# Project V: The Completeness of the Fourier Series in $L^2$.

Mou Minghao

November 28, 2021

## Contents

# 1 Background Knowledge

## 1.1 Introduction

Before we go into the discussion of completeness of Fourier Series in $L^2$, we need first introduce some background knowledge.

The main goal in this preliminary section is to show that $(L^2(\mathbb{T}), \langle \cdot, \cdot \rangle)$ *is a Hilbert space.* Here, $\mathbb{T} = \frac{\mathbb{R}}{2\pi\mathbb{Z}}$ stands for the one-dimensional torus, $L^2(\mathbb{T})$ is the collection of all Lebesgue-measurable functions defined on $\mathbb{T}$ and $\langle \cdot, \cdot \rangle$ is the inner product we defined on $L^2(\mathbb{T})$. In the followings, all those terminologies will be explicitly defined.

# Why do we need a Hilbert Space?

- The concept of a Hilbert space is a natural generalization of that of a Euclidean space. Here, the dot product is replaced by the notion of an inner product. Just like in the finite dimensional case, a Hilbert space $V$ also comes with the concept of a *basis*. Given a collection of vectors $\{v_n\}$, a natural generalization of the definition of a basis is that any vector $f \in V$ can be written as a linear combination of the $v_n$'s as $f = \sum_{n=1}^{\infty} c_n v_n$. However, this involves an infinite sum, and there would be an issue with convergence. Indeed, the meaning of the equal sign should be interpreted as *any vector $f$ can be approximated arbitrarily well by finite linear combinations of the $v_n$'s.*

- Our goal in this report is to show the completeness of Fourier series in $L^2(\mathbb{T})$. In other words, we want to show that the base functions (or, the eigenfunctions of the eigenvalue problem) $\{e_n\}, e_n = \frac{1}{\sqrt{2\pi}} e^{inx}$ forms a basis for $L^2(\mathbb{T})$. Since a lot of functions are in $L^2(\mathbb{T})$, then we can write those functions as a finite sum of $e_n's$, this will give a good approximation if $n$ is large. Moreover, we can also write any $f \in L^2(\mathbb{T})$ as an infinite sum of $e_n's$ if the convergence is not a problem.

## 1.2 Definitions

In the Euclidean space, we have dot product. In order to deal with general spaces, we shall need the following concept: inner product.

**Definition 1.1 (Inner Product)** A rule

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}.$$

is called an *inner product* if the followings are satisfied:

- $\langle \cdot, \cdot \rangle$ is linear in its first argument,

$$\langle af_1 + bf_2, g\rangle = a\langle f_1, g\rangle + b\langle f_2, g\rangle$$

- $\langle f, g\rangle = \overline{\langle g, f\rangle}$.

- for any $f \neq 0$, $\langle f, f\rangle > 0$.

Given a space of vectors, we would like to find a way to measure the distance of those vectors, just like what we do in $\mathbb{R}^n$ by the usual Euclidean norm.

**Definition 1.2 (Norm)** A function $f : V \to [0, \infty)$, denoted by $f(v) = \|v\|$ is called a *norm* if

- $\|v\| = 0 \Leftrightarrow v = 0$.

- $\|v_1 + v_2\| \leq \|v_1\| + \|v_2\|,\ \forall v_1, v_2 \in V$.

- $\alpha \in \mathbb{C}, v \in V \to \|\alpha v\| = |\alpha|\|v\|$.

**Remark.** Given an inner product, it naturally induces a norm,

$$\|x\| = \langle x, x\rangle^{1/2}.$$

After we have a method to measure the distance, we would like to focus on spaces with good properties, that is, if a sequence of vectors in this space is Cauchy, then they will approach to a vector still in this space.

**Definition 1.3 (Banach Space)** A *Banach Space* is a complete normed linear space $(V, \|\cdot\|)$. (i.e. For any Cauchy sequence $\{v_n\}_{n=1}^{\infty} \subset V$, $\{v_n\}_{n=1}^{\infty} \to v \in V$.)

Below we generalize the concept of inner product space.

**Definition 1.4 (Hilbert Space)** Let $(V, \langle \cdot, \cdot \rangle)$ be a complex vector space. It is called a *Hilbert Space* if $(V, \|\cdot\|)$ is a Banach Space.

**Definition 1.5 (Basis)** *Let $V$ be a Hilbert space with an inner product $\langle \cdot, \cdot \rangle$ and the induced norm $\| \cdot \|$. We say $\{v_n\}$ forms a basis for $V$ if for any $\epsilon > 0$, there exists a $N \in \mathbb{N}$ such that for $\forall n \geq N$,*

$$\|f - \sum_{k=1}^{n} c_n v_n\| < \epsilon.$$

It is relatively obvious to show that $L^2(\mathbb{T})$ is a vector space by the linearity of integration. So, we equip this vector space with an inner prduct defined below.

In order to construct a Hilbert space, we define a specific inner product for $L^2(\mathbb{T})$.

**Definition 1.6 (Inner Product on $L^2(\mathbb{T})$)** for $f$ and $g$, both in $L^2(\mathbb{T})$, define the inner product as

$$\langle f, g \rangle \triangleq \int_{\mathbb{T}} f\bar{g} d\mu.$$

**Definition 1.7** A function $f : \mathbb{R} \to \mathbb{C}$ is said to be of period $2\pi$ if

$$f(x + 2\pi) = f(x), \ \forall x \in \mathbb{R}.$$

**Remark 1.** The choice of $2\pi$ for the period is simply for convenience. In fact, for a general periodic function, a rescaling of the independent variable can be performed to reduce its period to $2\pi$ and then the our following arguments can be applied as well.

**Remark 2.** Although we could represent a $2\pi$-periodic function by a function $f$ defined on a closed interval $[a, a + 2\pi]$ (here the choice of $a$ is arbitrary) such that $f(a) = f(a + 2\pi)$, we do not adopt this setting. Instead, we will take the modern setting that indentifying a $2\pi$-periodic function on $\mathbb{R}$ by a function on the circle, or the one-dimensional torus, $\mathbb{T} = \mathbb{R}/(2\pi\mathbb{Z})$, which we define by identifying points in $\mathbb{R}$ that differ by $2n\pi$ for some $n \in \mathbb{Z}$.

To be precise, we define a binary relation $\sim$ on $\mathbb{R}$. $x \sim y$ if $x - y = 2n\pi$, for some $n \in \mathbb{Z}$. It is easy to show that $\sim$ is indeed an equivalence relation on $\mathbb{R}$. Then, $\mathbb{T}$ denotes the collection of all equivalence classes of $\sim$.

## 1.3 Construction of $L^2(\mathbb{T})$

Next, we define the central staff where our discussion focus on, $L^2(\mathbb{T})$. There are two methods to construct $L^2(\mathbb{T})$:

**Method 1.** First, consider $C(\mathbb{T})$, the collection of all continuous functions defined on $\mathbb{T}$, then $L^2(\mathbb{T})$ is the completion[1] of $C(\mathbb{T})$ with respect to the $L^2$ norm

$$\|f\| = \left( \int_{\mathbb{T}} |f|^2 dx \right)^{1/2}.$$

Here the integral of $\mathbb{T}$ is the usual Riemann integral with respect to $x$ taken over any interval with length $2\pi$.

**Method 2.** Since we can identify $\mathbb{T}$ as any interval with length $2\pi$ in $\mathbb{R}$, it enables us to equip $\mathbb{T}$ with a $\sigma$-algebra $M(\mathbb{T})$ (the collection of all Lebesgue measure sets, defined by the *Caratheodory's Criterion*) and the Lebesgue measure $\lambda$ (The restriction of outer measure on $M(\mathbb{T})$). This is justified by the Measure Extension Theorem. Then, In this case, $\|f\| \triangleq \left( \int_{\mathbb{T}} |f|^2 d\mu \right)^{1/2}$. If we want to specify $M(\mathbb{T})$ and $\mu$, we use the notation $L^2(\mathbb{T}, M(\mathbb{T}), \mu)$. In short, $L^2(\mathbb{T}, M(\mathbb{T}), \mu)$ denotes the collection of all equivalence classes of Lebesgue measurable, square integrable functions from $\mathbb{T}$ to $\mathbb{C}$ with respect to the equivalence relation $\mu$ almost-everywhere.

In the following, we take the second construction method and for notational simplicity, we use $L^2(\mathbb{T})$ rather than $L^2(\mathbb{T}, M(\mathbb{T}), \mu)$.

In order to show $(L^2(\mathbb{T}), \langle\cdot,\cdot\rangle)$ is a Hilbert space, we need to show that $(L^2(\mathbb{T}), \|\cdot\|)$ is a Banach space, where $\|\cdot\|$ is the norm induced by the inner product. This is a direct consequence of the Riesz-Fischer Theorem.

**Theorem 1.1 (Riesz-Fischer Theorem)** Let $E \subset \mathbb{R}$ be measurable and $1 \leq p \leq \infty$. Then $L^p(E)$ is a Banach space[2].

We finish this section by arriving at the conclusion:

$$(L^2(\mathbb{T}), \langle\cdot,\cdot\rangle) \text{ is a Hilbert space.}$$

[1] For the definition and validity of completion of a normed space, please see completion.

[2] The proof for one dimensional case can be found in *Real Analysis* by Royden Fitzpatrick, the $4^{th}$ edition, section 7.3. One-dimensional proof is sufficient for your treatment.

# 2 Completeness of Fourier Series in $L^2$

## 2.1 What is Completeness?

Suppose we are given the collection of functions $\left\{e_n = \frac{1}{\sqrt{2\pi}}e^{inx}\right\} \subset L^2(\mathbb{T})$ and an arbitrary $f \in L^2(\mathbb{T})$. Let's take the inner product of f with $e_n$, this gives a collection of complex numbers $\{\langle f, e_n\rangle\}$.

**Definition 2.1 (Parseval's Identity)**

$$\sum_{n=1}^{\infty} \langle f, e_n\rangle^2 = \|f\|^2.$$

Then, we can define what is called the completeness of a collection of functions.

**Definition 2.2 (Completeness in $L^2$)** An infinite collection of functions $\{x_n\}$ is said to be *complete* in $L^2$ if for any $f \in L^2$, the Parseval's identity holds.

Hence, to show that $\{e_n\}$ is complete in $L^2$, we only need to prove the collection of functions $\{e_n(x)\}_{n=1}^{\infty}$, $e_n = \frac{1}{\sqrt{2\pi}}e^{inx}$ forms an orthonormal basis of $L^2(\mathbb{T})$.

## 2.2 Proof of Completeness

Let us first introduce some basic definitions.

**Definition 2.3 (Orthogonality)** $f, g$ in a Hilbert space are said to be *orthogonal* if

$$\langle f, g\rangle = 0.$$

**Notation.** if $f$ and $g$ are orthogonal, we denote $f \perp g$.

**Definition 2.4 (Orthonormal)** A sequence $\{x_n\}$ in an inner product space is said to be *orthonormal* if they are orthogonal to each other (i.e. $x_i \perp x_j, \forall i \neq j$) and $\|x_n\| = 1, \forall n$.

**Definition 2.5 (Fourier Coefficients)** The *Fourier Coefficients* of a vector $x \in V$ with respect to an orthonormal collection $B$ is the set $\{\langle x, b\rangle : b \in B\}$.

**Proposition 2.1** $\{e_n\}_{n=1}^{\infty}$ is orthonormal.

*Proof.*

$$\langle e_n, e_m \rangle = \int_{\mathbb{T}} \overline{\frac{1}{\sqrt{2\pi}} e^{imx}} \frac{1}{\sqrt{2\pi}} e^{inx} d\mu = \frac{1}{2\pi} \int_0^{2\pi} e^{i(n-m)x} dx = \begin{cases} 1 & m = n \\ 0 & m \neq n \end{cases}$$

□

We are half done for our objective. The remaining is to show that $\{e_n\}_{n=1}^{\infty}$ is complete. That is, for any $f \in L^2(\mathbb{T})$, we have the following identity

$$\|f\|^2 = \sum_{n=1}^{\infty} |A_n|^2.$$

Where $A_n = \langle f, e_n \rangle$.

Not surprisingly, if the given collection of vectors is $\{e_n\}$, then the coefficients that give the best approximation are the *Fourier Coefficients*.

**Theorem 2.1 (Least Square Approximation)** Let $(V, \langle \cdot, \cdot \rangle)$ be a Hilbert space. Let $\{x_n\} \subset V$ be a collection of orthogonal vectors. $N \in \mathbb{N}$ is fixed. If $\|f\| < \infty$, then the Fourier coefficients minimize

$$\|f - S_N(f)\|.$$

*Proof. The error term is given by*

$$E_N^2 = \|f - \sum_{n=1}^{N} c_n x_n\|^2 = \int_{\mathbb{T}} |f - \sum_{n=1}^{N} c_n x_n|^2 d\mu.$$

*Expanding the square, we have*

$$E_N^2 = \|f\|^2 - \sum_{n=1}^{N} c_n \langle f, x_n \rangle - \sum_{n=1}^{N} \overline{c_n} \langle x_n, f \rangle + \sum_{n=1}^{N} \sum_{m=1}^{N} c_n \overline{c_m} \langle x_n, x_m \rangle.$$

*By the orthogonality,*

$$\sum_{n=1}^{N} \sum_{m=1}^{N} c_n \overline{c_m} \langle x_n, x_m \rangle = \sum_{n=1}^{N} c_n^2 \|x_n\|^2.$$

*By completing the square, we have*

$$E_N^2 = \sum_{n=1}^{N} \|x_n\|^2 \left[c_n - \frac{\langle f, x_n\rangle}{\|x_n\|^2}\right]^2 + \|f\|^2 - \sum_{n=1}^{N} \frac{\langle f, x_n\rangle}{\|x_n\|^2}.$$

*It is obvious that the above is minimized if and only if* $c_n = \frac{\langle f, x_n\rangle}{\|x_n\|^2}$*, which is indeed the Fourier coefficient if* $\{x_n\}$ *is chosen to be orthonormal.* $\square$

Directly, we derive a useful corollary from the above theorem.

**Corollary 2.1 (Bessel's Inequality)** Let $\{x_n\}$ be orthonormal in an inner product space $(V, \langle\cdot,\cdot\rangle)$. Then, for any $x \in V$,

$$\sum_{n=1}^{\infty} \langle x, x_n\rangle^2 \leq \|x\|^2. \tag{1}$$

*Proof.* *In the proof of least square approximation, if we substitute* $c_n$ *to be Fourier coefficients, then*

$$0 \leq E_N^2 = \|f\|^2 - \sum_{n=1}^{N} \langle f, x_n\rangle, \ \forall N \in \mathbb{N}.$$

*This is equivalent to*

$$\|f\|^2 \geq \sum_{n=1}^{N} \langle f, x_n\rangle, \ \forall N \in \mathbb{N}.$$

*By sending* $N \to \infty$,

$$\|f\|^2 \geq \sum_{n=1}^{\infty} \langle f, x_n\rangle.$$

$\square$

We present the main theorem here. The proof is hard and is based on several lemmas.

**Theorem 2.2** $\{e_n\}_{n=1}^{\infty}$ is an orthonormal basis for $L^2(\mathbb{T})$.

**Lemma 2.1 (Pythagorean Theorem)** If $\{v_n\}_{n=1}^{\infty}$ is a sequence of orthogonal vectors in Hilbert space $V$, then

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

*Proof.*

$$\|\sum_{n=1}^{\infty} v_n\|^2 = \langle\sum_{n=1}^{\infty} v_n, \sum_{m=1}^{\infty} v_m\rangle = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\langle v_n, v_m\rangle = \sum_{n=1}^{\infty}\langle v_n, v_n\rangle = \sum_{n=1}^{\infty}\|v_n\|^2.$$

□

**Remark.** This is the generalization of the so-called *Pythagorean Theorem* one might have learned in primary school geometry. An important thing should be mentioned here is the proof implicitly applies *the continuity of inner product*[3].

**Lemma 2.2 ($C^1 \Rightarrow$ Pointwise Convergence)** The classical Fourier series converges to $f(x)$ pointwisely on $(a, b)$ provided that $f(x) \in C^1[a, b]$.

*Proof.* *We assume that $l = \pi$, which can easily be arranged through a change of scale. Thus the Fourier series is*

$$f(x) = \frac{1}{2}A_0 + \sum_{n=1}^{\infty}(A_n \cos nx + B_n \sin nx).$$

*The $N$th partial sum is*

$$S_N(x) = \frac{1}{2}A_0 + \sum_{n=1}^{N}(A_n \cos nx + B_n \sin nx).$$

*Plug the coefficients $A_n, B_n$ into the partial sum, we have*

$$S_N(x) = \int_{-\pi}^{\pi} K_N(x-y)f(y)\frac{dy}{2\pi},$$

*where $K_N(\theta) = 1 + 2\sum_{n=1}^{N}\cos n\theta$. Note that*

$$K_N(\theta) = \frac{\sin(N+\frac{1}{2})\theta}{\sin\frac{1}{2}\theta}.$$

*This can be proved by complexification. Hence,*

$$S_N(x) - f(x) = \int_{-\pi}^{\pi} K_N(\theta)[f(x+\theta) - f(x)]\frac{d\theta}{2\pi},$$

[3] For an explanation and a proof of continuity of inner product, please see continuity of inner product.

*or*

$$S_N(x) - f(x) = \int_{-\pi}^{\pi} g(\theta) sin[(N+\frac{1}{2})\theta]\frac{d\theta}{2\pi},$$

*where*

$$g(\theta) = \frac{f(x+\theta)-f(x)}{\sin\frac{1}{2}\theta}.$$

*Notice that the functions*

$$\phi_N(\theta) = \sin[(N+\frac{1}{2})\theta],\ N \in \mathbb{N}$$

*forms an orthogonal set. Therefore, by Bessel's inequality,*

$$\sum_{N=1}^{\infty} \frac{|\langle g, \phi_N\rangle|^2}{\|\phi_N\|^2} \le \|g\|^2.$$

*Since* $\|\phi_N\|^2 = \pi$ *and* $\|g\| < \infty$, *we have* $\langle g, \phi_N\rangle \to 0$ *and the proof is complete.* □

**Lemma 2.3 ($C^1 \Rightarrow$ Uniform Convergence)** Suppose $f \in C^1(\mathbb{T})$ s.t. B.C., then $S_N(f)$ converges uniformly to $f$ on $\mathbb{T}$.

*Proof.* *We only prove this when the collection of orthonormal vectors is given by* $\{e_n\}$. *The Fourier coefficient is given by*

$$\begin{aligned} c_n &= \int_{\mathbb{T}} \frac{1}{\sqrt{2\pi}} f e^{inx} d\mu \\ &= \frac{1}{in\sqrt{2\pi}} \int_0^{2\pi} f e^{inx} dx \\ &= \frac{1}{in} \int_{\mathbb{T}} \frac{1}{\sqrt{2\pi}} f' e^{inx} d\mu = \frac{1}{in} c_n', \end{aligned}$$

*where* $c_n'$ *stands for the Fourier coefficient of* $f'$. *Note that since* $f \in C^1(\mathbb{T})$, *we know that* $S_N(f) \to f$ *pointwisely. Also, since* $f \in C^1(\mathbb{T})$, *we have* $\|f\| = \left(\int_{\mathbb{T}} |f|^2 d\mu\right)^2 < \infty$.

$$\begin{aligned} \sup|f - S_N(f)| &\le \sup \sum_{n=N+1}^{\infty} |c_n e_n| \\ &= \frac{1}{n\sqrt{2\pi}} \sum_{n=N+1}^{\infty} |c_n'| \\ &\le \left(\sum_{n=N+1}^{\infty} \frac{1}{2n^2\pi}\right)^{1/2} \left(\sum_{n=N+1}^{\infty} |c_n'|^2\right)^{1/2} \to 0,\ as\ N \to \infty. \end{aligned}$$

*This is because* $\sum_{n=1}^{\infty}\frac{1}{2n^2\pi}<\infty$ *and by Bessel's inequality 2.1,* $\sum_{n=1}^{\infty}|c_n'|^2<\|f\|^2<\infty$. *This shows that* $S_N(f)\to f$ *uniformly.*

**Remark.** The requirement can be weaken to $f$ is continuous but $f'$ is with countably many discontinuities, since a function with countably many discontinuities is still Riemann-integrable.

**Lemma 2.4 (Approximation)** $C_0^\infty(\mathbb{T})$ is dense in $L^2(\mathbb{T})$.

*Proof.* *The proof is given in the appendix 5.1.* □

**Remark.** Suppose $U\subset V$ is a subspace of $V$ and $V$ is equipped with a metric $d$. By denseness we mean that for any $v\in V$ and $\forall\epsilon>0$, there exists a $u\in U$ such that $d(u,v)<\epsilon$. A concrete example is $\mathbb{Q}$ *is dense in* $\mathbb{R}$, *where* $\mathbb{R}$ *is equipped with the usual Euclidean norm* $|\cdot|$.

*Proof of the main theorem.*
Fix a $\epsilon>0$. Note that by *lemma 2.4*, we can find a $g\in C_0^\infty(\mathbb{T})$ such that

$$\|f-g\|<\frac{\epsilon}{3}.$$

The distance between $f$ and $S_N(f)$ can be decomposed into three parts, we try to control each part respectively.

$$\|f-S_N(f)\|=\|f-g\|+\|g-S_N(g)\|+\|S_N(g)-S_N(f)\|.$$

Since $g$ is infinitely differentiable, by invoking *lemma 2.3*, we know that $S_N(g)\to g$ uniformly, thus, there exists a $K\in\mathbb{N}$ such that for $\forall N\geq K$,

$$\begin{aligned}\|g-S_N(g)\|&=\left(\int_{\mathbb{T}}|g-S_N(g)|^2d\mu\right)^{1/2}\\&\leq\left(\sup_{[0,2\pi]}|S_N(g)-g|^2\right)^{1/2}\left(\int_0^{2\pi}1dx\right)^{1/2}\\&=\sqrt{2\pi}\left(\sup_{[0,2\pi]}|S_N(g)-g|^2\right)^{1/2}<\frac{\epsilon}{3}.\end{aligned}$$

For the third term, since $S_N(f-g) \perp [(f-g) - S_N(f-g)]$, by *lemma 2.1*, we have

$$\|f-g\| = \|S_N(f-g)\| + \|(f-g) - S_N(f-g)\|$$

This implies

$$\|S_N(f) - S_N(g)\| = \|S_N(f-g)\| \leq \|f-g\| < \frac{\epsilon}{3}.$$

Thus, for any $\epsilon > 0$, there exists a $K \in \mathbb{N}$ such that for all $N \geq K$,

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

This is equivalent to

$$\lim_{N\to\infty} \|f - S_N(f)\| = 0.$$

□

As we have stated before, an infinite collection of orthogonal functions is said to be complete in $L^2$ if the Parseval's identity holds. By the main theorem, we know that $\{e_n\}$ forms an orthonormal basis for $L^2(\mathbb{T})$. If we are given $f \in L^2(\mathbb{T})$, then $f = \sum_{n=1}^{\infty} c_n e_n$, where $c_n$'s are some coefficients to be determined. Finally, we prove the Parseval's identity as a corollary of the previous theorem.

**Corollary 2.2 (Parseval's Identity)** If $f \in L^2(\mathbb{T})$ and $\{x_n\}$ is an orthonormal basis, then the following equation holds

$$\sum_{n=1}^{\infty} \langle f, x_n \rangle^2 = \|f\|^2$$

*Proof.* *By the previous theorem, we know that if $f \in L^2(\mathbb{T})$, then*

$$\lim_{N\to\infty} E_N = \lim_{N\to\infty} \|f - S_N(f)\| = 0.$$

*Thus,*

$$\sum_{n=1}^{\infty} \langle f, x_n \rangle^2 = \|f\|^2.$$

□

**Remark.** The Parseval's identity equates the mean-square norm of the function with a corresponding norm of its Fourier coefficients.

Finally, we give two interesting results based on *Bessel's inequality* and *Parseval's identity.*

Since the terms of a converging series tend to 0, we have the the following results.

**Corollary 2.3 (Riemann-Lebesgue Lemma)** If $f \in L^2(\mathbb{T})$, then $\hat{f}(n) \to 0$ as $|n| \to \infty$.

An equivalent reformulation of this proposition is if $f$ is integrable on $[0, 2\pi]$, then

$$\int_0^{2\pi} f(x)\sin(Nx)dx \to 0, \quad as\ N \to \infty.$$

and

$$\int_0^{2\pi} f(x)\cos(Nx)dx \to 0, \quad as\ N \to \infty.$$

The last lemma is a more general version of the Parseval's identity.

**Lemma 2.5** Suppose $F$ and $G$ are both in $L^2(\mathbb{T})$ with

$$F \sim \sum a_n e^{inx} \quad and \quad G \sim \sum b_n e^{inx}.$$

Then

$$\frac{1}{2\pi}\int_0^{2\pi} F(x)\overline{G(x)}dx = \sum_{n=-\infty}^{\infty} a_n \overline{b_n}.$$

*Proof. The proof follows from Parseval's identity and the fact that*

$$(F, G) = \frac{1}{4}[\|F + G\|^2 - \|F - G\|^2 + i(\|F + iG\|^2 - \|F - iG\|^2)],$$

*which holds in every Hermitian inner product space.* □

## 2.3 Conclusions

- **What is the completeness of Fourier series in $L^2$?**

  This means that the collection of functions $\{e_n\}$ forms an orthonormal basis in $L^2$. That is to say, any function in $L^2$ can be well approximated by a finite linear combination of $e_n$'s.

- **Why is the result useful?**

  The completeness assures us a certain kind of convergence under a very weak assumption on $f$. Under the usual Riemann integral, $f$ is either continuous or piecewise continuous will be good enough. Also, we can use the more general Lebesgue integral, this enables the theorem to be true for much more functions, not even need to be piecewise continuous. Also, the knowledge of the derivative is not required.

## *What's the usefulness in PDE?*

For typical wave or heat equations with initial conditions and boundary conditions.

$$\begin{cases} \partial_{tt} u = c^2 \Delta u \\ u(\boldsymbol{x}, 0) = \phi \,;\, u_t(\boldsymbol{x}, 0) = \psi \\ u|_{\partial\Omega} = 0 \end{cases} \quad ; \quad \begin{cases} \partial_t u = k \Delta u \\ u(\boldsymbol{x}, 0) = \phi \\ u|_{\partial\Omega} = 0 \end{cases}$$

In one-dimensional case, we can use the separation of variables to solve the problems and get a series solution. By the initial condition, we can get the coefficients. If those initial conditions are of good property (here, we mean $\phi, \psi \in L^2$), then we can guarantee that $\phi$ and $\psi$ can truly can written in a Fourier series form and the solution derived by the separation of variables method gives the correct solution.

## 3 Relation with Uniform Convergence and Point-wise Convergence

We have seen three different convergence up to now. Here is a brief summary of pointwise convergence, uniform convergence and $L^2$ convergence. Here, we fix a domain $\Omega$.

- **Pointwise Convergence**: $\sum_{n=1}^{\infty} f_n$ is said to *converge pointwisely* to $f$ if for any $x \in \Omega$, $\left\{\sum_{n=1}^{N} f_n(x)\right\} \to f(x)$. Note that for each fixed $x$, $\left\{\sum_{n=1}^{N} f_n(x)\right\}$ is just a sequence of real (complex) numbers.

- **Uniform Convergence**: $\sup_{x\in\Omega} |\sum_{n=1}^{N} f_n(x) - f(x)| \to 0$, as $n \to \infty$. This means that the maximum difference approches 0.

- $L^2$ **Convergence**: $\int_\Omega |\sum_{n=1}^{N} f_n - f| \to 0$, as $n \to \infty$.

Before we go through all the proofs and examples, let us summarize all the results here.

- On a finite domain, uniform convergence implies pointwise convergence.

- $L^2$ convergence does not imply uniform convergence.

- $L^2$ convergence does not imply pointwise convergence.

**Proposition 3.1** *On a domain $\Omega$ with finite measure, uniform convergence implies $L^2$ convergence.*

*Proof.* *Suppose $S_N(f) \to f$ uniformly on $\Omega$, then*

$$\|f - S_N(f)\|^2 \leq \int_\Omega |S_N(f) - f|^2 d\mu \leq \sup_\Omega |f - S_N(f)|^2 \mu(\Omega) \to 0, \text{ as } N \to \infty.$$

However, $L^2$ convergence does not imply uniform convergence.

**Example 3.1 ($L^2$ convergence does not imply uniform convergence)** *Let $f_n(x) = (1-x)x^{n-1}$ on the interval $0 < x < 1$. The partial sum is*

$$S_N(f) = \sum_{n=1}^{N} (x^{n-1} - x^n) = 1 - x^N \to 1, \text{ as } N \to \infty.$$

*Note that*

$$\int_0^1 |x^N|^2 dx = \frac{1}{2N+1} \to 0, \text{ as } N \to \infty.$$

*However, it does not converge uniformly to* $1$*. The reason is as follows: fix* $\epsilon = 1/2$*, For each* $N \in \mathbb{N}$*, choose* $x \in [\left(\frac{1}{2}\right)^{1/N}, 1)$*, then we have*

$$\sup_{x\in(0,1)} |S_N(f) - f| = \sup_{x\in(0,1)} |x^N| \geq \frac{1}{2} = \epsilon.$$

**Example 3.2 ($L^2$ convergence does not imply pointwise convergence)** *Let*

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

*In the interval* $(0, l)$*,* $\sum_{n=1}^{N} f_n \to 0$ *pointwisely. However,*

$$\int_0^l |\sum_{n=1}^{N} f_n(x)|^2 dx = N \int_0^{Nl} \frac{1}{(1+y^2)^2} dy \to +\infty, \text{ as } N \to \infty.$$

# 4 Gibbs Phenomenon

The *Gibbs Phenomenon* is what happens to Fourier series at jump discontinuities. The key is that *it happens when the convergence is only pointwise, not uniform.* Gibbs showed that near the jump discontinuity, the partial sum $S_N(f)$ always differs from $f$ by an *overshoot* of about 9 percent. Though the width of the overshoot goes to 0 as $N$ goes to $\infty$, the extra height remains at 9 percent of the jump. That is to say

$$\lim_{N\to\infty} |S_N(f)(x) - f(x)| \neq 0,$$

when $x$ is near the jump discontinuity. However, for $x$ that does not jump, $S_N(f)(x) \to f(x)$.

We now use a concrete example to illustrate the Gibbs phenomenon.

**Example 4.1**

$$f(x) = \begin{cases} \frac{1}{2} & 0 < x < \pi \\ -\frac{1}{2} & -\pi < x < 0 \end{cases}$$

*$f$ has the Fourier series*

$$\sum_{n=1,odd}^{\infty} \frac{2}{n\pi} sin(nx).$$

*(Figures 1,2,3,4,5 (see appendix 5.2) are sketches of $S_N$ when $N = 5, 16, 30, 50, 100$ respectively.)*

*Note that the partial sum can be rewritten as*

$$S_N(x) = (\int_0^{\pi} - \int_{-\pi}^{0}) K_N(x-y)\frac{dy}{4\pi} = (\int_0^{\pi} - \int_{-\pi}^{0}) \frac{sin[(N+\frac{1}{2})(x-y)]}{sin(\frac{1}{2}(x-y))}\frac{dy}{4\pi}.$$

*Let $M \triangleq N + \frac{1}{2}$. In the first integral, set $\theta = M(x-y)$. In the second integral, set $\theta = M(y-x)$. Then*

$$S_N(x) = (\int_{-Mx}^{Mx} - \int_{M\pi - Mx}^{M\pi + Mx}) \frac{\sin\theta}{2M\sin(\theta/2M)}\frac{d\theta}{2\pi}.$$

*We are interested in what happens near the jump, i.e. near $x = 0$. Note that the first integral in the last expression is larger since it has a smaller denominator. By first derivative test,* *the first integral is maximized when* $\sin Mx = 0$. *So we set* $x = \frac{\pi}{M}$. *Then we have*

$$S_N(\frac{\pi}{M}) = (\int_{-\pi}^{\pi} - \int_{M\pi-\pi}^{M\pi+\pi}) \frac{\sin\theta}{2M\sin(\theta/2M)} \frac{d\theta}{2\pi}.$$

*Note also*

$$\frac{\pi}{4} < [1 - \frac{1}{M}]\frac{\pi}{2} \leq \frac{\theta}{2M} \leq [1 + \frac{1}{M}]\frac{\pi}{2} < \frac{3\pi}{4}.$$

*for* $M > 2$. *Hence* $\sin(\theta/2M) > 1/\sqrt{2}$, *so that the second integral is less than*

$$\int_{M\pi-\pi}^{M\pi+\pi} 1 \cdot [\frac{2M}{\sqrt{2}}]^{-1} \frac{d\theta}{2\pi} = \frac{1}{\sqrt{2}M},$$

*which tends to zero as* $M \to \infty$.

*On the other hand, inside the first integral we have* $|\theta| \leq \pi$ *and*

$$2M \sin\frac{\theta}{2M} \to \theta, \text{ uniformly in } -\pi < \theta < \pi \text{ as } M \to \infty.$$

*Hence,*

$$\lim_{N\to\infty} \int_{-\pi}^{\pi} \frac{\sin\theta}{d\theta} 2\pi \simeq 0.5 + \boldsymbol{1 \ * \ 0.09}.$$

*This is the Gibbs's 9 percent overshoot.*

# 5 Appendix

## 5.1 Proof of Lemma 2.4

To prove this approximation theorem, we first define the *Schwarz Space*.

**Definition 5.1 (Schwarz Space)** The class $\mathbb{S}$ consists of functions $f$ on $\mathbb{R}^d$ that are smooth (indefinitely differentiable) and such that for each multi-index $\alpha$ and $\beta$, the function $x^\alpha(\frac{\partial}{\partial x})^\beta$ is bounded on $\mathbb{R}^d$.[4]

**Theorem 5.1** The space $\mathbb{S}(\mathbb{R}^d)$ is dense in $L^2(\mathbb{R}^d)$. In other words, given any $f \in L^2(\mathbb{R}^d)$, there exists a sequence $\{f_n\} \subset \mathbb{S}(\mathbb{R}^d)$ such that

$$\|f - f_n\|_{L^2(\mathbb{R}^d)} \to 0, \ as \ n \to \infty$$

*Proof.* *For the proof of the lemma, we fix $f \in L^2(\mathbb{R}^d)$ and $\epsilon > 0$. Then, for each $M > 0$, we define*

$$g_M(x) = \begin{cases} f(x) & |x| \leq M \ and \ |f(x)| \leq M, \\ 0 & Otherwise. \end{cases}$$

*Then, $|f(x) - g_M(x)| \leq 2|f(x)|$, hence $|f(x) - g_M(x)|^2 \leq 4|f(x)|^2$, and since $g_M(x) \to f(x)$ as $M \to \infty$ for almost every $x$, then dominated convergence theorem guarantees that for some $M$, we have*

$$\|f - g_M\|_{L^2(\mathbb{R}^d)} < \epsilon.$$

*We write $g = g_M$, note that this function is bounded and supported on a bounded set, and observe that it now suffices to approximate $g$ by functions in the Schwartz space. TO achieve this goal, we use a method called* ***regularization****, which consists of smoothing $g$ by convolving it with an approximation of the identity. Consider a function $\psi(x)$ on $\mathbb{R}^d$ with the following properties:*

- *$\psi$ is smooth.*
- *$\psi$ is supported in the unit ball.*
- *$\psi \geq 0$.*

[4]Recall that $x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_d^{\alpha_d}$ and $(\frac{\partial}{\partial x})^\beta = (\frac{\partial}{\partial x_1})^{\beta_1}(\frac{\partial}{\partial x_2})^{\beta_2} \cdots (\frac{\partial}{\partial x_d})^{\beta_d}$.

- $\int_{\mathbb{R}^d} \psi(x)dx = 1$.

*For instance, one can take*

$$\psi(x) = \begin{cases} ce^{-\frac{1}{1-|x|^2}}, \ |x| < 1, \\ 0, \ |x| \geq 1. \end{cases}$$

*where the constanct $c$ is chosed so that the forth property is satisfied.*
*Next, we consider the approximation to the identity defined by*

$$K_\delta(x) = \delta^{-d}\psi(x/\delta).$$

*The key observation is that $g * K_\delta$ belongs to $\mathbb{S}(\mathbb{R}^d)$, with this convolution in fact bounded and supported on a fixed bounded set, uniformly in $\delta$ (assuming for example that $\delta \leq 1$). Indeed, we may write*

$$(g * K_\delta)(x) = \int g(y)K_\delta(x-y)dy = \int g(x-y)K_\delta(y)dy.$$

*We note that since $g$ is supported on some bounded set and $K_\delta$ vanishes outside the ball of radius $\delta$, the function $g * K_\delta$ is supported in some fixed bounded set independent of $\delta$. Also, the functionn $g$ is bounded by construction, hence*

$$|(g * K_\delta(x))| \leq \int |g(x-y)||K_\delta(y)|dy \leq \sup_{z\in\mathbb{R}^d} |g(z)|.$$

*which shows that $g * K_\delta$ is also uniformly bounded in $\delta$. Moreover, from the first integral expression for $g * K_\delta$ above, one may differentiate under the integral sign to see that $g * K_\delta$ is smooth and all of its derivatives have support in some fixed bounded set.*
*The proof of the lemma will be complete if we can show that $g * K_\delta$ converges to $g$ in $L^2(\mathbb{R}^d)$. Now Theorem 2.1 in Chapter 3*[5] *guarantees that for almost every $x$, the quantity $|(g*K_\delta)(x)-g(x)|^2$ converges to $0$ as $\delta$ tends to $0$. An application of the bounded convergence theorem yields*

$$\|(g * K_\delta) - g\|^2_{L^2(\mathbb{R}^d)} \to 0, \ as \ \delta \to 0.$$

[5] Since the proof requires a lot of prerequisite knowledge and might be quite lengthy, please refer to *Real analysis* by Stein.

*In particular,* $\|(g*K_\delta)-g\|_{L^2(\mathbb{R}^d)} < \epsilon$ *for an appropriate* $\delta$ *and hence* $\|f-g*K_\delta\|_{L^2(\mathbb{R}^d)} < 2\epsilon$, *and choosing a sequence of* $\epsilon$ *tending to zero gives the construction of the desired sequence* $\{f_n\}$. $\square$

**Remark.** In the proof of the main theorem in this text, we pick a function in $C_0^\infty(\mathbb{T})$. Indeed, we could choose one from $\mathbb{S}(\mathbb{T})$, which can also achieve the same effect.

## 5.2 Figures

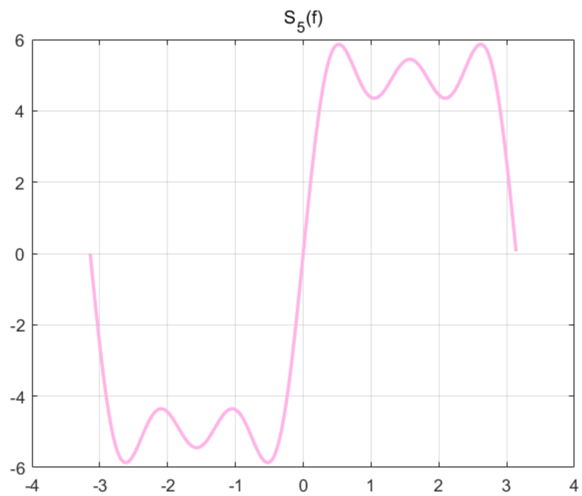


Figure 1: $S_N(f),\ N = 5$

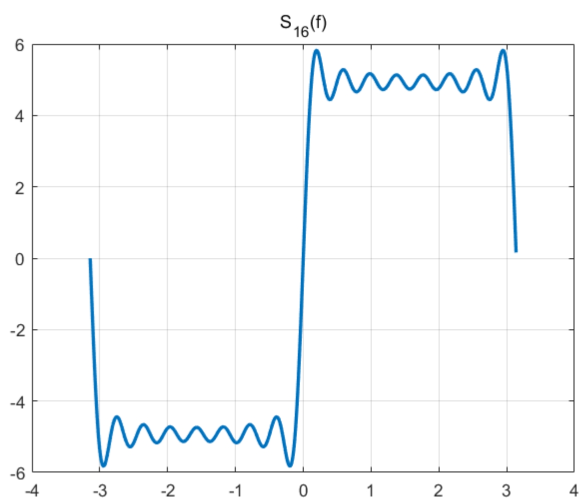


Figure 2: $S_N(f),\ N = 16$

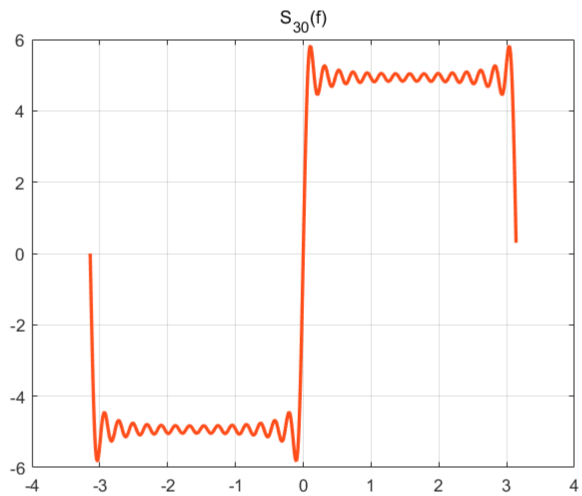


Figure 3: $S_N(f),\ N = 30$

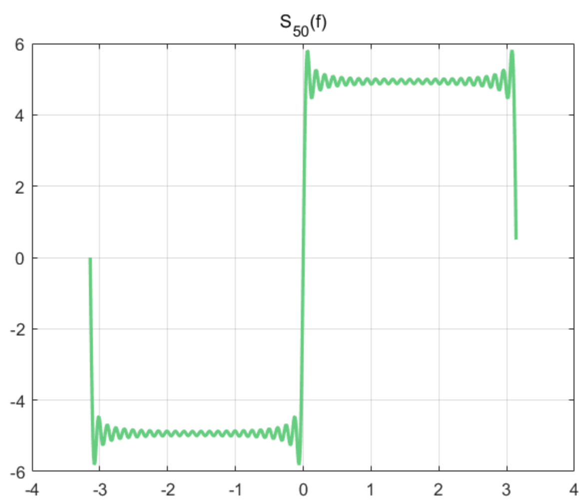


Figure 4: $S_N(f),\ N = 50$

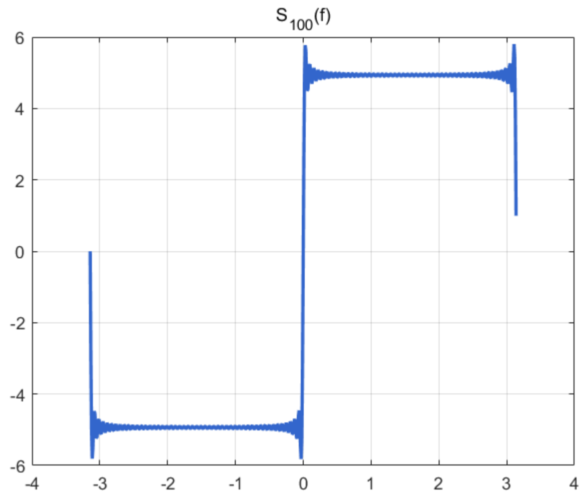


Figure 5: $S_N(f),\ N = 100$

## Reference Materials

[1] An online note from UCD.

[2] An online note on $L^2$ convergence.

[3] An online note form The University of Arizona.

[4] An online note from MIT openwares.

[5] Royden, Halsey Lawrence, and Patrick Fitzpatrick. *Real Analysis*. Vol. 32. New York: Macmillan, 1988.

[6] Strauss, Walter A. *Partial differential equations: An introduction*. John Wiley & Sons, 2007.

[7] Athreya, Krishna B., and Soumendra N. Lahiri. *Measure theory and probability theory*. Vol. 19. New York: Springer, 2006.

[8] Bogachev, Vladimir Igorevich, and Maria Aparecida Soares Ruas. *Measure theory*. Vol. 1. Berlin: springer, 2007.

[9] Stein, Elias M., and Rami Shakarchi. *Fourier analysis: an introduction*. Vol. 1. Princeton University Press, 2011.

[10] Stein, Elias M., and Rami Shakarchi. *Real analysis*. Princeton University Press, 2009.