Fourier Analysis and the Theory of Distributions

Kexing Ying

February 19, 2022

Contents

1 Orthonormal System

We will in this section recall some results about orthonormal systems in Euclidean spaces[1] and generalize them to complex spaces.

1.1 Euclidean Space

Definition 1.1. A system of nonzero vectors $\{X_\alpha\} \subseteq R$ where R is an Euclidean space is called orthogonal if $\langle X_\alpha, X_\beta \rangle = 0$ for all $\alpha \neq \beta$.

In addition, if for all α, $\langle X_\alpha, X_\alpha \rangle = 1$, we say the system is orthonormal.

Clearly, given an orthogonal system $\{X_\alpha\}$, we may normalize the vector such that $\{X_\alpha / \|X_\alpha\|\}$ is an orthonormal system. Furthermore, recall that a system of orthogonal vectors is linearly independent.

Definition 1.2. A complete (i.e. the smallest closed subspace containing the system is R) orthogonal system $\{X_\alpha\} \subseteq R$ is said to be an orthogonal basis of R.

Some important spaces we shall study in this course include $\mathbb{R}^2$ (equipped with the Euclidean norm), l_2, $\mathcal{C}([-\pi, \pi])$ (the space of continuous functions on $[-\pi, \pi]$ equipped with the L_2 norm).

Proposition 1.1. Let R be a separable Euclidean space. Then any orthogonal system in R is countable.

Proof. By normalizing, we may assume the system $\{X_\alpha\}$ is orthonormal. Then, for $\alpha \neq \beta$,

$$\|X_\alpha - X_\beta\|^2 = \|X_\alpha\|^2 - 2\langle X_\alpha, X_\beta \rangle + \|X_\beta\|^2 = \|X_\alpha\|^2 + \|X_\beta\|^2 = 2.$$

Then, $B_{1/2}(X_\alpha) \cap B_{1/2}(X_\beta) = \emptyset$ for all $\alpha \neq \beta$. Thus, if the system is not countable, we have found a uncountable number of disjoint open balls, contradicting the separability of R. $\quad\square$

Proposition 1.2. Let $f_1, f_2, \cdots$ be a linearly independent system in a Euclidean space R. Then, there exists an orthonormal system $\phi_1, \phi_2, \cdots$ such that

$$\phi_n = a_{n_1} f_1 + \cdots + a_{n_n} f_n$$

and

$$f_n = b_{n_1} \phi_1 + \cdots + b_{n_n} \phi_n$$

for some $a_{n_k}, b_{n_k} \in \mathbb{R}$ and $a_{n_n}, b_{n_n} \neq 0$. Furthermore, the system $\phi_1, \phi_2, \cdots$ is uniquely determined up to a multiplication by ± 1.

Proof. Use Gram-Schmidt. $\quad\square$

Corollary 0.1. A separable Euclidean space R possesses an orthonormal basis.

Proof. Simply obtain the orthonormal system corresponding to the countable dense system of R. The resulting system is complete since the two systems have the same linear closure. $\quad\square$

---

[1] In this course, we shall call real inner product spaces Euclidean spaces.

Definition 1.3 (Fourier Coefficients). Let $\phi_1, \phi_2, \cdots$ be an orthonormal system in R and let $f \in R$. Consider the sequence $c_k = \langle f, \phi_k \rangle$ for all $k = 1, 2, \cdots$. Then c_k are called the coordinates or Fourier coefficients of f with respect to the system $\{\phi_k\}$ and $\sum_{k=1}^{\infty} c_k \phi_k$ is called the Fourier series of f.

Note that this series in the definition is a formal series as we do not yet know whether or not the series converges.

In the finite case, it is not difficult to see that the sequence α_k for $k = 1, \cdots, n$ which minimizes $\|f - S_n^{(\alpha)}\|$ where $S_n^{(\alpha)} := \sum_{k=1}^{n} \alpha_k \phi_k$ is the Fourier coefficients. Indeed, we have

$$\|f - S_n^{(\alpha)}\|^2 = \langle f, f \rangle - 2\langle f, S_n^{(\alpha)} \rangle + \langle S_n^{(\alpha)}, S_n^{(\alpha)} \rangle$$
$$= \|f\|^2 - 2\sum \alpha_k c_k + \sum \alpha_k^2$$
$$= \|f\|^2 - \sum c_k^2 + \sum (\alpha_k - c_k)^2.$$

Hence, $\|f - S_n^{(\alpha)}\|$ is minimized when $\alpha_k = c_k$ for all $k = 1, \cdots, n$. With this in mind, choosing α to be the Fourier coefficients, we have

$$\|f - S_n^{(c)}\| = \|f\|^2 - \sum_{k=1}^{n} c_k^2.$$

Geometrically, $f - S_n^{(\alpha)}$ is orthogonal to the subspace generated by $\phi_1, \cdots, \phi_n$ if and only if $\alpha = c$.

Furthermore, by noting $0 \le \|f - S_n^{(c)}\| = \|f\|^2 - \sum_{k=1}^{n} c_k^2$, we have

$$\sum_{k=1}^{n} c_k^2 \le \|f\|^2 < \infty,$$

and hence, taking $n \to \infty$, we have $\sum_{k=1}^{\infty} c_k^2$ exists and is bounded above by $\|f\|^2$. This inequality is known as the Bessel inequality.

Definition 1.4 (Closed Orthonormal System). The orthonormal system $\{\phi_k\}$ is closed if for any $f \in R$, we have
$$\sum_{k=1}^{\infty} c_k^2 = \|f\|^2.$$
This property is called the Parseval equality.

Again, by observing $\|f - S_n^{(c)}\| = \|f\|^2 - \sum_{k=1}^{n} c_k^2$, the system is closed if and only if for any f, the partial sums of the Fourier series converge to f, i.e. $f = \sum_{k=1}^{\infty} c_k \phi_k$.

Proposition 1.3. In a separable Euclidean space R, an orthonormal system is complete if and only if it is closed.

Proof. Suppose first that $\{\phi_k\}$ is closed. Then, for all $f \in R$, $f = \sum_{k=1}^{\infty} c_k \phi_k$. Thus, the finite linear combinations of $\{\phi_k\}$ is dense in R and thus, $\{\phi_k\}$ is complete.

On the other hand, suppose that $\{\phi_k\}$ is complete (it is countable as R is separable), for any $f \in R$, there exists some α^k such that $\|f - S_\infty^{(\alpha^k)}\| \to 0$. As we have seen, for any partial sum $S_n^{(\alpha^k)}$, we have $\|f - S_n^{(c)}\| \le \|f - S_n^{(\alpha^k)}\|$ and so,

$$\|f - S_\infty^{(c)}\| \le \|f - S_\infty^{(\alpha^k)}\| \to 0$$

implying $\|f - S_\infty^{(c)}\| = 0$ and the system is closed. $\qquad\square$

Proposition 1.4. Given $f, g \in R$ and a closed orthonormal system $\{\phi_k\}$,

$$\langle f, g \rangle = \sum_{k=1}^{\infty} c_k d_k$$

where $(c_k), (d_k)$ are the Fourier coefficients of f and g with respect to $\{\phi_k\}$ respectively.

Proof. We have, by Parseval's identity, $\|f\|^2 = \sum c_k^2$, $\|g\|^2 = \sum d_k^2$ and $\|f + g\|^2 = \sum (c_k + d_k)^2 = \sum c_k^2 + 2 \sum c_k d_k + \sum d_k^2$, we have

$$\sum c_k^2 + 2 \sum c_k d_k + \sum d_k^2 = \|f + g\|^2 = \|f\|^2 + 2\langle f, g \rangle + \|g\|^2.$$

Thus, cancelling using $\|f\|^2 = \sum c_k^2$ and $\|g\|^2 = \sum d_k^2$, we have $\langle f, g \rangle = \sum_{k=1}^{\infty} c_k d_k$ as required. $\qquad\square$

In the case the system is only orthogonal but not necessary orthonormal, we may normalize the Fourier coefficients, i.e. given an orthogonal system $\{\phi_k\}$, we have $\{\phi/\|\phi_k\|\}$ is an orthonormal system, and so, we define

$$c_k = \left\langle f, \frac{\phi_k}{\|\phi_k\|} \right\rangle = \frac{1}{\|\phi_k\|} \langle f, \phi_k \rangle.$$

Similarly, the Fourier series of f is becomes

$$\sum_{k=1}^{\infty} c_k \frac{\phi_k}{\|\phi_k\|} = \sum \frac{\langle f, \phi_k \rangle}{\|\phi_k\|^2} \phi_k.$$

Substituting this definition of the Fourier coefficients into the Bessel inequality, we obtain

$$\sum_{k=1}^{\infty} \frac{\langle f, \phi_k \rangle^2}{\|\phi_k\|^2} \le \|f\|^2,$$

for any orthogonal system $\{\phi_k\}$.

Theorem 1 (Riesz)**.** Let $\{\phi_k\}$ be a orthonormal system in a complete Euclidean space R (i.e. a real Hilbert space) and let $c \in \ell_2$ (i.e. $\sum_{k=1}^{\infty} c_k^2 < \infty$). Then, there exists some $f \in R$ such that $c_k = \langle f, \phi_k \rangle$ and Parseval's identity holds, i.e.

$$\sum_{k=1}^{\infty} c_k^2 = \|f\|^2.$$

Proof. Let $f_n := \sum_{k=1}^{n} c_k \phi_k$. Then, by definition, we have $c_k = \langle f_n, \phi_k \rangle$ for all $k = 1, \cdots, n$. Then, for all $p \geq 1$, we have

$$\|f_{n+p} - f_n\|^2 = \|c_{n+1}\phi_{n+1} + \cdots + c_{n+p}\phi_{n+p}\|^2 = \sum_{k=n+1}^{n+p} c_k^2.$$

Now, as $\sum c_k^2 < \infty$, we have $\{f_n\}$ is Cauchy, and thus, as R is complete, there exists some $f \in R$ such that $f_n \to f$. Thus, by noting,

$$\langle f, \phi_k \rangle = \langle f_n \phi_k \rangle + \langle f - f_n, \phi_k \rangle = c_k + \langle f - f_n, \phi_k \rangle,$$

where $\langle f - f_n, \phi_k \rangle \to 0$ as $n \to \infty$ since $|\langle f - f_n, \phi_k \rangle| \leq \|f - f_n\|\|\phi_k\|$ by the Cauchy-Schwarz inequality, we have $c_k = \langle f, \phi_k \rangle$.

Finally, Parseval's identity, follows as $\| \cdot \|^2$ is continuous in a normed space. $\qquad\square$

Let us recall the following result from functional analysis.

Proposition 1.5. Any separable Hilbert space is isomorphic to ℓ_2 (thus, any two separable Hilbert spaces are isomorphic).

Proof. Let H be a separable Hilbert space and choose $\{\phi_k\}$ a complete orthonormal system (which exists as H is separable). Then, for any $f \in H$, we map f to the sequence corresponding to its Fourier coefficients, i.e.

$$\psi : f \mapsto (c_1, c_2, \cdots)$$

which is well-defined by Bessel's inequality. On the other hand, by Riesz's theorem, for any $x \in \ell_2$, $\sum x_k^2 < \infty$ and so, there exists a unique $f \in H$, such that $\psi(f) = x$. Thus, as ψ is clearly linear (as the inner products are linear with respect to the left component), we have the isomorphism between H and ℓ_2. $\qquad\square$

1.2 Complex Inner Product Space

We will in the course take the complex inner product to be anti-linear in the second component. As promised earlier, most definitions can be generalized from the real case to the complex directly.

Definition 1.5 (Fourier Coefficients)**.** Let R be a complex inner product space. Then, for an orthonormal system (ϕ_n) and $f \in R$, we define its Fourier coefficients to be $c_k := \langle f, \phi_k \rangle$ for all $k = 1, \cdots, n$. Similarly, we define the Fourier series of f to be the formal series $\sum_{k=1}^{\infty} c_k \phi_k$.

Going through the same argument as the real case, we obtain the complex version of Bessel's inequality.

Proposition 1.6 (Bessel's Inequality)**.** Given an orthonormal system (ϕ_n) and $f \in R$, we have

$$\sum_{k=1}^{\infty} |c_k|^2 \leq \|f\|^2.$$

Going through all proved theorems for real spaces, we find they also hold for complex spaces (with trivial modifications).

2 Fourier Series

2.1 Trigonometric Series

We will consider the space $L_2[-\pi, \pi]$ (the space of square-integrable functions from $[-\pi, \pi]$ quotiented by the a.e.-equal equivalence relation equipped with the inner product $\langle f, g \rangle :=$ $\int_{[-\pi,\pi]} fg\,\mathrm{d}\lambda$), and the trigonometric system

$$\{\mathbf{1}, \cos(nx), \sin(nx) \mid n = 1, 2, \cdots\}.$$

It is not difficult to see that this system is orthogonal, but in fact, it is also complete. Indeed, completeness follows by the Weierstrass approximation theorem for trigonometric polynomials (we will discuss this later / recall the Stone-Weierstrass theorem and observe that the trigonometric system seperates points).

Nonetheless, this system is not orthonormal, and thus, we normalise the system such that the system becomes

$$\left\{ \frac{1}{\sqrt{2\pi}}\mathbf{1}, \frac{1}{\sqrt{\pi}}\cos(nx), \frac{1}{\sqrt{\pi}}\sin(nx) \mid n = 1, 2, \cdots \right\}.$$

Hence, the Fourier series of an element $f \in L_2[-\pi, \pi]$ becomes the famous formula

$$\frac{a_0}{2} + \sum_{k=1}^{\infty}(a_k \cos kx + b_k \sin kx),$$

where

$$a_k := \frac{1}{\pi}\int_{[-\pi,\pi]} f(x)\cos(kx)\lambda(\mathrm{d}x), b_k := \frac{1}{\pi}\int_{[-\pi,\pi]} f(x)\sin(kx)\lambda(\mathrm{d}x).$$

By recalling the above theory, the n-th partial sum of this series provides the best (in L_2 metric) approximation of f among all trigonometric polynomials of degree n. Hence, as the trigonometric system is complete, Parseval's identity holds, and so,

$$\|f - S_n\|_2 \to 0, \text{ as } n \to \infty.$$

By observing that $e^{ix} = \cos x + i\sin x$, we may rewrite the Fourier series can be written in the complex form. In particular, $L_2[-\pi, \pi]$ has the orthogonal system $\{e^{inx} \mid n \in \mathbb{Z}\}$ and the Fourier series of $f \in L_2[-\pi, \pi]$ is given by

$$\sum_{n=-\infty}^{\infty} c_n e^{inx}, \text{ where } c_n = \frac{1}{2\pi}\int_{[-\pi,\pi]} f(x)e^{-inx}\lambda(\mathrm{d}x).$$

- Since a function on $[-\pi, \pi]$ can be extended to $\mathbb{R}$ by periodicity, we can, instead of functions on $[-\pi, \pi]$ consider periodic functions with period 2π on $\mathbb{R}$.

- Since $\cos nx, \sin nx$ are bounded functions, the integrals defining the trigonometric Fourier coefficients exists for any function in $L_1[-\pi, \pi]$, i.e. if $f \in L_1[-\pi, \pi]$, then

$$\int f \cos(nx), \int f \sin(nx) < \int |f| < \infty.$$

- $L_2[-\pi, \pi] \subseteq L_1[-\pi, \pi]$ by Hölder's inequality and thus, with the above remark in mind, the definition of Fourier series is also well-defined for any integrable functions (though convergence is much opaque in this case).

While the Fourier series of f converges to f in L_2 though it is not clear that the Fourier series converges point-wise to f (it might be interesting to recall that convergence in L_p implies convergence in measure and the existence of a subsequence which converges almost everywhere).

Consider the partial sum of the Fourier series of $f \in L_2[-\pi, \pi]$,

$$S_n(x) = \frac{a_0}{2} + \sum_{k=1}^{n}(a_k \cos kx + b_k \sin kx)$$

$$= \frac{1}{\pi} \int_{[-\pi,\pi]} f(t) \left(\frac{1}{2} + \sum_{k=1}^{n}(\cos kx \cos kt + \sin kx \sin kt) \right) \lambda(\mathrm{d}t)$$

$$= \frac{1}{\pi} \int_{[-\pi,\pi]} f(t) \left(\frac{1}{2} + \sum_{k=1}^{n} \cos k(t - x) \right) \lambda(\mathrm{d}t).$$

By noting the identity

$$\frac{1}{2} + \sum_{k=1}^{n} \cos ku = \frac{\sin \frac{2n+1}{2}u}{2 \sin \frac{u}{2}},$$
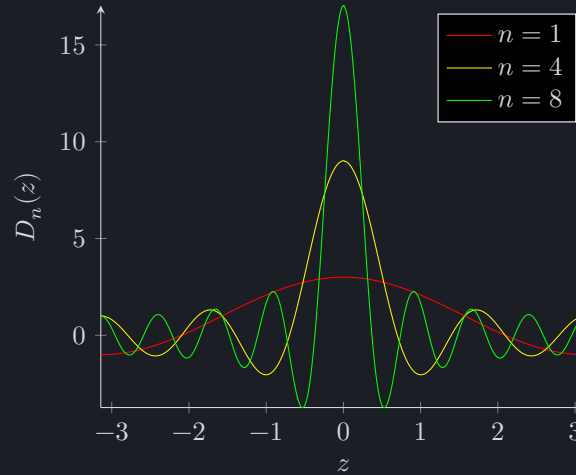
we obtain

$$S_n(x) = \frac{1}{\pi} \int_{[-\pi,\pi]} f(t) \frac{\sin \frac{2n+1}{2}(t - x)}{2 \sin \frac{t-x}{2}} \lambda(\mathrm{d}t).$$

Finally, by noting the periodicity of f, by change of variable $z = t - x$, we obtain

$$S_n(x) = \int_{[-\pi,\pi]} f(x + z) D_n(z) \lambda(\mathrm{d}z), \text{ where } D_n(z) := \frac{1}{2\pi} \frac{\sin \frac{2n+1}{2}z}{\sin \frac{z}{2}}$$

and D_n is known as the Dirichlet kernel. We remark that the Dirichlet kernel $D_n(z)$ tends to $\frac{2n+1}{2\pi}$ as $z \to 0$ and rapidly osculates for large n though this does not impact our calculation as we are dealing with a point which has measure 0.

By observing the graph of the Dirichlet kernel, in some heuristic sense, we note that $D_n(z) \to \delta(z)$ for some function where δ is 0 at all points but $z = 0$ while $\int_{[-\epsilon,\epsilon]} \delta d\lambda = 1$ for all $\epsilon > 0$. Such an function cannot exist, however it motivates the second part of the course - the theory of distributions.

2.2   Conditions for Point-wise Convergence

We observe that $\|D_n\|_1 = 1$ and so we may write

$$S_n(x) - f(x) = \int_{[-\pi,\pi]} (f(x+z) - f(x)) D_n(z) \lambda(\mathrm{d}z).$$

Clearly, $S_n(x) \to f(x)$ as $n \to \infty$ if and only if the right hand side of the above tends to 0.

Lemma 2.1 (Riemann-Lebesgue). If $\phi \in L_1[a,b]$ for some $a < b$, then both

$$\int_{[a,b]} \phi(x) \sin(\gamma x) \lambda(\mathrm{d}x), \int_{[a,b]} \phi(x) \cos(\gamma x) \lambda(\mathrm{d}x)$$

tends to 0 as $\gamma \to \infty$.

Proof. We will prove the statement for the sin case. We observe that if ϕ is continuously differentiable, by integration by parts, we have

$$\int_{[a,b]} \phi(x) \sin(\gamma x) \lambda(\mathrm{d}x) = \left[-\phi(x) \frac{\cos \gamma x}{\gamma} \right]_a^b + \int_{[a,b]} \phi'(x) \frac{\cos \gamma x}{\gamma} \lambda(\mathrm{d}x),$$

which tends to 0 as $\gamma \to \infty$ (we note that ϕ' is continuous on a compact set, and hence bounded above). Now, in the general case, we observe that continuously differentiable functions are everywhere dense in $L_1[a,b]$, and so, for every $\epsilon > 0$, there exists some continuously differentiable $\phi_\epsilon(x)$, such that

$$\int_{[a,b]} |\phi - \phi_\epsilon| \mathrm{d}\lambda = \|\phi - \phi_\epsilon\|_1 < \frac{\epsilon}{2}.$$

Hence,

$$\left| \int_{[a,b]} \phi(x) \sin(\gamma x) \lambda(\mathrm{d}x) \right| \le \left| \int_{[a,b]} (\phi(x) - \phi_\epsilon(x)) \sin(\gamma x) \lambda(\mathrm{d}x) \right| + \left| \int_{[a,b]} \phi_\epsilon(x) \sin(\gamma x) \lambda(\mathrm{d}x) \right|$$

$$\le \|\phi - \phi_\epsilon\|_1 + \left| \int_{[a,b]} \phi_\epsilon(x) \sin(\gamma x) \lambda(\mathrm{d}x) \right|$$

$$< \frac{\epsilon}{2} + \left| \int_{[a,b]} \phi_\epsilon(x) \sin(\gamma x) \lambda(\mathrm{d}x) \right|.$$

Since ϕ_ϵ is continuously differentiable, the last term tends to 0 as $\gamma \to \infty$ implying it is less than $\epsilon/2$ for sufficiently large γ. Thus, we have established the required limit. $\square$

Corollary 1.1. If $f \in L_1[-\pi, \pi]$, then, its Fourier coefficients $a_k, b_k \to 0$ as $k \to \infty$.

With the above in mind, we will now provide a sufficient condition for convergence of Fourier series at a point x.

Theorem 2. If $f \in L_1[-\pi, \pi]$ and for any $x \in [-\pi, \pi]$, there exists some $\delta > 0$,

$$\int_{[-\delta,\delta]} \left| \frac{f(x+t) - f(x)}{t} \right| \lambda(\mathrm{d}t) < \infty$$

exists (this is called Dini's condition), then $S_n(x) \to f(x)$ as $n \to \infty$.

Proof. We observe

$$S_n(x) - f(x) = \int_{[-\pi,\pi]} (f(x+z) - f(x))D_n(z)\lambda(\mathrm{d}z)$$

$$= \frac{1}{2\pi} \int_{[-\pi,\pi]} \frac{f(x+z) - f(x)}{z} \frac{z}{\sin \frac{z}{2}} \sin\left(\frac{2n+1}{2}z\right) \lambda(\mathrm{d}z).$$

Then, if Dini's condition is satisfied, as $\sin \frac{z}{2}$ is bounded on $[-\pi, \pi]$, we have

$$\frac{f(x+z) - f(x)}{z} \frac{z}{\sin \frac{z}{2}} \in L_1[-\pi, \pi]$$

and hence, by the Riemann-Lebesgue lemma, $S_n(x) - f(x) \to 0$ as $n \to \infty$. $\qquad \square$

We observe an trivial sufficient condition for which Dini's condition holds. In particular, the Dini condition holds if f is continuous at x and its derivative at x exists (or the limit of the derivative exists from the right and the left).

Suppose on the other hand that f has a discontinuity of the first kind at x (i.e. both the limit from the left and the right exists at x). Then, the argument in the proof remains to hold if we replace the Dini condition by

$$\int_{[-\delta,0]} \left| \frac{f(x+t) - f(t)}{t} \right| \lambda(\mathrm{d}t) < \infty \text{ and } \int_{[0,\delta]} \left| \frac{f(x+t) - f(t)}{t} \right| \lambda(\mathrm{d}t) < \infty.$$

Then, denoting $f(x+0), f(x-0)$ the limit of f at x from the right and the left respectively, we have

$$S_n(x) - \frac{f(x+0) + f(x-0)}{2} =$$
$$\int_{[-\pi,0]} (f(x+z) - f(x-0))D_n(z)\lambda(\mathrm{d}z) + \int_{[0,\pi]} (f(x+z) - f(x+0))D_n(z)\lambda(\mathrm{d}z)$$

Hence, the $S_n(x)$ converges to the average of the limit f at x from the left and from the right.

Let us summarise the above in the following statement.

Proposition 2.1. If f is a bounded function of period 2π with only discontinuities of the first kind, and also possesses at each point left and right derivatives, Then, its Fourier series converges everywhere and its sum equals $f(x)$ at points of continuity and equals

$$\frac{f(x+0) + f(x-0)}{2}$$

at points of discontinuity.

We remark there exists continuous functions whose Fourier series diverge at some points. More curiously, there exists L_1 functions which diverge at all points.

2.3 Continuous Functions

Suppose now f is a continuous function with period 2π on $\mathbb{R}$, then it is uniquely determined by its Fourier series (exercise). Nonetheless, as we have seen, the Fourier series of f does not necessarily equal to f at every point. However, we can still reconstruct f from its Fourier series. Consider the partial sums

$$S_k(x) = \frac{a_0}{2} + \sum_{j=1}^{k} a_j \cos(jx) + b_j \sin(jx).$$

Then, we define the Feje's sums

$$\sigma_n(x) := \frac{1}{n}(S_0(x) + S_1(x) + \cdots S_{n-1}(x)),$$

we have the following theorem.

Definition 2.1 (Fejer's Kernel)**.** Fejer's kernel is defined to be the function

$$\Phi_n(z) = \frac{1}{2\pi n} \left(\frac{\sin(nz/2)}{\sin(z/2)} \right)^2$$

for all $n \geq 0$.

Lemma 2.2. Denoting Φ_n as Fejer's kernel, we have

- $\Phi_n \geq 0$;
- $\int_{[-\pi,\pi]} \Phi_n \, d\lambda = 1$;
- for all $\delta \in (0, \pi]$, $\int_{[-\pi,-\delta]} \Phi_n \, d\lambda = \int_{[\delta,\pi]} \Phi_n \, d\lambda \to 0$ as $n \to \infty$.

Proof. Exercise. $\qquad\square$

Theorem 3 (Fejer's)**.** If f is a continuous function with period 2π, then the sequence (σ_n) of its Fejer's's sums converges to f uniformly on $\mathbb{R}$.

Proof. Recall that

$$S_k(x) = \frac{1}{2\pi} \int_{[-\pi,\pi]} f(x+z) D_k(z) \lambda(\mathrm{d}z),$$

and so,

$$\sigma_n(x) = \frac{1}{2\pi n} \int_{[-\pi,\pi]} f(x+z) \sum_{k=0}^{n-1} D_k(z) \lambda(\mathrm{d}z).$$

Then, using the identity that

$$\sum_{k=0}^{n-1} \sin(2k+1)u = \frac{\sin^2(nu)}{\sin u},$$

10

we obtain

$$\sigma_n(x) = \int_{[-\pi,\pi]} f(x+z)\Phi_n(z)\lambda(\mathrm{d}z)$$

where Φ_n is Fejer's kernel. Now, by nothing that f is continuous and periodic, f is bounded and uniformly continuous on $\mathbb{R}$. Thus, there exists some $M > 0$, such that for all $\sup_x |f(x)| \leq M$ and for all $\epsilon > 0$, there exists some $\delta \in (0, \pi]$ such that $|f(x) - f(y)| < \frac{\epsilon}{2}\left(\int_{[-\pi,\pi]} \Phi_n \mathrm{d}\lambda\right)^{-1}$ for all $|x-y| < 2\delta$. Now,

$$f(x) - \sigma_n(x) = \int_{[-\pi,\pi]} (f(x) - f(x+z))\Phi_n(z)\lambda(\mathrm{d}z)$$

$$= \left(\int_{[-\pi,-\delta]} + \int_{[-\delta,\delta]} + \int_{[\delta,\pi]}\right)(f(x) - f(x+z))\Phi_n(z)\lambda(\mathrm{d}z).$$

Then, by the above lemma, there exists some N such that for all $n \geq N$, $\int_{[-\pi,-\delta]} \Phi_n \mathrm{d}\lambda < \frac{\epsilon}{8M}$ and so, for all $n \geq N$, we have $\int_{[-\pi,-\delta]} (f(x) - f(x+z))\Phi_n(z)\lambda(\mathrm{d}z) \leq \int_{[-\pi,-\delta]} 2M\Phi_n(z)\lambda(\mathrm{d}z) < 2M\frac{\epsilon}{8M} = \epsilon/4$

$$f(x) - \sigma_n(x) < \frac{\epsilon}{4} + \frac{\epsilon}{4} + \int_{[-\delta,\delta]} (f(x) - f(x+z))\Phi_n \mathrm{d}\lambda$$

$$< \frac{\epsilon}{2} + \frac{\epsilon}{2}\left(\int_{[-\pi,\pi]} \Phi_n \mathrm{d}\lambda\right)^{-1} \int_{[-\delta,\delta]} \Phi_n \mathrm{d}\lambda$$

$$\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Hence, $|f(x) - \sigma_n(x)| = f(x) - \sigma_n(x) < \epsilon$ for all $x \in \mathbb{R}$ implying uniform convergence as required. □

While we had used the Weierstass approximation to obtain the above, Fejer's theorem also implies Weierstass approximation theorem straight away. Nonetheless, while the Stone-Weierstass theorem is more general, Fejer's theorem provides an explicit construction of such an series.

Corollary 3.1 (Weierstass Approximation Theorem)**.** Any continuous periodic function is a limit of a uniformly convergent sequence of trigonometric polynomials.

Corollary 3.2. The trigonometric system $\{1, \sin nx, \cos nx \mid n = 1, 2, \cdots\}$ is complete in $L_2[-\pi, \pi]$ as uniform convergence implies convergence in L_2.

We remark that convergence in $L_p[-\pi, \pi]$ implies the existence of a subsequence which converges almost everywhere and hence, by Egorov's theorem, there exists a subsequence which converges uniformly on an arbitrarily large space.

By recalling that uniform convergence is definitionally equal to the supremum norm, Fejer's theorem tells us that for all $f \in \mathcal{C}_\infty[-\pi, \pi]$, the sequence of its Fejer's sums converges f in $\mathcal{C}_\infty[-\pi, \pi]$. Furthermore, if $f \in L_1[-\pi, \pi]$, then $\sigma_n \to f$ in L_1.

Corollary 3.3. For all $f \in L_1[-\pi, \pi]$, it is (a.e.) uniquely determined by its Fourier coefficients.

Proof. Suppose $f, g \in L_1[-\pi, \pi]$ have the same Fourier coefficients. Then, f and g have the same Fejer's sums. But, as L_1 is a metric space, the Fejer's sums has a unique limit and hence, $f = g$ in L_1. $\qquad\square$

3 Fourier Transform

Thus far, we have considered periodic functions in L_1 is represented uniquely by its Fourier series. We would now like to generalize this theory to non-periodic functions.

As an intuition (we will formally justify this later), consider the case first that f is periodic on $[-l, l]$ for some $l > 0$. From the exercise sheet, we were able to extend the Fourier series and conclude that, if $f \in L_1$ and satisfies Dini's condition at each point of $[-l, l]$, then

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cos\left(\frac{k\pi}{l}x\right) + b_k \sin\left(\frac{k\pi}{l}x\right),$$

where

$$a_k = \frac{1}{l} \int_{[-l,l]} f(t) \cos\left(\frac{k\pi}{l}t\right) \lambda(\mathrm{d}t) \text{ and } b_k = \frac{1}{l} \int_{[-l,l]} f(t) \sin\left(\frac{k\pi}{l}t\right) \lambda(\mathrm{d}t).$$

Thus,

$$f(x) = \frac{1}{2l} \int_{[-l,l]} f \mathrm{d}\lambda + \frac{1}{l} \sum_{k=1}^{\infty} f(t) \cos\left(\frac{k\pi}{l}(t - x)\right) \lambda(\mathrm{d}t).$$

Hence, setting $y_k = \pi k/l$, $\Delta y = \pi/l$ and taking $l \to \infty$, we expect

$$f(x) = \frac{1}{\pi} \int_{(0,\infty)} \lambda(\mathrm{d}y) \int_{\mathbb{R}} f(t) \cos(y(t - x)) \lambda(\mathrm{d}t).$$

This equation is known as the Fourier integral. To see the similarity of this with the Fourier series, we may write the Fourier integral as

$$f(x) = \int_{(0,\infty)} (a_y \cos(yx) + b_y \sin(y_x)) \lambda(\mathrm{d}y)$$

where

$$a_y = \frac{1}{\pi} \int_{\mathbb{R}} f(t) \cos(yt) \lambda(\mathrm{d}t) \text{ and } b_y = \frac{1}{\pi} \int_{\mathbb{R}} f(t) \sin(yt) \lambda(\mathrm{d}t).$$

Theorem 4. Let $f \in L_1$ and satisfy Dini's condition at every point. Then for all $x \in \mathbb{R}$, we have

$$f(x) = \frac{1}{\pi} \int_{(0,\infty)} \lambda(\mathrm{d}y) \int_{\mathbb{R}} f(t) \cos(y(t - x)) \lambda(\mathrm{d}t).$$

Proof. Let us denote

$$J_x(A) := \frac{1}{\pi} \int_{(0,A)} \lambda(\mathrm{d}y) \int_{\mathbb{R}} f(t) \cos(y(t - x)) \lambda(\mathrm{d}t),$$

and we will show $J_x(A) \to f(x)$ as $A \to \infty$. By assumption, $f \in L_1$ and so, J_x converges absolutely, and hence, by Fubini's theorem,

$$J_x(A) = \frac{1}{\pi} \int_{\mathbb{R}} \lambda(\mathrm{d}t) f(t) \int_{(0,A)} \cos(y(t - x)) \lambda(\mathrm{d}y) = \frac{1}{\pi} \int_{\mathbb{R}} \lambda(\mathrm{d}t) f(t) \frac{\sin A(t - x)}{t - x}.$$

By change of variables $z = t - x$, we obtain

$$J_x(A) = \frac{1}{\pi} \int_{\mathbb{R}} f(x+z) \frac{\sin Az}{z} \lambda(\mathrm{d}z).$$

Now, by observing that $\frac{1}{\pi} \int_{\mathbb{R}} \frac{\sin Az}{z} \lambda(\mathrm{d}z) = 1$, we have

$$\begin{aligned}
J_x(A) - f(x) &= \frac{1}{\pi} \int_{\mathbb{R}} \frac{f(x+z) - f(x)}{z} \sin Az \lambda(\mathrm{d}z) \\
&= \frac{1}{\pi} \int_{[-N,N]} \frac{f(x+z) - f(x)}{z} \sin Az \lambda(\mathrm{d}z) \\
&+ \frac{1}{\pi} \int_{\mathbb{R}\setminus[-N,N]} \frac{f(x+z)}{z} \sin Az \lambda(\mathrm{d}z) \\
&- \frac{f(x)}{\pi} \int_{\mathbb{R}\setminus[-N,N]} \frac{\sin Az}{z} \lambda(\mathrm{d}z)
\end{aligned}$$

for all $N > 0$. By observing that the last two integrand are bounded, both integrals tends to 0 as $N \to \infty$. On the other hand, the first integral tends to 0 as $A \to \infty$ for fixed N by the Riemann-Lebesgue lemma. With this in mind, let $\epsilon > 0$, then, there exists some $N > 0$ such that

$$\left| \frac{1}{\pi} \int_{\mathbb{R}\setminus[-N,N]} \frac{f(x+z)}{z} \sin Az \lambda(\mathrm{d}z) \right|, \left| \frac{f(x)}{\pi} \int_{\mathbb{R}\setminus[-N,N]} \frac{\sin Az}{z} \lambda(\mathrm{d}z) \right| < \frac{\epsilon}{3}.$$

Then, by the Riemann-Lebesgue lemma (and Dini's condition), there exists some B, such that for all $A \geq B$,

$$\left| \frac{1}{\pi} \int_{[-N,N]} \frac{f(x+z) - f(x)}{z} \sin Az \lambda(\mathrm{d}z) \right| < \frac{\epsilon}{3}.$$

Hence, for all $A \geq B$, we have

$$\begin{aligned}
|J_x(A) - f(x)| &\leq \left| \frac{1}{\pi} \int_{[-N,N]} \frac{f(x+z) - f(x)}{z} \sin Az \lambda(\mathrm{d}z) \right| \\
&+ \left| \frac{1}{\pi} \int_{\mathbb{R}\setminus[-N,N]} \frac{f(x+z)}{z} \sin Az \lambda(\mathrm{d}z) \right| \\
&- \left| \frac{f(x)}{\pi} \int_{\mathbb{R}\setminus[-N,N]} \frac{\sin Az}{z} \lambda(\mathrm{d}z) \right| < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon,
\end{aligned}$$

implying $J_x(A) \to f(x)$ as $A \to \infty$ as required. $\qquad\square$

Similar to the finite case, we may write the Fourier integral using the complex exponential function. In particular, as cos is even, we have

$$f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} \lambda(\mathrm{d}y) \int_{\mathbb{R}} f(t) \cos(y(t-x)) \lambda(\mathrm{d}t).$$

On the other hand, as sin is odd,

$$\frac{1}{2\pi}\int_{\mathbb{R}}\lambda(\mathrm{d}y)\int_{\mathbb{R}}f(t)\sin(y(t-x))\lambda(\mathrm{d}t)=0.$$

Summing the two integrals, we obtain that, if $f \in L_1$ and satisfies the Dini condition, then

$$f(x)=\frac{1}{2\pi}\int_{\mathbb{R}}\lambda(\mathrm{d}y)\int_{\mathbb{R}}f(t)e^{-iy(t-x)}\lambda(\mathrm{d}t).$$

Definition 3.1 (Fourier Transform). Let $f \in L_1$. We define the Fourier transform of f to be

$$g(y)=\mathcal{F}[f](y):=\int_{\mathbb{R}}f(t)e^{-iyt}\lambda(\mathrm{d}t).$$

Proposition 3.1 (Inversion of the Fourier Transform). In the case that f satisfies Dini's condition at some point x, then

$$f(x)=\frac{1}{2\pi}\int_{\mathbb{R}}\mathcal{F}[f](y)e^{iyx}\lambda(\mathrm{d}y).$$

Proof. See above. □

3.1 Properties of the Fourier Transform

From the above process, we observe that the Fourier transform is very similar to the Fourier series in which the Fourier transform is analogous to the Fourier coefficients while the inversion formula is analogous to the Fourier series both of which holds under Dini's condition.

Theorem 5. Let $f \in L_1$ and suppose $\mathcal{F}[f]=0$, then $f=0$ almost everywhere.

Thus, similar to the case of the Fourier series, the Fourier transform (a.e.) uniquely determines a function.

Proof. We note that we have not required Dini's condition and thus, we may not use the inversion formula.

By change of variable, we note that $\int f(x+t)e^{-iyx}\lambda(\mathrm{d}x)=0$ for all t,y, and let us define

$$\phi_{\mu}(x):=\int_{(0,\mu)}f(x+t)\lambda(\mathrm{d}t)$$

for some $\mu > 0$. As ϕ_{μ} is integrable, by Fubini's theorem, it follows that $\mathcal{F}[\phi_{\mu}]=0$. Furthermore, as ϕ_{μ} is an integral function, it is absolutely continuous on any finite interval, and hence, has a derivative a.e. and so, it satisfies Dini's condition a.e. Hence, by the Fourier inversion formula, we have

$$\phi_{u}(x)=\frac{1}{2\pi}\int_{\mathbb{R}}\mathcal{F}[\phi_{\mu}](y)e^{iyx}\lambda(\mathrm{d}y)=0,$$

almost everywhere. But, as ϕ_{μ} is continuous (as its absolutely continuous), and so $\phi_{\mu}=0$ everywhere. Thus, it follows that $f=0$ a.e. □

Lemma 3.1. Let $f, f_n \in L_1$ for $n = 1, 2, \cdots$ such that $f_n \to f$ in L_1. Then, $g_n(y) := \mathcal{F}[f_n](y) \to \mathcal{F}[f](y)$ uniformly on $\mathbb{R}$.

Proof. Follows as

$$|g_n - \mathcal{F}[f]| \leq \int_{\mathbb{R}} |f_n - f| d\lambda.$$

$\square$

Lemma 3.2. Let $f \in L_1$, then $\mathcal{F}[f]$ is a bounded continuous function and $\mathcal{F}[f](y) \to 0$ as $|y| \to \infty$.

Proof. Again, boundedness follows from the bound $|\mathcal{F}[f]| \leq \|f\|_1 < \infty$.

For the other two properties, we will first prove for indicator functions. Consider

$$\mathcal{F}[\mathbf{1}_{[a,b]}](y) = \int_{[a,b]} e^{-iyx} \lambda(dx) = \frac{e^{-iyb} - e^{-iya}}{-iy}$$

which is continuous and tends to 0 as $|y| \to \infty$.

By the linearity of the Lebesgue integral, and hence, the Fourier transform is linear, we have that the Fourier transform of any simple function is also continuous and tends to 0 as $|y| \to \infty$. Then, as the simple function is dense in L_1, for any $f \in L_1$, there exists a sequence (f_n) of simple functions such that $f_n \to f$ in L_1. This, by the above lemma, we obtain $\mathcal{F}[f_n] \to \mathcal{F}[f]$ uniformly. Hence, $\mathcal{F}[f]$ is continuous and tends to 0 as $|y| \to \infty$. $\square$

Lemma 3.3. Let $f \in L_1$ be differentiable with L_1 derivative f' and suppose f is absolutely continuous on any finite integral. Then,

$$\mathcal{F}[f'] = iy \cdot \mathcal{F}[f].$$

Proof. Since f is a.c.

$$f(x) = f(0) + \int_{[0,x]} f' d\lambda.$$

Thus, as $f' \in L_1$, $\lim_{x \to \pm\infty} f(x)$ exists. Now since $f \in L_1$, $\lim_{x \to \pm\infty} f(x) = 0$ and so, by integration by parts,

$$\mathcal{F}[f'] = \int f'(x) e^{-iyx} \lambda(dx) = iy \int f(x) e^{-iyx} \lambda(dx) = iy\mathcal{F}[f]$$

as required. $\square$

Corollary 5.1. In the case that f is n-times differentiable and $f^{(k)} \in L_1$ for all $k = 0, \cdots, n$, then applying the above lemma n-times, we have

$$\mathcal{F}[f^{(n)}] = (iy)^n \cdot \mathcal{F}[f],$$

and $|\mathcal{F}[f]| = |\mathcal{F}[f^{(n)}]|/|y|^n \leq C/|y|^n \to 0$ as $|y| \to \infty$ as $\mathcal{F}[f^{(n)}]$ is bounded.

From the above corollary, we see that, similar to the Fourier series, $\mathcal{F}[f]$ decays faster at infinity the smoother it is.

Corollary 5.2. If $f, f', f'' \in L_1$, then $\mathcal{F}[f] \in L_1$.

Proof. We have $|\mathcal{F}[f]| \leq C/|y|^2$ for some constant C and so,

$$\int_{\mathbb{R}} |\mathcal{F}[f]| \mathrm{d}\lambda \leq \int_{\mathbb{R}} \frac{C}{|y|^2} \lambda(\mathrm{d}y) < \infty.$$

$\square$

The opposite is also true, namely, the faster f decays at infinity, the smoother $\mathcal{F}[f]$ is.

Lemma 3.4. Let $f(x), xf(x) \in L_1$, then, $\mathcal{F}[f]$ is differentiable and

$$D_y \mathcal{F}[f](y) = \mathcal{F}[-ixf(x)]$$

On the other hand, if $f(x), xf(x), \cdots, x^p f(x) \in L_1$, then $\mathcal{F}[f]$ is p-times differentiable and

$$D_y^p \mathcal{F}[f](y) = \mathcal{F}[(-ix)^p f(x)].$$

Proof. It suffices to show the first statement as the second follows by induction. But this is clear by differentiating with respect to y under the integral (where we may differentiate under the integral as $|xf(x)e^{-iyx}| \leq |xf(x)|$ where $xf(x) \in L_1$ and so, we may differentiate by the dominated convergence theorem). $\square$

Corollary 5.3. If $x^p f(x) \in L_1$ for all p. Then, $\mathcal{F}[f]$ is infinitely differentiable.

Corollary 5.4. If $e^{\delta|x|} f(x) \in L_1$ for some $\delta > 0$, then $\int f(x) e^{ix\zeta} \lambda(\mathrm{d}x)$ uniformly converges for all $\zeta = y + iz$, $|z| < \delta$. Therefore,

$$g(\zeta) := \int f(x) e^{ix\zeta}$$

is an analytic function on the neighbourhood $\mathbb{R} \times (-\delta, \delta)$.

3.2   Schwartz Functions

Let us now consider the space of functions which are invariant under the Fourier transform. This will lead us to define the Schwartz functions.

Since the smoothness and decay at infinity swap after an application of the Fourier transform (F.T.), we can easily indicate a class of function which are invariant under F.T.

Definition 3.2 (Schwartz Function)**.** We define the space S_∞ as the space of functions on $\mathbb{R}$ which are infinitely differentiable and such that for all $p, q = 0, 1, \cdots$, there exists some constant $C_{p,q}^f$ such that

$$|x^p f^{(q)}(x)| < C_{p,q}^f$$

for all $x \in \mathbb{R}$. If $f \in S_\infty$, then we call f a Schwartz function.

Proposition 3.2. If $f \in S_\infty$, then $g := \mathcal{F}[f] \in S_\infty$.

Proof. Since $|x^p f^{(q)}(x)| < x^{-2} C^f_{p+2,q}$ and so, is integrable for all p, q. Hence, $g = \mathcal{F}[f]$ is bounded and infinitely differentiable. Moreover, since $f^{(q)} \in L_1$, $g(y)$ tends to 0 at infinity faster than $1/|y|^q$ for all q, and so, it remains to show the same for $g^{(p)}$. But this follows by the identity

$$\mathcal{F}[((-ix)^p f)^{(q)}] = (iy)^q \mathcal{F}[(-ix)^p f] = (iy)^q g^{(p)}(y).$$

and that $g^{(p)}$ is bounded for all p. Hence, it follows $\mathcal{F}[f] \in S_\infty$. $\qquad\square$

Since Schwartz functions are infinitely differentiable, Dini's condition hold and thus, the Fourier inverse transform holds and the map $f \mapsto \mathcal{F}[f]$ is a bijection.

3.3 Convolution and Heat Equation

Definition 3.3 (Convolution). Let $f_1, f_2 \in L_1$, the convolution of f_1 with f_2 is defined to be

$$f(x) = f_1 * f_2(x) := \int_{\mathbb{R}} f_1(y) f_2(x - y) \lambda(\mathrm{d}y).$$

Proposition 3.3. If $f_1, f_2 \in L_1$, then $f_1 * f_2 \in L_1$.

Proof. By Fubini's theorem, we have

$$\int |f_1 * f_2(x)| \lambda(\mathrm{d}x) \leq \int \int |f_1(y)||f_2(x - y)| \lambda(\mathrm{d}y)\lambda(\mathrm{d}x)$$

$$= \int |f_1(y)| \int |f_2(x - y)| \lambda(\mathrm{d}x)\lambda(\mathrm{d}y) < \infty.$$

$\qquad\square$

Theorem 6. If $f_1, f_2 \in L_1$, then

$$\mathcal{F}[f_1 * f_2] = \mathcal{F}[f_1]\mathcal{F}[f_2].$$

Proof. Using Fubini's theorem and the change of variable $t = x - y$, we have

$$\mathcal{F}[f_1 * f_2] = \int f_1 * f_2(x)e^{-izx}\lambda(\mathrm{d}x) = \int e^{-izx} \int f_1(y)f_2(x - y)\lambda(\mathrm{d}y)\lambda(\mathrm{d}x)$$

$$= \int f_1(y) \int f_2(x - y)e^{-izx}\lambda(\mathrm{d}x)\lambda(\mathrm{d}y) = \int f_1(y)e^{-izy} \int f_2(t)e^{-izt}\lambda(\mathrm{d}t)\lambda(\mathrm{d}y)$$

$$= \mathcal{F}[f_2] \int f_1 e^{-izy}\lambda(\mathrm{d}y) = \mathcal{F}[f_1]\mathcal{F}[f_2].$$

$\qquad\square$

We note that since the derivative under F.T. becomes multiplication, the Fourier transform is useful for solving differential equations.

Consider a linear differential equation with constant coefficients for $f(x)$

$$f^{(n)} + a_1 f^{(n-1)} + \cdots + a_{n-1}f' + a_n f = \phi(x)$$

under the Fourier transform is mapped to (denoting $\mathcal{F}[f] = g$,

$$(iy)^n g(y) + a_1 (iy)^{n-1} g(y) + \cdots + a_{n-1} iy g(y) + a_n g(y) = \mathcal{F}[\phi](y)$$

if $f, \phi \in L_1$. Thus, solving for g, one may apply the inverse Fourier transform to obtain f.

However, we already know how to solve linear differential equations with matrix exponentials, and so, Fourier transforms are more useful for solving partial differential equations. To illustrate this, let us consider the heat equation

$$D_t u = D_x^2 u$$

where $u = u(x,t)$ is a function for $x \in \mathbb{R}, t \geq 0$. The heat equation models the evolution of the temperature over time in 1-D. We are to be provided the initial temperature distribution $u_0(x) = u(x,0)$ and we are asked to find $u(x,t)$ for all $t > 0$.

Theorem 7. Suppose that $u_0, u_0', u_0'' \in L_1$, if there exist a solution u of the heat equation such that

- $u(x,t), D_x u(x,t), D_x^2 u(x,t) \in L_1$ for all $t \geq 0$;

- for all $T \geq 0$, there exists some $f_T \in L_1$ such that for all $0 \leq t \leq T$,

$$|D_t u(x.t)| \leq f_T(x),$$

then this solution is unique (I'm claiming this, is this true?) and can be found explicitly.

Proof. Applying F.T. to the right hand side of the heat equation, we obtain

$$\mathcal{F}[D_t u](y) = -y^2 \mathcal{F}_t[u](y)$$

where we denote $\mathcal{F}_t[u](y)$ the Fourier transform of $u(t, \cdot)$ for all $t \geq 0$. Then, by the second condition, the the dominated convergence theorem,

$$\mathcal{F}[D_t u](y) = \int \frac{\partial u}{\partial t} e^{-iyx} \lambda(\mathrm{d}x) = \frac{\partial}{\partial t} \int u e^{-iyx} \lambda(\mathrm{d}x) = D_t \mathcal{F}_t[u](y).$$

Hence, we have

$$D_t \mathcal{F}_t[u](y) = -y^2 \mathcal{F}_t[u](y).$$

Therefore, as $\mathcal{F}_t[u](y)$ is Lipschitz (as it is continuously differentiable by condition 1), this IVP has a unique solution. Furthermore, we may find this solution explicitly by integrating both sides resulting in

$$\mathcal{F}_t[u](y) = e^{-y^2 t} \mathcal{F}_0[u] = e^{-y^2 t} \mathcal{F}[u_0].$$

Thus, we may find u by taking the inverse Fourier transform. In particular, we have

$$e^{-y^2 t} = \mathcal{F}\left[\frac{1}{2\sqrt{\pi t}} e^{-x^2/4t}\right]$$

and hence,

$$u(x,t) = \mathcal{F}^{-1}[e^{-y^2 t} \mathcal{F}[u_0]] = \mathcal{F}^{-1}\left[\mathcal{F}\left[\frac{1}{2\sqrt{\pi t}} e^{-x^2/4t}\right] \mathcal{F}[u_0]\right]$$

$$= \mathcal{F}^{-1}\left[\mathcal{F}\left[\frac{1}{2\sqrt{\pi t}} e^{-x^2/4t} * u_0(x)\right]\right] = \frac{1}{2\sqrt{\pi t}} e^{-x^2/4t} * u_0(x).$$

By unfolding the definition,

$$u(x,t) = \frac{1}{2\sqrt{\pi t}}e^{-x^2/4t} * u_0(x) = \frac{1}{2\sqrt{\pi t}}\int e^{-y^2/4t}u_0(x-y)\lambda(\mathrm{d}y)$$

which is known as a Poisson integral. $\qquad\square$

3.4 Fourier Transform in L_2

For convenience and generality, we will consider the complex L_2 in this section. Recall that for $f \in L_1[-\pi,\pi]$, we defined the complex Fourier coefficients

$$c_n := \frac{1}{2\pi}\int_{[-\pi,\pi]} f(x)e^{-inx}\lambda(\mathrm{d}x).$$

By Hölder's inequality, we have $L_q \subseteq L_p$ for all $p \leq q$ in a finite measure space, and thus we have $L_2[-\pi,\pi] \subseteq L_1[-\pi,\pi]$. Thus, c_n is defined for any $f \in L_2[-\pi,\pi]$. Moreover, the map $f \in L_2[-\pi,\pi]$ to its Fourier coefficients is a linear operator from $L_2[-\pi,\pi]$ to ℓ_2 satisfying Parseval's identity.

The same is not true for general measure spaces and the Fourier transform is not necessarily defined for L_2 functions. Therefore, we will need to extend the definition of Fourier transforms.

Theorem 8 (Plancherel's Theorem)**.** For any $f \in L_2$, the integral

$$g_N(y) := \int_{[-N,N]} f(x)e^{-iyx}\lambda(\mathrm{d}x)$$

is L_2 for any N. Furthermore, g_N converges in L_2 to some $g \in L_2$ as $N \to \infty$. We call this g the Fourier transform of f and Parseval's identity holds, namely

$$\int_{\mathbb{R}} |g(y)|^2\lambda(\mathrm{d}y) = 2\pi \int_{\mathbb{R}} |f(x)|^2\lambda(\mathrm{d}x).$$

Thus, if $f \in L_1 \cap L_2$, then g coincides with $\mathcal{F}[f]$.

Proof. We will first prove Plancherel's theorem for Schwartz functions and then use the fact that such functions are dense in L_2 (as simple functions can be approximated by Schwartz functions) to complete the proof.

Let $f_1, f_2 \in S_\infty$ and denote g_1, g_2 their Fourier transforms (which also belong in S_∞). Then, by the inverse Fourier transform and the Fubini theorem,

$$\int_{\mathbb{R}} f_1\overline{f_2}\mathrm{d}\lambda = \int \frac{1}{2\pi}\int g_1(y)e^{iyx}\lambda(\mathrm{d}(y))\overline{g_2(y)}\lambda(\mathrm{d}(x))$$

$$= \frac{1}{2\pi}\int g_1(y)\overline{\int f_2(x)e^{-iyx}\lambda(\mathrm{d}x)}\lambda(\mathrm{d}y)$$

$$= \frac{1}{2\pi}\int g_1(y)\overline{g_2(y)}\lambda(\mathrm{d}y).$$

Hence, the Parseval's identity holds by setting $f_1 = f_2$.

Now, let $f \in L_2$ such that $f(x) = 0$ for all $x \notin [-a, a]$ for some $a > 0$. Then, $f \in L_2[-a, a]$ and as $[-a, a]$ has finite measure, $f \in L_1[-a, a]$ and so, as f is identically zero outside $[-a, a]$, $f \in L_1$. Hence, the Fourier transform is defined.

Now since S_∞ is dense, we may take a sequence (f_n) of Schwarz functions which are 0 outside $[-a, a]$ and $f_n \to f$ in L_2. Hence, by the same argument $f_n \to f$ in L_1. Therefore, $g_n := \mathcal{F}[f_n]$ converges to g uniformly. Moreover, (g_n) is Cauchy and hence convergent to g in L_2. Indeed, as $g_n - g_m \in S_\infty$, and by Parseval's identity for S_∞, we have

$$\int |g_n - g_m|^2 \mathrm{d}\lambda = 2\pi \int |f_n - f_m|^2 \mathrm{d}\lambda,$$

from which the Cauchy property follows. Parseval's identity follows similarly by taking the limit of $\|f_n\|_2^2 = \frac{1}{2\pi} \|g_n\|_2^2$ as $n \to \infty$ on both sides.

Finally, for any $f \in L_2$, define $f_N := \mathbf{1}_{[-N,N]} f$. Clearly, by dominated convergence, $\|f - f_N\|_2 \to 0$ as $N \to \infty$. Since $f_N \in L_1$, its Fourier transform exists and is given by

$$g_N(y) = \int f_N(x) e^{-iyx} \lambda(\mathrm{d}y) = \int_{[-N,N]} f(x) e^{-iyx} \lambda(\mathrm{d}x).$$

Then, as Parseval holds for functions vanishing outside compact interval, we have $\|f_N - f_M\|_2^2 = \frac{1}{2\pi} \|g_N - g_M\|_2^2$ and hence, (g_n) is Cauchy and thus, converges to some $g \in L_2$. Similar to above, we can take the limit $\|f_n\|_2^2 = \frac{1}{2\pi} \|g_n\|_2^2$ as $n \to \infty$ on both sides to obtain Parseval's identity. $\square$

Corollary 8.1. Parseval's identity implies that for any $f_1, f_2 \in L_2$,

$$\int f_1 \overline{f_2} \mathrm{d}\lambda = \frac{1}{2\pi} \int g_1 \overline{g_2} \mathrm{d}\lambda.$$

Proof. Try using the polarization identity:

$$\langle u, v \rangle = \frac{1}{4} (\|u + v\|^2 - \|u - v\|^2 + i\|u + iv\|^2 - i\|u - iv\|^2).$$

$\square$

3.5 Laplace Transform

In the case that a function is neither L_1, we may not in general use the Fourier transform directly. Instead, we would like to consider an analogue of the Fourier transform on of which is known as the Laplace transform. In particular, the idea is that some nonintegrable functions f becomes integrable by multiplying it with $e^{-\gamma x}$ for some $\gamma \in \mathbb{R}$. Namely, we take

$$g(s) := \int_{\mathbb{R}} f(x) e^{-isx} \lambda(\mathrm{d}x) = \int_{\mathbb{R}} f(x) e^{-iyx} e^{-zx} \lambda(\mathrm{d}x)$$

where $s = y - iz$.

Suppose f satisfy the Dini condition and $f(x) = 0$ for all $x < 0$ and $|f(x)| < ce^{\gamma_0 x}$ for all $x \geq 0$ for some $c, \gamma_0 > 0$, then, we say $f \in L$. We note that for $s = y + iz$, we have

$$g(s) = \int_{\mathbb{R}} f(x) e^{-isx} = \int_{[0,\infty)} f(x) e^{zx} e^{-iyx} \lambda(\mathrm{d}x)$$

21

converges uniformly for $z < -\gamma_0$ and so, g is defined on the half plane $\mathbb{R} + i(-\infty, -\gamma_0)$ and is analytic on said half plane.

For fixed $z < -\gamma_0$, we note that g is simply the Fourier transform of the function $f(x)e^{zx}$ and so, as f satisfy the Dini condition, we may apply the inversion formula, namely

$$f(x)e^{zx} = \frac{1}{2\pi} \int_{\mathbb{R}} g(s)e^{iyx}\lambda(\mathrm{d}y),$$

so (noting that the integrals are principle),

$$f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} g(s)e^{-zx+iyx}\lambda(\mathrm{d}y) = \frac{1}{2\pi} \int_{iz-\infty}^{iz+\infty} g(s)e^{isx}\lambda(\mathrm{d}s).$$

Now, changing the variable with $p = is$ and denote $\Phi(p) = g(s)$ and $w = -z$, we obtain,

$$\Phi(p) = \int_{[0,\infty)} f(x)e^{-px}\lambda(\mathrm{d}x)$$

and (where the integral is principle),

$$f(x) = \frac{1}{2\pi i} \int_{w-i\infty}^{w+i\infty} \Phi(p)e^{px}\lambda(\mathrm{d}x) \equiv \frac{1}{2\pi i} \lim_{T\to\infty} \int_{w-iT}^{w+iT} \Phi(p)e^{px}\lambda(\mathrm{d}p),$$

where $w > \gamma_0$. This motivates the definition of the Laplace transform.

Definition 3.4 (Laplace Transform). Given $f \in L$, we define the Laplace transform of f to be

$$\Phi(p) = \int_{[0,\infty)} f(x)e^{-px}\lambda(\mathrm{d}x).$$

From the above argument, we have that Φ is analytic for $\mathrm{Re}(p) > \gamma_0$.

The Laplace transform has applications to solutions of differential equations. Consider the IVP

$$y^{(n)} + a_1 y^{(n-1)} + \cdots + a_n y = b(x), a_1, \cdots, a_n \in \mathbb{C},$$

with initial condition $y(0) = y_0, y'(0) = y_1, \cdots, y^{(n-1)}(0) = y_{n-1}$. Furthermore, assume $b \in L$ and we will look for a solution y such that $y, y', \cdots, y^{(n)} \in L$.

Let $\mathcal{Y}(p), B(p)$ be the Laplace transform of y and b respectively. Then, by integration by parts, we observe

$$\int_{[0,\infty)} y'(x)e^{-px}\lambda(\mathrm{d}x) = y(x)e^{-px}|_0^\infty + p \int_{[0,\infty)} y(x)e^{-px}\lambda(\mathrm{d}x) = p\mathcal{Y}(p) - y_0.$$

By induction, we have

$$\int_{[0,\infty)} y^{(n)}(x)e^{-px}\lambda(\mathrm{d}x) = p^n\mathcal{Y}(p) - y_{n-1} - py_{n-2} - \cdots - p^{n-1}y_0.$$

Thus, after applying the Laplace transform to the differential equation, we obtain

$$Q(p) + R(p)\mathcal{Y}(p) = B(p)$$

where $R(p) = p^n + a_1 p^{n-1} + \cdots + a_n$ and $Q(p)$ is a polynomial of degree $n-1$ with coefficients in our initial conditions. Hence, we conclude,

$$\mathcal{Y}(p) = \frac{B(p) - Q(p)}{R(p)}$$

implying

$$y(x) = \frac{1}{2\pi i} \int_{w-i\infty}^{w+i\infty} \frac{B(p) - Q(p)}{R(p)} e^{ipx} \lambda(\mathrm{d}p),$$

by applying the inverse Laplace transform.

3.6 Fourier-Stieltjes Transform

Let us now return to the Fourier transforms. Recall the Stieltjes integral and we observe that, for $f \in L_1$, we may write its Fourier transform as a Lebesgue-Stieltjes integral,

$$\mathcal{F}[f](y) = \int_{\mathbb{R}} e^{-iyx} f(x) \lambda(\mathrm{d}x) = \int_{\mathbb{R}} e^{-iyx} \mathrm{d}F(x)$$

where $F(x) = \int_{(-\infty, x]} f(t) \lambda(\mathrm{d}t)$. Note that F is absolutely continuous with bounded variation on $\mathbb{R}$. On the other hand, the last integral remains to be well-defined for any function of bounded variation.

Let us first recall some definitions required of the Lebesgue-Stieltjes integral.

Definition 3.5 (Bounded Variation on an Interval)**.** A function F is said to have bounded variation on $[a, b]$ if the variation of F,

$$V_a^b F := \sup_{P \in \mathcal{P}} \sum_{i=1}^{n_P} |f(x_i) - f(x_{i-1})| < \infty,$$

where $\mathcal{P}$ is the set of all finite divisions of $[a, b]$. Namely, $P \in \mathcal{P}$ denotes the points

$$a_0 = x_0 \leq x_1 \leq \cdots \leq x_{n_P} = b.$$

We observe that, if F is non-decreasing, then the sum of any partition is telescoping, and thus, $V_a^b F = F(b) - F(a)$ for any interval $[a, b]$.

Definition 3.6 (Bounded Variation on $\mathbb{R}$)**.** A function F is said to have bounded variation on $\mathbb{R}$ if the limit

$$V_{-\infty}^{\infty} F := \lim_{\substack{a \to -\infty; \\ b \to \infty}} V_a^b F$$

exists and is finite.

Proposition 3.4. If F is a function of bounded variation on $\mathbb{R}$, then, there exists non-decreasing functions F_1, F_2 such that $F = F_1 - F_2$. Moreover, F is differentiable almost everywhere and can be written as

$$F = \phi + \psi + y$$

where ϕ is absolutely continuous (and $F' = \phi'$ a.e.), ψ is singular continuous (and $\psi' = 0$ a.e.), and y is a jump function (with $y' = 0$ a.e.).

With this in mind, we recall that the Lebesgue-Stieltjes measure can be defined for all functions of bounded variation and the Lebesgue-Stieltjes integral is simply the Lebesgue integral taken with respect to that measure.

Definition 3.7 (Lebesgue-Stieltjes Measure of an Interval)**.** Let F be a function of bounded variation on $[a,b]$ and suppose $F = F_1 - F_2$ where F_1, F_2 are non-decreasing. Then, the Lebesgue-Stieltjes measure of F on the interval $[a,b]$ is defined to be the unique measure obtained from the Caratheodory extension of the outer measure

$$w((s,t]) := F_1(t) - F_1(s) + F_2(t) - F_2(s).$$

Definition 3.8 (Fourier-Stieltjes Transform)**.** The integral

$$g(y) := \int_{\mathbb{R}} e^{-iyx} \mathrm{d}F(x)$$

where F is a function of bounded variation on $\mathbb{R}$, is called the Fourier-Stieltjes transform of F.

Lemma 3.5. The Fourier-Stieltjes transform (F.S.T.) of F, g is bounded and continuous on $\mathbb{R}$.

Proof. We have $|g(y)| \leq \int |e^{-iyx}| \mathrm{d}F(x) = \lambda_F(\mathbb{R})$. Now, since $\lambda_{V^x_{-\infty}F} \geq \lambda_F$ where $V^x_{-\infty}F$ is non-decreasing, by the continuity of measures from below,

$$\lambda_F(\mathbb{R}) = \lim_{x \to \infty} \lambda_F((-\infty, x]) \leq \lim_{x \to \infty} \lambda_{V^x_{-\infty}F}((-\infty, x]) = \lim_{x \to \infty} V^x_{-\infty}F = V^\infty_{-\infty}F.$$

Hence, $|g|$ is bounded by $V^\infty_{-\infty}F$.

Secondly, we observe

$$|g(y_1) - g(y_2)| \leq \int_{[-N,N]} |e^{-iy_1 x} - e^{-iy_2 x}| \mathrm{d}V^x_{-\infty}F + \int_{|x| \geq N} |e^{-iy_1 x} - e^{-iy_2 x}| \mathrm{d}V^x_{-\infty}F.$$

Thus, taking N sufficiently large, the second term is arbitrarily small (independent of y_1, y_2), and thus, as the first term tends to 0 at $|y_1 - y_2| \to 0$, we conclude g is continuous. $\qquad \square$

Unlike the Fourier transform, the Fourier-Stieltjes transform of a function does not necessarily vanish at $\pm\infty$. Indeed, defining $F = \mathbf{1}_{[0,\infty)}$, the Lebesgue-Stieltjes measure of F is $\lambda_F = \delta_0$ (Dirac measure at 0), and so, the F.S.T. of F is

$$g(y) = \int_{\mathbb{R}} e^{-iyx} \mathrm{d}F(x) = \int_{\mathbb{R}} e^{-iyx} \mathrm{d}\delta_0 = 1.$$

Recall for L_1 functions, we defined the convolution of two functions f_1, f_2 to be $f_1 * f_2(x) = \int_{\mathbb{R}} f_1(x-y) f_2(x) \lambda(\mathrm{d}y)$. We will define a similar definition for functions of bounded variations.

Let $F(x) = \int_{(-\infty, x]} f(t) \lambda(\mathrm{d}t)$ and $F_i(x) = \int_{(-\infty, x]} f_i(t) \lambda(\mathrm{d}t)$ for $i = 1, 2$, where $f = f_1 * f_2$, by the absolute integrability of f, f_1, f_2 and Fubini's theorem, we have

$$F(x) = \int_{(-\infty, x]} f(t) \lambda(\mathrm{d}t) = \int_{(-\infty, x]} \int_{\mathbb{R}} f_1(t-y) f_2(y) \lambda(\mathrm{d}y) \lambda(\mathrm{d}t)$$

$$= \int_{\mathbb{R}} \left(\int_{(-\infty, x]} f_1(t-y) \lambda(\mathrm{d}t) \right) f_2(y) \lambda(\mathrm{d}y) = \int_{\mathbb{R}} F_1(x-y) \mathrm{d}F_2(y).$$

This motivates for the definition of the convolution of two functions of bounded variation.

Definition 3.9 (Convolution of Functions of Bounded Variation). Given F_1, F_2 are functions of bounded variation, we define the convolution of F_1 with F_2 to be

$$F_1 * F_2(x) = \int_{\mathbb{R}} F_1(x - y) \mathrm{d}F_2(y).$$

Proposition 3.5. If F_1, F_2 are functions of bounded variation on $\mathbb{R}$, then so is $F_1 * F_2$ a function of bounded variation on $\mathbb{R}$.

Proof. Consider

$$|F(x_1) - F(x_2)| \leq \int_{\mathbb{R}} |F_1(x_1 - y) - F_1(x_2 - y)| \mathrm{d}V_{-\infty}^y F_2.$$

So,

$$\sum_i |F(x_i) - F(x_{i-1})| \leq \sum_i \int_{\mathbb{R}} |F_1(x_i - y) - F_1(x_{i-1} - y)| \mathrm{d}V_{-\infty}^y F_2$$

$$= \int_{\mathbb{R}} \left(\sum_i |F_1(x_i - y) - F_1(x_{i-1} - y)| \right) \mathrm{d}V_{-\infty}^y F_2$$

$$\leq \int_{\mathbb{R}} V_{-\infty}^\infty F_1 \mathrm{d}V_{-\infty}^y F_2 = V_{-\infty}^\infty F_1 V_{-\infty}^\infty F_2.$$

Hence, $V_{-\infty}^\infty F \leq V_{-\infty}^\infty F_1 V_{-\infty}^\infty F_2 < \infty$ implying F has bounded variation as required. $\square$

Theorem 9. Let F be the convolution of the function F_1, F_2 of bounded variations on $\mathbb{R}$ and let g, g_1, g_2 be their Fourier-Stieltjes transforms respectively. Then, $g = g_1 \cdot g_2$.

Proof. Let $F = F_1 * F_2$ and consider the division

$$a = x_0 \leq x_1 \leq \cdots \leq x_n = b$$

of the interval $[a, b]$. Then, for any y, as e^{-iyx} is continuous, the Lebesgue-Stieltjes integral of e^{-iyx} coincides with the Riemann-Stieltjes integral. In particular,

$$\int_{[a,b]} e^{-iyx} \mathrm{d}F(x) = \lim_{\max \Delta x_k \to 0} \sum_{k=1}^n e^{iyx_k} (F(x_k) - F(x_{k-1}))$$

$$= \lim_{\max \Delta x_k \to 0} \int_{\mathbb{R}} \sum_{k=1}^n e^{iy(x_k - t)} (F(x_k - t) - F(x_{k-1} - t)) e^{-iyt} \mathrm{d}F_2(t)$$

$$= \int_{\mathbb{R}} \left(\int_{a-y}^{b-y} e^{-iyx} \mathrm{d}F_1(x) \right) e^{-iyt} \mathrm{d}F_2(t)$$

$$= \left(\int_{a-y}^{b-y} e^{-iyx} \mathrm{d}F_1(x) \right) \left(\int_{\mathbb{R}} e^{-iyt} \mathrm{d}F_2(t) \right)$$

where the third equality holds by dominated convergence. Hence, taking the limit $a \to \infty, b \to -\infty$, we obtain

$$g(y) = \int_{\mathbb{R}} e^{-iyx} \mathrm{d}F(x) = \left(\int_{\mathbb{R}} e^{-iyx} \mathrm{d}F_1(x) \right) \left(\int_{\mathbb{R}} e^{-iyt} \mathrm{d}F_2(t) \right) = g_1(y)g_2(y)$$

as required. $\qquad\square$

This fact is applicable in probability theory and especially useful when dealing with characteristic functions. In particular, if ξ, η are two independent random variables and F_1, F_2 are their distributions, i.e. $F_1(x) = \xi_* \mathbb{P}((-\infty, x])$ and $F_2(x) = \eta_* \mathbb{P}((-\infty, x])$. Then, $\xi + \eta$ has distribution $F_1 * F_2$ (see probability theory).

The characteristic function of a random variable X is defined to be the F.S.T. of its distribution. Namely, if X has distribution F, then its characteristic function is defined to be

$$g(y) := \int_{\mathbb{R}} e^{-iyx} \mathrm{d}F(x).$$

Therefore, a corollary of the above theorem is that the characteristic function of $\xi + \eta$ is simply the product of the characteristic functions of the two.

From this, one may obtain the distribution of $\xi + \eta$ by applying the inverse Fourier-Stieltjes transform which formula is on the exercise sheet.

4 Linear Functional on Normed Space

4.1 Review

We will first review some definitions and results from functional analysis.

Definition 4.1 (Functional). Let L be a linear space (not necessarily topological) over $\mathbb{C}$ (or $\mathbb{R}$), a linear functional is simply a function $f : L \to \mathbb{C}$ (or $\mathbb{R}$).

The functional is said to be linear if for all $x, y \in L$, $\alpha \in \mathbb{C}$ (or $\mathbb{R}$),

$$f(x + y) = f(x) + f(y); f(\alpha x) = \alpha f(x).$$

If a linear space has a compatible topology, namely addition and scalar multiplication are continuous, then we have a topological linear space and we may talk about notions of continuity of the linear functionals. We will in this course restrict to the case where the norm is generated by a norm on the linear space.

Definition 4.2. A functional f on a normed linear space E is continuous if for all $x_0 \in E$, $\epsilon > 0$, there exists an open neighbourhood U of x_0 such that for all $x \in U$,

$$|f(x) - f(x_0)| < \epsilon.$$

Proposition 4.1. A linear functional on a finite dimensional normed linear space is continuous.

Proposition 4.2. A linear functional is continuous if and only if it is continuous at 0.

Proposition 4.3. A linear functional is continuous if and only if there is a neighbourhood of 0 for which it is bounded.

Corollary 9.1. A linear function is continuous if and only if it is bounded on the unit ball.

With the above propositions in mind, we see that if f is continuous, $\sup_{\|x\| \leq 1} |f(x)| < \infty$ and so, we may define a norm on the space of linear functionals by taking

$$\|f\| := \sup_{\|x\| \leq 1} |f(x)|.$$

Equivalently, by considering for all $\|x\| \neq 0$, $\|x/\|x\|\| = 1$ and $f(x/\|x\|) = f(x)/\|x\|$, we have

$$\|f\| = \sup_{x \neq 0} \frac{|f(x)|}{\|x\|}.$$

With this, we have the inequality $|f(x)| \leq \|f\|\|x\|$ for all $x \in E$.

Proposition 4.4. If E is an inner product space, then for all $a \in E$, $a \neq 0$, defining the functional $f : E \to \mathbb{C}$ by $f(x) := \langle x, a \rangle$, f is a continuous linear functional with norm $\|a\|$.

Definition 4.3. Let L be a linear space and let p be a functional on L. Then, p is said to be convex (or a sub-linear functional) if $p(x + y) \leq p(x) + p(y)$, and $p(\alpha x) = |\alpha| p(x)$ for all $x, y \in L, \alpha \in \mathbb{C}$.

Theorem 10 (Hahn-Banach)**.** Let L be a linear space and let p be a sub-linear functional on L. Then, if f_0 is a linear functional on $M \leq L$ where M is a subspace of L such that $|f_0(x)| \leq p(x)$ for all $x \in M$. Then, there exists a linear functional f on L such that $f|_M = f_0$ and $|f(x)| \leq p(x)$ for all $x \in L$.

It is clear that the norm is a sub-linear functional on a normed linear space and so, we may apply Hahn-Banach to this special case.

Corollary 10.1. Let E be a normed linear space and let f_0 be a continuous linear functional on the subspace $E_0 \subseteq E$, then there exists a continuous linear functional f on E such that

- $f(x) = f_0(x)$ for all $x \in E_0$;

- $\|f\| = \|f_0\|$.

Proof. Let $c := \|f_0\|$. Then, defining $p(x) := c\|x\|$, by Hahn-Banach we have the extension of f_0 such that $f|_{E_0} = f_0$ and $|f(x)| \leq c\|x\|$ and so,

$$\|f\| = \sup_{x \neq 0} \frac{|f(x)|}{\|x\|} \leq c.$$

Yet, $\|f\| \geq \|f_0\| = c$, we have $\|f\| = \|f_0\|$ as required. $\qquad\square$

Definition 4.4. Given a linear space L, we define its adjoint (or dual) L^* as the space of all linear functionals on L equipped with point-wise addition and scalar multiplication.

It is easy to check that, if E is a normed linear space, then E^* is also a normed linear space. The topology induced from this norm is called the strong topology on E^*.

Theorem 11. The dual space $(E^*, \|\cdot\|)$ is complete.

Proof. Let $(f_n) \subseteq E^*$ be Cauchy. Then, f_n converges point-wise as $\mathbb{C}$ (or $\mathbb{R}$) is complete. Thus, defining $f(x) := \lim_{n \to \infty} f_n(x)$, it remains to show $f \in E^*$.

By the algebra of limits, f is linear, and so, it suffices to show it is bounded. Indeed, as (f_n) is Cauchy, there exists some N such that for all $n, m \geq N$, $\|f_n - f_m\| < 1$. Thus, for all $\|x\| = 1$,
$$|f_n(x) - f_m(x)| \leq \|f_n - f_m\| < 1.$$

Now, by taking $m \to \infty$, we have $|f_n(x) - f(x)| \leq 1$ implying $\|f\| \leq \|f_n\| + 1 < \infty$. Hence, $f \in E^*$ and E^* is complete as required. $\qquad\square$

We note that the dual is complete irrespective of the completeness of E.

Proposition 4.5. If E is not complete, denoting $\overline{E}$ the completion of E, E^* and $\overline{E}^*$ are isometric (isomorphic in the linear space sense and preserves the norm).

Proof. Since $E \subseteq \overline{E}$ is everywhere dense in $\overline{E}$, and $f \in E^*$ can be extended uniquely to a continuous functional $\overline{f}$ in $\overline{E}^*$ and $\|f\| = \|\overline{f}\|$. Thus, by considering $f \mapsto \overline{f}$ and its inverse $\overline{f} \mapsto \overline{f}|_E$, we have our isometry. $\qquad\square$