) &= \sum_{i=1}^n\sum_{j=1}^m m_{i,j}(x_i-x_{i-1})(y_j-y_{j-1}). \end{aligned}$$

Hence

$$\begin{aligned} U(f,\mathcal{P})-L(f,\mathcal{P}) &= \sum_{i=1}^n\sum_{j=1}^m (M_{i,j}-m_{i,j})(x_i-x_{i-1})(y_j-y_{j-1}) \\ &\leq \sum_{i=1}^n\sum_{j=1}^m \frac{\epsilon}{(b-a)(d-c)}(x_i-x_{i-1})(y_j-y_{j-1}) \\ &= \epsilon. \end{aligned}$$

Therefore, as $\epsilon > 0$ was arbitrary, f is Riemann integrable. ∎

Proof of Fubini's Theorem (Theorem C.1.1). First, it is necessary to demonstrate that all of the integrals in the statement of the theorem are well-defined.

Since f is Riemann integrable,

$$\iint_{[a,b]\times[c,d]} f(x,y)\,dA$$

is well-defined. Moreover, since

$$x\mapsto f(x,y_0) \qquad \text{and} \qquad y\mapsto f(x_0,y)$$

are continuous on $[a, b]$ and $[c, d]$ respectively for all $x_0 \in [a, b]$ and $y_0 \in [c, d]$, we obtain that

$$\int_a^b f(x, y_0)\, dx \quad \text{and} \quad \int_c^d f(x_0, y)\, dy$$

are well-defined for all $x_0 \in [a, b]$ and $y_0 \in [c, d]$. Thus, to see that

$$\int_c^d \int_a^b f(x, y)\, dx\, dy \quad \text{and} \quad \int_a^b \int_c^d f(x, y)\, dy\, dx$$

are well-defined, it suffices to show that

$$y \mapsto \int_a^b f(x, y)\, dx \quad \text{and} \quad x \mapsto \int_c^d f(x, y)\, dy$$

are continuous on $[c, d]$ and $[a, b]$ respectively.

To see that the first is continuous, let $\epsilon > 0$. Since f is uniformly continuous on $[a, b] \times [c, d]$, there exists a $\delta > 0$ such that if $(x_1, y_1), (x_2, y_2) \in [a, b] \times [c, d]$ are such that $|x_1 - x_2| < \delta$ and $|y_1 - y_2| < \delta$, then

$$|f(x_1, y_1) - f(x_2, y_2)| < \frac{\epsilon}{b - a}.$$

Therefore, for all $y_1, y_2 \in [c, d]$ with $|y_1 - y_2| < \delta$ we have that

$$\begin{aligned} \left| \int_a^b f(x, y_1)\, dx - \int_a^b f(x, y_2)\, dx \right| &= \left| \int_a^b f(x, y_1) - f(x, y_2)\, dx \right| \\ &\leq \int_a^b |f(x, y_1) - f(x, y_2)|\, dx \\ &\leq \int_a^b \frac{\epsilon}{b - a}\, dx \\ &= \epsilon. \end{aligned}$$

Thus $y \mapsto \int_a^b f(x, y)\, dx$ is uniformly continuous on $[c, d]$. Similarly $x \mapsto \int_c^d f(x, y)\, dy$ is uniformly continuous on $[a, b]$.

To see that

$$\iint_{[a,b]\times[c,d]} f(x, y)\, dA = \int_c^d \int_a^b f(x, y)\, dx\, dy,$$

let $\epsilon > 0$ be arbitrary. Since f is Riemann integrable on $[a, b] \times [c, d]$, there exists a partition $\mathcal{P}$ of $[a, b] \times [c, d]$ such that if $\mathcal{P} = \{\{(x_i, y_j)\}_{i=0}^n\}_{j=0}^m$ where

$$\begin{aligned} a = x_0 < x_1 < \cdots < x_n = b \\ c = y_0 < y_1 < \cdots < y_m = d \end{aligned}$$

and for all $(i, j) \in \{1, \ldots, n\} \times \{1, \ldots, m\}$, we let

$$\begin{aligned} M_{i,j} &= \sup(\{f(x, y) \mid (x, y) \in [x_{i-1}, x_i] \times [y_{j-1}, y_j]\} \\ m_{i,j} &= \inf(\{f(x, y) \mid (x, y) \in [x_{i-1}, x_i] \times [y_{j-1}, y_j]\}, \end{aligned}$$

then with

$$U(f,\mathcal{P}) = \sum_{i=1}^{n}\sum_{j=1}^{m} M_{i,j}(x_i - x_{i-1})(y_j - y_{j-1})$$
$$L(f,\mathcal{P}) = \sum_{i=1}^{n}\sum_{j=1}^{m} m_{i,j}(x_i - x_{i-1})(y_j - y_{j-1})$$

we have that

$$L(f,\mathcal{P}) \leq \iint_{[a,b]\times[c,d]} f(x,y)\,dA \leq U(f,\mathcal{P}) < L(f,\mathcal{P}) + \epsilon.$$

Since for all $(i,j) \in \{1,\ldots,n\} \times \{1,\ldots,m\}$ we have

$$m_{i,j} \leq f(x,y) \leq M_{i,j}$$

for all $(x,y) \in [x_{i-1}, x_i] \times [y_{j-1}, y_j]$, we obtain that

$$m_{i,j}(x_i - x_{i-1}) \leq \int_{x_{i-1}}^{x_i} f(x,y)\,dx \leq M_{i,j}(x_i - x_{i-1})$$

for all $y \in [y_{j-1}, y_j]$ and $(i,j) \in \{1,\ldots,n\} \times \{1,\ldots,m\}$. Therefore, for any fix $j \in \{1,\ldots,m\}$ and $y \in [y_{j-1}, y_j]$, we obtain by summing over i that

$$\begin{aligned}\sum_{i=1}^{n} m_{i,j}(x_i - x_{i-1}) &\leq \sum_{i=1}^{n}\int_{x_{i-1}}^{x_i} f(x,y)\,dx\\ &= \int_a^b f(x,y)\,dx\\ &\leq \sum_{i=1}^{n} M_{i,j}(x_i - x_{i-1}).\end{aligned}$$

Thus, by integrating all three expressions over $y \in [y_{j-1}, y_j]$ for all $j \in \{1,\ldots,m\}$, we obtain that

$$\begin{aligned}\sum_{i=1}^{n} m_{i,j}(x_i - x_{i-1})(y_j - y_{j-1}) &\leq \int_{y_{j-1}}^{y_j}\int_a^b f(x,y)\,dx\,dy\\ &\leq \sum_{i=1}^{n} M_{i,j}(x_i - x_{i-1})(y_j - y_{j-1}).\end{aligned}$$

Thus, by summing over all $j \in \{1, \dots, m\}$, we obtain that

$$\begin{aligned} L(f, \mathcal{P}) &= \sum_{j=1}^{m} \sum_{i=1}^{n} m_{i,j}(x_i - x_{i-1})(y_j - y_{j-1}) \\ &\leq \sum_{j=1}^{m} \int_{y_{j-1}}^{y_j} \int_a^b f(x, y)\, dx\, dy \\ &= \int_a^b \int_a^b f(x, y)\, dx\, dy \\ &\leq \sum_{j=1}^{m} \sum_{i=1}^{n} M_{i,j}(x_i - x_{i-1})(y_j - y_{j-1}) \\ &= U(f, \mathcal{P}). \end{aligned}$$

Therefore, since both

$$\iint_{[a,b]\times[c,d]} f(x, y)\, dA \quad \text{and} \quad \int_c^d \int_a^b f(x, y)\, dx\, dy$$

are in the interval

$$[L(f, \mathcal{P}), U(f, \mathcal{P})]$$

and

$$U(f, \mathcal{P}) - L(f, \mathcal{P}) < \epsilon,$$

we obtain that

$$\left| \iint_{[a,b]\times[c,d]} f(x, y)\, dA - \int_c^d \int_a^b f(x, y)\, dx\, dy \right| < \epsilon.$$

Therefore, as $\epsilon > 0$ was arbitrary, we obtain that

$$\iint_{[a,b]\times[c,d]} f(x, y)\, dA = \int_c^d \int_a^b f(x, y)\, dx\, dy.$$

The proof that

$$\iint_{[a,b]\times[c,d]} f(x, y)\, dA = \int_a^b \int_c^d f(x, y)\, dy\, dx$$

is similar. ∎

C.2 Leibniz Integral Rule

Another useful result from multi-variate calculus is the ability to differentiate under the integral sign.

Theorem C.2.1 (Leibniz Integral Rule). *Let $f : [a,b] \times [c,d] \to \mathbb{R}$ be such that $\frac{\partial f}{\partial y}$ is continuous on $[a,b]\times[c,d]$. Then*

$$\frac{d}{dy}\left(\int_a^b f(x,y)\,dx\right) = \int_a^b \frac{\partial f}{\partial y}(x,y)\,dx$$

on (c,d).

Proof. Since $\frac{\partial f}{\partial y}$ is continuous on $[a,b]\times[c,d]$, f is continuous on $[a,b]\times[c,d]$ and thus both integrals in the statement of the theorem are well-defined.

To see the desired result, fix $y_0 \in (c,d)$. Consider the function $F : [c,d] \to \mathbb{R}$ defined by

$$F(t) = \int_c^t \int_a^b \frac{\partial f}{\partial y}(x,y)\,dx\,dy$$

for all $t \in [c,d]$. Note F is well-defined since $\frac{\partial f}{\partial y}$ is continuous and

$$y \mapsto \int_a^b \frac{\partial f}{\partial y}(x,y)\,dx$$

is continuous by Fubini's Theorem (Theorem C.1.1). Moreover, by the Fundamental Theorem of Calculus, we know that

$$F'(t) = \int_a^b \frac{\partial f}{\partial y}(x,t)\,dx$$

for all $t \in (c,d)$. Thus, it suffices to show that

$$\left(\frac{d}{dy}\left(\int_a^b f(x,y)\,dx\right)\right)(y_0) = F'(y_0).$$

Notice if $h > 0$ is such that $y_0 + h \in [c,d]$, then by Fubini's Theorem (Theorem C.1.1) and the Fundamental Theorem of Calculus, we have that

$$\begin{aligned}
\frac{F(y_0+h) - F(y_0)}{h} &= \frac{1}{h}\left(\int_c^{y_0+h}\int_a^b \frac{\partial f}{\partial y}(x,y)\,dx\,dy - \int_c^{y_0}\int_a^b \frac{\partial f}{\partial y}(x,y)\,dx\,dy\right) \\
&= \frac{1}{h}\left(\int_{y_0}^{y_0+h}\int_a^b \frac{\partial f}{\partial y}(x,y)\,dx\,dy\right) \\
&= \frac{1}{h}\int_a^b \int_{y_0}^{y_0+h} \frac{\partial f}{\partial y}(x,y)\,dy\,dx \\
&= \int_a^b f(x,y_0+h) - f(x,y_0)\,dx \\
&= \frac{1}{h}\left(\int_a^b f(x,y_0+h)\,dx - \int_a^b f(x,y_0)\,dx\right).
\end{aligned}$$

Thus

$$\lim_{h\searrow 0}\frac{1}{h}\left(\int_a^b f(x,y_0+h)\,dx - \int_a^b f(x,y_0)\,dx\right) = \lim_{h\searrow 0}\frac{F(y_0+h)-F(y_0)}{h} = F'(y_0).$$

Since similar arguments show

$$\lim_{h\nearrow 0}\frac{1}{h}\left(\int_a^b f(x,y_0+h)\,dx - \int_a^b f(x,y_0)\,dx\right) F'(y_0),$$

the result follows. ∎

C.3 Laplace's Equation in Polar Coordinates

One result necessary for applications to the steady-state heat equation is the following version of the Laplacian in polar coordinates.

Theorem C.3.1. *If f is a continuous function on the closed unit disk centred at the origin that is twice continuously differentiable on the open unit centred at the origin, then, using polar coordinates,*

$$\Delta f = \frac{\partial^2 f}{\partial r^2} + \frac{1}{r}\frac{\partial f}{\partial r} + \frac{1}{r^2}\frac{\partial f^2}{\partial \theta^2}.$$

Proof. Recall if (x, y) is a point in the closed unit disk and

$$x = r\cos(\theta) \quad \text{and} \quad y = r\sin(\theta)$$

are the polar coordinates, then

$$\begin{aligned}
\frac{\partial x}{\partial r} &= \cos(\theta),\\
\frac{\partial y}{\partial r} &= \sin(\theta),\\
\frac{\partial x}{\partial \theta} &= -r\sin(\theta), \text{ and}\\
\frac{\partial y}{\partial \theta} &= r\cos(\theta).
\end{aligned}$$

To show the desired formula holds, let us compute

$$\frac{\partial^2 f}{\partial r^2}, \quad \frac{\partial f}{\partial r}, \quad \text{and} \quad \frac{\partial f^2}{\partial \theta^2}$$

using the Chain Rule. Indeed

$$\frac{\partial f}{\partial r} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial r} = \cos(\theta)\frac{\partial f}{\partial x} + \sin(\theta)\frac{\partial f}{\partial y}.$$

Therefore,

$$\begin{aligned}
\frac{\partial^2 f}{\partial r^2} &= \frac{\partial}{\partial r}\left(\cos(\theta)\frac{\partial f}{\partial x} + \sin(\theta)\frac{\partial f}{\partial y}\right) \\
&= \cos(\theta)\frac{\partial}{\partial r}\left(\frac{\partial f}{\partial x}\right) + \sin(\theta)\frac{\partial}{\partial r}\left(\frac{\partial f}{\partial y}\right) \\
&= \cos(\theta)\left(\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right)\frac{\partial x}{\partial r} + \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)\frac{\partial y}{\partial r}\right) \\
&\quad + \sin(\theta)\left(\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)\frac{\partial x}{\partial r} + \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right)\frac{\partial y}{\partial r}\right) \\
&= \cos(\theta)\left(\cos(\theta)\frac{\partial^2 f}{\partial x^2} + \sin(\theta)\frac{\partial^2 f}{\partial y \partial x}\right) \\
&\quad + \sin(\theta)\left(\cos(\theta)\frac{\partial^2 f}{\partial x \partial y} + \sin(\theta)\frac{\partial^2 f}{\partial y^2}\right) \\
&= \cos^2\frac{\partial^2 f}{\partial x^2} + 2\cos(\theta)\sin(\theta)\frac{\partial^2 f}{\partial x \partial y} + \sin^2(\theta)\frac{\partial^2 f}{\partial y^2}.
\end{aligned}$$

Similarly

$$\frac{\partial f}{\partial \theta} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial \theta} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial \theta} = -r\sin(\theta)\frac{\partial f}{\partial x} + r\cos(\theta)\frac{\partial f}{\partial y}.$$

Therefore

$$\begin{aligned}
\frac{\partial^2 f}{\partial \theta^2} &= \frac{\partial}{\partial \theta}\left(-r\sin(\theta)\frac{\partial f}{\partial x} + r\cos(\theta)\frac{\partial f}{\partial y}\right) \\
&= -r\cos(\theta)\frac{\partial f}{\partial x} - r\sin(\theta)\frac{\partial}{\partial \theta}\left(\frac{\partial f}{\partial x}\right) - r\sin(\theta)\frac{\partial f}{\partial y} + r\cos(\theta)\frac{\partial}{\partial \theta}\left(\frac{\partial f}{\partial y}\right) \\
&= -r\cos(\theta)\frac{\partial f}{\partial x} - r\sin(\theta)\left(\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right)\frac{\partial x}{\partial \theta} + \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)\frac{\partial y}{\partial \theta}\right) \\
&\quad - r\sin(\theta)\frac{\partial f}{\partial y} + r\cos(\theta)\left(\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)\frac{\partial x}{\partial \theta} + \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right)\frac{\partial y}{\partial \theta}\right) \\
&= -r\cos(\theta)\frac{\partial f}{\partial x} - r\sin(\theta)\left(-r\sin(\theta)\frac{\partial^2 f}{\partial x^2} + r\cos(\theta)\frac{\partial^2 f}{\partial y \partial x}\right) \\
&\quad - r\sin(\theta)\frac{\partial f}{\partial y} + r\cos(\theta)\left(-r\sin(\theta)\frac{\partial^2 f}{\partial x \partial y} + r\cos(\theta)\frac{\partial^2 f}{\partial y^2}\right) \\
&= -r\cos(\theta)\frac{\partial f}{\partial x} - r\sin(\theta)\frac{\partial f}{\partial y} \\
&\quad + r^2\sin^2(\theta)\frac{\partial^2 f}{\partial x^2} - 2r\cos(\theta)\sin(\theta)\frac{\partial^2 f}{\partial x \partial y} + r^2\cos^2(\theta)\frac{\partial^2 f}{\partial y^2}.
\end{aligned}$$

Therefore

$$\begin{aligned}
&\frac{\partial^2 f}{\partial r^2} + \frac{1}{r}\frac{\partial f}{\partial r} + \frac{1}{r^2}\frac{\partial f^2}{\partial \theta^2} \\
&= \left(\cos^2 \frac{\partial^2 f}{\partial x^2} + 2\cos(\theta)\sin(\theta)\frac{\partial^2 f}{\partial x \partial y} + \sin^2(\theta)\frac{\partial^2 f}{\partial y^2}\right) \\
&\quad + \frac{1}{r}\left(\cos(\theta)\frac{\partial f}{\partial x} + \sin(\theta)\frac{\partial f}{\partial y}\right) \\
&\quad + \frac{1}{r^2}\left(-r\cos(\theta)\frac{\partial f}{\partial x} - r\sin(\theta)\frac{\partial f}{\partial y}\right) \\
&\quad + \frac{1}{r^2}\left(r^2\sin^2(\theta)\frac{\partial^2 f}{\partial x^2} - 2r\cos(\theta)\sin(\theta)\frac{\partial^2 f}{\partial x \partial y} + r^2\cos^2(\theta)\frac{\partial^2 f}{\partial y^2}\right) \\
&= (\cos^2(\theta) + \sin^2(\theta))\left(\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}\right) \\
&= \Delta f
\end{aligned}$$

as desired. ∎

Index