

Introduction to Functional Analysis

Notes taken by Runqiu Ye
Carnegie Mellon University

Spring 2025

Contents

1	Banach space theory	3
1.1	Quotient spaces, Baire category and uniform boundedness	3
2	Hilbert space theory	4
2.1	Basic Hilbert space theory	4
2.2	Orthonormal bases and Fourier Series	7
2.3	Fejer's theorem and convergence of Fourier series	10
2.4	Minimizers, orthogonal complements, and the Riesz representation theorem	14

1 Banach space theory

1.1 Quotient spaces, Baire category and uniform boundedness

Theorem. Let $\|\cdot\|$ be a **seminorm** on a vector space V . If we define $E = \{v \in V : \|v\| = 0\}$, then E is a subspace of V , and the function on V/E defined by

$$\|v + E\| = \|v\|$$

for any $v + E \in V/E$ defines a **norm**.

Theorem (Baire Category Theorem). Let M be a complete metric space, and let $\{C_n\}_{n=0}^\infty$ be a collection of closed subsets of M such that $M = \bigcup_{n=0}^\infty C_n$. Then at least one of the C_n contains an open ball $B(x, r) = \{y \in M : d(x, y) < r\}$.

Theorem (Uniform Boundedness Theorem). Let B be Banach space and V a normed vector space. Let $\{T_n\}_{n=0}^\infty$ be a sequence in $\mathcal{B}(B, V)$. Then if for all $b \in B$ we have $\sup_n \|T_n b\| < \infty$ (that is, this sequence is pointwise bounded), then $\sup_n \|T_n\| < \infty$ (the operator norms are bounded).

Proof. For each $k \in \mathbb{N}$, define

$$C_k = \left\{ b \in B : \|b\| \leq 1, \sup_{n \in \mathbb{N}} \|T_n b\| \leq k \right\}.$$

This set is closed for each $k \in \mathbb{N}$, but by assumption, we have

$$\{b \in B : \|b\| \leq 1\} = \bigcup_{k=0}^\infty C_k.$$

The left hand side is a closed subset of B , and is thus a complete metric space. By Baire Category Theorem, there exists $k \in \mathbb{N}$ such that C_k contains an open ball $B(b_0, \delta_0)$. Then, if $b \in B(b_0, \delta_0)$, we have $b_0 + b \in B(b_0, \delta_0)$ and thus

$$\sup_{n \in \mathbb{N}} \|T_n(b_0 + b)\| \leq k.$$

It follows that

$$\sup_{n \in \mathbb{N}} \|T_n b\| \leq \sup_{n \in \mathbb{N}} \|T_n(b_0 + b)\| + \sup_{n \in \mathbb{N}} \|T_n b_0\| \leq 2k.$$

Suppose $\|b\| = 1$, then $\frac{\delta_0}{2}b \in B(b_0, \delta_0)$ and thus for all $n \in \mathbb{N}$, we have

$$\left\| T_n \left(\frac{\delta_0}{2} b \right) \right\| \leq 2k.$$

Therefore,

$$\sup_{n \in \mathbb{N}} \|T_n\| \leq \frac{4k}{\delta_0}.$$

□

2 Hilbert space theory

2.1 Basic Hilbert space theory

Definition (Pre-Hilbert space). A **pre-Hilbert** space H is a vector space over \mathbb{C} with a **Hermitian inner product**, which is a map $\langle \cdot, \cdot \rangle : H \times H \rightarrow \mathbb{C}$ satisfying the following properties.

1. For all $\lambda_1, \lambda_2 \in \mathbb{C}$ and $v_1, v_2, w \in H$, we have

$$\langle \lambda_1 v_1 + \lambda_2 v_2, w \rangle = \lambda_1 \langle v_1, w \rangle + \lambda_2 \langle v_2, w \rangle.$$

2. For all $v, w \in H$, we have $\langle v, w \rangle = \overline{\langle w, v \rangle}$.

3. For all $v \in H$, we have $\langle v, v \rangle \geq 0$, with equality if and only if $v = 0$.

Definition. Let H be a pre-Hilbert space. For all $v \in H$, we define

$$\|v\| = \langle v, v \rangle^{\frac{1}{2}}.$$

Theorem (Cauchy-Schwarz inequality). Let H be a pre-Hilbert space. For all $u, v \in H$, we have

$$|\langle u, v \rangle| \leq \|u\| \|v\|.$$

Proof. Define $f(t) = \|u + tv\|^2$. Notice that

$$\begin{aligned} f(t) &= \langle u + tv, u + tv \rangle \\ &= \langle u, u \rangle + t^2 \langle v, v \rangle + t \langle u, v \rangle + t \langle v, u \rangle \\ &= \|u\|^2 + t^2 \|v\|^2 + 2t \operatorname{Re}(\langle u, v \rangle). \end{aligned}$$

This implies that

$$0 \leq f(t_{\min}) = \|u\|^2 - \frac{\operatorname{Re}(\langle u, v \rangle)^2}{\|v\|^2}.$$

It follows that

$$|\operatorname{Re}(\langle u, v \rangle)| \leq \|u\| \|v\|.$$

This is almost what we want. To finish up, first note that if $\langle u, v \rangle = 0$ then there is nothing to prove, so suppose $\langle u, v \rangle \neq 0$, and define

$$\lambda = \frac{\overline{\langle u, v \rangle}}{|\langle u, v \rangle|}.$$

Note that we have $|\lambda| = 1$ and we have the chain of equalities of real numbers:

$$|\langle u, v \rangle| = \lambda \langle u, v \rangle = \langle \lambda u, v \rangle = \operatorname{Re} \langle \lambda u, v \rangle \leq \|\lambda u\| \|v\|.$$

However, $\|\lambda u\| = \|u\|$, so the proof is complete. □

Theorem. If H is a pre-Hilbert space, then $\|\cdot\|$ is a norm on H .

Proof. Note that

$$\|v\| = 0 \iff \langle v, v \rangle = 0 \iff v = 0.$$

Now if $\lambda \in \mathbb{C}$ and $v \in H$, then

$$\langle \lambda v, \lambda v \rangle = \lambda \bar{\lambda} \langle v, v \rangle = |\lambda|^2 \|v\|^2.$$

Therefore, $\|\lambda v\| = |\lambda| \|v\|$.

Finally, let $u, v \in H$, then

$$\begin{aligned}\|u + v\|^2 &= \langle u + v, u + v \rangle \\ &= \|u\|^2 + \|v\|^2 + 2 \operatorname{Re} \langle u, v \rangle \\ &\leq \|u\|^2 + \|v\|^2 + 2 |\langle u, v \rangle| \\ &\leq \|u\|^2 + \|v\|^2 + 2 \|u\| \|v\| \\ &= (\|u\| + \|v\|)^2.\end{aligned}$$

This completes the proof. \square

Theorem. If $u_n \rightarrow u$ and $v_n \rightarrow v$ in a pre-Hilbert space H , then $\langle u_n, v_n \rangle \rightarrow \langle u, v \rangle$.

Proof. If $u_n \rightarrow u$ and $v_n \rightarrow v$, then $\|u_n - u\| \rightarrow 0$ and $\|v_n - v\| \rightarrow 0$. It follows that

$$\begin{aligned}|\langle u_n, v_n \rangle - \langle u, v \rangle| &= |\langle u_n - u, v_n \rangle - \langle u, v - v_n \rangle| \\ &\leq |\langle u_n - u, v_n \rangle| + |\langle u, v - v_n \rangle| \\ &\leq \|u_n - u\| \|v_n\| + \|u\| \|v - v_n\| \\ &\leq \|u_n - u\| \sup_{k \in \mathbb{N}} \|v_k\| + \|u\| \|v - v_n\| \\ &\rightarrow 0\end{aligned}$$

as $n \rightarrow \infty$. This completes the proof. \square

Definition (Hilbert space). A **Hilbert space** is a pre-Hilbert space that is complete with respect to the norm $\|\cdot\| = \langle \cdot, \cdot \rangle^{\frac{1}{2}}$.

Example. Some examples of Hilbert spaces:

- $\mathbb{C}^n = \{z = (z_1, \dots, z_n) : z_j \in \mathbb{C}\}$ with $\langle z, w \rangle = \sum_j z_j \overline{w_j}$ is a Hilbert space.
- $\ell^2 = \left\{a = \{a_k\}_{k=0}^\infty : a_k \in \mathbb{C}, \sum_{k=0}^\infty |a_k|^2 < \infty\right\}$ with $\langle a, b \rangle = \sum_{k=0}^\infty a_k \overline{b_k}$ is a Hilbert space.
- If $E \subset \mathbb{R}$ is measurable, then $L^2(E) = \left\{f : E \rightarrow \mathbb{C}, \int_E |f|^2 < \infty\right\}$ with $\langle f, g \rangle = \int_E f \overline{g}$ is a Hilbert space.

We will show that each separable Hilbert space is isometrically isomorphic to either \mathbb{C}^n or ℓ^2 .

Now we have seen that ℓ^2 and L^2 spaces are Hilbert spaces. This is expected since the definition of the inner product in these spaces uses the fact that they are ℓ^2 or L^2 . A natural question then is whether other ℓ^p or L^p spaces are also Hilbert spaces with respect to some inner product? It turns out there is a simple way to decide whether a norm comes from an inner-product, and thus whether a Banach space is a Hilbert space.

Theorem (Parallelogram Law). If H is a pre-Hilbert space, then for all $u, v \in H$, we have

$$\|u + v\|^2 + \|u - v\|^2 = 2(\|u\|^2 + \|v\|^2).$$

In addition, if H is a normed vector space satisfying this equality, then H is a pre-Hilbert space.

Using the previous theorem, we can verify that ℓ^p and L^p with $p \neq 2$ are **not** Hilbert spaces.

Definition (Orthogonal). If H is a pre-Hilbert space, $u, v \in H$ are **orthogonal** if $\langle u, v \rangle = 0$. We denote this as $u \perp v$.

Definition (Orthonormal sets). If H is a pre-Hilbert space, a subset $\{e_\lambda\}_{\lambda \in \Lambda} \subset H$ is **orthonormal** if for all $\lambda \in \Lambda$, we have $\|e_\lambda\| = 1$ and $\lambda_1 \neq \lambda_2$ implies $e_{\lambda_1} \perp e_{\lambda_2}$.

Remark. we will mainly be interested in the case where we have a countable orthonormal set.

Example. The set $\left\{\frac{e^{inx}}{\sqrt{2\pi}}\right\}_{n \in \mathbb{Z}}$ as elements in $L^2([-\pi, \pi])$ is an orthonormal subset of $L^2([-\pi, \pi])$. Indeed, for any $m, n \in \mathbb{Z}$, we have

$$\int_{-\pi}^{\pi} e^{imx} \overline{e^{inx}} = \int_{-\pi}^{\pi} e^{i(m-n)x} = \begin{cases} 2\pi & (m = n), \\ 0 & (m \neq n). \end{cases}$$

Therefore, $\left\langle \frac{e^{inx}}{\sqrt{2\pi}}, \frac{e^{imx}}{\sqrt{2\pi}} \right\rangle = \delta_{mn}$, and $\left\{\frac{e^{inx}}{\sqrt{2\pi}}\right\}_{n \in \mathbb{Z}}$ is an orthonormal subset of $L^2([-\pi, \pi])$.

Theorem (Bessel). If $\{e_n\}_{n=0}^{\infty}$ is countable orthonormal subset of a pre-Hilbert space H , then for all $u \in H$, we have

$$\sum_{n=0}^{\infty} |\langle u, e_n \rangle|^2 \leq \|u\|^2.$$

Proof. We first do the finite case. Suppose $\{e_n\}_{n=1}^N$ is an orthonormal subset of H . Then,

$$\begin{aligned} \left\| \sum_{n=0}^N \langle u, e_n \rangle e_n \right\|^2 &= \left\langle \sum_{n=0}^N \langle u, e_n \rangle e_n, \sum_{n=0}^N \langle u, e_n \rangle e_n \right\rangle \\ &= \sum_{n=0}^N \sum_{m=1}^N \langle u, e_n \rangle \overline{\langle u, e_m \rangle} \langle e_n, e_m \rangle \\ &= \sum_{n=0}^N |\langle u, e_n \rangle|^2. \end{aligned}$$

Also,

$$\begin{aligned} \left\langle u, \sum_{n=0}^N \langle u, e_n \rangle e_n \right\rangle &= \sum_{n=0}^N \overline{\langle u, e_n \rangle} \langle u, e_n \rangle \\ &= \sum_{n=0}^N |\langle u, e_n \rangle|^2. \end{aligned}$$

Therefore,

$$\begin{aligned} 0 &\leq \left\| u - \sum_{n=0}^N \langle u, e_n \rangle e_n \right\|^2 \\ &= \|u\|^2 + \left\| \sum_{n=0}^N \langle u, e_n \rangle e_n \right\|^2 - 2 \operatorname{Re} \left\langle u, \sum_{n=0}^N \langle u, e_n \rangle e_n \right\rangle \\ &= \|u\|^2 - \sum_{n=0}^N |\langle u, e_n \rangle|^2, \end{aligned}$$

as desired.

For the infinite case, just take the limit as $N \rightarrow \infty$. □

Definition (Maximal orthonormal subset). An orthonormal subset $\{e_\lambda\}_\lambda$ of a pre-Hilbert space is **maximal** if $u \in H$ and $\langle u, e_\lambda \rangle = 0$ for all $\lambda \in \Lambda$ implies that $u = 0$.

Theorem. Every non-trivial pre-Hilbert space has a maximal orthonormal subset.

This can be proved using Zorn's Lemma. We will prove something less strong but often equally useful by hand, without applying Zorn's Lemma.

Theorem. Every non-trivial separable pre-Hilbert space has a countable maximal orthonormal subset.

Proof. Use the Gram-Schmidt process. Let $\{v_j\}_{j=0}^{\infty}$ be a countable dense subset of H where $v_0 \neq 0$. Claim that for any $n \in \mathbb{N}$, there exists $m(n) \leq n$ and an orthonormal subset $\{e_1, \dots, e_{m(n)}\}$ such that

1. $\text{span}\{e_1, \dots, e_{m(n)}\} = \text{span}\{v_1, \dots, v_n\}$.
2. If $v_n \in \text{span}\{v_1, \dots, v_{n-1}\}$, we have

$$\{e_1, \dots, e_{m(n)}\} = \{e_1, \dots, e_{m(n-1)}\} \cup \emptyset.$$

Otherwise, we have

$$\{e_1, \dots, e_{m(n)}\} = \{e_1, \dots, e_{m(n-1)}\} \cup e_{m(n)}$$

for some $e_{m(n)} \in H$.

Prove this by induction. For the base case, let $e_1 = \frac{v_1}{\|v_1\|}$. For the inductive step, suppose the claim holds for $n = k$. If $v_{k+1} \in \text{span}\{v_1, \dots, v_k\}$, then

$$\text{span}\{e_1, \dots, e_{m(k)}\} = \text{span}\{v_1, \dots, v_k\} = \text{span}\{v_1, \dots, v_{k+1}\}.$$

Now suppose $v_{k+1} \notin \text{span}\{v_1, \dots, v_k\}$. Define

$$w_{k+1} = v_{k+1} - \sum_{j=1}^{m(k)} \langle v_{k+1}, e_j \rangle e_j.$$

Note that $w_{k+1} \neq 0$ and define $e_{m(k+1)} = \frac{w_{k+1}}{\|w_{k+1}\|}$. Then, $\|e_{m(k+1)}\| = 1$ and for all $1 \leq l \leq m(k)$,

$$\begin{aligned} \langle e_{m(k+1)}, e_l \rangle &= \frac{1}{\|w_{k+1}\|} \left\langle v_{k+1} - \sum_{j=1}^{m(k)} \langle v_{k+1}, e_j \rangle e_j, e_l \right\rangle \\ &= \frac{1}{\|w_{k+1}\|} (\langle v_{k+1}, e_l \rangle - \langle v_{k+1}, e_l \rangle) \\ &= 0. \end{aligned}$$

Therefore, $e_{m(k+1)}$ is the desired vector we want and we have completed the proof for the claim.

Now let

$$S = \bigcup_{n=0}^{\infty} \{e_1, \dots, e_{m(n)}\}.$$

Then S is a countable orthonormal subset of H . Now we show S is maximal. Suppose $u \in H$ and $\langle u, e_l \rangle = 0$. Since $\{v_j\}_{j=0}^{\infty}$ is dense in H , there exists $\{v_{j(k)}\}_{k=0}^{\infty}$ such that $v_{j(k)} \rightarrow u$ as $k \rightarrow \infty$. By our claim, we know $v_{j(k)} \in \text{span}\{e_1, \dots, e_{m(j(k))}\}$. By Bessel's inequality,

$$\|v_{j(k)}\|^2 = \sum_{l=1}^{m(j(k))} |\langle v_{j(k)}, e_l \rangle|^2 = \sum_{l=1}^{m(j(k))} |\langle v_{j(k)} - u, e_l \rangle|^2 \leq \|v_{j(k)} - u\|^2,$$

where for the first equality we used the fact that $v_{j(k)} \in \text{span}\{e_1, \dots, e_{m(j(k))}\}$. Since $v_{j(k)} \rightarrow u$ as $k \rightarrow \infty$, the inequality implies that $\|v_{j(k)}\| \rightarrow 0$ as $k \rightarrow \infty$ and thus $\|u\| = 0$, showing that S is indeed a maximal orthonormal subset of H . \square

Corollary. ℓ^2 and L^2 have countable maximal orthonormal subset since they are both separable.

2.2 Orthonormal bases and Fourier Series

Definition (Orthonormal basis). Let H be a Hilbert space. An **orthonormal basis** of H is a countable maximal orthonormal subset $\{e_n\}_n$ of H .

Theorem. If $\{e_n\}_{n=0}^{\infty}$ is an orthonormal basis in Hilbert space H , then for all $u \in H$, we have

$$\sum_{n=0}^{\infty} \langle u, e_n \rangle e_n = u.$$

This is the Fourier-Bessel series.

This tells us we can write each element in H as a infinite linear combination of the orthonormal basis.

Proof. We first prove the sequence of partial sums $\{\sum_{n=0}^m \langle u, e_n \rangle e_n\}_m$ is Cauchy. Let $\varepsilon > 0$. By Bessel's inequality, we have

$$\sum_{n=0}^{\infty} |\langle u, e_n \rangle|^2 \leq \|u\|^2 < \infty.$$

Therefore, there exists $M \in \mathbb{N}$ such that $N \geq M$ implies $\sum_{n=N+1}^{\infty} |\langle u, e_n \rangle|^2 < \varepsilon^2$. Then for all $m > l \geq M$, we have

$$\left\| \sum_{n=0}^m \langle u, e_n \rangle e_n - \sum_{n=0}^l \langle u, e_n \rangle e_n \right\|^2 \leq \sum_{n=l+1}^m |\langle u, e_n \rangle|^2 \leq \sum_{n=l+1}^{\infty} |\langle u, e_n \rangle|^2 < \varepsilon^2.$$

Therefore, the sequence of partial sum is Cauchy. Since H is complete, there exists $\bar{u} \in H$ such that $\bar{u} = \sum_{n=0}^{\infty} \langle u, e_n \rangle e_n$. It remains to show that $\bar{u} = u$. By continuity of inner-product, for all $l \in \mathbb{N}$, we have

$$\begin{aligned} \langle u - \bar{u}, e_l \rangle &= \lim_{m \rightarrow \infty} \left\langle u - \sum_{n=0}^m \langle u, e_n \rangle e_n, e_l \right\rangle \\ &= \lim_{m \rightarrow \infty} \left[\langle u, e_l \rangle - \sum_{n=0}^m \langle u, e_n \rangle \langle e_n, e_l \rangle \right] \\ &= 0. \end{aligned}$$

Since $\{e_n\}_{n=0}^{\infty}$ is maximal, this implies that $u - \bar{u} = 0$ and the proof is complete. \square

Theorem. Let H be a Hilbert space. If H has an orthonormal basis, then H is separable.

Proof. Suppose $\{e_n\}_{n=0}^{\infty}$ is an orthonormal basis for H . Then

$$S = \bigcup_{m \in \mathbb{N}} \left\{ \sum_{n=0}^m q_n e_n : q_n \in \mathbb{Q} + i\mathbb{Q} \right\}$$

is a countable set. Also, by the previous theorem, S is dense in H . \square

Remark. Let H be a Hilbert space. H is separable if and only if H has an orthonormal basis.

Theorem (Parseval's identity). If H is a Hilbert space and $\{e_n\}_{n=0}^{\infty}$ is a countable orthonormal basis, then for all $u \in H$, we have

$$\sum_n |\langle u, e_n \rangle|^2 = \|u\|^2$$

Proof. We have $u = \sum_n \langle u, e_n \rangle e_n$. This implies that

$$\begin{aligned} \|u\|^2 &= \lim_{m \rightarrow \infty} \left\langle \sum_{n=0}^m \langle u, e_n \rangle e_n, \sum_{l=0}^m \langle u, e_l \rangle e_l \right\rangle \\ &= \lim_{m \rightarrow \infty} \sum_{n=0}^m \sum_{l=0}^m \langle u, e_n \rangle \overline{\langle u, e_l \rangle} \langle e_n, e_l \rangle \\ &= \lim_{m \rightarrow \infty} \sum_{n=0}^m |\langle u, e_n \rangle|^2 \\ &= \sum_{n=0}^{\infty} |\langle u, e_n \rangle|^2. \end{aligned}$$

\square

Theorem. If H is an infinite dimensional separable Hilbert space, then H is isometrically isomorphic to ℓ^2 . That is, there exists bijective bounded linear map $T : H \rightarrow \ell^2$ such that for all $u, v \in H$, we have

$$\|Tu\|_{\ell^2} = \|u\|_H \quad \text{and} \quad \langle Tu, Tv \rangle_{\ell^2} = \langle u, v \rangle_H.$$

Proof. Since H is separable, there exists an orthonormal basis $\{e_n\}_{n=0}^\infty$. For all $u \in H$, the previous theorem gives

$$\|u\| = \left(\sum_{n=0}^{\infty} |\langle u, e_n \rangle|^2 \right)^{\frac{1}{2}}.$$

Define $T : H \rightarrow \ell^2$ by

$$Tu = \{\langle u, e_n \rangle\}_{n=0}^\infty \in \ell^2.$$

It is easy to check that T is the desired isometric isomorphism. \square

Next we use the theories we learned in a more concrete setting — the Fourier series.

Theorem. The subset $\left\{ \frac{e^{inx}}{\sqrt{2\pi}} \right\}_{n \in \mathbb{Z}}$ is an orthonormal subset of $L^2([-\pi, \pi])$.

Definition. Let $f \in L^2([-\pi, \pi])$. Then the n -th **Fourier coefficient** of f is

$$\widehat{f}(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-int} dt.$$

The N -th **Fourier sum** of f is

$$S_N f(x) = \sum_{|n| \leq N} \widehat{f}(n) e^{inx} = \sum_{|n| \leq N} \left\langle f, \frac{e^{int}}{\sqrt{2\pi}} \right\rangle \frac{e^{inx}}{\sqrt{2\pi}}.$$

The **Fourier series** of f is the formal series $\sum_{n \in \mathbb{Z}} \widehat{f}(n) e^{-inx}$.

The natural question now is whether we have for all $f \in L^2([-\pi, \pi])$,

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

That is, whether we have the following convergence in L^2 .

$$\lim_{N \rightarrow \infty} \|f - S_N f\|_2 = 0.$$

This question is then equivalent to whether $\left\{ \frac{e^{inx}}{\sqrt{2\pi}} \right\}_{n \in \mathbb{Z}}$ is maximal in $L^2([-\pi, \pi])$. That is, whether $\widehat{f}(n) = 0$ for all $n \in \mathbb{N}$ implies $f = 0$.

The answer to the question is yes, but it is going to take some work. We first do some simple calculation.

Theorem. For all $f \in L^2([-\pi, \pi])$ and for all $N \in \mathbb{N}$, we have

$$S_N f(x) = \int_{-\pi}^{\pi} D_N(x-t) f(t) dt,$$

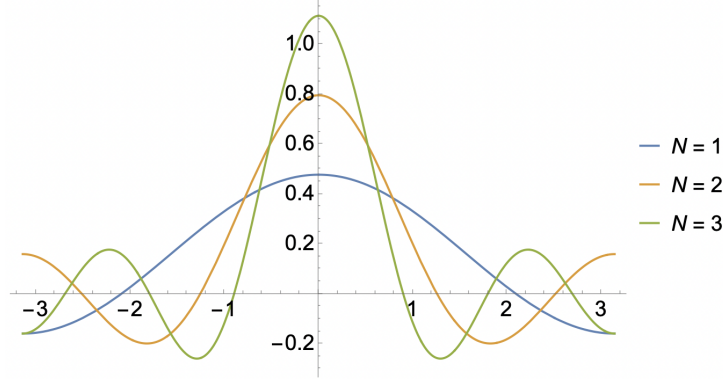
where

$$D_N(x) = \begin{cases} \frac{2N+1}{2\pi} & (x = 0) \\ \frac{\sin(N+\frac{1}{2})x}{2\pi \sin \frac{x}{2}} & (x \neq 0) \end{cases}$$

is the **Dirichlet kernel**. Figure 1 shows a plot of $D_N(x)$ on $[-\pi, \pi]$ for $N = 1, 2, 3$. Note that D_N is a smooth function.

Proof. If $f \in L^2([-\pi, \pi])$, we have

$$\begin{aligned} S_N f(x) &= \sum_{|n| \leq N} \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-int} dt \right) e^{inx} \\ &= \int_{-\pi}^{\pi} f(t) \left(\frac{1}{2\pi} \sum_{|n| \leq N} e^{in(x-t)} \right) dt. \end{aligned}$$

Figure 1: Plot of Dirichlet kernel $D_N(x)$ on $[-\pi, \pi]$ for $N = 1, 2, 3$.

Let $D_N(x) = \frac{1}{2\pi} \sum_{|n| \leq N} e^{inx}$. Then for $x \neq 0$, we have

$$\begin{aligned}
 D_N(x) &= \frac{1}{2\pi} \sum_{n=-N}^N e^{-inx} \\
 &= \frac{1}{2\pi} e^{-iNx} \sum_{n=0}^{2N} (e^{ix})^n \\
 &= \frac{1}{2\pi} e^{-iNx} \frac{1 - e^{i(2N+1)x}}{1 - e^{ix}} \\
 &= \frac{1}{2\pi} \frac{e^{i(N+\frac{1}{2})x} - e^{-i(N+\frac{1}{2})x}}{e^{\frac{ix}{2}} - e^{-\frac{ix}{2}}} \\
 &= \frac{1}{2\pi} \frac{2i \sin(N + \frac{1}{2})x}{2i \sin \frac{x}{2}} \\
 &= \frac{1}{2\pi} \frac{\sin(N + \frac{1}{2})x}{\sin \frac{x}{2}},
 \end{aligned}$$

as desired. For $x = 0$, we also clearly have $D_N(0) = \frac{(2N+1)}{2\pi}$. The proof is thus complete. \square

Definition. If $f \in L^2([-\pi, \pi])$, we define the N -th **Cesaro-Fourier mean** of f by

$$\sigma_N f(x) = \frac{1}{N+1} \sum_{k=0}^N S_k f(x).$$

The idea behind defining the Cesaro mean is that if the original sequence converges, the Cesaro mean also converge to the same limit. However, Cesaro have even better property — the Cesaro mean can converge even if the original sequence does not converge. Therefore, it has better convergence properties and hopefully we can show it converge to f in L^2 more easily. The goal now is then to show

$$\|\sigma_N f - f\|_2 \rightarrow 0 \text{ as } N \rightarrow \infty.$$

This would tell us if all Fourier coefficients are zero, then the Cesaro means are zero, and the limit above would tell us f is zero.

2.3 Fejer's theorem and convergence of Fourier series

In this section, we will show that if $f \in L^2([-\pi, \pi])$, then $\|\sigma_N f - f\|_2 \rightarrow 0$ as $N \rightarrow \infty$.

First we will rewrite the Cesaro Fourier mean, just like what we did for the partial Fourier sum using the Dirichlet kernel.

Theorem. For any $f \in L^2([-\pi, \pi])$, we have

$$\sigma_N f(x) = \int_{-\pi}^{\pi} K_N(x-t) f(t) dt,$$

where

$$K_N(x) = \begin{cases} \frac{N+1}{2\pi} & (x = 0) \\ \frac{1}{2\pi(N+1)} \frac{\sin^2 \frac{N+1}{2} x}{\sin^2 \frac{x}{2}} & (x \neq 0) \end{cases}$$

is the Fejer kernel.

Moreover, we have

1. $K_N(x) \geq 0$, $K_N(x) = K_N(-x)$, and $K_N(x)$ is 2π periodic.
2. $\int_{-\pi}^{\pi} K_N(t) dt = 1$.
3. If $\delta \in (0, \pi]$, then for all $\delta \leq |x| \leq \pi$, we have

$$|K_N(x)| \leq \frac{1}{2\pi(N+1) \sin^2 \frac{\delta}{2}}.$$

A plot for $K_N(x)$ is shown in Figure 2.

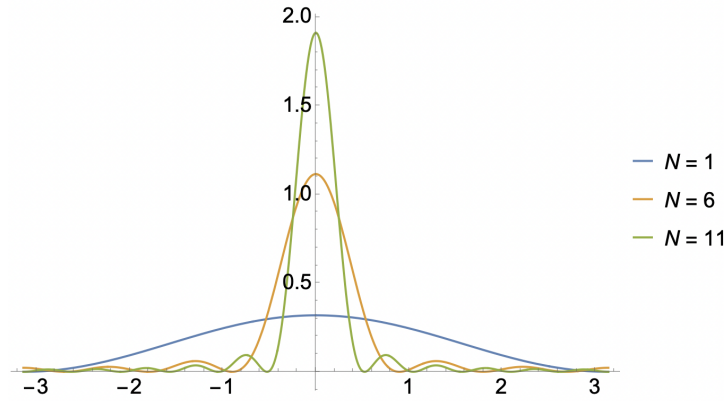


Figure 2: Plot of Dirichlet kernel $D_N(x)$ on $[-\pi, \pi]$ for $N = 1, 6, 11$.

Note that $K_N(x)$ is concentrated at 0 when N is very large. In this case, we have

$$\begin{aligned} \sigma_N f(x) &= \int_{-\pi}^{\pi} K_N(x-t) f(t) dt \\ &\approx f(x) \int_{-\pi}^{\pi} K_N(t) dt \\ &= f(x). \end{aligned}$$

This provides a rough intuition behind the Fejer kernel. The fact that K_N is non-negative makes a huge difference compared to the Dirichlet kernel, since it gives much better properties.

Proof. Recall that

$$S_k f(x) = \int_{-\pi}^{\pi} D_k(x-t) f(t) dt,$$

where

$$D_k(t) = \begin{cases} \frac{2N+1}{2\pi} & (t = 0), \\ \frac{1}{2\pi} \frac{\sin((N+\frac{1}{2})t)}{\sin \frac{t}{2}} & (t \neq 0). \end{cases}$$

It follows that

$$\begin{aligned}\sigma_N f(x) &= \frac{1}{N+1} \sum_{k=0}^N S_k f(x) \\ &= \int_{-\pi}^{\pi} \frac{1}{N+1} \sum_{k=0}^N D_k(x-t) f(t) dt.\end{aligned}$$

Then for $x \neq 0$, we have

$$\begin{aligned}K_N(x) &= \frac{1}{N+1} \sum_{k=0}^N D_k(x) \\ &= \frac{1}{2\pi(N+1)} \frac{1}{2\sin^2 \frac{x}{2}} \sum_{k=0}^N 2 \sin \frac{x}{2} \sin \left(k + \frac{1}{2}\right) x \\ &= \frac{1}{2\pi(N+1)} \frac{1}{2\sin^2 \frac{x}{2}} \sum_{k=0}^N [\cos kx - \cos(k+1)x] \\ &= \frac{1}{2\pi(N+1)} \frac{1}{\sin^2 \frac{x}{2}} \frac{1 - \cos(N+1)x}{2} \\ &= \frac{1}{2\pi(N+1)} \frac{\sin^2 \frac{N+1}{2} x}{\sin^2 \frac{x}{2}}.\end{aligned}$$

It follows immediately that $K_N(x) \geq 0$, $K_N(x)$ is even and 2π periodic.

For property 2, note that for all k ,

$$\int_{-\pi}^{\pi} D_k(t) dt = \int_{-\pi}^{\pi} \sum_{n=-k}^k e^{int} dt = 1.$$

Then,

$$\int_{-\pi}^{\pi} K_N(t) dt = \frac{1}{N+1} \sum_{k=0}^N \int_{-\pi}^{\pi} D_k(t) dt = 1,$$

as desired.

For property 3, let $\delta \in (0, \pi]$. Note that $\sin^2 \frac{x}{2}$ is even and increasing on $[0, \pi]$. It follows that $\delta \leq |x| \leq \pi$ implies $\sin^2 \frac{x}{2} \geq \sin^2 \frac{\delta}{2}$. Therefore,

$$K_N(x) \leq \frac{1}{2\pi(N+1)} \frac{\sin^2 \frac{N+1}{2} x}{\sin^2 \frac{\delta}{2}} \leq \frac{1}{2\pi(N+1) \sin^2 \frac{\delta}{2}}.$$

□

Since the continuous functions that vanishes at both end points is dense in $L^2([-\pi, \pi])$, it make sense to first prove the theorem for continuous functions. We have the following theorem by Fejer.

Theorem (Fejer). If $f \in C([-\pi, \pi])$ is 2π -periodic, $f(\pi) = f(-\pi)$, then $\sigma_N f \rightarrow f$ uniformly on $[-\pi, \pi]$.

Proof. First we extennd f by periodicity to all of \mathbb{R} . Then $f \in C(\mathbb{R})$, 2π -periodic. This implies that f is uniformly continuous and bounded.

Let $\varepsilon > 0$. Since f is uniformly continuous, there exists $\delta > 0$ such that $|y - z| < \delta$ implies that $|f(y) - f(z)| < \frac{\varepsilon}{2}$. Choose $M \in \mathbb{N}$ such that

$$\frac{2 \|f\|_{\infty}}{(N+1) \sin^2 \frac{\delta}{2}} < \frac{\varepsilon}{2}.$$

for all $N \geq M$. Also, since f and K_N are both 2π -periodic, we have

$$\sigma_N f(x) = \int_{-\pi}^{\pi} K_N(x-t)f(t) dt = \int_{x-\pi}^{x+\pi} K_N(\tau)f(x-\tau) d\tau = \int_{-\pi}^{\pi} K_N(\tau)f(x-\tau) d\tau.$$

Then for all $N \geq M$ and for all $x \in [-\pi, \pi]$, we have

$$\begin{aligned} |\sigma_N f(x) - f(x)| &= \left| \int_{-\pi}^{\pi} K_N(t)f(x-t) dt - \int_{-\pi}^{\pi} K_N(t)f(x) dt \right| \\ &\leq \int_{-\pi}^{\pi} K_N(t) |f(x-t) - f(x)| dt \\ &\leq \int_{|t| < \delta} K_N(t) |f(x-t) - f(x)| dt + \int_{\delta \leq |t| \leq \pi} K_N(t) |f(x-t) - f(x)| dt \\ &\leq \frac{\varepsilon}{2} \int_{|t| < \delta} K_N(t) dt + 2 \|f\|_{\infty} \int_{\delta \leq |t| \leq \pi} \frac{1}{2\pi(N+1) \sin^2 \frac{\delta}{2}} dt \\ &\leq \frac{\varepsilon}{2} + \frac{2 \|f\|_{\infty}}{(N+1) \sin^2 \frac{\delta}{2}} \\ &\leq \varepsilon. \end{aligned}$$

□

Remark. The same proof can be modified if instead of $K_N(x) \geq 0$, we have

$$\sup_{N \in \mathbb{N}} \int_{-\pi}^{\pi} |K_N(x)| dx < \infty.$$

Note that

$$\int_{-\pi}^{\pi} |D_N(x)| dx \sim \log N,$$

so we cannot reproduce the proof using Dirichlet kernel.

We only need some last bit of information to conclude the answer of our main question.

Theorem. For all $f \in L^2([-\pi, \pi])$, we have $\|\sigma_N f\|_2 \leq \|f\|_2$.

Proof. Suppose first the $f \in C([-\pi, \pi])$ and 2π -periodic. Then $\sigma_N f(x) = \int_{-\pi}^{\pi} K_N(t)f(x-t) dt$. It follows that

$$\begin{aligned} \int_{-\pi}^{\pi} |\sigma_N f(x)|^2 dx &= \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} f(x-s) \overline{f(x-t)} K_N(s) K_N(t) ds dt dx \\ &= \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} K_N(s) K_N(t) \left[\int_{-\pi}^{\pi} f(x-s) \overline{f(x-t)} dx \right] ds dt \\ &\leq \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} K_N(s) K_N(t) \|f(\cdot - s)\|_2 \|f(\cdot - t)\|_2 ds dt \\ &\leq \|f\|_2^2 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} K_N(s) K_N(t) ds dt \\ &= \|f\|_2^2, \end{aligned}$$

where we used Cauchy-Schwarz inequality. This implies that $\|\sigma_N f\|_2 \leq \|f\|_2$.

Now for the general case, by density there exists sequence $\{f_n\}_{n=0}^{\infty}$ of 2π -periodic continuous function that $\|f_n - f\|_2 \rightarrow 0$. Then, $\|\sigma_N f_n - \sigma_N f\| \rightarrow 0$ as $n \rightarrow \infty$. Therefore,

$$\|\sigma_N f\|_2 = \lim_{n \rightarrow \infty} \|\sigma_N f_n\|_2 \leq \lim_{n \rightarrow \infty} \|f_n\|_2 = \|f\|_2.$$

□

Theorem. For all $f \in L^2([-\pi, \pi])$, we have $\|\sigma_N f - f\|_2 \rightarrow 0$ as $N \rightarrow \infty$. In particular, as a immediate corollary, if $\widehat{f}(n) = 0$ for all $n \in \mathbb{Z}$, then $f = 0$.

Proof. Let $f \in L^2([-\pi, \pi])$ and $\varepsilon > 0$. Again by density there exists $g \in C([-\pi, \pi])$ 2π -periodic such that $\|f - g\|_2 \leq \frac{\varepsilon}{3}$. Since $\sigma_N g \rightarrow g$ uniformly on $[-\pi, \pi]$, there exists $M \in \mathbb{N}$ such that for all $N \geq M$ and all $x \in [-\pi, \pi]$, we have

$$|\sigma_N g(x) - g(x)| < \frac{\varepsilon}{3\sqrt{2\pi}}.$$

Then for all $N \geq M$,

$$\begin{aligned} \|\sigma_N f - f\|_2 &\leq \|\sigma_N(f - g)\|_2 + \|\sigma_N g - g\| + \|g - f\|_2 \\ &\leq 2\|f - g\|_2 + \left(\int_{-\pi}^{\pi} |\sigma_N g(x) - g(x)|^2 dx \right)^{\frac{1}{2}} \\ &\leq \varepsilon. \end{aligned}$$

□

Remark. We have shown that for all $f \in L^2([-\pi, \pi])$, $\|S_N f - f\|_2 \rightarrow 0$. This does not say $S_N f$ converge to f almost everywhere. However, by a theorem by Carleson, for all $f \in L^2([-\pi, \pi])$, we actually do have $S_N f \rightarrow f$ almost everywhere. Also, for all $1 < p < \infty$, $\|S_N f - f\|_p \rightarrow 0$. This is not true for $p = 1$ or $p = \infty$.

2.4 Minimizers, orthogonal complements, and the Riesz representation theorem