

Convergence of Fourier series and Kolmogorov's theorem

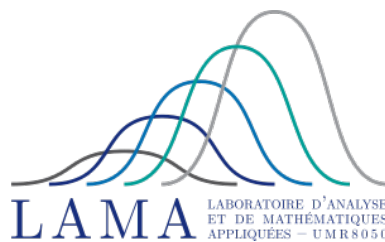
Master 1 Internship

Tooryanand Seetohul

Supervisor: Dr Chenmin Sun



**Université
de Rennes**



August 29, 2023

Acknowledgements

I would like to thank Dr Chenmin Sun for supervising me and helping me understand various aspects of harmonic analysis during this internship. I am grateful for his advice and guidance since I consider this internship as I move forward into my second year of masters and continue in my mathematical career.

Introduction

This document is a compilation of the materials covered during my Master 1 internship under Dr. Chenmin Sun. The goal of this internship was to understand Kolmogorov's proof for almost-everywhere divergence of L^1 Fourier series as well as to bridge my gap of knowledge in the field of Harmonic Analysis.

The basis of the internship was Volume 1 of the book Classical and Multilinear Harmonic Analysis by Camil Muscalu and Wilhelm Schlag [1]. The document closely follows the contents of Chapter 1, 2, 5 and 6.1 of the book, and it holds most of the proofs that I presented to Dr. Sun during our weekly sessions. Chapter 5 pertains to probability theory and it was reviewed to help consolidate my lack of mastery of probability theory.

Note that some theorems are not explicated since I had either already encountered it in my master 1 courses or the demonstrations are already fairly explicit in the book.

Finally, the last section details of the proof proposed by Kolmogorov in 1923 for an almost everywhere divergence of L^1 Fourier series.

1 Fourier series: Convergence and summability

1.1 Basics: Partial Sum and Dirichlet Kernel

Given any measure $\mu \in \mathcal{M}(\mathbb{T})$, the Fourier series of μ is given by

$$\sum_{n=-\infty}^{\infty} \hat{\mu}(n) e(nx)$$

where $\hat{\mu}(n) := \int_0^1 \mu(x) e(-nx) dx$ and is called the n^{th} Fourier coefficient of μ . For any $\mu(x) = f(x) dx$ with $f \in L^1(\mathbb{T})$, we write $\hat{f}(n)$ instead of $\hat{\mu}(n)$.

Note the orthogonality relation which will prove to be important in studying Fourier coefficients:

$$\int_{\mathbb{T}} e_n(x) \overline{e_m(x)} dx = \delta_0(n - m)$$

To study the Fourier series of f , we define the Dirichlet Kernel $D_N(x)$.

The partial sum of the Fourier series of a function $f \in L^1(\mathbb{T})$ is given by :

$$\begin{aligned} S_N f(x) &= \sum_{n=-N}^N \hat{f}(n) e(nx) \\ &= \sum_{n=-N}^N \int_0^1 f(y) e(-ny) dy e(nx) \\ &= \int_0^1 \sum_{n=-N}^N f(y) e(-ny) e(nx) dy \\ &= \int_0^1 f(y) \sum_{n=-N}^N e(n(x-y)) dy \\ &= \int_0^1 f(y) \sum_{n=-N}^N D_N(x-y) dy \\ &= D_N * f(x) \end{aligned}$$

with the $D_N(x) = \sum_{n=-N}^N e(nx)$.

To better understand the properties of the Dirichlet Kernel, we derive the form of the function and give its properties.

Theorem 1 (Dirichlet's Kernel). • For any integer $N \geq 0$,

$$D_N(x) = \frac{\sin((2N+1)\pi x)}{\sin(\pi x)}$$

- The pointwise bounds

$$|D_N(x)| \leq C \min(N, \frac{1}{|x|})$$

holds true for any integer $N \geq 1$ and an absolute constant C .

- There exist $L^1(\mathbb{T})$ bounds:

$$A^{-1} \log(N) \leq \|D_N(x)\|_{L^1(\mathbb{T})} \leq A \log N$$

for $N \geq 2$ and another absolute constant A .

Proof. Rearranging the terms of the finite sum, one has $D_N(x) = \sum_{n=-N}^N e(nx) = 1 + \sum_{n=1}^N [e(nx) + e(-nx)]$. Using the identity $e^{i\theta} + e^{-i\theta} = 2\cos(\theta)$, we have $D_N(x) = 1 + \sum_{n=1}^N 2\cos(2n\pi x)$. Multiplying the latter relationship throughout with $\sin(\pi x)$ to have

$$\sin(\pi x)D_N(x) = \sin(\pi x) + \sum_{n=1}^N 2\sin(\pi x)\cos(2n\pi x)$$

Using the trigonometric identity $2\sin(\theta)\cos(\alpha) = \sin(\theta + \alpha) + \sin(\theta - \alpha)$, we have:

$$\begin{aligned} \sin(\pi x)D_N(x) &= \sin(\pi x) + \sum_{n=1}^N [\sin((2n+1)\pi x) + \sin((1-2n)\pi x)] \\ &= \sin(\pi x) + \sum_{n=1}^N [\sin((2n+1)\pi x) - \sin((2n-1)\pi x)] \quad (\sin(-t) = -\sin(t)) \end{aligned}$$

Noticing the resulting telescoping sum, we have:

$$= \sin((2N+1)\pi x)$$

and the result follows.

Now, we adress the pointwise bounds.

Fix an infinitesimal $\epsilon > 0$.

Observe that $\sin(\pi x) = 0$ for $x = 0, 1$, and consider the region around $x = 0$ and $x = 1$ given by $(0, \epsilon) \cup (1 - \epsilon, 1)$. In this region, $\sin(x) \approx x$ and $\sin((2N+1)\pi x) \approx (2N+1)\pi x$ such that

$$D_N(x) = \frac{(2N+1)\pi x}{\pi x} = \frac{(2N+1)\pi x}{\pi x} = 2N+1 < C_1 N$$

for a constant C_1 .

Now, consider $(0, \epsilon)^c \cap (1 - \epsilon, 1)^c$. We will use these two following inequalities:

$$\sin((2N+1)\pi x) \leq 1$$

Using the concavity of the sin function, one has $|\sin(\pi x)| \geq \frac{|\pi x|}{2}$ such that $\frac{1}{|\sin(\pi x)|} \leq \frac{2}{\pi|x|}$. Combining both inequalities, we have

$$\frac{\sin((2N+1)\pi x)}{\sin(\pi x)} \leq \frac{1}{\sin(\pi x)} \leq C_2 \frac{1}{|x|}$$

for a choice of constant C_2 . Take $C = \max(C_1, C_2)$ to have the required pointwise inequality $|D_N(x)| \leq C \min(N, \frac{1}{|x|})$.

For the L^1 bounds, we work with the domain $[-\pi, \pi]$ instead of $[0, 1]$ by using the diffeomorphism $x \mapsto 2\pi x - \pi$ as well as the relationship $\sin(x) = -\sin(x - \pi)$ to have

$$\int_{[0,1]} \frac{|\sin((2N+1)\pi x)|}{|\sin(\pi x)|} dx = \frac{1}{2\pi} \int_{[-\pi, \pi]} \frac{|\sin((N + \frac{1}{2})y)|}{|\sin(\frac{y}{2})|} dy$$

We can now benefit from the symmetry of the $|\sin(t)|$ function over the interval $t \in [-\pi, \pi]$ to have

$$\frac{1}{2\pi} \int_{[-\pi, \pi]} \frac{|\sin((N + \frac{1}{2})y)|}{|\sin(\frac{y}{2})|} dy = \frac{1}{\pi} \int_{[0, \pi]} \frac{|\sin((N + \frac{1}{2})y)|}{|\sin(\frac{y}{2})|} dy$$

We can now prove the lower bound. Divide the interval $[0, \pi]$ into $\bigcup_{k=0}^{2N} [\frac{\pi k}{2N+1}, \frac{\pi(k+1)}{2N+1}]$. Since $|\sin(\frac{y}{2})| \leq |\frac{y}{2}|$, we have:

$$\sum_{k=0}^{2N} \frac{1}{\pi} \int_{[\frac{\pi k}{2N+1}, \frac{\pi(k+1)}{2N+1}]} \frac{|\sin((N + \frac{1}{2})y)|}{|\sin(\frac{y}{2})|} dy \leq \sum_{k=0}^{2N} \frac{2}{\pi} \int_{[\frac{\pi k}{2N+1}, \frac{\pi(k+1)}{2N+1}]} \frac{|\sin((N + \frac{1}{2})y)|}{|y|} dy$$

On $[\frac{k\pi}{2N+1}, \frac{(k+1)\pi}{2N+1}]$, we have $\frac{1}{|y|} \geq \frac{(2N+1)}{\pi(k+1)}$ which gives us the following inequality:

$$\sum_{k=0}^{2N} \frac{2}{\pi} \int_{[\frac{\pi k}{2N+1}, \frac{\pi(k+1)}{2N+1}]} \frac{|\sin((N + \frac{1}{2})y)|}{|y|} dy \geq \sum_{k=0}^{2N} \frac{2(2N+1)}{\pi(k+1)} \int_{[\frac{\pi k}{2N+1}, \frac{\pi(k+1)}{2N+1}]} |\sin((N + \frac{1}{2})y)| dy$$

For a fixed $k \in \mathbb{Z}$,

$$\begin{aligned} \int_{[\frac{\pi k}{2N+1}, \frac{\pi(k+1)}{2N+1}]} |\sin((N + \frac{1}{2})y)| dy &= 2 \\ \sum_{k=0}^{2N} \frac{2(2N+1)}{\pi(k+1)} \int_{[\frac{\pi k}{2N+1}, \frac{\pi(k+1)}{2N+1}]} |\sin((N + \frac{1}{2})y)| dy &= \sum_{k=0}^{2N} \frac{4(2N+1)}{\pi(k+1)} \\ &= \frac{4(2N+1)}{\pi} \sum_{k=0}^{2N} \frac{1}{(k+1)} \\ &\geq A_1 \log(N) \end{aligned}$$

for an appropriate constant A_1 .

For the upper L^1 bounds, it suffices to use the pointwise upper bounds of the Dirichlet's kernel calculated before.

$$\begin{aligned} \frac{1}{2\pi} \int_{[-\pi, \pi]} \frac{|\sin((N + \frac{1}{2})y)|}{|\sin(\frac{y}{2})|} dy &= \frac{1}{\pi} \int_{[0, \pi]} \frac{|\sin((N + \frac{1}{2})y)|}{|\sin(\frac{y}{2})|} dy \\ &\leq \frac{1}{\pi} \int_{[0, \frac{\pi}{5N}]} N dy + \frac{1}{\pi} \int_{[\frac{\pi}{5N}, \pi]} \frac{1}{|x|} dy \\ &\leq A_2 \log N \end{aligned}$$

for an appropriate constant A_2 . Choose $A = \max(A_1, A_2)$ to have the appropriate absolute constant.

□

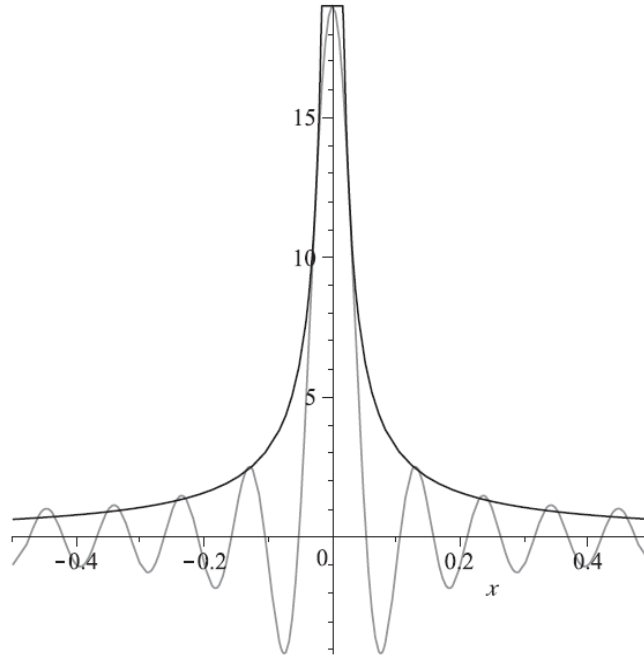


Figure 1.1. The Dirichlet kernel D_N and the upper envelope $\min(2N + 1, |\pi x|^{-1})$ for $N = 9$. See Exercise 1.1.

Figure 1: $D_N(x)$ for $N = 9$

The main goal of this internship is to study the convergence of the Fourier series and hence the convergence of $S_N f(x)$. Note that we already have these answers respective to different norms and classes of functions.

Theorem 2. If $f \in C^\alpha(\mathbb{T})$ with $0 < \alpha \leq 1$, then $\|S_N f - f\|_\infty \rightarrow 0$ as $N \rightarrow \infty$.

The heavily oscillatory nature from Figure 1 and the lower L_1 bounds from Theorem 1 of the Dirichlet Kernel hints towards the fact that the Dirichlet kernel might not be a good kernel. However, if one were to cheat a little, we can obtain a kernel with nice properties.

1.2 Fejér Kernel $K_N(x)$

Consider the Fourier coefficient of the Dirichlet's kernel: $\hat{D}_N(k) = \sum_{n=-N}^N \int_0^1 e((n-k)x)$. Using the orthogonality relation, one has $\hat{D}_N(k) = 1$ if $|k| \leq n$ and 0 if $|k| > n$.

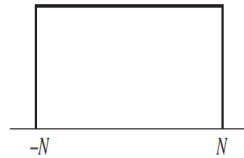


Figure 2: Fourier transform of $D_N(x)$

To get around the discountinuity, consider the Cesàro means $\sigma_N f := \frac{1}{N} \sum_{n=0}^{N-1} S_n f(x) = \frac{1}{N} \sum_{n=0}^{N-1} D_n * f(x)$ which gives the following kernel

$$K_N(x) := \frac{1}{N} \sum_{n=0}^{N-1} D_n(x)$$

Theorem 3. *The Fejér kernel has the following properties:*

$$\begin{aligned}\hat{K}_N(k) &= \left(1 - \frac{|k|}{N}\right)^+ \\ K_N(x) &= \frac{1}{N} \left(\frac{\sin(N\pi x)}{\sin(\pi x)} \right)^2 \\ 0 &\leq K_N(x) \leq CN^{-1} \min(N^2, x^{-2})\end{aligned}$$

Proof. Observe that $\hat{K}_N(k) = \frac{1}{N} \sum_{n=0}^{N-1} \hat{D}_n(k)$. Now, if $|k| > (N-1)$, then $\hat{D}_n(k) = 0$.

If $0 < k \leq (N-1)$, then $\hat{D}_n(k) = 1$ for each $n \geq k$ in the above sum for $(N-k)$ times. If $(1-N) \leq k \leq 0$, then $\hat{D}_n(k) = 1$ for each $0 \leq k$ in the sum for $N - (N-k) = k$ times. Hence, we end up with the relation $\hat{K}_N(k) = \left(1 - \frac{|k|}{N}\right)^+$

Replacing the closed form for the Dirichlet kernel into the Cesaro means, one has

$$\begin{aligned}K_N(x) &= \frac{1}{N} \sum_{n=0}^{N-1} \frac{\sin((2n+1)\pi x)}{\sin(\pi x)} \\ &= \frac{1}{N \sin(\pi x)} \sum_{n=0}^{N-1} \sin((2n+1)\pi x) \\ &= \frac{1}{N(\sin(\pi x))^2} \sum_{n=0}^{N-1} \sin(\pi x) \sin((2n+1)\pi x)\end{aligned}$$

Using the trigonometric identity, $\sin(\theta) \sin(\alpha) = \frac{\cos(\theta+\alpha) - \cos(\theta-\alpha)}{2}$, we have

$$\frac{1}{N(\sin(\frac{\pi}{2}x))^2} \sum_{n=0}^{N-1} \sin((2n+1)\pi x) \sin(\pi x) = \frac{1}{2N(\sin(\pi x))^2} \sum_{n=0}^{N-1} [\cos(2(n+1)\pi x) - \cos(2n\pi x)]$$

Observing the telescoping nature of the sum, one concludes that

$$\frac{1}{2N(\sin(\frac{\pi}{2}x))^2} \sum_{n=0}^{N-1} \sin((2n+1)\pi x) \sin(\pi x) = \frac{1 - \cos(2N\pi x)}{2N(\sin(\pi x))^2}$$

Using the double angle formula

$$= \frac{1}{N} \left[\frac{\sin(2N\pi x)}{\sin(\pi x)} \right]^2$$

The lower bound for the pointwise Féjer kernel function is evident. For the upper bound, the arguments are completely analogous to proving the pointwise bounds of the Dirichlet kernel's.

□

Definition 1 (Approximate identity). *The family $\{\Phi\}_{N=1}^{\infty} \subset L^{\infty}(\mathbb{T})$ forms an approximate identity provided that*

1. (A1) $\int_0^1 \Phi_N(x) dx = 1$ for all N .
2. (A2) $\sup_N \int_0^1 |\Phi_N(x)| dx < \infty$
3. (A3) for all $\delta > 0$ one has $\int_{|x|>\delta} |\Phi_N(x)| dx \xrightarrow{0}$ as $N \rightarrow \infty$

It is clear from Theorem 1 that the Dirichlet Kernel does not form an approximate identity. However, Féjer kernel does form an approximate identity.

Lemma 1. $\{F_N(x)\}_{N=1}^{\infty}$ forms an approximate identity.

Proof. • Property (A1) is easily verified by direct integration of Fejér kernel.

- Property (A2) follows from (A1) with the observation that $K_N(x) > 0$ for all $x \in \mathbb{T}$.
- For property (A3), observe that for $|x| > \delta$, one has the pointwise majoration $|K_N(x)| \leq \frac{C}{Nx^2}$ given by Theorem 3. Hence,

$$\int_{\delta}^1 K_N(x) dx \leq \int_{\delta}^1 \frac{C}{Nx^2} dx = \frac{C}{N} \left(\frac{1}{\delta} - 1 \right)$$

which tends to zero as $N \rightarrow \infty$.

□

Proposition 1. For any approximate identity $\{\Phi_N\}_{N=1}^{\infty}$ one has the following:

1. If $f \in C(\mathbb{T})$, then $\|\Phi_N * f - f\|_{\infty} \rightarrow 0$ as $N \rightarrow \infty$
2. If $f \in L^p(\mathbb{T})$, where $1 \leq p < \infty$, then $\|\Phi_N * f - f\|_p \rightarrow 0$ as $N \rightarrow \infty$
3. For any measure $\mu \in \mathbb{T}$, one has:

$$\Phi_N * \mu \rightharpoonup \mu$$

as $N \rightarrow \infty$ in the weak-* sense.

Definition 2 (Trigonometric polynomial). A trigonometric polynomial is a series given by $\sum_n a_n e(nx)$ in which only finitely many a_n s are non-zero.

Note that the N^{th} Fourier series of a function (and hence its convolution with $K_N(x)$) are trigonometric polynomials.

Corollary 1. The exponential family $\{e(nx)\}_{n \in \mathbb{Z}}$ satisfies the following properties.

- (i) The trigonometric polynomials are dense in $C(\mathbb{T})$ in the uniform topology and in $L^p(\mathbb{T})$ for any $1 \leq p < \infty$
- (ii) (Parseval's identity) For any $f \in L^2(\mathbb{T})$,

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

- (iii) The exponentials $\{e(nx)\}_{n \in \mathbb{Z}}$ form an orthogonal basis in $L^2(\mathbb{T})$
- (iv) For all $f, g \in L^2(\mathbb{T})$, one has Parseval's identity,

$$\int_{\mathbb{T}} f(x) \overline{g(x)} dx = \sum_{n \in \mathbb{Z}} \hat{f}(n) \overline{\hat{g}(n)}$$

Note that the fact that $K_N(x)$ is an approximate identity provides another proof for the Stone-Weierstrass theorem.

Corollary 2 (Riemann-Lebesgue Lemma). If $f \in L^1(\mathbb{T})$, then $\hat{f}(n) \rightarrow 0$ as $n \rightarrow \infty$.

Proof. From Lemma 1, we have that the Fejér kernel is an approximating identity and hence, there exists $N \in \mathbb{N}$ such that $\|\sigma_N f - f\|_1 < \epsilon$ for a fixed $\epsilon > 0$. For $|n| > N$, we have $\hat{K}_N = (1 - \frac{|n|}{N})^+ = 0$ and $\sigma_N \hat{f} = \hat{K}_N * \hat{f}(k) = \hat{K}_N(k) \hat{f}(k) = 0$ such that $|\hat{f}(n)| = |\sigma_N \hat{f}(n) - \hat{f}(n)| \leq \|\sigma_N f - f\|_1 < \epsilon$. □

1.3 Bernstein's inequality

Definition 3 (de la Vallée Poussin Kernel). *The de la Vallée Poussin kernel $V_n(x)$ is given by*

$$V_n(x) := [1 + e(nx) + e(-nx)] K_n(x)$$

Lemma 2. *The Fourier transform of the de la Vallée Poussin is given by:*

$$\hat{V}_n(j) = 1 \text{ if } |j| \leq n$$

Proof. For $|j| < n$, we have

$$\begin{aligned} \hat{V}_n &= \int_{\mathbb{T}} e(-jx) [1 + e(nx) + e(-nx)] K_n(x) dx \\ &= \int_{\mathbb{T}} e(-jx) K_n(x) dx + \int_{\mathbb{T}} e(-(j-n)x) K_n(x) dx + \int_{\mathbb{T}} e(-(n+j)x) K_n(x) dx \\ &= \hat{K}_n(j) + \hat{K}_n(j-n) + \hat{K}_n(n+j) \\ &= \frac{j}{n} + \frac{N-j}{N} \\ &= 1 \end{aligned}$$

both when $j \geq 0$ or $j < 0$.

□

Note also that

$$\begin{aligned} K'_n(x) &= \frac{1}{n} \sum_{j=0}^n D'_j(x) \\ &= \frac{1}{n} \sum_{j=1}^n \left[\sum_{k=-j}^j 2\pi i j e(jx) \right] \\ &\leq \frac{2\pi i}{n} \sum_{j=1}^n n^2 \\ &\leq C_1 n \end{aligned}$$

Then, we have $V'_n(x) = 2\pi i n (e(nx) - e(-nx)) K_n(x) + (e(nx) + e(-nx)) K'_n(x) \leq C_2 n + C_1 n = (C_2 + C_1) n \leq C n$

Proposition 2 (Bernstein's inequality). *Let f be a trigonometric polynomial with $\hat{f}(k) = 0$ for all $|k| > n$. Then,*

$$\|f'\|_p \leq C n \|f\|_p$$

for any $1 \leq p \leq \infty$ where C is an absolute constant.

Proof. The de La Vallée Poussin kernel is built from the Fejér kernel and hence $V_n * f$ forms a trigonometric polynomial as well. To assert that $V_n * f = f$, we simply note that they have the same Fourier coefficients using the previous remarks.

$$V_n \hat{*} f(j) = \hat{V}_n(j) * \hat{f}(j) \begin{cases} \hat{f}(j) & \text{if } |j| \leq n \\ 0 & \text{if } |j| > n \end{cases}$$

Hence, we have $f = V_n * f$ and $f' = V_n' * f$. Using Young's inequality and the previous remarks, we have

$$\|f'\|_p = \|V_n' * f\|_p \leq \|V_n'\|_p \|f\|_p \leq Cn \|f\|_p$$

□

1.4 Convex Fourier Coefficients

A sequence $\{a_n\}_n$ is said to be convex if $a_{n+1} + a_{n-1} - 2a_n \geq 0$ for all $n \geq 0$.

Lemma 3. *Let $\{a_n\}_{n \in \mathbb{Z}}$ be an even and convex sequence of non-negative numbers that tends to zero. Then, there exists $f \in L^1(\mathbb{T})$ with $f > 0$ and $\hat{f}(n) = a_n$.*

For $\phi \in C^2$ and $x\phi'(x) \xrightarrow{x \rightarrow \infty} 0$ and $\phi(x) \xrightarrow{x \rightarrow \infty} 0$, we have by integration by parts:

$$\int_x^\infty \phi(y)(y-x)dy = \phi(x)$$

In the discrete setting, this condition translates to:

$$\begin{aligned} n(a_n - a_{n-1}) &\xrightarrow{n \rightarrow \infty} 0 \text{ and } a_n \xrightarrow{n \rightarrow \infty} 0 \\ \sum_{n>m} (a_{n+1} + a_{n-1} - 2a_n)(n-m) &= \sum_{n>m} n(a_{n+1} + a_{n-1} - 2a_n) \frac{(n-m)}{n} \\ &= a_m \end{aligned}$$

$$\begin{aligned} \sum_{n>m} n(a_{n+1} + a_{n-1} - 2a_n) \frac{(n-m)}{n} &= a_m \\ \sum_{n>m} n(a_{n+1} + a_{n-1} - 2a_n) \hat{K}_n(m) &= a_m \end{aligned}$$

Proof. Define $f := \sum_{n>m} n(a_{n+1} + a_{n-1} - 2a_n)K_n$.

It suffices to use the pointwise upper bound of K_n , $n(a_n - a_{n+1}) \rightarrow 0$ and that $\sum_{n=1}^N n(a_{n+1} + a_{n-1} - 2a_n) = a_0 - a_N - N(a_N - a_{N+1}) \rightarrow a_0$ to prove that $f \in L^1(\mathbb{T})$.

□

Proposition 3 ("Converse" to Bernstein). *Suppose $f \in L^1(\mathbb{T})$ and $\hat{f}(j) = 0$ for all j such that $|j| < n$. Then,*

$$\|f''\|_p \geq Cn^2 \|f\|_p$$

holds for all $1 \leq p \leq \infty$. Note that C is independent of $n \in \mathbb{Z}^+$, f and p .

Proof. Take $a_n = \frac{1}{n^2}$ for $n \in \mathbb{Z}$. Note that $a_n \xrightarrow{n \rightarrow \infty} 0$ and $n(a_n - a_{n-1}) \xrightarrow{n \rightarrow \infty} 0$ and $\{a_n\}_n$ forms a convex sequence:

$$\frac{1}{(n+1)^2} + \frac{1}{(n-1)^2} - \frac{2}{n^2} > 0$$

By Lemma 3, we can construct g such that $\hat{g}(n) = a_n \mathbf{1}_{\{n>j\}}$.

Now, observe that, for $j \geq n$, we have

$$\hat{f}''(j) = (2\pi i j)^2 \hat{f}(j) = 4\pi^2 j^2 f(j)$$

Consider $g * f''(x)$. We have $g * \hat{f}''(j) = \hat{g}(j)\hat{f}''(j) = \frac{1}{4\pi^2 j^2}(4\pi^2 j^2)\hat{f}(j) = \hat{f}(j)$

Then, $g * f''(x) = f(x)$ since f is a trigonometric polynomial. We have by, Young's inequality:

$$\|f\|_p = \|g * f''\|_p \leq \|g\|_1 \|f''\|_p \leq \frac{C}{n^2} \|f''\|_p$$

and the inequality follows. □

1.5 Smoothness and Fourier coefficients

Definition 4 (Sobolev Space). *For any $s \in \mathbb{R}$, define the Hilbert space $H^s(\mathbb{T})$ by means of the norm*

$$\|f\|_{H^s(\mathbb{T})}^2 := |\hat{f}(0)|^2 + \sum_{n \in \mathbb{Z}} |n|^{2s} |\hat{f}(n)|^2$$

For $f \in C^1(\mathbb{T})$:

$$\hat{f}'(n) := 2\pi i n \hat{f}(n)$$

by integration by parts and

$$\hat{f}(n) = O\left(\frac{1}{n}\right)$$

and

$$C^1(\mathbb{T}) \hookrightarrow \mathbb{H}'(\mathbb{T})$$

Exercise 1. *Let $f \in L^1(\mathbb{T})$. Show that $f \in C^\infty(\mathbb{T})$ if and only if $\hat{f}$ decays rapidly. i.e. for every $M \geq 1$, one has $\hat{f}(n) = O(|n|^{-M})$ as $|n| \rightarrow \infty$.*

Proof. If $f \in C^M(\mathbb{T})$, then $f^{(n)}(n) = (2\pi i n)^m \hat{f}(n)$ by the previous remark.

If $\hat{f}(n)$ decays rapidly, then $\sum_{n \in \mathbb{Z}} \hat{f}(n)e(nx)$ converges such that $S_N f \xrightarrow{N \rightarrow \infty} f$ almost everywhere with respect to the L^∞ norm. Define f with its Fourier series expansion and differentiate to conclude that $f \in C^\infty$. □

1.6 Sobolev Spaces and Embeddings

This section answers the question about the minimal value of $\alpha > 0$ such that $C^\alpha(\mathbb{T}) \hookrightarrow \mathbb{A}(\mathbb{T})$. To this end, we answer the question for which α does $C^\alpha(\mathbb{T}) \hookrightarrow \mathbb{H}^s(\mathbb{T})$?

Theorem 4 (Bernstein's Theorem). *For any $1 \geq \alpha > 0$, one has $C^\alpha(\mathbb{T}) \hookrightarrow \mathbb{H}^\beta(\mathbb{T})$ for arbitrary $0 < \beta < \alpha$. In particular, for any $f \in C^\alpha(\mathbb{T})$ with $\alpha \geq \frac{1}{2}$ one has*

$$\sum_{n \in \mathbb{Z}} |\hat{f}(n)| < \infty$$

and thus $C^\alpha(\mathbb{T}) \hookrightarrow \mathbb{A}(\mathbb{T})$ for any $\alpha > \frac{1}{2}$.

We do not prove the theorem here since the reader will encounter explicit demonstrations in the book.

2 Harmonic Functions; Poisson kernel

Denote $\Omega \subset \mathbb{R}^2$ as an open and connected domain.

Definition 5 (Harmonic function). *A function $u \in C^2(\mathbb{T})$ is said to be harmonic if $\Delta u = 0$ on Ω .*

Definition 6 (Holomorphic Functions). *For a complex-valued function F , we can decompose it into the imaginary and real parts: $F = u + iv$*

Holomorphic functions are functions F such that u and v are both harmonic. The necessary and sufficient conditions for holomorphy is given by the Cauchy-Riemann equations:

$$\begin{aligned} u_x - v_y &= 0 \\ u_y + v_x &= 0 \end{aligned}$$

Lemma 4. *If u is harmonic on a simply-connected domain Ω , then there exists a holomorphic function F on Ω such that $\text{Re}(F) = u$*

Sketch of proof. If F is holomorphic, then it would satisfy the Cauchy-Riemann equations:

$$\begin{aligned} F(z) &= u(x, y) + iv(x, y) \\ u_x &= v_y \text{ and } u_y = -v_x \end{aligned}$$

$$\begin{pmatrix} v_x \\ v_y \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} u_x \\ u_y \end{pmatrix}$$

To represent $v(x, y)$, we let $z \in \mathbb{C}$ be arbitrary and fix a point z_0

Let $\gamma : [0, 1] \mapsto \Omega$ be the point joining z_0 to z .

$$v(x, y) - v(x_0, y_0) = \int_0^1 \frac{dV(\gamma(t))}{dt} dt = \int_0^1 (-u_y, u_x) \cdot \gamma'(t) dt$$

which yields a holomorphic function that respects the Cauchy-Riemann conditions. □

2.1 Mean-Value Property and Maximum principle

Theorem 5 (Gauss-Green Theorem). *Suppose U is a bounded subset of $\mathbb{R}^2$ and ∂U is $\mathbb{C}^1$. For $u \in C'(\overline{U})$, we have:*

$$\int_U \text{div}(u) dx = \int_{\partial U} u \cdot v dS$$

where dS is a measure on ∂U and v is an outward-pointing unit vector.

Theorem 6 (Mean value property). *Let u be harmonic on some domain $\Omega \subset \mathbb{R}^2$. Then, for every $z \in \Omega$ and for all $0 < r < \text{dist}(z, \partial\Omega)$, one has*

$$u(z) = \int_{\mathbb{T}} u(z + re(\theta)) d\theta = \frac{1}{\pi r^2} \int_{z+r\partial\mathbb{D}} (\Delta u) dm$$

Proof.

$$\frac{d}{dr} \int_{\mathbb{T}} u(z + re(\theta)) d\theta = \frac{1}{2\pi r} \int_{z+r\partial\mathbb{D}} \partial_n u(w) \sigma(dw) \quad (\sigma(dw) : \text{measure on the border})$$

Using Gauss-Green Theorem

$$= \frac{1}{2\pi r} \int_{z+r\partial\mathbb{D}} (\Delta u) dm$$

Since u is harmonic on $\mathbb{D}$, this implies $\Delta u = 0$, and hence, $\frac{1}{2\pi r} \int_{z+r\partial\mathbb{D}} (\Delta u) dm = \frac{d}{dr} \int_{\mathbb{T}} u(z + re(\theta)) d\theta = 0$

Consider the circular means $M(r) := \int_{\mathbb{T}} u(z + re(\theta)) d\theta = \frac{1}{2\pi r} \int_{|w-z|=r} (\Delta u) dm$. By the previous computations, $M(r)$ is a constant with respect to r .

Hence, we have $M(r) \xrightarrow{r \rightarrow 0} u(z)$. Integrating $su(z)$ with respect to s on $0 < s < r$ gives us the required formula for solid disks.

□

Theorem 7 (Maximum Value Principle). *Let u be a harmonic function on $\Omega \subset \mathbb{R}^2$. If u attains an extremum in Ω , then u must be constant.*

2.2 Poisson Kernel

In this section, we address the question of a kernel in the complex plane. Denote the disk $\{z \in \mathbb{C} \mid |z| < 1\}$ by $\mathbb{D}$ and its boundary $\partial\mathbb{D}$ as $\mathbb{T}$.

Set $z = e(\theta)$

Suppose we are given a continuous function f on $\mathbb{T}$ and we wish to find the harmonic function u extending f on $\mathbb{D}$:

$$\begin{aligned} \Delta u &= 0 \text{ in } \mathbb{D} \\ u &= f \text{ in } \mathbb{T} \end{aligned}$$

For a pointwise solution that extends f , we need to affirm that $f \in C^2(\mathbb{D}) \cap C(\overline{\mathbb{D}})$

$$z = x + iy = re(\theta)$$

$$z^n = r^n e(n\theta)$$

Using the Chain Rule, the Laplacian in polar coordinates can be expressed:

$$u_{xx} + u_{yy} = u_{rr} + \frac{1}{r} u_r + \frac{1}{r^2} u_{\theta\theta}$$

We have:

$$\begin{aligned}
\Delta z^n &= n(n-1)r^{(n-2)}e(n\theta) + \frac{n}{r}r^{(n-1)}e(n\theta) - \frac{n^2}{r^2}r^n e(n\theta) \\
&= (n^2 - n)r^{(n-2)}e(n\theta) + nr^{(n-2)}e(n\theta) - n^2r^{n-2}e(n\theta) \\
&= (n^2 - n)r^{(n-2)}e(n\theta) + (-n^2 + n)r^{(n-2)}e(n\theta) \\
&= 0
\end{aligned}$$

The result is analogous for $\bar{z}$.

Hence, we define:

$$u(z) := \sum_{n=0}^{\infty} \hat{f}(n)z^n + \sum_{n=-\infty}^{-1} \hat{f}(n)z^{|n|}$$

which satisfies the differential equation conditions.

We then have: $u(e(\theta)) = \sum_{n=-\infty}^{\infty} \hat{f}(n)e(n\theta) = f(\theta)$ since f is the restriction of a holomorphic function u to $\mathbb{T}$.

Note that $\hat{f}(n) = \int_{\mathbb{T}} e(-n\varphi)f(\varphi)d\varphi$. Replacing in the proposed solution of u , we have:

$$u(re(\theta)) = \int_{\mathbb{T}} \sum_{n \in \mathbb{Z}} r^{|n|} e(n(\theta - \varphi)) d\varphi$$

Definition 7 (Poisson Kernel). *We define the Poisson Kernel as*

$$P_r(\theta) = \sum_{n \in \mathbb{Z}} r^{|n|} e(n\theta)$$

such that $u(re(\theta)) = (P_r * f)(\theta)$ for $0 \leq r < 1$

Exercise 2. For $0 \leq r < 1$, we can express the Poisson kernel as:

$$P_r(\theta) := \sum_{n \in \mathbb{Z}} r^{|n|} e(n\theta) = \frac{1 - r^2}{1 - 2r \cos(2\pi\theta) + r^2}$$

Observe that:

$$\begin{aligned}
\sum_{n \in \mathbb{Z}} r^{|n|} e(n\theta) &= 1 + \sum_{n \in \mathbb{N}} r^{|n|} [e(n\theta) + e(-n\theta)] \\
&= 1 + \sum_{n \in \mathbb{N}} 2r^{|n|} \cos(2\pi n\theta)
\end{aligned}$$

We use the following identity: $\frac{1}{(1-z)} = \sum_{n \in \mathbb{N}} z^n = \sum_{n \in \mathbb{N}} r^n e(n\theta)$ to have:

$$\begin{aligned}
\frac{(1+z)}{(1-z)} &= (1+z) \frac{1}{(1-z)} \\
&= (1+z) \sum_{n \in \mathbb{N}} z^n \\
&= \sum_{n \in \mathbb{N}} z^n + \sum_{n \in \mathbb{N}} z^{n+1} \\
&= 1 + \sum_{n \in \mathbb{N}} z^{n+1} + \sum_{n \in \mathbb{N}} z^{n+1} \\
&= 1 + 2 \sum_{n \in \mathbb{N}} z^{n+1} \\
&= 1 + 2 \sum_{n \geq 1} r^n e(n\theta)
\end{aligned}$$

Hence, we have:

$$\begin{aligned}
\operatorname{Re} \left(\frac{(1+z)}{(1-z)} \right) &= 1 + 2 \sum_{n \geq 1} r^n \cos(n\theta) \\
&= \frac{(1-r^2)}{(1-2r \cos(2\pi\theta) + r^2)}
\end{aligned}$$

2.3 The Poisson kernel as an approximate identity

Notice that

- (i) $P_r(\theta)$ is a harmonic function.
- (ii) $(P_r * \mu)(\theta)$ defines a harmonic function on $\mathbb{D}$ for $\mu \in \mathcal{M}(\mathbb{T})$:

$$\begin{aligned}
(P_r * \mu)(\theta) &= \int_{\mathbb{T}} P_r(\theta - \varphi) \mu(d(\varphi)) \\
&= \int_{\mathbb{T}} 1 \mu(d(\varphi)) + 2 \sum_{n \geq 0} r^n \int_{\mathbb{T}} \cos(n(\theta - \varphi)) \mu(d(\varphi)) \\
&= 1
\end{aligned}$$

Since constant functions are harmonic functions, we are done.

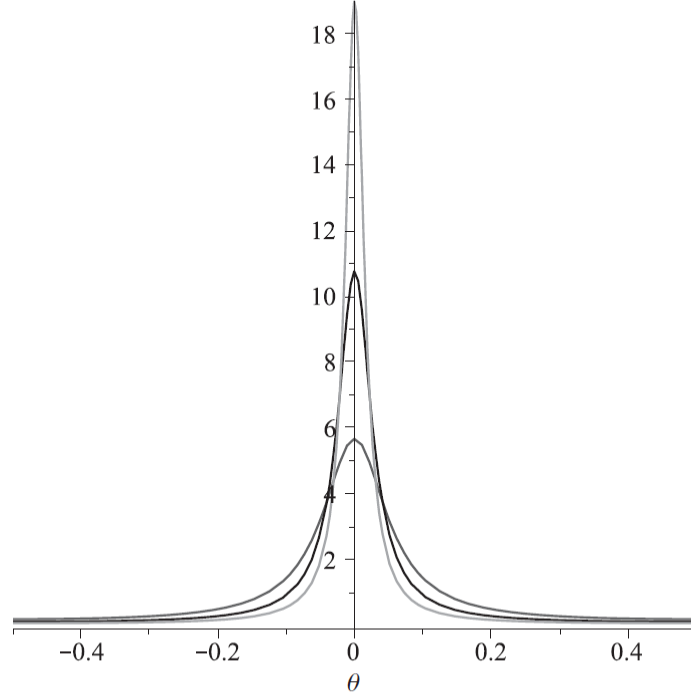


Figure 2.1. Graphs of P_r for $r = 0.70, 0.83, 0.90$.

Figure 3: The Poisson Kernel

Exercise 3. Check that $\{P_r\}_{0 < r < 1}$ is an approximate identity. The role of $N \in \mathbb{Z}^+$ is played by $0 < r < 1$ and $N \rightarrow \infty$ is replaced by $r \rightarrow 1^-$.

- (A1) $\int_0^1 P_r(\theta) d\theta = 1$ for all $0 \leq r < 1$

One needs only expand $P_r(\theta)$ into $\sum_{n \in \mathbb{Z}} r^{|n|} e(n\theta)$ and the result follows from integration.

- (A2) $\sup_r \int_0^1 |P_r(\theta)| d\theta < \infty$
- (A3) For all $\delta > 0$, one has $\int_{|\theta| > \delta} |P_r(\theta)| d\theta \rightarrow 0$ as $r \rightarrow 1$.

We integrate $P_r(\theta)$ over $[-\pi, \pi]$ instead of $[0, 1]$ in order to benefit from the symmetry of the $\cos$ function.

$$P_r(\theta) = \frac{1 - r^2}{(1 - 2r \cos(\theta) + r^2)}$$

For $0 < \delta < \theta \leq \pi$, one has:

$$\begin{aligned} |2r \cos(\delta)| &> |2r \cos(\theta)| \\ 1 - 2r \cos(\delta) + r^2 &< 1 - 2r \cos(\theta) + r^2 \\ \left| \frac{1 - r^2}{1 - 2r \cos(\theta) + r^2} \right| &< \left| \frac{1 - r^2}{1 - 2r \cos(\delta) + r^2} \right| \end{aligned}$$

We then have $\int_\delta^\pi \left| \frac{1 - r^2}{1 - 2r \cos(\theta) + r^2} \right| d\theta \leq \int_\delta^\pi \left| \frac{1 - r^2}{1 - 2r \cos(\delta) + r^2} \right| d\theta \leq \left| \frac{1 - r^2}{1 - 2r \cos(\delta) + r^2} \right| |\pi - \delta|$. The latter tends to zero as $r \rightarrow 1^-$. One can use the symmetry of $\cos$ to find the same inequality for $-\pi < -\delta < \theta < 0$.

2.4 The boundary Value problem

Notation for this section: $F_r(\theta) = F(re(\theta))$

Lemma 5. • (i) If $F \in C(\overline{\mathbb{D}})$ and $\Delta(F) = 0$, then $F_r = P_r * F_1$ for any $0 \leq r < 1$.

- (ii) If $\Delta F = 0$ on $\mathbb{D}$, then $F_{rs} = P_r * F_s$ for $0 \leq r < s < 1$
- (iii) As a function of $r \in (0, 1)$, the norms $\|F_r\|_p$ are non-decreasing for any $1 \leq p < \infty$

Proof. (i) Let $u(re(\theta)) := (P_r * F_1)(\theta)$ where $F_1(\theta) = F(e(\theta)) = F|_{\partial\mathbb{D}}$. By lemma ..., u is a harmonic function on $\mathbb{D}$.

By approximate identity, we have $\|u_r - F_1\|_\infty \xrightarrow{r \rightarrow 1} 0$.

Hence, u agrees with F_r on the boundary. Since both u and F are harmonic functions on $\overline{\mathbb{D}}$, we have $u - F$ is a harmonic function on $\partial\mathbb{D}$.

On $\partial\mathbb{D}$, $u - F = 0$ and by the maximum principle, this holds true on $\mathbb{D}$.

Hence, we have $u = F_r$ on $\mathbb{D}$.

(ii) Take $s\mathbb{D} \rightarrow \mathbb{D}$ and apply (i).

(iii) Using (ii) and Young's inequality, we have:

$$\|F_{rs}\|_p = \|P_r * F_s\|_p \leq \|P_r\|_1 \|F_s\|_p = \|F_s\|_p$$

□

Definition 8 ("Little" hardy space). For $1 \leq p < \infty$, we define:

$$h^p(\mathbb{D}) := \{u : \overline{\mathbb{D}} \mapsto \mathbb{C} \mid \sup_{0 < r < 1} \int_{\mathbb{T}} |u(re(\theta))|^p d\theta < \infty\}$$

and replace the condition to $\sup_{\theta \in \mathbb{T}} |u(re(\theta))|$ for $p = \infty$

Recall: Banach-Alaogh-Bourbaki Theorem

Given a normed vector space E , the closed unit ball of the dual E' is compact for the weak-* topology.

Lemma 6. Let $F \in h^1(\mathbb{D})$. Then, there exists a unique measure $\mu \in \mathcal{M}(\mathbb{T})$ such that $F_r = P_r * \mu$

Proof. If $F \in h^1(\mathbb{D})$ such that $F_{r_j} \xrightarrow{*} u$ for some $u \in \mathcal{M}(\mathbb{T})$.

Then, for $0 < r < 1$ and by lemma 5, we have:

$$P_r * F_{r_j}(\theta) = F_{rr_j}(\theta)$$

Note that if $F_{r_j} \xrightarrow{*} u$ implies pointwise convergence of F_{r_j} to u .

Hence, we have:

$$\lim_{j \rightarrow \infty} P_r * F_{r_j}(\theta) = \lim_{j \rightarrow \infty} F_{rr_j}(\theta) = F_r(\theta)$$

and we are done. □

Theorem 8. There is a one-to-one correspondence between $h^1(\mathbb{D})$ and $\mu \in \mathbb{M}(\mathbb{T})$, given $\mu \in \mathbb{M}(\mathbb{T}) \mapsto F_r(\theta) := (P_r * \mu)(\theta)$. Furthermore,

$$\|\mu\| = \sup_{0 < r < 1} \|F_r\| = \lim_{r \rightarrow 1} \|F_r\|_1$$

and one has the following:

(i) The measure μ is absolutely continuous with respect to the Lebesgue measure if and only if $\{F_r\}$ converges in $L^1(\mathbb{T})$. If this is the case, then $d\mu = fd\theta$, where f is the L^1 limit of F_r .

(ii) The following are equivalent for $1 < p \leq \infty$:

$$d\mu = fd\theta \iff \{F_r\}_{0 < r < 1} \text{ is } L^p\text{-bounded} \iff \{F_r\} \text{ converges in } L^p \text{ if } 1 < p < \infty$$

and in the weak-* sense in L^∞ if $p = \infty$ as $r \rightarrow 1$

(iii) That the above function f is continuous $\iff F$ extends to a continuous function on $\overline{\mathbb{D}} \iff F_r$ converges uniformly as $r \rightarrow 1^-$.

Proof. For the one-to-one correspondence, we note that the surjection is given by Lemma 6. There exists $F_r = P_r * \mu$ by lemma ... Using Young's inequality, we have

$$\lim_{r \rightarrow 1} \|F_r\|_1 = \lim_{r \rightarrow 1} \|P_r * \mu\|_1 \leq \|P_r\|_1 \|\mu\|_{\mathcal{M}} = \|\mu\|_{\mathcal{M}}$$

Since $F_r \xrightarrow{*} u$, we use Fatou's lemma to have:

$$\|\mu\|_{\mathcal{M}} \leq \liminf_{r \rightarrow 1} \|F_r\|_1$$

The two inequalities above establish the injection needed.

□

2.5 The Hardy-Littlewood maximal function

In this section, we explore the question of almost-everywhere convergence of $P_r * f$ to f as $r \rightarrow 1$.

If $f \in L^1(\mathbb{T})$, we have $\lim_{n \rightarrow \infty} g_n = f$ by density of $C(\mathbb{T})$.

We have the two limits: $\|P_r * g_n - g_n\|_{L^\infty} \xrightarrow{r \rightarrow 1} 0$ and $\|P_r * (f - g_n)\|_{L^\infty} \xrightarrow{n \rightarrow \infty} 0$

We need a way to interchange the limits $r \rightarrow 1$ and $n \rightarrow \infty$.

To this end, we define the Hardy-Littlewood Maximal Function $M(f)$.

Definition 9 (Hardy-Littlewood Maximal Function). *For any measure $\mu \in \mathcal{M}(\mathbb{T})$:*

$$(Mu)(\theta) := \sup \left\{ \frac{u(I)}{|I|} : \theta \in I \right\}$$

If $\mu(dx) = f(x)dx$, then

$$(Mf)(\theta) := \sup \left\{ \frac{1}{|I|} \int_{\varphi \in I} |f(\varphi)| d\varphi : \theta \in I \right\}$$

Definition 10 (Weak- L^p). *Given (X, μ) a measurable space and $g : X \mapsto \mathbb{C}$ a measurable function. For $1 \leq p < \infty$, g is said to be weak- L^p if*

$$\|g\|_{\mathcal{L}(L^p, L^\infty)}^p := \sup_{\lambda > 0} (\lambda^p |\{x \in X : |g(x)| > \lambda\}|) < \infty$$

Equivalently, there exists a constant C such that for all $\lambda > 0$, we have

$$|\{\theta \in \mathbb{T} : |g(\theta)| > \lambda\}| \leq \frac{C}{\lambda^p}$$

Lemma 7 (Wiener's Covering Lemma). *If $\{B_i\}_{i=1}^N$ is a finite collection of balls, then there exists a finite sub-collection $\{B_{j_k}\}_{k=1}^m$ such that*

$$\bigcup_{j=1}^N B_j \subset \bigcup_{k=1}^m 3B_{j_k}$$

Proposition 4. *The Hardy-Littlewood maximal operator M satisfies:*

i M is bounded from L^1 to weak- L^1 , i.e.

$$|\{\theta \in \mathbb{T} | M(f)(\theta) > \lambda\}| \leq \frac{3\|f\|_1}{\lambda}$$

for $\lambda > 0$

ii For any $1 < p \leq \infty$, M is bounded on L^p

Proof. (i) Since the Lebesgue measure is regular, we need only prove the claim for a compact measurable set K . Fix $\lambda > 0$ and let $K \subset \{\theta \in \mathbb{T} | Mf(\theta) > \lambda\}$

Since $\mathbb{T}$ is compact, there exists an open cover $\{I_j\}_{j=1}^N$ such that $\int_{I_j} |f(\varphi)| d\varphi > \lambda |I_j|$

Using Wiener's Covering Lemma, we can find pairwise disjoint arcs such that $\{J_l\}_l^L$ such that $\bigcup_{i=1}^N I_i \subset \bigcup_{l=1}^L 3J_l$. Hence, we have:

$$|K| \leq \left| \bigcup_{l=1}^L 3J_l \right| \leq 3 \sum_{l=1}^L |J_l| \leq \frac{3}{\lambda} \sum_{l=1}^L \int_{J_l} |f(\varphi)| d\varphi \leq \frac{3}{\lambda} \|f\|_1$$

(ii) This is immediate from the Marcinkiewicz's interpolation theorem. □

2.6 Almost-everywhere convergence

Definition 11 (Radially bounded). *Let $\{\Phi_n\}_{n \geq 0}$ be an approximate identity on $\mathbb{T}$. It is radially bounded if there exists functions $\{\Psi_n\}_{n \geq 0}$ on $\mathbb{T}$ such that (A1) – (A3) hold, and additionally:*

$$(A4) : |\Phi_n(\theta)| \leq \Psi_n(\phi) \text{ for } 0 \leq \phi \leq \theta \leq \frac{1}{2}$$

and $\sup \|\Psi_n\|_1 < \infty$

Note that the Fejér, Poisson and de la Vallée Poussin kernels satisfy (A4).

Lemma 8. *If $\{\Phi_n\}_{n=1}^\infty$ satisfies (A4) then for any $f \in L^1(\mathbb{T})$, one has*

$$\sup_n |(\Phi_n * f)(\theta)| \leq \sup_n \|\Phi_n\|_1 Mf(\theta)$$

for all $\theta \in \mathbb{T}$.

Proof. Denote I_k as the interval $[\frac{1}{2^{k+1}}, \frac{1}{2^k}]$. Note that $|I_k| = \frac{1}{2^{k+1}}$

Then we have:

$$\begin{aligned}
\int_0^{\frac{1}{2}} \Phi_n(x-y)f(y)dy &\leq \int_0^{\frac{1}{2}} \Psi_n(x-y)f(y)dy \\
&\leq \sum_{k=1}^{\infty} \int_{I_k} \psi_n(x-y)f(y)dy \\
&\leq \|\Psi_n\|_{\infty} \sum_{k=1}^{\infty} \int_{I_k} f(y)dy \\
&\leq \|\Psi_n\|_1 \sum_{k=1}^{\infty} \int_{I_k} f(y)dy \\
&\leq \|\Psi_n\|_1 \sum_{k=1}^{\infty} \frac{|I_k|}{|I_k|} \int_{I_k} f(y)dy \\
&\leq \|\Psi_n\|_1 \sum_{k=1}^{\infty} I_k M(f)(y) \\
&\leq \|\Psi_n\|_1 M(f)(y) \sum_{k=1}^{\infty} I_k \\
&= \frac{\|\Psi_n\|_1 M(f)(y)}{2}
\end{aligned}$$

We can then use the evenness of Φ_n to conclude. □

Theorem 9. *If $\{\Phi_n\}_{n=1}^{\infty}$ satisfies (A1) – (A4) then, for any $f \in L^1(\mathbb{T})$, one has $\Phi_n * f \rightarrow f$ almost everywhere as $n \rightarrow \infty$.*

Proof. Pick an $\epsilon > 0$ and let $g \in C(\mathbb{T})$ with $\|f - g\|_1 < \epsilon$.

By considering the term $(\Phi_n * g)(\theta) - (\Phi_n * g)(\theta) + g(\theta) - g(\theta)$, we have, using the sub-additivity of the measure:

$$\begin{aligned}
|\{\theta \in \mathbb{T} | \limsup_{n \rightarrow \infty} |(\Phi_n * f)(\theta) - f(\theta)| > \sqrt{\epsilon}\}| &\leq |\{\theta \in \mathbb{T} | \limsup_{n \rightarrow \infty} |(\Phi_n * h)(\theta)| > \frac{\sqrt{\epsilon}}{2}\}| + |\{\theta \in \mathbb{T} | |h(\theta)| > \frac{\sqrt{\epsilon}}{2}\}| \\
&\leq |\{\theta \in \mathbb{T} | \sup_n |(\Phi_n * h)(\theta)| > \frac{\sqrt{\epsilon}}{2}\}| + |\{\theta \in \mathbb{T} | |h(\theta)| > \frac{\sqrt{\epsilon}}{2}\}| \\
&\leq |\{\theta \in \mathbb{T} | CMh(\theta) > \frac{\sqrt{\epsilon}}{2}\}| + |\{\theta \in \mathbb{T} | |h(\theta)| > \frac{\sqrt{\epsilon}}{2}\}| \quad (\text{Lemma 8}) \\
&\leq C\sqrt{\epsilon} \quad (\text{proposition 4})
\end{aligned}$$

□

3 Introduction to Probability Theory

Definition 12 (Probability Space). *A probability space is a measure space Ω with its σ - algebra Σ , and a positive measure $\mathbb{P}$ such that $\mathbb{P}(\Omega) = 1$.*

Definition 13 (Random Variable). *A random variable X is a real or complex valued measurable function on Ω .*

Definition 14 (Independence). *Events A and B are independent if and only if $\mathbb{P}(A \cap B) = \mathbb{P}(A) \mathbb{P}(B)$.*

More generally, finitely many σ -algebras Σ_i of Σ are called independent if, for any $A_j \in \Sigma_j$, one has

$$\mathbb{P}\left(\bigcap_j A_j\right) = \prod_j \mathbb{P}(A_j)$$

Finitely many random variables $\{X_j\}_j$ are independent if the σ -algebras $\{X^{-1}(\mathcal{B}_j)\}_j$, where $\{\mathcal{B}_j\}_j$ are the Borel sets of $\mathbb{R}$ or $\mathbb{C}$, are independent.

Definition 15 (Rademacher function). *Define the Rademacher functions by:*

$$r_j(t) := \text{sgn}(\sin(2\pi 2^j t)), 0 < t < 1, j = 0, 1, 2, \dots$$

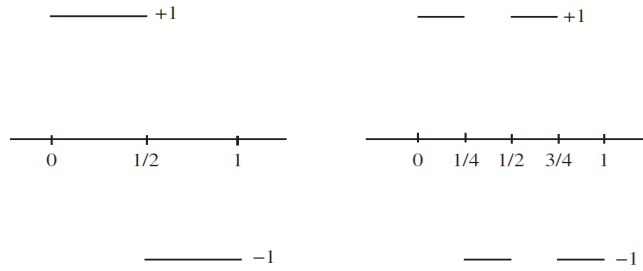


Figure 5.1. The first two Rademacher functions.

3.1 Expectation and Variance

Definition 16 (Expectation and Variance). *The expectation of a random variable X is defined as*

$$\mathbb{E}[X] := \int_{\Omega} X d\mathbb{P}$$

The variance of a random variable X is defined as

$$\mathbb{V}[X] := \mathbb{E}[X - \mathbb{E}[X]]^2 = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$$

Definition 17 (Convergence in Probability ($\xrightarrow{\mathbb{P}}$)). *A sequence $\{X_n\}_n$ of random variables converges in probability towards a random variable X if for all $\epsilon > 0$,*

$$\lim_{n \rightarrow \infty} \mathbb{P}(|X_n - X| > \epsilon) = 0$$

Definition 18 (Almost-sure convergence). *A sequence $\{X_n\}_n$ of random variables converges almost surely towards a random variable X if for all $\epsilon > 0$,*

$$\mathbb{P}\left(\lim_{n \rightarrow \infty} |X_n - X| > \epsilon\right) = 0$$

3.2 Sum of independent variables

In this section, we prove the Weak and Strong Law of Large Numbers for L^1 , L^2 and L^p random variables for $p > 2$.

Theorem 10 (Weak Law of Large numbers for L^2 Variables). *Let $\{X_j\}_{j=1}^\infty$ be an i.i.d sequence of L^2 variables. Then*

$$\frac{1}{N} \sum_{j=1}^N X_j \xrightarrow{\mathbb{P}} \mathbb{E}[X_1]$$

Proof. Assume $\mathbb{E}[X_1] = 0$ since we can rescale any variable with finite expectation to zero with a translation and obtain the same results. Define

$$S_N = \sum_{j=1}^N X_j$$

To prove convergence in probability, we need to show that $\mathbb{P}(\frac{1}{N}|S_N| > \epsilon) \xrightarrow{N \rightarrow \infty} 0$. Since X_j for $j \in \{1, \dots, N\}$ are L^2 variables, this statement is equivalent to proving $\mathbb{P}(S_N^2 > \epsilon^2 N^2) \xrightarrow{N \rightarrow \infty} 0$.

Observe that the X_j s are mutually orthogonal:

$$\mathbb{E}[X_i X_j] = \mathbb{E}[X_i] \mathbb{E}[X_j]$$

using the definition of independence. This results in the fact that $\mathbb{E}[S_N^2] = \mathbb{E}[\sum_{i=1}^N X_i^2] = \sum_{i=1}^N \mathbb{E}[X_i^2] = N \mathbb{E}[X_1^2]$. Hence,

$$\begin{aligned} \mathbb{P}(|S_N| > \epsilon N) &= \mathbb{P}(S_N^2 > \epsilon^2 N^2) \\ &\leq \frac{\mathbb{E}[S_N^2]}{\epsilon^2 N^2} && \text{(Using Markov's inequality)} \\ &= \frac{\mathbb{E}[X_1^2]}{\epsilon^2 N} \\ &\xrightarrow{N \rightarrow \infty} 0 \end{aligned}$$

□

Exercise 4. *Derive the Weierstrass approximation theorem from the previous proposition. Given any $f \in C([0, 1])$, define the Bernstein polynomial*

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

for $0 \leq x \leq 1$. Interpret this as $\mathbb{E}[f(\frac{S_n}{n})]$ for a suitable i.i.d. sequence and use the previous proposition to conclude your derivation.

Fix $\epsilon > 0$. For $i \in \llbracket 1, \dots, N \rrbracket$, we define the i.i.d random L^2 variables $\{X_i\}_i$ such that $X_i \sim \text{Ber}(x)$. This implies that $S_N \sim \text{Bin}(N, x)$, and using the previous theorem, we have

$$\mathbb{P}\left(\left|\frac{S_N}{N} - x\right| > \epsilon\right) \leq \frac{\mathbb{E}[X_1^2]}{\epsilon^2 N} = \frac{x}{\epsilon^2 N} \xrightarrow{N \rightarrow \infty} 0$$

Since f is continuous, we can assert

$$\lim_{N \rightarrow \infty} \mathbb{P}\left(\left|f\left(\frac{S_N}{N}\right) - f(x)\right| > \epsilon\right) = 0$$

Hence, we have for any fixed $x \in [0, 1]$, we have $\lim_{n \rightarrow \infty} \mathbb{E}[\mathbf{1}_{\{|f(\frac{S_N}{N}) - f(x)| > \epsilon\}} |f(\frac{S_N}{N}) - f(x)|] = 0$

Now, observe that

$$\mathbb{E}[|f(\frac{S_N}{N}) - x|] = \mathbb{E}[\mathbf{1}_{\{|f(\frac{S_N}{N}) - f(x)| < \epsilon\}} |f(\frac{S_N}{N}) - f(x)|] + \mathbb{E}[\mathbf{1}_{\{|f(\frac{S_N}{N}) - f(x)| > \epsilon\}} |f(\frac{S_N}{N}) - f(x)|]$$

Choose N large enough to have $\mathbb{E}[\mathbf{1}_{\{|f(\frac{S_N}{N}) - f(x)| > \epsilon\}} |f(\frac{S_N}{N}) - f(x)|] \leq \delta$ for an infinitesimal $\delta > 0$.

We then have:

$$\begin{aligned} \mathbb{E}[|f(\frac{S_N}{N}) - f(x)|] &= \mathbb{E}[\mathbf{1}_{\{|f(\frac{S_N}{N}) - f(x)| < \epsilon\}} |f(\frac{S_N}{N}) - f(x)|] + \mathbb{E}[\mathbf{1}_{\{|f(\frac{S_N}{N}) - f(x)| > \epsilon\}} |f(\frac{S_N}{N}) - f(x)|] \\ &\leq \mathbb{E}[\mathbf{1}_{\{|f(\frac{S_N}{N}) - f(x)| < \epsilon\}} |f(\frac{S_N}{N}) - f(x)|] + \delta \\ &\leq \epsilon \mathbb{P}(|f(\frac{S_N}{N}) - f(x)| < \epsilon) \\ &\leq \epsilon \end{aligned}$$

and the Weierstrass approximation theorem follows.

Lemma 9 (Borel-Cantelli Lemma). *Let $\{A_j\}_{j=1}^{\infty}$ be an infinite sequence of events. Then*

$$\sum_{j=1}^{\infty} \mathbb{P}(A_j) < \infty \Rightarrow \mathbb{P}(A_j \text{ occurs infinitely often}) = 0$$

Now, assume in addition that the A_j are independent. Then,

$$\sum_{j=1}^{\infty} \mathbb{P}(A_j) = \infty \Rightarrow \mathbb{P}(A_j \text{ occurs infinitely often}) = 1$$

Proof. Interpreting the statements as set-theoretic statements, we have $\mathbb{P}(A_j \text{ occurs infinitely often}) = \mathbb{P}(\limsup_{j \rightarrow \infty} A_j)$

Recall:

$$\begin{aligned} \limsup_{n \rightarrow \infty} A_n &= \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k \\ \limsup_{n \rightarrow \infty} A_n &\in \bigcup_{m \geq n} A_m \text{ for all } n \geq 0. \end{aligned}$$

$$\mathbb{P}(\limsup_{n \rightarrow \infty} A_n) \leq \mathbb{P}(\bigcup_{m \geq n} A_m) \leq \sum_{m \geq n} \mathbb{P}(A_m)$$

Since $\sum_{n=0}^{\infty} \mathbb{P}(A_n)$ is convergent by hypothesis, we have $\sum_{m \geq n} \mathbb{P}(A_m) \xrightarrow{n \rightarrow \infty} 0$ and the first statement follows.

Recall:

$$\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k)^c = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k^c = \liminf_{n \rightarrow \infty} A_k^c$$

$$\mathbb{P}(\limsup_{n \rightarrow \infty} A_n) = 1 - \mathbb{P}(\liminf_{n \rightarrow \infty} A_n^c)$$

We prove that $\mathbb{P}(\liminf_{n \rightarrow \infty} A_n^c) = 0$ so that the second statement of the claim follows.

$$\begin{aligned}
\mathbb{P}(\liminf_{n \rightarrow \infty} A_n^c) &= \mathbb{P}\left(\bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k^c\right) \\
&\text{using Monotone Convergence Theorem} \\
&= \lim_{n \rightarrow \infty} \mathbb{P}\left(\bigcap_{k=n}^{\infty} A_k^c\right) \\
&\text{using Monotone Convergence Theorem} \\
&= \lim_{n \rightarrow \infty} \lim_{q \rightarrow \infty} \mathbb{P}\left(\bigcap_{k=n}^q A_k^c\right) \\
&\text{By independence of the } A_j \text{ events} \\
&= \lim_{n \rightarrow \infty} \lim_{q \rightarrow \infty} \prod_{k=n}^q \mathbb{P}(A_k^c) \\
&= \lim_{n \rightarrow \infty} \lim_{q \rightarrow \infty} \prod_{k=n}^q (1 - \mathbb{P}(A_k)) \\
&\text{For } x > 0, \text{ we have } 1 - x \leq e^{-x} \\
&\leq \lim_{n \rightarrow \infty} \lim_{q \rightarrow \infty} \prod_{k=n}^q e^{-\mathbb{P}(A_k)} \\
&= \lim_{n \rightarrow \infty} \lim_{q \rightarrow \infty} e^{-\sum_{k=n}^q \mathbb{P}(A_k)}
\end{aligned}$$

The last term tends to zero since $\sum_{j=1}^{\infty} e^j$ diverges. □

Exercise 5. Let $\{X_n\}_{n=1}^{\infty}$ be an i.i.d sequence with $\mathbb{P}(\{X_1 > \lambda\}) = e^{-\lambda}$ for all $\lambda > 0$. Show that $\limsup_{n \rightarrow \infty} (\log n)^{-1} X_n = 1$ almost surely.

Proof. Fix $0 < \delta \leq 1$

Take $\{A_j\}$ as the events where $\frac{X_j}{(\log(j))} \leq 1 - \delta$

Take $\{B_k\}$ as the events where $\frac{X_k}{(\log(k))} \geq 1 + \delta$ for $j, k \in \mathbb{N}$

We prove that the events B_k occurs finitely many times.

$$\begin{aligned}
\sum_{k=1}^{\infty} \mathbb{P}(B_k) &= \sum_{k=1}^{\infty} \mathbb{P}(X_k \geq (\log(k))(1 + \delta)) \\
&= \sum_{k=1}^{\infty} e^{-\log(k)(1+\delta)} \\
&= \sum_{k=1}^{\infty} \left(\frac{1}{k}\right)^{(1+\delta)} \\
&< \infty
\end{aligned}$$

Hence, we have that B_k occurs only finitely many times by the previous lemma.

Note that we also have:

$$\begin{aligned}
\sum_{k=1}^{\infty} \mathbb{P}(A_k) &= \sum_{k=1}^{\infty} \mathbb{P}(X_k \geq \log(1+k)(1+\delta)) \\
&= \sum_{k=1}^{\infty} e^{-\log(k)(1-\delta)} \\
&= \sum_{k=1}^{\infty} \left(\frac{1}{k}\right)^{(1-\delta)} \\
&= \infty
\end{aligned}$$

Hence, the events A_k happen infinitely often and the hypothesis follows. $\square$

3.3 Strong Law of Large Numbers

Theorem 11 (Strong Law of Large Numbers for L^4 variables.). *Let $\{X_j\}_{j=1}^{\infty}$ be an i.i.d sequence of L^4 variables.*

Then, $\frac{1}{N} \sum_{j=1}^N X_j \rightarrow \mathbb{E}[X_1]$ almost surely.

Proof. Assume that $\mathbb{E}[X_1] = 0$, for otherwise we can translate the random variables by their expectation and obtain the same result.

We wish to show that

$$\mathbb{P}\left(\lim_{N \rightarrow \infty} \left| \frac{1}{N} \sum_{j=1}^N X_j \right| > \epsilon\right) = \mathbb{P}\left(\lim_{N \rightarrow \infty} |S_N| > N\epsilon\right) = 0$$

Since the X_j s are L^4 variables, we can prove that $\mathbb{P}(|S_N|^4 > \epsilon^4 N^4) = 0$ so that the claim follows.

•

$$\mathbb{E}[|S_N|^4] = \sum_{a,b,c,d=1}^N \mathbb{E}[X_a \dots X_d]$$

- For $a = b = c \neq d$, we have $\mathbb{E}[X_a^3 X_d] = \mathbb{E}[X_a^3] \mathbb{E}[X_d] = 0$
- For $a, b \neq c, d$, we have $\mathbb{E}[X_a^2 X_c^2] < \infty$

The order of permututation for indices $a, b \neq c, d$ is of the order N^2 such that $\mathbb{E}[|S_N|^4] < CN^2$ for some constant C .

Define A_N as the event that $A_N = \{|S_N|^4 > \epsilon^4 N^4\}$

We then use Markov's inequality such that

$$\mathbb{P}(A_N) \leq \frac{\mathbb{E}[|S_N|^4]}{\epsilon^4 N^4} \leq \frac{C}{\epsilon N^2}$$

and the claim follows from the Borel-Cantelli lemma. $\square$

3.4 Sub-Gaussian Bounds

Lemma 10. *Let $r_j(t)$ be Rademacher functions as described in Definition 15. For $N \in \mathbb{N}$ and $\{a_j\}_{j=1}^N \subset \mathbb{C}$, one has*

$$\mathbb{P}\left(\left\{\sum_{j=1}^N |r_j(t)a_j| > \lambda \left(\sum_{j=1}^N |a_j|^2\right)^{\frac{1}{2}}\right\}\right) \leq 4e^{-\frac{\lambda^2}{2}}$$

for all $\lambda > 0$.

Proof. Suppose $a_j \in \mathbb{R}$ since the following results hold by duplication for the real and imaginary part of any complex number.

Define S_N as $\sum_{j=1}^N r_j a_j$.

$$\begin{aligned}\mathbb{E}[e^{tS_N}] &= \mathbb{E}[e^{t \sum_{j=1}^N r_j a_j}] \\ &= \prod_{j=1}^N \mathbb{E}[e^{tr_j a_j}] \\ &= \prod_{j=1}^N \cosh(ta_j)\end{aligned}$$

We use the fact that $\cosh(x) \leq e^{\frac{x^2}{2}}$ for $x \in \mathbb{R}$ to get the inequality

$$\mathbb{E}[e^{tS_N}] \leq \prod_{j=1}^N e^{\frac{t^2 a_j^2}{2}} = e^{\sum_{j=1}^N \frac{t^2 a_j^2}{2}}$$

Take $\sigma^2 = \sum_{j=1}^N a_j^2$, and using Markov's inequality, we have:

$$\begin{aligned}\mathbb{P}(S_N > \lambda\sigma) &= \mathbb{P}(e^{tS_N} > e^{t\lambda\sigma}) \\ &\leq \frac{\mathbb{E}[e^{tS_N}]}{e^{t\lambda\sigma}} \\ &\leq e^{\frac{t^2 \sigma^2}{2}} e^{-\lambda\sigma t}\end{aligned}$$

The function $\frac{t^2 \sigma^2}{2} - \lambda\sigma t$ is minimised by $t = \frac{\lambda}{\sigma}$ such that the last inequality is bound by $e^{-\frac{\lambda^2}{2}}$. The same computations prove that $\mathbb{P}(\{S_N < -\sigma\lambda\}) \leq 2e^{-\frac{\lambda^2}{2}}$

For $a_j \in \mathbb{C}$, we repeat with the decomposition of a_j into the real and imaginary parts:

$$\mathbb{P}(\{|S_N| > \lambda\sigma\}) \leq 4e^{-\frac{\lambda^2}{2}}$$

□

As a consequence of this lemma, we present *Khinchine's inequality* which proves to be essential in establishing the existence of an L^1 function with almost-everywhere divergent Fourier Series.

Lemma 11 (Khinchine's inequality). *For any $1 \leq p < \infty$, there exists constants A_p and B_q such that*

$$B_q^{-1} \left(\sum_{j=1}^N a_j^2 \right)^{\frac{p}{2}} \leq \mathbb{E}[|\sum_{j=1}^N r_j a_j|^p] \leq A_p \left(\sum_{j=1}^N a_j^2 \right)^{\frac{p}{2}}$$

for any $N \in \mathbb{N}$ and $\{a_j\}_{j=1}^N \subset \mathbb{C}$.

Proof. We take $\sum_{j=1}^N a_j^2 = 1$ since the results that follow will hold true up to scaling.

Define $S_N := \sum_{j=1}^N a_j r_j$. We have:

$$\begin{aligned}\mathbb{E}[|S_N|^p] &= \int_0^\infty \mathbb{P}(|S_N| > \lambda) p \lambda^{p-1} d\lambda \\ &\quad \text{using the previous lemma} \\ &\leq \int_0^\infty 4e^{-\frac{\lambda^2}{2}} p \lambda^{p-1} d\lambda \\ &:= A_p < \infty\end{aligned}$$

For the lower bound, we make use of the upper bound. Note that

$$\mathbb{E}[S_N^2] = \mathbb{E}\left(\sum_{j=1}^N r_j a_j\right)^2 = \sum_{j=1}^N a_j^2 = 1$$

by considering that $\mathbb{E}[r_j] = 0$ and $\mathbb{E}[r_j^2] = 1$ for any $j \in \mathbb{N}$.

Using Hölder's inequality, we have:

$$1 = \mathbb{E}[|S_N|^2] \leq (\mathbb{E}[|S_N|^p])^{\frac{1}{p}} (\mathbb{E}[|S_N|^q])^{\frac{1}{q}} \leq B_q (\mathbb{E}[|S_N|^p])^{\frac{1}{p}}$$

and we are done. □

Exercise 6. By explicit expansion and integration for even integers p , prove that

$$\int_0^1 \left| \sum_{j=1}^N a_j r_j(t) \right|^p \leq C(p) \left(\sum_{j=1}^N |a_j|^2 \right)^{\frac{p}{2}}$$

Then, recover the general case of $1 \leq p < \infty$. Show that $C(p) \leq C_0 \sqrt{p}$ with an absolute constant C_0 and that it is optimal.

Since p is even, denote $p = 2k$ for $k \in \mathbb{N}$.

Consider the multinomial expansion of

$$\left(\sum_{j=1}^N a_j r_j(t) \right)^p = \sum_{\{k_1 + \dots + k_N = p\}} \binom{p}{k_1, \dots, k_N} a_1^{k_1} \dots a_N^{k_N} r_1(t)^{k_1} \dots r_N(t)^{k_N}$$

Since $\mathbb{E}[r_j^\alpha] = 0$ if α is odd, we only consider the even powers of the sum for our purpose.

$$\begin{aligned}\sum_{\substack{\{k_1 + \dots + k_N = p\} \\ k_j \text{ even}}} \binom{p}{k_1, \dots, k_N} a_1^{k_1} \dots a_N^{k_N} r_1(t)^{k_1} \dots r_N(t)^{k_N} &= \sum_{\substack{\{k_1 + \dots + k_N = p\} \\ k_j \text{ even}}} \frac{(2k)!}{(2z_1)! \dots (2z_N)!} a_1^{k_1} \dots a_N^{k_N} r_1(t)^{k_1} \dots r_N(t)^{k_N} \\ &\quad \text{for } k_j = 2z_j \text{ where } z_j \in \mathbb{N}\end{aligned}$$

Use the inequality $n! \geq 2^n$, proved by induction, to have the following inequality:

$$\begin{aligned} \binom{2n}{n} &\geq 1 \\ \frac{2n!}{(n!)(n!)} &\geq 1 \\ \frac{2n!}{(n!)} &\geq (n!) \\ \frac{2n!}{(n!)} &> 2^n \\ (2n)! &\geq (n!)2^n \end{aligned}$$

Hence, we have from the previous inequality:

$$\begin{aligned} &\sum_{\substack{\{k_1+\dots+k_N=p\} \\ k_j \text{ even}}} \frac{(2k)!}{(2z_1)! \dots (2z_N)!} a_1^{k_1} \dots a_N^{k_N} r_1(t)^{k_1} \dots r_N(t)^{k_N} \\ &\leq \frac{(2k)!}{2^{z_1+\dots+z_k}} \sum_{2z_1+\dots+2z_N=p} \frac{1}{(z_1)! \dots (z_N)!} a_1^{2z_1} \dots a_N^{2z_N} r_1(t)^{k_1} \dots r_N(t)^{k_N} \\ &= \frac{(2k)!}{2^k} \sum_{2z_1+\dots+2z_N=p} \frac{1}{(z_1)! \dots (z_N)!} a_1^{2z_1} \dots a_N^{2z_N} r_1(t)^{k_1} \dots r_N(t)^{k_N} \\ &= \frac{(2k)!}{2^k (k!)} \sum_{2z_1+\dots+2z_N=p} \frac{(k!)}{(z_1)! \dots (z_N)!} a_1^{2z_1} \dots a_N^{2z_N} r_1(t)^{k_1} \dots r_N(t)^{k_N} \\ &= \frac{(2k)!}{2^k k!} \left(\sum_{i=1}^n a_i^2 \right)^k \end{aligned}$$

which gives us the appropriate constant A_p for Khinchine's inequality.

Since $\frac{(2k)!}{k!} = (k+1)! \dots (2k)! < (2k)^k$ and hence we have $A_p \leq k^k = 2^{\frac{p}{2}}$

For p odd, we have, we have $p = 2n + 1$ and hence $2n < p < 2n + 2$ for $n \in \mathbb{N}$.

We have

$$\int_0^1 \left| \sum_{j=1}^N a_j r_j(t) \right|^p dt = \int_0^1 \left| \sum_{j=1}^N a_j r_j(t) \right|^{2n} \left| \sum_{j=1}^N a_j r_j(t) \right| dt$$

Use Cauchy-Schwartz inequality to have the inequality:

$$\|S_N\|_p \leq (n)^{\frac{1}{2}} \left(\sum_{i=1}^N a_i^2 \right)^{\frac{1}{2}} \leq \left(\frac{2n+1}{2} \right)^{\frac{1}{2}} \left(\sum_{i=1}^N a_i^2 \right)^{\frac{1}{2}}$$

3.5 Central Limit Theorem

Now, we show that the estimates from Khinchin's inequality are actually sharp ones.

We start by stating some facts.

Denote Fourier transform of any $\mathbb{R}$ -valued random variable X (also called the characteristic function of X) by $\mathbb{E}(e^{-itX}) = \int e^{-itx} F(dx)$, where F is the distribution of X .

If X and Y are independent, then

$$\mathbb{E}[e^{-itX}] \mathbb{E}[e^{-itY}]$$

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

Theorem 12 (Central Limit Theorem). *Let $\{X_j\}_{j=1}^\infty$ be i.i.d L^2 variables with $\mathbb{E}[X_1] = 0$ and $\mathbb{E}[X_1^2] = 1$. Then, for all $\lambda_1 < \lambda_2$, one has:*

$$\mathbb{P}(\{\lambda_1\sqrt{n} < \sum_{j=1}^N X_j < \lambda_2\sqrt{n}\}) \xrightarrow{n \rightarrow \infty} (2\pi)^{\frac{1}{2}} \int_{\lambda_1}^{\lambda_2} e^{-\frac{u^2}{2}} du$$

Proof. Let $\mathbb{E}[e^{i\xi X_1}] := \chi(\xi)$. Note that since $\mathbb{E}[X_1] = 0$, we have $\chi'(0) = \mathbb{E}[X_1] = 0$ and $\chi''(0) = \mathbb{E}[X_1^2] = \mathbb{V}[X_1] = 1$

Then,

$$\begin{aligned} \mathbb{E}[e^{i\xi \frac{S_N}{\sqrt{N}}}] &= \prod_{j=1}^N \mathbb{E}[e^{\frac{i\xi X_j}{\sqrt{N}}}] \\ &= \mathbb{E}[e^{\frac{i\xi X_1}{\sqrt{N}}}]^N \\ &= \chi(\frac{\xi}{\sqrt{N}})^N \end{aligned}$$

Hence, expanding $\chi(\frac{\xi}{\sqrt{N}})^N$ using the Taylor's theorem yields $(1 - \frac{\xi^2}{2N} + o(\xi^2 N^{-1}))^N$ which tends to $e^{-\frac{\xi^2}{2}}$ as $N \rightarrow \infty$.

Begin Block here In the Schwarz space: $\|x^\alpha f^{(\beta)}\|_\infty < \infty$, Fourier inversion establishes a bijection:

$$\begin{aligned} \hat{\phi}(\xi) &= \int_{\mathbb{R}} e^{-i\xi x} \phi(x) dx \\ \phi(x) &= \frac{1}{2\pi} \int_{\mathbb{R}} e^{i\xi x} \hat{\phi}(\xi) d\xi \end{aligned}$$

End Block here

Since $e^{-\frac{\xi^2}{2}}$ is the Fourier transform of the Gaussian distribution, we need only use the remark on Fourier inversion to assert our claim. $\square$

3.6 Strong Law of Large Numbers for L^1 variables

To prove the Strong Law of Large Numbers, we need the following Kolmogorov's estimate:

Lemma 12 (Kolmogorov's estimate). *Let $\{X_j\}_{j=1}^\infty$ be independent L^2 variables with $\mathbb{E}[X_j] = 0$. Then*

$$\mathbb{P}(\{\max_{1 \leq l \leq n} |S_l| \geq \epsilon\}) \leq \epsilon^{-2} \sum_{j=1}^n X_j^2$$

for all $\epsilon > 0$.

Proof. For simplicity, we take the variables to be real-valued. Define

$$A_l := \{\max_{1 \leq j < l} |S_j| \leq \epsilon, |S_l| \geq \epsilon\} = \{\max_{1 \leq j < l} |S_j|^2 \leq \epsilon^2, |S_l|^2 \geq \epsilon^2\}$$

for $1 \leq l \leq n$

Note we have

$$\begin{aligned} \mathbb{E}[|S_l|] &= \mathbb{E}[\mathbf{1}_{\{A_l\}} |S_l|] + \mathbb{E}[\mathbf{1}_{\{A_l\}^c} |S_l|] \\ &\leq \mathbb{E}[\mathbf{1}_{\{A_l\}} |S_l|] + \epsilon^2 \mathbb{P}(\{A_l\}^c) \end{aligned}$$

Also, we have using the independence of S_l and $S_n - S_l$:

$$\begin{aligned}
\mathbb{E}[\mathbf{1}_{\{A_l\}} S_n^2] &= \mathbb{E}[\mathbf{1}_{\{A_l\}} ((S_n - S_l) + S_l)^2] \\
&= \mathbb{E}[\mathbf{1}_{\{A_l\}} S_l^2] + \mathbb{E}[\mathbf{1}_{\{A_l\}} ((S_n - S_l) S_l)] + \mathbb{E}[\mathbf{1}_{\{A_l\}} ((S_n - S_l)^2)] \\
&\geq \mathbb{E}[\mathbf{1}_{\{A_l\}} S_l^2] + \mathbb{E}[\mathbf{1}_{\{A_l\}} S_l] \mathbb{E}[(S_n - S_l)] \\
&= \mathbb{E}[\mathbf{1}_{\{A_l\}} S_l^2]
\end{aligned}$$

since $\mathbb{E}[(S_n - S_l)] = 0$

Now, note that $\left\{ \max_{1 \leq l \leq n} |S_l| > \epsilon \right\} \subset \bigcup_{l=1}^n A_l$ such that $\mathbb{P}(\left\{ \max_{1 \leq l \leq n} |S_l| > \epsilon \right\}) \leq \sum_{l=1}^n \mathbb{P}(A_l)$

Using Markov's inequality, we have:

$$\begin{aligned}
\sum_{l=1}^n \mathbb{P}(A_l) &\leq \epsilon^{-2} \sum_{l=1}^n \int_{A_l} S_l^2 d\mathbb{P} \\
&= \epsilon^{-2} \sum_{l=1}^n \int_{A_l} S_n^2 d\mathbb{P} && \text{(by previous remark)} \\
&\leq \epsilon^{-2} \mathbb{E}[S_n^2]
\end{aligned}$$

□

Lemma 13. *Let $\{X_n\}_{n=1}^\infty$ be independent of L^2 variables with both $\sum_{n=1}^\infty X_n$ and $\sum_{n=1}^\infty \mathbb{V}(X_n)$ convergent. Then, $\sum_{n=1}^\infty X_n$ converges almost surely.*

Proof. Assume that $\mathbb{E}[X_n] = 0$ for all n . For $\epsilon > 0$, we can use the previous lemma to have the inequality:

$$\mathbb{P}(\left\{ \sup_{N \leq n \leq m} \left| \sum_{j=n}^m X_j \right| > \epsilon \right\}) \leq \epsilon^{-2} \sum_{n=N}^\infty \mathbb{V}(X_n)$$

and the claim follows. □

We now set to prove the Strong Law of Large Numbers. We begin with a preliminary lemma.

Lemma 14. *Suppose we have an arbitrary sequence of $\{X_j\}_{j=1}^\infty$ of independent L^2 random variables that satisfies $\mathbb{E}[X_n] = 0$ and $n^{-2} \mathbb{V}[X_n] < \infty$, then these variables obey the Strong Law of Large numbers.*

Proof. For $\epsilon > 0$, we define for $k \geq 0$,

$$A_k := \{|S_n| > \epsilon n \mid \text{for some } 2^k \leq n < 2^{k+1}\}$$

Using Markov's inequality, we have:

$$\begin{aligned}
\mathbb{P}(A_k) &\leq \frac{\mathbb{E}[|S_n|^2]}{\epsilon^2 n^2} && (\text{for } 2^k \leq n < 2^{k+1}) \\
&\leq \frac{\mathbb{E}[|S_n|^2]}{\epsilon^2 2^{2k}} \\
&\leq \epsilon^{-2} 2^{-2k} \mathbb{E}[|S_n|^2] \\
&\leq \epsilon^{-2} 2^{-2k} \sum_{n=1}^{2^{k+1}} \mathbb{E}[X_n^2] && (\text{Using the independence of } \{X_j\}_j) \\
&= \epsilon^{-2} 2^{-2k} \sum_{n=1}^{2^{k+1}} \mathbb{V}[X_n]
\end{aligned}$$

Hence, we have:

$$\begin{aligned}
\sum_{k=0}^{\infty} \mathbb{P}(A_k) &\leq \epsilon^{-2} \sum_{k=1}^{\infty} 2^{-2k} \sum_{n=1}^{2^{k+1}} \mathbb{V}[X_n] \\
&= \epsilon^{-2} \sum_{k=1}^{\infty} \frac{\mathbb{V}[X_n]}{2^{2(k-1)}} \\
&\quad \text{Using Comparison test and the inequality } n > 2^k \\
&\leq C \epsilon^{-2} \sum_{n=1}^{\infty} \frac{\mathbb{V}[X_n]}{n^2}
\end{aligned}$$

Hence, we have almost sure convergence of $\{X_j\}_j$ using Borel-Cantelli, and the claim follows. $\square$

Theorem 13. *The Strong Law of Large Numbers holds for L^1 variables.*

Proof. Suppose we are given a sequence of L^1 random variables $\{X_n\}_n$.

Define $Y_n := X_n \mathbf{1}_{\{|X_n| < n\}}$. Note that Y_n is an L^2 random variable.

Since $\mathbb{E}[Y_n] < n < N$, we have

$$\frac{1}{N} \sum_{n=1}^N \mathbb{E}[Y_n] \rightarrow 0$$

for $N \rightarrow \infty$.

Observe that

$$\begin{aligned}
\sum_{n=1}^{\infty} \frac{\mathbb{V}[Y_n]}{n^2} &\leq C \sum_{n=1}^{\infty} \frac{\sum_{k=0}^{n-1} (k+1)^2 \mathbb{P}(k \leq |X_1| < k+1)}{n^2} \\
&\leq C_2 \sum_{k=0}^{\infty} (k+1) \mathbb{P}(k \leq |X_1| < k+1) \\
&\leq C_3 \mathbb{E}[|X_1|]
\end{aligned}$$

Hence, by lemma 14, Y_n follows the Strong Law of Large Numbers.

For each $n \in \mathbb{N}$, we have $\{X_n \neq Y_n\} \subset \{X_n > n\}$ and we have:

$$\sum_{n=1}^{\infty} \mathbb{P}(\{X_n \neq Y_n\}) \leq \sum_{n=1}^{\infty} \mathbb{P}(|X_1| > n) \leq \mathbb{E}[|X_1|]$$

such that $X_n = Y_n$ almost always by Borel-Cantelli lemma. Therefore, the claim holds true. $\square$

3.7 Stopping Times

Definition 19 (Filtration). Let $\{\mathcal{F}_n\}_{n=1}^\infty$ be a sequence of sub- σ algebras. Then, $\{\mathcal{F}_n\}_{n=1}^\infty$ is a filtration if and only if $\mathcal{F}_n \subset \mathcal{F}_{n+1}$ for all $n \in \mathbb{N}$

We consider $(\mathcal{F}_n)_{n=1}^\infty$ to be fixed in any future theorems and lemmas in this section.

Definition 20 (Stopping Time). A random variable T is a stopping time relative to $\mathcal{F}_n$ if and only if $\{T = j\} \in \mathcal{F}_j$ for $j \geq 1$.

Example:

Suppose given $\{X_n\}_n^\infty$ random variables and $\mathcal{F}_n$ the σ -algebra generated by $\{X_j\}_{j=1}^\infty$ [also known as the natural filtration of $\{X_n\}_n$]. Then, we define $T = \min(i \geq 0, S_N \geq C)$

$$\{T = j\} \subset \left\{ \sum_{n=1}^{j-1} X_n \leq C \right\} \subset \bigcup_{n=1}^{j-1} \mathcal{F}_n \subset \mathcal{F}_j$$

Equivalent Statement:

A random variable T is a stopping time relative to $\mathcal{F}_n$ if and only if $\{T \leq j\} \in \mathcal{F}_j$ for $j \geq 1$.

Notice that $\{T = n\} = \{T \leq n\} \setminus \{T \leq n-1\} \in \mathcal{F}_n$ since $\{T = n\} \in \mathcal{F}_n$ and $\{T \leq n\} \in \mathcal{F}_{n-1} \subset \mathcal{F}_n$.

3.8 Wald's identity

Theorem 14 (Wald's identity). Suppose $\{X_n\}_{n=1}^\infty$ is an i.i.d L^1 sequence and let T be its associated stopping time with $\mathbb{E}[T] < \infty$. Then $\mathbb{E}[S_T] = \mathbb{E}[T]\mathbb{E}[X_1]$

Now, suppose that $\mathbb{E}[X_1] = 0$ and that X_1 is an L^2 variable. Then, $\mathbb{E}[S_T^2] = \mathbb{E}[T]\mathbb{E}[X_1^2]$

Proof. Note that $S_T = \sum_{k=1}^\infty X_k \mathbf{1}_{T \geq k}$.

Assume that $\mathbb{E}[X_1] \geq 0$.

$$\mathbf{1}_{\{T \geq k\}} = 1 - \mathbf{1}_{\{T < k\}}$$

$\{T < k\}$ is independent of X_k [or more specifically $\sigma(X_k)$] since T is not a function of $X_k, X_{k+1}, \dots$

$$\begin{aligned} \mathbb{E}[S_T] &= \sum_{k=1}^\infty \mathbb{E}[X_k \mathbf{1}_{\{T \geq k\}}] \\ &= \sum_{k=1}^\infty \mathbb{E}[X_k] \mathbb{E}[\mathbf{1}_{\{T \geq k\}}] && \text{(Independence of } \{T < k\}) \\ &= \sum_{k=1}^\infty \mathbb{E}[X_1] \mathbb{P}(\{T \geq k\}) \\ &= \mathbb{E}[X_1] \sum_{k=1}^\infty \mathbb{P}(\{T \geq k\}) \\ &= \mathbb{E}[X_1] \mathbb{E}[T] \end{aligned}$$

For the case where X_j is not a positive random variable, we need only decompose into $X_j = X_j^+ - X_j^-$ $\square$

$\wedge$ - and, $\vee$ - or

$T \wedge n$ is a stopping time.

$$\begin{aligned} \{T \wedge n \leq k\} &= (\{T \wedge n \leq k\} \cap \{T \leq k\}) \cup (\{T \wedge n \leq k\} \cap \{T > k\}) \\ &= (\{T \wedge n \leq k\} \cap \{T \leq k\}) \cup (\{n \leq k\} \cap \{T > k\}) \\ &= \{T \leq k \wedge n\} \cup \{\emptyset \cap T > k\} \\ &\subset \mathcal{F}_{n-k} \subset \mathcal{F}_n \end{aligned}$$

Consider $S_{T \wedge n}$ for $n \geq 1$. Then,

$$\begin{aligned} S_{T \wedge n}^2 &= \left(\sum_{k=1}^n X_k \mathbf{1}_{\{T \geq k\}} \right)^2 \\ &= \left(\sum_{k=1}^{n-1} X_k \mathbf{1}_{\{T \geq k\}} + X_n \mathbf{1}_{\{T \geq n\}} \right)^2 \\ &= \left(\sum_{k=1}^{n-1} X_k \mathbf{1}_{\{T \geq k\}} \right)^2 + 2X_n \mathbf{1}_{\{T \geq n\}} \sum_{k=1}^{n-1} X_k \mathbf{1}_{\{T \geq k\}} + X_n^2 \mathbf{1}_{\{T \geq n\}} \end{aligned}$$

Using the independence of X_n from $X_k \{T \geq k\}$ and $\{T \geq n\}$, and the fact that $\mathbb{E}[X_n] = 0$ to have:

$$\begin{aligned} \mathbb{E}[S_{T \wedge n}^2] &= \mathbb{E}[S_{T \wedge (n-1)}^2] + \mathbb{E}[X_n^2 \mathbf{1}_{T \geq n}] \\ &= \mathbb{E}[S_{T \wedge (n-1)}^2] + \mathbb{E}[X_1^2 \mathbf{1}_{T \geq n}] \\ &= \mathbb{E}[S_{T \wedge (n-1)}^2] + \mathbb{E}[X_1^2] \mathbb{P}(T \geq n) \end{aligned}$$

Since $\sum_{n=1}^{\infty} \mathbb{P}(T \geq n) = \mathbb{E}[T]$, we have by summing with respect to n :

$$\lim_{n \rightarrow \infty} \sum_{n=1}^N \mathbb{E}[S_{T \wedge n}^2 - S_{T \wedge (n-1)}^2] = \mathbb{E}[X_1^2] \mathbb{E}[T]$$

Since $\sum_{n=1}^N \mathbb{E}[S_{T \wedge n}^2 - S_{T \wedge (n-1)}^2]$ is a telescopic sum, we end up with:

$$\lim_{n \rightarrow \infty} \sum_{n=1}^N \mathbb{E}[S_{T \wedge n}^2] = \mathbb{E}[X_1^2] \mathbb{E}[T]$$

Use Monotone Convergence Theorem to assert the claim.

Exercise 7. By considering the Rademacher functions, and with $T = n$ the first time at which $S_n > 0$, show that one cannot remove the assumption $\mathbb{E}[T] < \infty$ from Theorem 14.

Hint: Find $\mathbb{P}(\{T = 2n\})$ by counting the paths in the plane of length $2n$, obtained by taking steps, in either direction $(1, 1)$ or the direction $(1, -1)$, which remains above the x -axis until the final step, when the x -axis is reached. This leads to the Catalan numbers C_n .

Proof. Define the stopping time $T = \inf\{n \in \mathbb{N} : S_n > 0\}$.

We now calculate the $\mathbb{P}(\{T = 2n\})$.

For the sum to reach zero or be under zero, for every n steps to the left, we need n steps to the right. There are $\binom{2n}{n}$ ways of choosing n lefts and n rights.

For every s steps to the left (-1) , we need $(s - 1)$ steps to the right $(+1)$ for $S_n < 0$ at $n = s$ and $T < \infty$. Hence, we need one less right step than left steps so that $S_n < 0$ for a $2n$ -step process. There are $\binom{2n}{n-1}$ such situations that arises in a $2n$ -steps process.

Substrating from $\binom{2n}{n}$, we get the Catalan numbers

$$C_n = \binom{2n}{n} - \binom{2n}{n-1} = \frac{1}{(n-1)} \binom{2n}{n}$$

We have that $E[X_1^2] = 1$ and $\mathbb{P}(T = 2n) = 0$. We have that $\mathbb{P}(T = 2n + 1) = \frac{1}{(n-1)} \binom{2n}{n} \left(\frac{1}{2}\right)^{(2n+1)}$

$$\begin{aligned} \mathbb{E}[T] &= \sum_{\substack{k=1 \\ k \text{ odd}}}^{\infty} k \mathbb{P}(T = k) \\ &= \sum_{n \in \mathbb{N}} (2n + 1) \mathbb{P}(T = 2n + 1) \\ &= \sum_{n \in \mathbb{N}} \frac{(2n + 1)}{(n - 1)} \frac{(2n)!}{(n!)^2} \left(\frac{1}{2}\right)^{2n+1} \end{aligned}$$

Stirling's Formula:

$$\begin{aligned} n! &\sim n^n e^{-n} \sqrt{2\pi n} \\ (n!)^2 &\sim n^{2n} e^{-2n} (2\pi n) = n^{(2n+1)} e^{-2n} (2\pi) \\ (2n)! &\sim 2^{2n} n^{2n} e^{-2n} \sqrt{4\pi n} \end{aligned}$$

Using Stirling's formula, we can conclude that $\mathbb{E}[T] = \infty$ Now, observe that:

$$\begin{aligned} \mathbb{E}[S_T^2] &= \mathbb{E}[S_T^2 \mathbf{1}_{\{T \geq n\}}] + \mathbb{E}[S_n^2 \mathbf{1}_{\{T < n\}}] \\ &= \mathbb{E}[S_n^2] + \mathbb{E}[S_T^2] \mathbb{P}(T < n) \\ &= \mathbb{E}[S_n^2] + \mathbb{E}[S_T^2] \mathbb{P}(T < n) \quad (\mathbb{E}[S_T^2] = 1) \\ &< \infty \end{aligned}$$

Hence, $\mathbb{E}[S_T^2] \neq \mathbb{E}[T] \mathbb{E}[X_1^2] = \infty$ □

3.9 Conditional Expectation

Definition 21 (Conditional probability). *For two events A and B , the probability of A given B is*

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

If the two events are independent, we have $\mathbb{P}(A|B) = \mathbb{P}(A)$

Definition 22 (Conditional Expectation). *Given a sub σ -algebra $\mathcal{G}$ of Ω , the conditional expectation of X given $\mathcal{G}$ noted $\mathbb{E}[X|\mathcal{G}]$ is the random variable Y such that almost surely:*

i Y is $\mathcal{G}$ -mesurable.

ii $\mathbb{E}[X\mathbf{1}_A] = \mathbb{E}[Y\mathbf{1}_A] = \mathbb{E}[\mathbb{E}[X|\mathcal{G}]\mathbf{1}_A]$ for $A \in \mathcal{G}$

Note that conditional expectation is also interpreted as orthogonal expectation on $\mathcal{G}$.

The conditional expectation $\mathbb{E}[X|\mathcal{F}]$ enjoys the following properties:

(i) If X is $\mathcal{F}$ -mesurable, then $\mathbb{E}[X|\mathcal{F}] = X$.

Proof. By definition, $\mathbb{E}[X|\mathcal{F}]$ is $\mathcal{F}$ -mesurable. Take $A \in \mathcal{F}$

$$\int_{\Omega} (\mathbb{E}[X|\mathcal{F}] - X)\mathbf{1}_{\{A\}} d\mathbb{P} = \int_A \mathbb{E}[X|\mathcal{F}] d\mathbb{P} - \int_{\Omega} X\mathbf{1}_{\{A\}} d\mathbb{P}$$

By definition of conditional expectation, we have $\int_{\Omega} X\mathbf{1}_{\{A\}} d\mathbb{P} = \int_{\Omega} \mathbb{E}[X|\mathcal{F}]\mathbf{1}_{\{A\}} d\mathbb{P}$ such that the last expression is equal to zero.

Hence $\mathbb{E}[X|\mathcal{F}] = X$ almost-surely.

□

(ii) If X is $\mathcal{F}$ -independent, then $\mathbb{E}[X|\mathcal{F}] = \mathbb{E}[X]$

Proof. Note that $\mathbb{E}[X]$ is a constant and constants are measurable such that $\mathbb{E}[X]$ is $\mathcal{F}$ -mesurable.

For $A \in \mathcal{F}$, we have:

$$\begin{aligned} \mathbb{E}[X\mathbf{1}_{\{A\}}] &= \mathbb{E}[X]\mathbb{E}[\mathbf{1}_{\{A\}}] && \text{(Independence of } X \text{ and } \mathcal{F}) \\ &= \mathbb{E}[\mathbb{E}[X]\mathbf{1}_{\{A\}}] \\ &= \mathbb{E}[\mathbb{E}[X|\mathcal{F}]\mathbf{1}_{\{A\}}] && \text{(Definition Conditional Expectation)} \end{aligned}$$

Hence $\mathbb{E}[X|\mathcal{F}] = \mathbb{E}[X]$ almost-everywhere.

□

(iii) If Y is $\mathcal{F}$ -mesurable and bounded (or of a definite sign if X is not in L^2), then $\mathbb{E}[XY|\mathcal{F}] = Y\mathbb{E}[X|\mathcal{F}]$

Proof. $Y\mathbb{E}[X|\mathcal{F}]$ is $\mathcal{F}$ -mesurable.

For $A \in \mathcal{F}$, take $Y = \mathbf{1}_{\{A\}}$

Pour tout $B \in \mathcal{F}$,

$$\begin{aligned} \mathbb{E}[Y\mathbb{E}[X|\mathcal{F}]\mathbf{1}_{\{B\}}] &= \mathbb{E}[\mathbb{E}[X|\mathcal{F}]\mathbf{1}_{\{A\}}\mathbf{1}_{\{B\}}] \\ &= \mathbb{E}[\mathbb{E}[X|\mathcal{F}]\mathbf{1}_{\{A \cap B\}}] \\ &= \mathbb{E}[X\mathbf{1}_{\{A \cap B\}}] && (A \cap B \in \mathcal{F}) \\ &= \mathbb{E}[X\mathbf{1}_{\{A\}}\mathbf{1}_{\{B\}}] \end{aligned}$$

Since it holds true for $Y = \mathbf{1}_{\{A\}}$, it holds true for any Y by Monotone Convergence Theorem or Dominated Convergence Theorem.

□

- (iv) If $X \geq 0$, then $\mathbb{E}[X|\mathcal{F}] \geq 0$
- (v) $|\mathbb{E}[X|\mathcal{F}]| \leq \mathbb{E}[|X||\mathcal{F}]$
- (vi) $\mathbb{E}[\cdot|\mathcal{F}]$ is linear.

(vii) (Conditional Jensen's inequality) If $\varphi \geq 0$ is convex on the real line, we have $\varphi(\mathbb{E}[X|\mathcal{F}]) \leq \mathbb{E}[\varphi(X)|\mathcal{F}]$

(vii) $\|\mathbb{E}[X|\mathcal{F}]\|_p \leq \|X\|_p$ for $1 \leq p \leq \infty$

Use Jensen's inequality with $x \mapsto |x|^p$

Lemma 15. *The analogues of the Monotone and Dominated Convergence Theorem as well as Fatou's lemma hold true for conditional expectation.*

Lemma 16. *Fix $1 \leq p < \infty$ and suppose given a sequence $\{X_j\}_j^n$ of independent L_1 variables with $\mathbb{E}[X_j] = 0$. Then, we have*

$$\sum_{j=1}^n X_j \in L^p \iff \left(\sum_{j=1}^n |X_j|^2\right)^{\frac{1}{2}} \in L^p \iff X_j \in L^p$$

Proof. $\sum_{j=1}^n X_j \in L^p \implies X_j \in L^p$

It suffices to consider the expansion of $(\sum_{j=1}^n X_j)^p$ as $(\sum_{j=1}^n X_j)^p = (X_i + \sum_{j \neq i}^n X_j)^p$ and use Fubini-Tonelli to assert the implication.

$$X_j \in L^p \implies \left(\sum_{j=1}^n |X_j|^2\right)^{\frac{1}{2}} \in L^p$$

Since $X_j^2 \in L^p$, then we can opine that $X_j \in L^{\frac{p}{2}}$ and the implication follows.

$$\left(\sum_{j=1}^n |X_j|^2\right)^{\frac{1}{2}} \in L^p \implies \sum_{j=1}^n X_j \in L^p \text{ in}$$

This is a direct consequence of Cauchy's inequality.

□

Proposition 5. *Let $\{X_j\}_{j=1}^\infty$ be independent L^1 variables with $\mathbb{E}[X_j] = 0$. Then, for every $1 \leq p \leq \infty$, one has*

$$A_p \left\| \left(\sum_{j=1}^n |X_j|^2\right)^{\frac{1}{2}} \right\|_p \leq \left\| \sum_{j=1}^n |X_j| \right\|_p \leq B_p \left\| \left(\sum_{j=1}^n |X_j|^2\right)^{\frac{1}{2}} \right\|_p$$

where A_p, B_p depends only on p .

Proof. Suppose the random variable X_j is symmetric.

We randomize the sum using Rademacher functions $\{r_j\}_j^n$ independent from $\{X_j\}$:

$$\mathbb{E} \left| \sum_{j=1}^n r_j X_j \right|^p = \mathbb{E} \left| \sum_{j=1}^n X_j \right|^p$$

Note that the expectation to the left is a "double" expectation over the two independent domain sets - r_j and X_j - random variables. It suffices to expand the sum with Newton's Multinomial Formula and the symmetry of X_j and r_j to assert the equality of above equality.

The claim then follows using Khinchine's inequality. Note that the term $(\sum_{j=1}^n |X_j|^2)^{\frac{1}{2}}$ is well-defined by the Lemma 16.

If X_j is not symmetric, we use a symmetrization argument: Take $\{X'_j\}_{j=1}^n$ another copy of $\{X_j\}_{j=1}^n$, but independent of it, and set $\tilde{X}_j := X_j - X'_j$ which is symmetric. The claim follows through the same previous construction. $\square$

3.10 Martingales

Definition 23 (Martingale). Suppose $\{X_n\}_{n=1}^\infty$ is a sequence of variables such that X_n is measurable with respect to $\mathcal{F}_n$. We say that $\{X_n\}_{n=1}^\infty$ is a martingale if:

- (i) $\mathbb{E}[|X_n|] < \infty$
- (ii) the sequence $(X_n)_{n=1}^\infty$ is $\mathcal{F}_n$ -mesurable.
- (iii) For all $n \geq 0$,

$$\mathbb{E}[X_{n+1}|\mathcal{F}_n] = X_n$$

We replace condition (iii) with $\mathbb{E}[X_{n+1}|\mathcal{F}_n] \geq X_n$ and $\mathbb{E}[X_{n+1}|\mathcal{F}_n] \leq X_n$ for sub-martingales and super-martingales.

Definition 24 (Predictable). A sequence $\{H_n\}_n$ of random variables is said to be predictable for $\{\mathcal{F}_n\}_{n=0}^\infty$ if for every $n \geq 0$, H_n is $\mathcal{F}_{n-1}$ -mesurable.

3.11 Properties of Martingales

If $(X_n)_{n \geq 0}$ is a martingale (respectively sub-martingale and super-martingale), then:

$$\begin{aligned}\mathbb{E}[X_{n+1}] &= \mathbb{E}[X_n] (= \mathbb{E}[X_0]) \\ \mathbb{E}[X_{n+1}] &\geq \mathbb{E}[X_n] (\geq \mathbb{E}[X_0]) \\ \mathbb{E}[X_{n+1}] &\leq \mathbb{E}[X_n] (\leq \mathbb{E}[X_0])\end{aligned}$$

Proof. It suffices to apply the definition of conditional expectation:

$$\mathbb{E}[X_{n+1}] = \mathbb{E}[\mathbb{E}[X_{n+1}|\mathcal{F}_n]] = \mathbb{E}[X_n]$$

for the martingales case. The case for sub-martingales and super-martingales is analogous. $\square$

Lemma 17 (Martingale and Jensen's inequality). If $(X_n)_{n \geq 0}$ is an $\mathcal{F}_n$ -martingale and $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ is a convex function such that $\mathbb{E}[|\varphi(X_n)|] < \infty$ for all $n \in \mathbb{N}$, then $\varphi(X_n)$ is a sub-martingale.

Proof. (Sketch of) For φ a convex and mesurable function, $\varphi(X_n)$ is an $\mathcal{F}_n$ -mesurable random variable. Using the Conditional Jensen Inequality, we have $\mathbb{E}[\varphi(X_{n+1})|\mathcal{F}_n] \geq \varphi(\mathbb{E}[X_{n+1}|\mathcal{F}_n]) = \varphi(X_n)$ $\square$

Proposition 6. Let $\{X_n\}_{n \geq 0}$ be an $\mathcal{F}_n$ -submartingale. If $(H_n)_n$ is a predictable, positive and bounded sequence, then $(H.X)_n$ defined by:

$$\begin{aligned}(H.X)_0 &= 0 \\ (H.X)_n &= \sum_{m=1}^n H_m(X_m - X_{m-1})\end{aligned}\tag{3.11}$$

is a sub-martingale. The same applies for martingale and super-martingale.

Proof. Note that $(H.X)_n$ is $\mathcal{F}_n$ -mesurable since it is a function of H_m and X_m for $m < n$.

Observe that

$$\begin{aligned}(H.X)_n &= \sum_{m=1}^n H_m(X_m - X_{m-1}) \\ &= H_n(X_n - X_{n-1}) + \sum_{m=1}^{n-1} H_m(X_m - X_{m-1}) \\ &= H_n(X_n - X_{n-1}) + (H.X)_{n-1}\end{aligned}$$

Then, we have:

$$\begin{aligned}\mathbb{E}[(H.X)_{n+1} | \mathcal{F}_n] &= \mathbb{E}[H_{n+1}(X_{n+1} - X_n) + (H.X)_n | \mathcal{F}_n] \\ &= H_{n+1}\mathbb{E}[X_{n+1} - X_n | \mathcal{F}_n] + (H.X)_n \\ &\text{since } H_{n+1} \text{ and } (H.X)_n \text{ is } \mathcal{F}_n\text{-mesurable}\end{aligned}$$

Since $\{X_n\}_n$ is a submartingale, we have $\mathbb{E}[X_{n+1} - X_n | \mathcal{F}_n] \geq 0$ (respectively $= 0$ and ≤ 0) and hence the claim follows. $\square$

Proposition 7. *Let $\{\mathcal{F}_n\}_n$ -filtration and T a stopping time. Suppose $\{X_n\}_{n=1}^\infty$ is a martingale. Then, $\{X_{T \wedge n}\}_n$ is a martingale (also called a Stopped Martingale).*

Mesurability:

Take $B \in \mathcal{B}(\mathbb{R})$. We have

$$\{X_{T \wedge n} \in B\} = \bigcup_{p=1}^{n-1} (\{X_p \in B, T = p\}) \bigcup \{X_n \in B, T = n\}$$

For $p \leq n$, $\{X_p \in B\} \subset \mathcal{F}_p \subset \mathcal{F}_n$

$$\{X_n \in B\} \subset \mathcal{F}_n$$

$$\{T < n\} = \{T < n\}^c \subset \mathcal{F}_n$$

Integrability is immediate by Dominated Convergence Theorem.

Martingale Property:

Note that $H_n = \mathbf{1}_{\{T > n\}}$ is a predictable sequence since $\{T > n\} = \{T \leq (n-1)\}^c$ is $\mathcal{F}_{n-1}$ -mesurable. We then

have:

$$\begin{aligned}
(H.X)_n &= \sum_{k=1}^n H_k(X_k - X_{k-1}) \\
&= \sum_{k=1}^n \mathbf{1}_{T \geq k}(X_k - X_{k-1}) \\
&= \sum_{k=1}^{T \wedge n} \mathbf{1}_{T \geq k}(X_k - X_{k-1}) \\
&\text{If } T < k', \text{ then } \mathbf{1}_{\{T \geq k'\}} = 0 \text{ for } n \geq k' > T \\
&= \sum_{k=1}^{T \wedge n} (X_k - X_{k-1}) \\
&\text{Telescoping Sum} \\
&= X_{T \wedge n} - X_0
\end{aligned}$$

Hence, we have $X_{T \wedge n} = (H.X)_n + X_0$ which proves that $X_{T \wedge n}$ is a martingale.

3.12 Doob's Convergence Theorem

Theorem 15. Let $\{X_n\}_{n=0}^\infty$ be a supermartingale that is L^1 bounded, i.e., such that $\sup_{n \geq 0} \mathbb{E}[X_n] < \infty$. Then, $X_n \rightarrow X_\infty$ almost surely, where X_∞ is real-valued. In particular, non-negative supermartingales converge almost surely to a finite limit.

The above theorem requires a heavily involved proof. We do not include it here, but a reader will encounter the proof in any standard Master 1 curriculum or probability theory textbook.

Proposition 8. Let $\{X_n\}_{n=0}^\infty$ be a submartingale, and let $M_n = \max_{1 \leq k \leq n} X_k$. Then,

$$\lambda \mathbb{P}(M_n \geq \lambda) \leq \mathbb{E}(X_n \mathbf{1}_{\{M_n \geq \lambda\}}) \leq \mathbb{E}[X_n^+]$$

Proof. We begin with a preliminary claim: $\mathbb{E}[X_n] \geq \mathbb{E}[X_{T \wedge n}]$

Define $Y_n := \mathbf{1}_{\{T < k \leq n\}}$ and note that Y_k is $\mathcal{F}_{k-1}$ -mesurable.

Define the following submartingale:

$$\begin{aligned}
Z_0 &:= 0 \\
Z_n &:= \sum_{k=1}^n Y_k(X_k - X_{k-1}) & (n \geq 1) \\
&= \sum_{k=1}^n \mathbf{1}_{\{T < k \leq n\}}(X_k - X_{k-1}) & (n \geq 1)
\end{aligned}$$

If $T = k' \geq n$, then $Z_n = 0$. If $1 < T = k' \leq n$, then

$$\begin{aligned}
\sum_{k=1}^n \mathbf{1}_{\{T < k \leq n\}}(X_k - X_{k-1}) &= \sum_{k=k'}^n (X_k - X_{k-1}) \\
&= X_n - X_{k'}
\end{aligned}$$

Hence, we have:

$$\begin{aligned} Z_n &:= \sum_{k=T \wedge n}^n (X_k - X_{k-1}) \\ &= X_n - X_{T \wedge n} \end{aligned}$$

We have $\mathbb{E}[Z_{n+1} | \mathcal{F}_{n+1}] \geq X_n - X_{T \wedge n}$ is a sub-martingale by Proposition 7.

The second inequality is immediate. We prove the first inequality.

Define the stopping time T as:

$$T := \inf\{k \geq 0 | X_k > \lambda\}$$

Define $A = \{M_n > \lambda\}$. Using the inequality, $\mathbb{E}[X_{T \wedge n}] \leq \mathbb{E}[X_n]$, we have:

$$\mathbb{E}[X_{T \wedge n} \mathbf{1}_{\{A\}}] + \mathbb{E}[X_{T \wedge n} \mathbf{1}_{\{A\}^c}] \leq \mathbb{E}[X_n \mathbf{1}_{\{A\}}] + \mathbb{E}[X_n \mathbf{1}_{\{A\}^c}]$$

On $\{A^c\}$, $T > n$ and hence $T \wedge n = n$ such that the inequality becomes:

$$\mathbb{E}[X_{T \wedge n} \mathbf{1}_{\{A\}}] \leq \mathbb{E}[X_n \mathbf{1}_{\{A\}}]$$

On A , $T \leq n$, $X_{T \wedge n} = X_T$ and $\{X_T \geq \lambda\}$ such that

$$\mathbb{E}[X_T \mathbf{1}_{\{A\}}] \leq \lambda \mathbb{E}[\mathbf{1}_{\{A\}}] = \lambda \mathbb{P}(M_n \geq \lambda)$$

and the claim follows. □

3.13 Doob's maximal inequality

Theorem 16 (Doob's maximal inequality). *Let $\{X_n\}_{n=1}^\infty$ be a martingale. $X_n^* := \sup_{0 \leq k \leq n} |X_k|$ and $1 < p < \infty$. Then*

$$\|X_n^*\|_p \leq \frac{p}{p-1} \|X_n\|_p$$

Proof.

$$\begin{aligned}
\|X_n^*\|_p^p &= \int_0^\infty \mathbb{P}(X_n^* > \lambda) p \lambda^{p-1} d\lambda \\
&= \int_0^\infty \lambda \mathbb{P}(X_n^* > \lambda) p \lambda^{p-2} d\lambda \\
&\leq p \int_0^\infty \mathbb{E}[X_n \mathbf{1}_{\{X_n^* > \lambda\}}] p \lambda^{p-2} d\lambda && \text{(Proposition 8)} \\
&= p \int_0^\infty \int_0^\infty x \mathbf{1}_{\{X_n^* > \lambda\}} dF(x) \lambda^{p-2} d\lambda \\
&= p \int_0^\infty x \int_0^\infty \mathbf{1}_{\{X_n^* > \lambda\}} \lambda^{p-2} d\lambda dF(x) && \text{(Fubini-Tonelli)} \\
&= p \int_0^\infty x \int_0^{X_n^*} \lambda^{p-2} d\lambda dF(x) \\
&= \frac{p}{p-1} \int_0^{X_n^*} x X_n^{*(p-1)} dF(x) \\
&= \frac{p}{p-1} \mathbb{E}[X_n X_n^{*(p-1)}] \\
&\leq \frac{p}{p-1} \mathbb{E}[X_n^p]^{\frac{1}{p}} \mathbb{E}[X_n^{*(\frac{(p-1)p}{p-1})}]^{\frac{p-1}{p}} && \text{(Using Hölder's inequality)} \\
&= \frac{p}{p-1} \mathbb{E}[X_n^p]^{\frac{1}{p}} \mathbb{E}[X_n^{*p}]^{\frac{p-1}{p}}
\end{aligned}$$

and we are done. $\square$

Theorem 17. Let $X := \{X_n\}_{n=0}^\infty$ be a martingale and $1 < p < \infty$. The following are equivalent:

- (i) X is L^p -bounded.
- (ii) X converges a.s. as well as in the L^p -sense.
- (iii) there exists $X_\infty \in L^p$ with $X_n = \mathbb{E}[X_\infty | \mathcal{F}_n]$ for all $n \geq 0$.

Proof. (Sketch of) (i) $\implies$ (ii)

Since an L^p -bounded variable is also L^1 -bounded variable, we can use Doob's inequality coupled with the Monotone Convergence theorem or Dominated Convergence theorem to assert (ii).

(ii) $\implies$ (iii)

Since $X_n \xrightarrow{L^1} X_\infty$, for $A \in \mathcal{F}_n$ and $k \geq n$,

$$|\mathbb{E}[\mathbf{1}_A X_k] - \mathbb{E}[\mathbf{1}_A X_\infty]| \leq \mathbb{E}[\mathbf{1}_A (X_k - X_\infty)] \leq \mathbb{E}[|X_k - X_\infty|]$$

The latter tends to zero as $k \rightarrow \infty$ and hence we can opine

$$\lim_{k \rightarrow \infty} \mathbb{E}[\mathbf{1}_A X_k] = \mathbb{E}[X_\infty \mathbf{1}_A]$$

By the definition of conditional expectation, we have $X_k = \mathbb{E}[X_\infty | \mathcal{F}_k]$

(iii) $\implies$ (ii) is immediate. $\square$

3.14 Kronecker's Lemma and Strong Law of Large Numbers

In this section, we derive the Strong Law of Large Numbers using Martingale Theory.

We start by the preliminary lemma: *Kronecker's lemma*

Lemma 18. *Let B be a Banach space and assume that $\{a_n\}_{n=1}^{\infty} \subset B$ is a sequence of vectors such that $\sum_{n=1}^{\infty} a_n$ converges in B . If $b_n > 0$ is increasing and tends to ∞ , then $\frac{1}{b} \sum_{k=1}^n a_k b_k \xrightarrow{n \rightarrow \infty} 0$*

Proof. Define $A_k = \sum_{n=k}^{\infty} a_n$ such that $a_k = A_k - A_{k+1}$.

We then have:

$$\begin{aligned} \sum_{k=1}^n a_k b_k &= \sum_{k=1}^n (A_k - A_{k+1}) b_k \\ &= \sum_{k=1}^n (A_k b_k) - \sum_{s=1}^n (A_{s+1} b_s) \\ &= A_1 b_1 - A_{n+1} b_n + \sum_{k=2}^n A_k b_k + \sum_{s=1}^n A_{s+1} b_s \quad (\text{Index switch from } s+1 \mapsto s) \\ &= A_1 b_1 - A_{n+1} b_n + \sum_{k=2}^n A_k b_k + \sum_{k=2}^n A_k b_{k-1} \end{aligned}$$

The latter converges by Cesàro's summation lemma.

□

Theorem 18 (Strong Law of Large Numbers). *Assume that $\{X_j\}_{j=1}^{\infty}$ are independent L^{2p} variables with $1 \leq p < \infty$ arbitrary but fixed, $\mathbb{E}[X_j] = 0$ and $\sup_j \mathbb{E}[|X_j|^{2p}] < \infty$. Then, with $S_{n,2} = \sum_{1 < i_1 \leq i_2 \leq n} X_{i_1} X_{i_2}$, one has:*

$$\frac{S_{n,2}}{\binom{n}{2}} \xrightarrow{n \rightarrow \infty} 0$$

almost-surely and in the L^p -sense.

Proof. Denote $S_n = \sum_{j=1}^n X_j$ and define the sequence $b_n := \binom{n}{2}$.

We define the sequence of random variables:

$$T_n := \sum_{m=2}^n X_m S_{m-1} b_m^{-1}$$

Note that T_n is a martingale with respect to the natural filtration and the independence of the X_j variables.

From the triangle inequality and Proposition 5, we have the following inequality:

$$\sum_{n \geq 2} \|T_n\|_p \leq \sum_{m=2}^{\infty} b_m^{-1} \|X_m\|_{2p} \|S_{m-1}\|_{2p} \leq \sum_{m=2}^{\infty} m^{-\frac{3}{2}} < \infty$$

We can then affirm that $\{T_n\}_n$ converges almost surely and in the L^p .

We can then use the previous lemma to deduce that

$$b_n^{-1} S_{n,2} \rightarrow 0$$

almost surely and in the L^p -sense.

□

4 Fourier series and Randomness

In this section, we explicit weak L^1 conditions needed to disprove the convergence of Fourier series of $L^1(\mathbb{T})$ functions.

We take a different direction than Section 6.1 of Muscalu and Schlag to present a less abstract and more adapted proof for the purpose of the internship. The following proof is based on the Terence Tao's lecture notes [2].

Recall:

We say that a function g has a weak L^1 bound if there exists a finite constant C such that

$$|\{\theta \in \mathbb{T} \mid |g(\theta)| > \lambda\}| \leq \frac{C}{\lambda}$$

for all $\lambda > 0$.

Recall: The Infinite Product Measure

Given the following measure spaces $(E_i, \mathcal{B}_i, \mu_i)_{i \in I}$ E_i a given space, $\mathcal{B}_i$ its σ -algebra, and μ_i its measure and I an index set. The product space is given as the triple $E = \prod_i E_i$, $\mathcal{B} = \sigma((\mathcal{B}_i)_{i \in I})$ - the smallest σ -algebra containing the $\prod_{i' \text{ finite}} \mathcal{B}_{i'} \times \prod_i E_i$. Define the product measure as $\mu(B) = \prod_{j \in J} \mu(B_j)$ for J finite. If $\mu_i(E_i) < \infty$, replace μ_i with $\tilde{\mu}_i = \frac{\mu_i}{\mu_i(E_i)}$ to obtain a probability space.

Theorem 19. *Paley-Zygmund Inequality* If $Z \geq 0$ is a random variable with finite variance, and if $0 \leq \theta \leq 1$, then

$$\mathbb{P}(Z > \theta \mathbb{E}[Z]) \geq (1 - \theta)^2 \frac{\mathbb{E}(Z)^2}{\mathbb{E}(Z^2)}$$

Proof. $\mathbb{E}[Z] = \mathbb{E}[Z \mathbf{1}_{Z < \theta \mathbb{E}[Z]}] + \mathbb{E}[Z \mathbf{1}_{Z > \theta \mathbb{E}[Z]}] \leq \theta \mathbb{E}[Z] + \mathbb{E}[Z \mathbf{1}_{Z > \theta \mathbb{E}[Z]}]$ Using Cauchy-Schwarz's inequality, $\mathbb{E}([Z] \mathbf{1}_{Z > \theta \mathbb{E}[Z]}) \leq (\mathbb{E}[Z^2])^{\frac{1}{2}} \mathbb{P}(Z > \theta \mathbb{E}[Z])^{\frac{1}{2}}$. Hence, $\mathbb{E}[Z] \leq \theta \mathbb{E}[Z] + (\mathbb{E}[Z^2])^{\frac{1}{2}} \mathbb{P}(Z > \theta \mathbb{E}[Z])^{\frac{1}{2}}$ and the inequality follows. $\square$

Lemma 19. Let $\{E_n\}_{n=1}^{\infty}$ be a sequence of measurable subsets of $\mathbb{T}$ such that

$$\sum_{n=1}^{\infty} |E_n| = \infty$$

Then, there exists a sequence $\{x_n\}_{n=1}^{\infty} \in \mathbb{T}$ such that

$$\sum_{n=1}^{\infty} \mathbf{1}_{\{E_n\}}(x + x_n) = \infty$$

for almost every $x \in \mathbb{T}$.

Proof. First, we take the infinite (probabilised) measure space $\Omega := \prod_{n=1}^{\infty} \mathbb{T}$. We will prove that the event $A_x = \{(x_1, \dots, x_n, \dots) \in \Omega \mid \sum_{n=1}^{\infty} \mathbf{1}_{\{E_n\}}(x + x_n) = \infty\}$ for every $x \in \mathbb{T}$ happens infinitely often.

Observe that

$$\begin{aligned}
A_x &= \{\{x_i\}_{i=1}^\infty \in \Omega \mid \sum_{n=1}^\infty \mathbf{1}_{\{E_n\}}(x + x_n) = \infty\} \\
&= \{\{x_i\}_{i=1}^\infty \in \Omega \mid x \text{ belongs to infinitely many } E_n(\cdot + x_n)\} \\
&\text{Using the limit interpretation of the above statement} \\
&= \{\{x_i\}_{i=1}^\infty \in \Omega \mid \bigcap_{N=1}^\infty \bigcup_{m=N}^\infty (x_m + E_m)\} \\
&= \{\{x_i\}_{i=1}^\infty \in \Omega \mid \bigcap_{N=1}^\infty A_N\} \quad (A_N := \bigcup_{m=N}^\infty (x_m + E_m)) \\
&= \bigcap_{N=1}^\infty A_N
\end{aligned}$$

Use DeMorgan's Law, $A_N^c = \{\{x_i\}_{i=1}^n \mid x \in \bigcap_{m=N}^\infty (x_m + E_m^c)\}$

Using the definition of the product measure, we have

$$\mathbb{P}(A_N^c) = \prod_{m=N}^\infty (1 - |E_m|) = 0$$

using the fact that $1 - y \leq e^{-y}$ for $y > 0$. Since $A_x^c = \bigcup_{N=1}^\infty A_N^c$, we have $\mathbb{P}(A_x^c) = 0$. Using the Fubini-Tonelli theorem, the lemma follows. $\square$

Lemma 19 proves that there exists deterministic choices of $x_i \in \mathbb{T}$ such that $x \in x_i + E_i$ for a fixed sequence of measurable subsets E_i .

Theorem 20. *There is an equivalence between:*

- *There exists a function $f \in L^1(\mathbb{T})$ such that $\sup_{N \geq 0} |S_N f(x)| = \infty$ for almost every $x \in \mathbb{T}$.*
- *The maximal inequality*

$$\sup_{\lambda > 0} |\{x \in \mathbb{T} : |S_N f(x)| \geq \lambda\}| \leq \frac{A \|f\|_1}{\lambda}$$

does not hold for all smooth $f \in \mathbb{T} \mapsto \mathbb{C}$ for all $\lambda > 0$.

Proof. Suppose there exists such an $f \in L^1(\mathbb{T})$ such that for $\lambda > 0$ and $N_0 \in \mathbb{N}$ such that for any $A > 0$, we have

$$|\{x \in \mathbb{T} : |S_N f(x)|_{0 \leq N \leq N_0} \geq \lambda\}| > \frac{A \|f\|_1}{\lambda}$$

Subject to a trivial re-scaling, we choose $\lambda = 1$. Denote $|\{x \in \mathbb{T} : |S_N f(x)|_{0 \leq N \leq N_0} \geq 1\}|$ by E .

However, E is not necessarily a set of measure 1, that is, the whole of $\mathbb{T}$. To show $\mathbb{T}$ -everywhere divergence of an L^1 function, we need to define a function over the whole of $\mathbb{T}$ using the property of f over E .

For this purpose, set $E_i = E$ in Lemma 19 to construct a set of measure 1 over which our constructed function's Fourier series will diverge. From our supposition, it satisfies the necessary condition for the existence of deterministic choices of $\{x_i\}_{i=1}^n$ such that for all $x \in \mathbb{T}$, and hence we have $x \in E + x_i$ for some $i \in \{1, \dots, n\}$. Using the definition of the product measure and Fubini-Tonelli as in Lemma 19, $|\bigcup_{i=1}^n E + x_i| = 1 - (1 - |E|)^n$.

Choose $n \in \mathbb{N}$ big enough and $n > \frac{1}{|E|}$ such that $|\bigcup_{i=1}^n E + x_i| \approx 1$.

Now, define the function $\tilde{f}$ over $\mathbb{T}$ using a randomised linear combination:

$$\tilde{f}(x) := \sum_{i=1}^n \epsilon_i f(x - x_i)$$

where ϵ_i are Rademacher functions.

Using the triangle inequality, we have the inequality $\mathbb{E}[\|\tilde{f}\|] \leq \sum_{i=1}^n \|f\|_1 = n\|f\|_1 \leq \frac{1}{|A|}$ since n is greater than $\frac{1}{|E|}$.

We would like to have deterministic choices of ϵ_i for which the Fourier series $\sup_{0 \leq N \leq N_0} (\tilde{f})$ is greater than 1 on the $\mathbb{T}$. To this end, we make use of the same probabilistic method as in Lemma 19.

By hypothesis, $1 \leq i \leq n$

$$\sup_{\substack{0 \leq N \leq N_0, \\ 1 \leq i \leq n}} |S_N f(x - x_i)| \geq 1$$

which is equivalent to the existence of $0 \leq N_x \leq N_0$ such that

$$\sup_{1 \leq i \leq n} |S_{N_x} f(x - x_i)| \geq 1$$

Denote $Q_x = \sum_{i=1}^n (|S_{N_x} f(x - x_i)|^2)^{\frac{1}{2}}$, and note that $Q_x \geq 1$ by the above statement. Khintchine's inequality asserts that

$$\begin{aligned} A_1 Q_x &\leq \mathbb{E}[|S_{N_x} \tilde{f}(x)|] \leq B_1 Q_x \\ (A_2 Q_x)^2 &\leq (\mathbb{E}[|S_{N_x} \tilde{f}(x)|^2]) \leq (B_2 Q_x)^2 \end{aligned}$$

for constants A_1, A_2, B_1 and B_2 .

From the above inequality, we assert that for certain constant θ , we have

$$\mathbb{P}(|S_{N_x} \tilde{f}(x)| > \theta A_1 Q_x) \geq \mathbb{P}(|S_{N_x} \tilde{f}(x)| > \theta \mathbb{E}[|S_{N_x} \tilde{f}(x)|])$$

Using Paley-Zygmund inequality, we have

$$\begin{aligned} \mathbb{P}(|S_{N_x} \tilde{f}(x)| > \theta \mathbb{E}[|S_{N_x} \tilde{f}(x)|]) &\geq (1 - \theta)^2 \frac{\mathbb{E}[|S_{N_x} \tilde{f}(x)|]^2}{\mathbb{E}[|S_{N_x} \tilde{f}(x)|^2]} \\ &\geq (1 - \theta)^2 \frac{A_1^2}{B_2^2} \end{aligned}$$

Choose θ such that $(1 - \theta)^2 \frac{A_1^2}{B_2^2} \approx 1$ and $\mathbb{P}(|S_{N_x} \tilde{f}(x)| > \theta A_1 Q_x) = \mathbb{P}(|S_{N_x} \tilde{f}(x)| \gtrsim 1)$. Hence, for any given sequence $\{\epsilon_i\}_{i=1}^n$, we have $\mathbb{P}(|S_{N_x} \tilde{f}(x)| \gtrsim 1) \geq 1$ and $\mathbb{P}(\sup_{1 \leq N \leq N_0} |S_N \tilde{f}(x)| \gtrsim 1) \gtrsim 1$.

Consider the infinite probability space $\Omega' = \prod_{i=1}^n \{-1, 1\} \times \Omega$ with the product measure, where Ω is defined as in Lemma 19. For any $x \in \mathbb{T}$, denote the event $A_x \subset \Omega$ by $\{\{\epsilon\}_{i=1}^n \times (x_i)_{i=1}^n \mid |S_{N_x} \tilde{f}(x)| = \sum_{i=1}^n \epsilon_i S_N f(x - x_i)| \gtrsim 1\} \geq 1\}$.

Since $\mathbb{P}(|S_{N_x} \tilde{f}(x)| \gtrsim 1) \gtrsim 1$, we have $\mathbb{P}(A_x) \gtrsim 1$ and hence, Fubini-Tonelli affirms that there exists deterministic choices of ϵ_i such that $|\{\sup_{1 \leq N \leq N_0} |S_N \tilde{f}(x)| \gtrsim 1\}| \approx 1$.

Now, we have all the ingredients necessary to construct the required $L^1(\mathbb{T})$ function f such that $\sup_{N \geq 0} S_N f = \infty$.

For every $A > 0$, we can find

- a smooth function f_A such that $\|f_A\|_1 \leq \frac{1}{A}$
- $N_A \in \mathbb{N}$
- $E_A \subset \mathbb{T}$ with $|E_A| \sim 1$ (with choices of $\{x_i\}_{i=1}^n$) such that

$$\sup_{0 \leq N \leq N_A} |S_N f(x)| \gtrsim 1$$

for $x \in E_A$

Now, we apply these criteria recurrently for $m = 1, 2, \dots$, we have:

1. $\|f^{(m)}\|_1 \leq \frac{1}{2^m}$ (by choosing $A^{(m)}$ appropriately)
2. $|E^{(m)}| \approx 1$
3. $\sup_{0 \leq N \leq N^{(m)}} |S_N f^{(m)}(x)| \geq m + \sum_{m' < m} \sup_{N \geq 0} \|S_N f^{(m')}\|_{L^\infty(\mathbb{T})}$ for $x \in E^{(m)}$
4. $\sup_{0 \leq N \leq N^{(m')}} \|S_N f^{(m')}\|_{L^\infty(\mathbb{T})} \leq \frac{1}{2^{m'}}$ for $m' < m$

Now, define $f := \sum_{m \geq 1} f^{(m)}$. By Condition 1, $f \in L^1(\mathbb{T})$ since

$$\|f\|_{L^1(\mathbb{T})} \leq \sum_{m \geq 1} \|f^{(m)}\|_1 \leq \sum_{m \geq 1} \frac{1}{2^m} < \infty$$

For any $x \in \mathbb{T}$, we have:

$$\sup_{0 \leq N \leq N^{(m)}} |S_N f| = \sup_{0 \leq N \leq N^{(m)}} \left| \sum_{j=1}^{\infty} S_N f^j \right|$$

Using reverse triangle inequality, we have:

$$\begin{aligned} &\geq \sup_{0 \leq N \leq N^{(m)}} \left| |S_N f^m| - \left| \sum_{j \neq m} S_N f^j \right| \right| \\ &= \sup_{0 \leq N \leq N^{(m)}} \left| S_N f^m - \sum_{j < m} S_N f^j - \sum_{j > m} S_N f^j \right| \end{aligned}$$

Replacing Condition 3:

$$\geq m - \sum_{j \geq m} S_N f^{(j)}$$

Replacing Condition 4:

$$\geq m - \sum_{j \geq m} \frac{1}{2^m}$$

This holds true for infinitely many m and hence $\sup_{N \geq 0} |S_N f(x)| = \infty$. □

4.1 Reasonant measures for the Dirichlet's Kernel

In this section, we build an explicit L^1 function that violates the weak L^1 -type condition so that its Fourier series diverges in $\mathbb{T}$ by the previous theorem. To this end, we make use of the highly oscillatory nature of the Dirichlet's kernel to create a measure μ that has the same peaks.

We start with the preliminary definition and lemma needed to define μ .

Definition 25 (Incommensurate). *We say that a vector $(\theta_1, \dots, \theta_d) \in \mathbb{T}$ is incommensurate if for any $(n_1, \dots, n_d) \in \mathbb{Z}^d \setminus \{0\}$, one has that $n_1\theta_1 + \dots + n_d\theta_d \notin \mathbb{Z}$*

Lemma 20. *The set of commensurate vectors in $\mathbb{T}^d$ have measure zero.*

Proof. Define A as $\{\theta = (\theta_1, \dots, \theta_n) | \theta \text{ is commensurate}\}$ and denote $B_i := \{\theta_i \in \mathbb{T} | n\theta_1 + \dots + n\theta_i + \dots + n\theta_d \in \mathbb{Z} | \{\theta_j\}_{j \neq i} \text{ fixed}\}$ for $n \in \mathbb{Z}$. Note that B_i has measure zero since $\mathbb{Z}$ is a measure zero set in $\mathbb{R}$.

Using Fubini-Tonelli theorem, one has

$$\lambda(A) = \int_{\mathbb{T}^{d-1}} \lambda(B_i) dx = 0.$$

□

Lemma 21. *Assume that $(\theta_1, \dots, \theta_d) \in \mathbb{T}^d$ is incommensurate. Then, the orbit*

$$\{(n\theta_1, \dots, n\theta_d) \mod \mathbb{Z}^d | n \in \mathbb{Z}\} \subset \mathbb{T}^d$$

is dense in $\mathbb{T}^d$.

Proof. To show that $\{n\theta = (n\theta_1, \dots, n\theta_d) \mod \mathbb{Z}^d | n \in \mathbb{Z}\}$ is dense, it suffices to prove that for any smooth function f over $\mathbb{T}^d$, we have:

$$\frac{1}{N} \sum_{n=1}^N f(n\theta_1, \dots, n\theta_d) \longrightarrow \int_{\mathbb{T}^d} f(x) dx = \hat{f}(0)$$

For otherwise, if a set were not to be dense, then it does not intersect another set A . One could build a smooth positive function g (e.g. with partitions of unity) such that on A , $g > 0$, and hence

$$\frac{1}{N} \sum_{n=1}^N g(n\theta_1, \dots, n\theta_d) = 0 \neq \int_{\mathbb{T}^d} g(x) dx$$

Since $f \in C^\infty(\mathbb{T})$, we can expand it into its Fourier series: $f(n\theta) = \sum_{v \in \mathbb{Z}^d} \hat{f}(v) e(n\theta.v)$

$$\begin{aligned} \frac{1}{N} \sum_{n=1}^N f(n\theta_1, \dots, n\theta_d) &= \frac{1}{N} \sum_{n=1}^N f(n\theta_1, \dots, n\theta_d) \\ &= \frac{1}{N} \sum_{n=1}^N \sum_{v \in \mathbb{Z}^d} \hat{f}(v) e(n\theta.v) \\ &= \sum_{v \in \mathbb{Z}^d} \hat{f}(v) \frac{1}{N} \sum_{n=1}^N e(n\theta.v) \end{aligned}$$

Using the geometric partial sum formula

$$= \hat{f}(0) + \sum_{v \in \mathbb{Z}^d \setminus \{0\}} \hat{f}(v) \frac{1}{N} \left[\frac{1 - e((N+1)\theta.v)}{1 - e(\theta.v)} \right]$$

By Exercise 1, the Fourier coefficients are rapidly decreasing in size since f is a smooth function. Since $|\frac{1}{N} \frac{1-e((N+1)\theta.v)}{1-e(\theta.v)}| \leq 1$ for all $N \geq 1$, we have

$$\sum_{v \in \mathbb{Z}^d \setminus \{0\}} \hat{f}(v) \frac{1}{N} \left[\frac{1-e((N+1)\theta.v)}{1-e(\theta.v)} \right] \rightarrow 0$$

such that $\frac{1}{N} \sum_{n=1}^N f(n\theta_1, \dots, n\theta_d) \rightarrow \hat{f}(0)$ □

Now, we construct the measure that violates the weak- $L^1(\mathbb{T})$ bound. For any $N \in \mathbb{N}$, we define

$$\mu_N := \frac{1}{N} \sum_{j=1}^N \delta_{\{x_{j,N}\}}$$

where $x_{j,N}$ is chosen to be close to $\frac{j}{N}$. Then, the Fourier series are given by:

$$(S_n \mu_N)(x) = \frac{1}{N} \sum_{j=1}^N \frac{\sin((2n+1)\pi(x - x_{j,N}))}{\sin(\pi(x - x_{j,N}))}$$

Theorem 21. *There exists a sequence $\{\mu_N\}_{N \geq 1}$ of probability measures on $\mathbb{T}$ with the property that, for each $N \geq 1$,*

$$\limsup_{n \rightarrow \infty} \frac{|S_n \mu_N(x)|}{\log(N)} > 0$$

for almost every $x \in \mathbb{T}$

Proof. For every $N \geq 1$, we can choose $x_{j,N} \in \mathbb{T}$ such that $|x_{j,N} - \frac{j}{N}| \leq \frac{1}{N^2}$ for $j \in \{1, \dots, N\}$ and such that $\{x_{j,N}\}_{j=1}^N$ is an incommensurate vector. By Lemma 20, this is possible since commensurate vectors have measure zero.

For any $x \in \mathbb{T}$, $x \in [\frac{i}{N}, \frac{i+1}{N}]$ for $i \in \{0, \dots, (N-1)\}$ and $|x - \frac{i}{N}| \leq \frac{1}{N}$. For the case where $j < i$, we have

$$\begin{aligned} |x - x_{j,N}| &\leq |x - \frac{i}{N}| + |\frac{i}{N} - x_{j,N}| \\ &\leq |x - \frac{i}{N}| + |\frac{i}{N} - \frac{(i-1)}{N}| + \dots + |\frac{(j+1)}{N} - \frac{(j)}{N}| + |\frac{j}{N} - x_{j,N}| \\ &\leq \frac{1}{N} + \frac{i-j-1}{N} + \frac{1}{N^2} \\ &\leq \frac{(i-j)}{N} + \frac{1}{N^2} \\ &\leq \frac{j}{N} \end{aligned}$$

By Lemma 21, we can argue that $\{(2n+1)(x - x_{j,N})\}_{j=1}^N \pmod{\mathbb{Z}^N} | n \in \mathbb{Z}\}$ is dense in $\mathbb{T}^N$ for almost every $x \in \mathbb{T}$.

It follows then that for almost every $x \in \mathbb{T}$, there are infinitely many choices of n such that

$$\sin((2n+1)\pi(x - x_{j,N})) \geq \frac{1}{2}$$

for all $1 \leq j \leq N$.

Using the inequality $|\sin(t)| \geq |t|$ for $t \in [0, \frac{\pi}{2}]$, we have for

$$|(S_n \mu_N)(x)| \geq \frac{1}{2N} \sum_{j=1}^N \frac{1}{|\sin(\pi(x - x_{j,N}))|} \geq \frac{\pi}{N} \sum_{j=1}^{\lfloor \frac{N}{2} \rfloor} \frac{1}{N} \geq C \log(\lfloor \frac{N}{2} \rfloor)$$

for some appropriate constant C . □

Combining Theorem 21 with Theorem 20 yields the existence of such a function whose Fourier series diverges.

5 Kolmogorov's proof [1923]

In his paper published in 1923, Kolmogorov did not make use of a weak $L^1(\mathbb{T})$ -bound inequality to prove almost-everywhere divergence of Fourier series of an $L^1(\mathbb{T})$ function [3]. Instead, he built an explicit function ¹.

First, we list the necessary ingredients necessary.

Construction of a sequence of function $\{\varphi_n\}_{n=1}^\infty$

1. $\varphi_n \geq 0$; $\int_0^{2\pi} \varphi_n(x) dx = 2$
2. $\sup_{N \geq 0} |S_N \varphi_n| < \infty$ for all n
3. For every $\varphi_n(x)$, we have $M_n > 0$, $E_n \subset \mathbb{T}$, and $q_n \in \mathbb{N}$ such that
 - (i) $\lim_{n \rightarrow \infty} M_n = \infty$
 - (ii) $\lim_{n \rightarrow \infty} E_n = 2\pi$
 - (iii) for $x \in E_n$, there exists a $q_n \in \mathbb{N}$ such that for $k \leq q_n$

$$|S_k \varphi_n(x)| \geq M_n$$

Construction of these functions will lead to an almost everywhere divergent Fourier series.

5.1 Construction of said functions $\{\varphi_n(x)\}_{n=1}^\infty$

To obtain $\varphi_n(x)$, we suppose, for a fixed n , that we are given a sequence of odd increasing sequence of natural numbers [The criteria are described further in the proof]

$$\lambda_1 = 1, \dots, \lambda_n$$

Choose for $1 \leq k \leq n$: $\{m_k\}_k$ a sequence of real numbers such that:

$$\begin{aligned} m_1 &= n \\ 2m_k + 1 &= \lambda_k(2n + 1) \end{aligned}$$

Note that $\{m_k\}_k$ is an increasing sequence since λ_k is an increasing sequence as well.

For $1 \leq k \leq n$, we choose

$$A_k = \frac{4\pi k}{(2n + 1)}$$

Hence, we obtain the intervals $\Delta_k := [A_k - \frac{1}{m_k^2}, A_k + \frac{1}{m_k^2}]$

¹It is to be noted that Kolmogorov proved this theorem at the young age of 19.

The intervals Δ_k s are mutually disjoint. Observe that

$$|A_k - A_{k+1}| = \frac{4\pi}{(2n+1)}$$

and the intervals Δ_k are intervals centered around A_k of radius $\frac{1}{m_k^2}$. We claim that for any $k \in \{1, \dots, n\}$, we have $\frac{1}{m_k^2} < \frac{1}{2} \frac{4\pi}{(2n+1)}$ such that the $\Delta_k \subset (\frac{\pi(k-1)}{(2n+1)}, \frac{\pi k}{(2n+1)})$ which are mutually disjoint sets. For any such k , we have:

$$\begin{aligned} m_k^2 &= \frac{(\lambda_k(2n+1) - 1)^2}{4} \\ m_k^2 &\geq \frac{(\lambda_k(2n+1))^2}{4} \\ m_k^2 &\geq \frac{\lambda_k^2(4n^2 + 4n + 1)}{4} \\ m_k^2 &\geq \frac{\lambda_k^2(4n^2)}{4} \\ m_k^2 &\geq \frac{\lambda_k^2(n^2)}{\pi} \\ m_k^2 &\geq \frac{\lambda_k^2(2n+1)}{2\pi} \end{aligned} \tag{*}$$

and the claim follows. Δ_k s are mutually disjoint.

Hence, we can now define the functions $\varphi_n(x)$:

$$\varphi_n(x) = \begin{cases} \frac{m_k^2}{n} (:= \varphi_n^{(k)}(x)) & \text{on } \Delta_k \text{ for } 1 \leq k \leq n \\ 0 & \text{on } [0, 2\pi] \setminus \bigcup_k \Delta_k \end{cases}$$

Condition (1):

By construction, $\varphi_n \geq 0$.

$$\begin{aligned} \int_0^{2\pi} \varphi_n(x) dx &= \sum_{k=1}^n \int_{\Delta_k} \frac{m_k^2}{n} dx \\ &= \sum_{k=1}^n \frac{2}{m_k^2} \frac{m_k^2}{n} \\ &= 2. \end{aligned}$$

Condition (2) is immediate since $\varphi_n(x)$ is a piecewise constant function.

Condition (3):

Define the segments $\{\sigma_k\}_{k=2}^n$ by

$$\sigma_k = [A_{k-1} + \frac{2}{n^2}, A_k - \frac{2}{n^2}]$$

Block 2

We claim that σ_k is in between Δ_{k-1} and Δ_k and hence mutually disjoint.

From (*), $m_k^2 \geq \lambda_k^2 n^2$ and hence, $m_k^2 \geq \frac{n^2}{2}$. We are then left with the inequality:

$$A_{k-1} - \frac{1}{m_{k-1}^2} \leq A_{k-1} + \frac{2}{n^2} \leq A_k - \frac{2}{n^2} \leq A_k - \frac{1}{m_k^2}$$

We note that, for any $k \in \{2, \dots, n\}$, we have $|\sigma_k| = A_k - \frac{2}{n^2} - A_{k-1} - \frac{2}{n^2} = A_k - A_{k-1} - \frac{4}{n^2} = \frac{4\pi}{(2n+1)} - \frac{4}{n^2}$.

Since σ_k s are mutually disjoint, we have $|\bigcup_{k=1}^n \sigma_k| = \sum_{k=1}^n |\sigma_k| = \frac{2\pi n}{(2n+1)} - \frac{4n}{n^2} \rightarrow 2\pi$.

For any $x \in \sigma_k$ and $\alpha \in \Delta_k$, one has $|\alpha - x| > 0$ such that

$$\frac{\varphi_n^{(k)}(\alpha)}{\sin(\frac{1}{2}(\alpha - x))}$$

is a well-defined continuous function.

By Riemann-Lebesgue, we have the following properties:

$$\lim_{m_k \rightarrow \infty} S_{m_k} \varphi_n^{(i)}(x) = \lim_{m_k \rightarrow \infty} \frac{1}{\pi} \int_0^{2\pi} \frac{\varphi_n^{(i)}(\alpha)}{\sin(\frac{1}{2}(\alpha - x))} \sin[(\frac{2m_k - 1}{2})(x - \alpha)] d\alpha = 0$$

for $i < k$.

[Condition here]:

Now, suppose that $\{\lambda_i\}_{i=1}^{k-1}$ is already defined and hence φ_n is defined on Δ_i for $i < k$.

Choose λ_k big enough such that m_k is big enough and that the m_k^{th} Fourier series of φ_n follows the following properties:

$$\frac{1}{\pi} \int_{\bigcup_i \Delta_i} \frac{\varphi_n(\alpha)}{\sin(\frac{1}{2}(\alpha - x))} \sin[(\frac{2m_k - 1}{2})(x - \alpha)] d\alpha < 1$$

by the previous consequence of the Riemann-Lebesgue lemma.

Now, consider the partial m_k^{th} Fourier series of φ_n on Δ_s for $s \geq k$:

$$\begin{aligned} \frac{1}{\pi} \int_{\Delta_s} \frac{m_s^2 \sin(\frac{2m_k - 1}{2}(\alpha - x))}{n \sin(\frac{1}{2}(\alpha - x))} d\alpha &= \frac{1}{\pi} \int_{\Delta_s} \frac{m_s^2 \sin(\frac{2m_k - 1}{2}(\alpha - x))}{n \sin(\frac{1}{2}(\alpha - x))} d\alpha - \frac{1}{\pi} \int_{\Delta_s} \frac{m_s^2 \sin(\frac{2m_k - 1}{2}(A_s - x))}{n \sin(\frac{1}{2}(A_s - x))} d\alpha \\ &+ \frac{1}{\pi} \int_{\Delta_s} \frac{m_s^2 \sin(\frac{2m_k - 1}{2}(A_s - x))}{n \sin(\frac{1}{2}(A_s - x))} d\alpha \\ &= \frac{1}{\pi} \int_{\Delta_s} \frac{m_s^2}{n} \left[\frac{\sin(\frac{2m_k - 1}{2}(\alpha - x))}{\sin(\frac{1}{2}(\alpha - x))} d\alpha - \frac{\sin(\frac{2m_k - 1}{2}(A_s - x))}{\sin(\frac{1}{2}(A_s - x))} \right] d\alpha \\ &+ \frac{1}{\pi} \int_{\Delta_s} \frac{m_s^2 \sin(\frac{2m_k - 1}{2}(\alpha - x))}{n \sin(\frac{1}{2}(\alpha - x))} d\alpha \end{aligned}$$

Note that the derivative of the Dirichlet's kernel $D_N(x) = \sum_{n=-N}^N e^{nx}$ is given by $\sum_{n=-N}^N n e^{nx}$ and its absolute value is bounded above by N^2 .

Notice also that $|A_s - \alpha| \leq \frac{2}{n^2} \leq \frac{1}{m_s^2}$ by the construction of $\{\sigma_j\}_j$ intervals.

Using the Mean-Value Theorem, we then have

$$\left| \frac{\sin(\frac{2m_k - 1}{2}(\alpha - x))}{\sin(\frac{1}{2}(\alpha - x))} - \frac{\sin(\frac{2m_k - 1}{2}(A_s - x))}{\sin(\frac{1}{2}(A_s - x))} \right| \leq |A_s - \alpha| m_k^2 \leq \frac{m_k^2}{m_s^2} \leq 1$$

by the construction of the increasing $\{m_j\}_j$ sequence.

Therefore, we have

$$\frac{1}{\pi} \int_{\Delta_s} \frac{m_s^2}{n} \left[\frac{\sin(\frac{2m_k-1}{2}(\alpha-x))}{\sin(\frac{1}{2}(\alpha-x))} - \frac{\sin(\frac{2m_k-1}{2}(A_s-x))}{\sin(\frac{1}{2}(A_s-x))} \right] d\alpha \leq \frac{1}{\pi} \int_{\Delta_s} \frac{m_s^2}{n} d\alpha \leq \frac{\pi}{n}$$

We have

$$\begin{aligned} \frac{1}{\pi} \int_{\Delta_s} \frac{m_s^2}{n} \frac{\sin(\frac{2m_k-1}{2}(\alpha-x))}{\sin(\frac{1}{2}(\alpha-x))} d\alpha &= \frac{1}{\pi} \int_{\Delta_s} \frac{m_s^2}{n} \left[\frac{\sin(\frac{2m_k-1}{2}(A_s-x))}{\sin(\frac{1}{2}(A_s-x))} \right] d\alpha + \frac{\tau\pi}{n} \\ &= \frac{2}{\pi n} \frac{\sin(\frac{2m_k-1}{2}(A_s-x))}{\sin(\frac{1}{2}(A_s-x))} + \frac{\tau\pi}{n} \end{aligned}$$

for $\tau \leq 1$

Now, observe that for $s > k$,

$$\begin{aligned} (A_s - A_k) &= \frac{4\pi(s-k)}{(2n+1)} \\ \frac{(2m_k+1)}{2}(A_s - A_k) &= \frac{(\lambda_k(2n+1))}{2}(A_s - A_k) \\ &\text{since } (2m_k+1) = \lambda_k(2n+1)) \\ &= (s-k)2\pi\lambda_k \end{aligned}$$

We have:

$$\begin{aligned} \sin(\frac{2m_k-1}{2}(A_s-x)) &= \sin(\frac{2m_k-1}{2}(A_s-A_k+A_k-x)) \\ &= \sin(\frac{2m_k-1}{2}((s-k)2\pi\lambda_k+A_k-x)) \\ &\text{Using the periodicity of the sin function} \\ &= \sin(\frac{2m_k-1}{2}(A_k-x)) \end{aligned}$$

Hence,

$$\frac{2}{\pi n} \frac{\sin(\frac{2m_k-1}{2}(A_s-x))}{\sin(\frac{1}{2}(A_s-x))} + \frac{\tau\pi}{n} \geq \frac{2}{\pi n} \left| \sin(\frac{2m_k-1}{2}(A_k-x)) \right| \frac{1}{|\sin(\frac{1}{2}(A_s-x))|}$$

From here, one need only notice that $|A_s - x| \leq |A_s - A_{k-1}| = \frac{4\pi(s-k-1)}{(2n+1)} < (s-k-1)\frac{2\pi}{n}$

Using the inequality $|\sin(t)| < |t|$, we have

$$\frac{2}{\pi n} \left| \sin(\frac{2m_k+1}{2}(A_k-x)) \right| \frac{1}{|\sin(\frac{1}{2}(A_s-x))|} \geq \frac{1}{\pi^2} \left| \sin(\frac{2m_k+1}{2}(A_k-x)) \right| \frac{1}{(s-k+1)}$$

Summing through values of s from k through n , we have an upper bound on the m_k^{th} Fourier series:

$$\frac{1}{\pi} \int_0^{2\pi} \varphi_n(x) \frac{\sin(\frac{2m_k-1}{2}(A_k-x))}{\sin(\frac{1}{2}(\alpha-x))} dx \geq \frac{1}{\pi^2} \left| \sin(\frac{(2m_k+1)}{2}(A_k-x)) \right| \sum_{s=k}^n \frac{1}{(s-k+1)} = \frac{1}{\pi^2} \left| \sin(\frac{(2m_k+1)}{2}(A_k-x)) \right| \sum_{r=1}^{n-k} \frac{1}{(r)}$$

We can now choose the M_n, q_n and E_n as required by Condition (3):

$$q_n = m_n$$

$$E_n = \bigcup_k \sigma_k$$

for k such that $n - k > \sqrt{n}$ and that

$$\frac{1}{\pi^2} |\sin(\frac{(2m_k + 1)}{2}(A_k - x))| > \frac{1}{\sqrt{\sum_{r=1}^{\lfloor \frac{n}{2} \rfloor} \frac{1}{r}}} := \frac{1}{N_n}$$

for $x \in \sigma_k$.

Hence, we have the inequality

$$\frac{1}{\pi} \int_0^{2\pi} \varphi_n(x) \frac{\sin(\frac{(2m_k - 1)}{2}(A_k - x))}{\sin(\frac{1}{2}(\alpha - x))} dx \geq \sqrt{\log(n)}$$

$$M_n = N_n$$

Note that one can prove that $\lim_{n \rightarrow \infty} |E_n| = 2\pi$ through analogous computations as in Block (2).

5.2 Existence of an $L^1(\mathbb{T})$ function with a.e divergent series

We now have all the ingredients necessary to demonstrate an $L^1(\mathbb{T})$ function whose Fourier series diverges almost everywhere.

Since $\{\varphi_n(x)\}_n$ are constructed functions satisfying Conditions (1), (2) and (3), we can inductively define an increasing subsequence $\{n_k\}_k$ such that:

(a) $M_{n_k} \geq 2^{2k}$.

We have $\frac{1}{\sqrt{M_{n_k}}} \leq \frac{1}{2^k}$ and hence

$$\sum_{k=1}^{\infty} \frac{1}{\sqrt{M_{n_k}}} \leq \sum_{k=1}^{\infty} \frac{1}{2^k} \leq 1$$

(b) $\sum_{s=1}^{(k-1)} \sup_{N \geq 0} |S_N \varphi_{n_s}| \leq \frac{1}{2} \sqrt{M_{n_k}}$

(c) $q_{n_i} < 2q_{n_i} + 1 \leq \frac{1}{2^k} \sqrt{M_{n_k}}$ for all $i < k$

Define

$$\Phi(x) = \sum_{k=1}^{\infty} \frac{1}{\sqrt{M_{n_k}}} \varphi_{n_k}(x)$$

Using the Monotone Convergence Theorem and the property (i) of $\varphi_i(x)$ for any $i \in \mathbb{N}$, $\Phi(x)$ remains an $L^1(\mathbb{T})$ function.

Now, consider for $x \in E_{n_k}$ with $N_k < q_{n_k}$:

$$S_{N_k}(x) = \sum_{i=1}^{\infty} \frac{1}{\sqrt{M_{n_i}}} S_{N_k} \varphi_{n_i}(x)$$

$$|S_{N_k} \Phi(x)| = \left| \sum_{i < k} \frac{1}{\sqrt{M_{n_i}}} S_{N_k} \varphi_{n_i}(x) + \frac{1}{\sqrt{M_{n_k}}} S_{N_k} \varphi_{n_k}(x) + \sum_{i > k} \frac{1}{\sqrt{M_{n_i}}} S_{N_k} \varphi_{n_i}(x) \right|$$

By condition (b), we have $\sum_{i < k} \frac{1}{\sqrt{M_{n_i}}} S_{N_k} \varphi_{n_i}(x) \leq \frac{1}{2} \sqrt{M_{n_k}}$.

By condition (c) and the (iii) condition of q_{n_k} , we have $\frac{1}{\sqrt{M_{n_k}}} S_{N_k} \varphi_{n_k}(x) \geq \sqrt{M_{n_k}}$.

We can then conclude that

$$\left| \sum_{i < k} \frac{1}{\sqrt{M_{n_i}}} S_{N_k} \varphi_{n_i}(x) + \frac{1}{\sqrt{M_{n_k}}} S_{N_k} \varphi_{n_k}(x) \right| \geq \frac{1}{2} \frac{1}{M_{n_k}}$$

Using the pointwise bound $D_N(x) \leq CN$ of the Dirichlet's kernel and condition i on φ_{n_k} , one has:

$$\sum_{i > k} S_{N_k} \varphi_{n_k} \frac{1}{\sqrt{M_{n_i}}} \leq \sum_{i > k} \frac{2CN_k}{\sqrt{M_{n_i}}}$$

Since $N_k < q_{n_k}$, using condition (c), we have $N_k \leq \frac{\sqrt{M_{n_i}}}{2^i}$ for $i > k$. Hence, we have:

$$\begin{aligned} \sum_{i > k} S_{N_k} \varphi_{n_k} \frac{1}{\sqrt{M_{n_i}}} &\leq \sum_{i > k} \frac{2CN_k}{\sqrt{M_{n_i}}} \\ &\leq 2C \sum_{i > k} \frac{1}{2^i} \\ &\leq 2C \frac{1}{2^k} \end{aligned}$$

Now we use the reverse triangle inequality to have:

$$\begin{aligned} |S_{N_k} \Phi(x)| &= \left| \sum_{i < k} \frac{1}{\sqrt{M_{n_i}}} S_{N_k} \varphi_{n_i}(x) + \frac{1}{\sqrt{M_{n_k}}} S_{N_k} \varphi_{n_k}(x) + \sum_{i > k} \frac{1}{\sqrt{M_{n_i}}} S_{N_k} \varphi_{n_i}(x) \right| \\ &\geq \left| \sum_{i < k} \frac{1}{\sqrt{M_{n_i}}} S_{N_k} \varphi_{n_i}(x) + \frac{1}{\sqrt{M_{n_k}}} S_{N_k} \varphi_{n_k}(x) \right| - \left| \sum_{i > k} \frac{1}{\sqrt{M_{n_i}}} S_{N_k} \varphi_{n_i}(x) \right| \\ &\geq \left| \frac{1}{2} M_{n_k} - \frac{6}{2^k} \right| \end{aligned}$$

The latter is a divergent sequence - which proves that the Fourier series of $\Phi(x)$ diverges.

Bibliography

- [1] Camil Muscalu and Wilhelm Schlag. *Classical and Multilinear Harmonic Analysis*. Vol. 1. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2013. DOI: 10.1017/CB09781139047081.
- [2] Terence Tao. *247B, notes 4: Almost everywhere convergence of Fourier series*. Aug. 2020. URL: <https://terrytao.wordpress.com/2020/05/14/247b-notes-4-almost-everywhere-convergence-of-fourier-series/>.
- [3] Andrey Kolmogoroff. “Une série de Fourier-Lebesgue divergente presque partout”. fre. In: *Fundamenta Mathematicae* 4.1 (1923), pp. 324–328. URL: <http://eudml.org/doc/213617>.