Mathematical Analysis: Introductory Functional Analysis.

Edward Chernysh[1]

Tuesday 8th August, 2017

[1]edward.chernysh at mail.mcgill.ca

Contents

1 Hilbert Spaces

We strive to provide a unified theory of Hilbert spaces from an analytic perspective. Hilbert spaces arise in many natural contexts, both in physics and in mathematics. A Hilbert space may be approached in one of two ways: as a vector space over the complex numbers $\mathbb{C}$ or over $\mathbb{R}$. In our case, we begin this chapter with a rigorous definition of Hilbert spaces over a field $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$.

We shall study the properties of general Hilbert spaces and construct a Hilbert space that is of particular interest. Afterwards, we shall introduce linear transformations, compact operators and isomorphisms of Hilbert spaces. At the end of this chapter, we prove the Spectral Theorem for Compact Operators.

1.1 Definition and Orthogonality

Definition. Let $\mathbb{K}$ be $\mathbb{R}$ or $\mathbb{C}$, and $\mathcal{H}$ be an inner-product space over $\mathbb{K}$. We shall say that $\mathcal{H}$ is a **Hilbert space** if and only if it is a separable Banach space with respect to the norm induced by the inner product $\langle \cdot, * \rangle : \mathcal{H} \times \mathcal{H} \to \mathbb{K}$.

Of course, this definition resembles that of a vector space since $\mathcal{H}$ is itself a vector space. The difference is mainly the point of view that we adopt. In the study of linear algebra, we look instead towards the representation properties of the vector space. Instead, we shall be mostly concerned with the *limiting properties* the space carries.

It should be evident from the very definition of the inner-product that $\langle \cdot, * \rangle$ is anti-linear in the second argument. Recall that a metric space (X, d) is separable provided it has a countable dense subset. Given two vectors $f, g \in \mathcal{H}$ we shall say that f and g are **orthogonal** (or perpendicular), written $f \perp g$, whenever $\langle f, g \rangle = 0$. Our first result is a generalization of the Pythagorean theorem to these spaces:

Lemma 1.1. *Let $\mathcal{H}$ be a complex Hilbert space and $f, g \in \mathcal{H}$ with $f \perp g$. Then $\|f + g\|^2 = \|f\|^2 + \|g\|^2$.*

PROOF. To see this we write:

$$\begin{aligned} \|f+g\|^2 = \langle f+g, f+g \rangle &= \langle f, f+g \rangle + \langle g, f+g \rangle \\ &= \langle f, f \rangle + \langle f, g \rangle + \langle g, f \rangle + \langle g, g \rangle \\ &= \|f\|^2 + \langle f, g \rangle + \langle g, f \rangle + \|g\|^2 \end{aligned}$$

Now, since $f \perp g$ we get $\langle f, g \rangle = \overline{\langle g, f \rangle} = 0$ which concludes the proof. □

Corollary 1.2 (Pythagorean Theorem). *Let $\mathcal{H}$ be a complex Hilbert space and assume $\{f_j\}_{j=1}^N$ is a family of pairwise orthogonal vectors in $\mathcal{H}$. Then,*

$$\left\|\sum\nolimits_{j=1}^N f_j\right\|^2 = \sum_{j=1}^N \|f_j\|^2 \tag{1}$$

Proof. We argue by induction on N. The case $N = 2$ is clear from the previous lemma, now assume (1) holds for N, we show the case $N+1$ follows. Certainly, if $\{f_j\}_{j=1}^{N+1}$ is pairwise orthogonal then so is $\{f_j\}_{j=1}^N$. Moreover, it is obvious that $\langle \sum_{j=1}^N f_j, f_{N+1}\rangle = 0$ by linearity. Whence we find:

$$\begin{aligned}
\left\|\sum\nolimits_{j=1}^{N+1} f_j\right\|^2 &= \left\langle \sum\nolimits_{j=1}^{N+1} f_j, \sum\nolimits_{j=1}^{N+1} f_j \right\rangle \\
&= \left\langle \sum\nolimits_{j=1}^{N} f_j, \sum\nolimits_{j=1}^{N+1} f_j \right\rangle + \left\langle f_{N+1}, \sum\nolimits_{j=1}^{N+1} f_j \right\rangle \\
&= \left\langle \sum\nolimits_{j=1}^{N} f_j, \sum\nolimits_{j=1}^{N} f_j \right\rangle + \left\langle \sum_{j=1}^{N} f_j, f_{N+1} \right\rangle + \left\langle f_{N+1}, \sum\nolimits_{j=1}^{N+1} f_j \right\rangle \\
&= \left\langle \sum\nolimits_{j=1}^{N} f_j, \sum\nolimits_{j=1}^{N} f_j \right\rangle + \left\langle f_{N+1}, \sum\nolimits_{j=1}^{N+1} f_j \right\rangle \\
&= \left\|\sum\nolimits_{j=1}^{N} f_j\right\|^2 + \langle f_{N+1}, f_{N+1}\rangle + \left\langle f_{N+1}, \sum\nolimits_{j=1}^{N} f_j \right\rangle \\
&= \left\|\sum\nolimits_{j=1}^{N} f_j\right\|^2 + \|f_{N+1}\|^2
\end{aligned}$$

as was to be shown. □

Let $\mathcal{H}$ be a complex Hilbert space and $\{e_k\}_{k\in\mathbb{N}}$ be a countable subset of $\mathcal{H}$. This set is said to be **orthonormal** provided for all indices $(k,\ell) \in \mathbb{N}^2$ one has:

$$\langle e_k, e_\ell\rangle = \begin{cases} 1 & k = \ell \\ 0 & \text{else} \end{cases}$$

Note that we may indeed assume the above set to be countable since $\mathcal{H}$ is separable. This same set is called a **Hilbert basis** for $\mathcal{H}$ if their linear combinations are dense in $\mathcal{H}$. These are sometimes called orthonormal bases. We shall now give a complete characterization of these bases for a Hilbert space $\mathcal{H}$ (over $\mathbb{C}$). First, we introduce notation. In the next theorem we shall write $\xi_k := \langle f, e_k\rangle \in \mathbb{C}$ and set $S_N(f) := \sum_{k=1}^N \xi_k e_k$ for $f \in \mathcal{H}$.

Theorem 1.3 (Characterization of Hilbert Bases). *Let $\mathcal{H}$ be a complex Hilbert* *space and $\{e_k\}_{k\in\mathbb{N}}$ an orthonormal subset of $\mathcal{H}$. The following statements are equivalent:*

(i) $\{e_k\}_{k\in\mathbb{N}}$ *is a Hilbert basis for* $\mathcal{H}$.

(ii) If $f \in \mathcal{H}$ *satisfies* $\langle f, e_j\rangle = 0$ *for all* $j \in \mathbb{N}$ *then* $f = 0$.

(iii) For all $f \in \mathcal{H}$ *the combination* $S_N(f) \to f$ *as* $N \to \infty$.

(iv) Parseval's identity holds true for all $f \in \mathcal{H}$*:*

$$\|f\|^2 = \sum_{k\in\mathbb{N}} |\xi_k|^2 \tag{$\mathfrak{P}$}$$

PROOF OF THEOREM. (i $\implies$ ii) Assume that $\langle f, e_j\rangle = 0$ for all j. Let ε positive be given; select a vector $g_\varepsilon \in \mathcal{H}$ where $g_\varepsilon = \sum_{k=1}^N \zeta_k e_k$ with $\|f - g_\varepsilon\| \leq \varepsilon$. Observe that by assumption on f one has by linearity of the Hermitian inner-product: $\langle f, g_\varepsilon\rangle = 0$. Therefore,

$$\|f - g_\varepsilon\|^2 = \langle f, f, \rangle - \underbrace{\langle f, g_\varepsilon\rangle}_{=0} - \underbrace{\langle g_\varepsilon, f\rangle}_{=0} + \langle g_\varepsilon, g_\varepsilon\rangle = \|f\|^2 + \|g_\varepsilon\|^2 \leq \varepsilon^2$$

It follows that $\|f\|^2 \leq \varepsilon^2$ and since $\varepsilon > 0$ was arbitrary we find that $\|f\| = 0$.

(i $\implies$ ii) Assume that $\langle f, e_j\rangle = 0$ for all j. Let ε positive be given; select a vector $g_\varepsilon \in \mathcal{H}$ where $g_\varepsilon = \sum_{k=1}^N \zeta_k e_k$ with $\|f - g_\varepsilon\| \leq \varepsilon$. Observe that by assumption on f one has by linearity of the Hermitian inner-product: $\langle f, g_\varepsilon\rangle = 0$. Therefore,

$$\|f\|^2 = \langle f, f, \rangle = \langle f, f - g_\varepsilon + g_\varepsilon\rangle = \underbrace{\langle f, g_\varepsilon\rangle}_{=0} + \langle f, f - g_\varepsilon\rangle = \langle f, f - g_\varepsilon\rangle$$

Now, using Cauchy-Schwarz we find that $\|f\|^2 \leq \|f\|\,\|f - g_\varepsilon\|$. Suppose now that $\|f\| \neq 0$, then we get $\|f\| \leq \|f - g_\varepsilon\| \leq \varepsilon$. Since $\varepsilon > 0$ was arbitrary we find that $\|f\| = 0$.

(ii $\implies$ iii) There are some preliminary "calculations" to be made. Fix a vector $f \in \mathcal{H}$ and define $S_N(f)$ as in (iii). We claim first that $f - S_N(f) \perp S_N(f)$ for all N sufficiently large. Indeed, to see this we write by the Pythagorean theorem (the $\{e_k\}_{k\in\mathbb{N}}$ are orthonormal)

$$\langle f - S_N(f), S_N(f)\rangle = \langle f, S_N(f)\rangle - \|S_N(f)\|^2 = \langle f, S_N(f)\rangle - \sum_{k=1}^N |\xi_k|^2$$

But now

$$\langle f, S_N(f)\rangle = \left\langle f, \sum_{k=1}^N \xi_k e_K \right\rangle = \sum_{k=1}^N \overline{\xi_k}\langle f, e_k\rangle = \sum_{k=1}^N |\xi_k|^2 ,$$

which proves that $f - S_N(f) \perp S_N(f)$ as was asserted. Therefore, applying the Pythagorean theorem proven in Corollary 1 we obtain that[1]

$$\|f\|^2 = \|f - S_N(f)\|^2 + \sum_{k=1}^N |\xi_k|^2 \geq \sum_{k=1}^N |\xi_k|^2 . \tag{2}$$

Now, passing to the limit in $N \to \infty$ in the righthand side of the equation above gives *Bessel's Identity*:

$$\sum_{k\in\mathbb{N}} |\xi_k|^2 \leq \|f\|^2 . \tag{3}$$

We now make the bold claim that $\{S_N(f)\}_{N\in\mathbb{N}}$ is Cauchy in $\mathcal{H}$. Certainly, from (3) we know that $\sum_k |\xi_k|^2 < \infty$ and so for all $N, M \in \mathbb{N}$, taking without harm $M > N$:

$$\|S_N(f) - S_M(f)\| \leq \sum_{k=N+1}^M |\xi_k|^2 \xrightarrow{N,M\to\infty} 0.$$

Since $\mathcal{H}$ is also a Banach space, there is a point, say, $g \in \mathcal{H}$ so that $S_N(f) \to g$ in norm as $N \to \infty$. We claim now that $f = g$. Indeed, it suffices by our assumption in (ii) to prove that $\langle f - g, e_j\rangle = 0$ for arbitrary j. Fix j and let $N \gg j$ be an integer. We note that, for N large,

$$\langle f - S_N(f), e_j\rangle = \langle f, e_j\rangle - \left\langle \sum_{k=1}^N \xi_k e_k, e_j \right\rangle = \langle f, e_j\rangle - \langle f, e_j\rangle = 0.$$

Which implies the desired result. Indeed, we have

$$\begin{aligned} |\langle f - g, e_j\rangle| &\leq |\langle f - S_N(f), e_j\rangle| + |\langle S_N(f) - g, e_j\rangle| \\ &\leq |\langle f - S_N(f), e_j\rangle| + \|S_N - g\| . \end{aligned}$$

Thus, passing to the limit in $|\langle f - g, e_j\rangle| \leq \|S_N - g\|$ we find that $f - g \perp e_j$ for all j whence $f = g$ as vectors.

(iii $\implies$ iv) We refer again to (2). As per our assumption we know that $S_N(f) \to f$ in norm as $N \to \infty$. Thus, taking the limit in (2) we get Parseval's identity in ($\mathfrak{P}$).

(iv $\implies$ i). To see this, we assume that Parseval's identity holds true. Now,

[1] Keep this equation in mind as it will be a central component of the remainder of the proof.

referring to (2) we find that $\|f - S_N\| \to 0$ as $N \to \infty$. Since S_N are linear combinations we have (i).

The theorem has now been proven.

□

We conclude this section with the observation that any Hilbert space $\mathcal{H}$ over $\mathbb{C}$ has a Hilbert basis. Indeed, since $\mathcal{H}$ is a vector space we may select a basis, say, $\mathcal{B}$. Now, to construct an orthonormal subset one needs only follow Gram-Schmidt.

1.2 Minimizers and Orthogonal Subspaces

The previous section was more algebraic than analytic. We defined the notion of a Hilbert space over a particular field $\mathbb{K}$ and developed the concept of a *Hilbert basis* for $\mathcal{H}$. We concluded this section by arguing for the existence of a basis (orthonormal) for this space $\mathcal{H}$.

Instead, this section is devoted to the *decomposition theory of Hilbert spaces*. For this section we fix a Hilbert space $\mathcal{H}$ over a field $\mathbb{K}$. In the like of the theory of vector spaces, we also wish to study linear operators and linear transformations between Hilbert spaces. For this reason, the following definition arises naturally:

Definition (Linear Transformation). Let $\mathcal{H}_1, \mathcal{H}_2$ be Hilbert spaces over $\mathbb{K}$. We say a mapping $\mathrm{L} : \mathcal{H}_1 \to \mathcal{H}_2$ is a **linear operator** provided:

$$\mathrm{L}\,(f + \alpha g) = \mathrm{L}(f) + \alpha \mathrm{L}(g), \quad \forall \alpha \in \mathbb{K},\ f, g \in \mathcal{H}_1$$

There are some remarks to be made about this definition. Primarily, we observe that any linear operator L fixes the zero vector. Indeed, to see this it suffices to write

$$\mathrm{L}(0) = \mathrm{L}(0 + 0) = \mathrm{L}(0) + \mathrm{L}(0)$$

In other-terms, a linear operator $\mathrm{L} : \mathcal{H}_1 \to \mathcal{H}_2$ is merely a homomorphism of vector spaces, where both vector spaces are Hilbert spaces. Naturally, we wish to extend this notion of a "homomorphism" to a "isomorphism".[2] This leads us unto the following definition:

Definition. A mapping $U : \mathcal{H}_1 \to \mathcal{H}_2$ between Hilbert spaces is said to be unitary if it is a linear operator, bijective and for all $f \in \mathcal{H}_1$ one has

$$\|U(f)\|_{\mathcal{H}_2} = \|f\|_{\mathcal{H}_1}$$

[2]One of our goals is to take linear algebra and reformulate it in analytic terms.

Namely, in the above we require that unitary mappings be *norm preserving*. These are the analytic analogues for isomorphisms of vector spaces. Note that much of what holds for linear transformations between vector spaces must also hold for linear operators. For instance, let $\mathrm{L} : \mathcal{H}_1 \to \mathcal{H}_2$ be a linear operator between vector spaces over a field $\mathbb{K}$. Then, $\ker \mathrm{L} = \{0\}$ if and only if L is injective. Certainly, if L is injective, since $0 \mapsto 0$ we find $\ker \mathrm{L} = \{0\}$. Conversely suppose that $\ker \mathrm{L} = \{0\}$. If $\mathrm{L}(f) = \mathrm{L}(g)$ then $\mathrm{L}(f - g) = 0$ so that $f - g \in \ker \mathrm{L}$ whence $f = g$.

In the likes of vector spaces, we consider subspaces of Hilbert spaces $\mathcal{H}$. Indeed, if $\mathcal{H}$ is a Hilbert space over $\mathbb{K}$ and $\mathcal{S} \subseteq \mathcal{H}$ is a vector subspace of $\mathcal{H}$, we shall write $\mathcal{S} < \mathcal{H}$ to say that $\mathcal{S}$ is a vector subspace of $\mathcal{H}$. Yet again, in the hopes of preserving properties of $\mathcal{H}$ we are led to giving the following definition:

Definition. If $\mathcal{H}$ is a Hilbert space over $\mathbb{K}$ and $\mathcal{S} < \mathcal{H}$, we shall say $\mathcal{S}$ is a **closed subspace** of $\mathcal{H}$ provided $\mathcal{S}$ is **topologically closed**[3] in $\mathcal{H}$. In this case, we shall write $\mathcal{S} \leq \mathcal{H}$.

Some theorems are now in order. It will turn out that the above definition gives us the necessary language to discuss the most powerful theorems of Hilbert space theory. We now give the following identity:

Proposition 1.4. *If $\mathcal{H}$ is a Hilbert space over $\mathbb{K}$ we have the parallelogram law:*

$$\|f + g\|^2 + \|f - g\|^2 = 2\left(\|f\|^2 + \|g\|^2\right) \tag{4}$$

PROOF. Let us now proceed by direct calculation. Write:

$$\begin{aligned} \|f + g\|^2 + \|f - g\|^2 &= \langle f + g, f + g\rangle + \langle f - g, f - g\rangle \\ &= \langle f, f\rangle + \langle g, f\rangle + \langle f, g\rangle + \langle g, g\rangle + \langle f, f\rangle \\ &\quad - \langle f, g\rangle - \langle g, f\rangle + \langle g, g\rangle \end{aligned}$$

which concludes this proof. □

Theorem 1.5 (Existence of Minimizers). *Let $\mathcal{H}$ be a Hilbert space over a field $\mathbb{K}$ and $\mathcal{S} \leq \mathcal{H}$. For each $f \in \mathcal{H}$ there exists $g_0 \in \mathcal{S}$ such that*

$$\|f - g_0\| = \inf_{g \in \mathcal{S}} \|f - g\|$$

and $(f - g_0) \perp \mathcal{S}$.[4]

[3] By topologicaly closed we mean closed in the topology of $\mathcal{H}$. Namely $\mathcal{S}$ is a closed subspace of $\mathcal{H}$ if and only if it contains all of its limit points in $\mathcal{H}$.

[4] By this we mean that $\langle f - g_0, g\rangle = 0$ for all vectors $g \in \mathcal{S}$.

PROOF. Let $d = \inf_{g \in \mathcal{S}} \|f - g\|$. We construct a sequence $\{g_n\}_{n \in \mathbb{N}} \subseteq \mathcal{S}$ such that $\|f - g_n\| \to d$ as n tends to infinity. First we want to show that the sequence converges to $g_0 \in \mathcal{S}$. To do show, we show that the sequence is Cauchy. Thence the completeness of $\mathcal{H}$ implies that the sequence converges and the some g_0 which is in $\mathcal{S}$ as it is a topologically closed subspace. Indeed, by the parallelogram law we have

$$\begin{aligned}
\|(f - g_n) + (f - g_m)\|^2 &+ \|(f - g_n) - (f - g_m)\|^2 \\
&= \|2f - (g_n + g_m)\|^2 + \|g_n - g_m\|^2 \\
&= 2\left(\|f - g_n\|^2 + \|f - g_m\|^2\right).
\end{aligned}$$

Rearranging the terms, we have

$$\begin{aligned}
\|g_n - g_m\|^2 &= 2\left(\|f - g_n\|^2 + \|f - g_m\|^2\right) - \|2f - (g_n + g_m)\|^2 \\
&= 2\left(\|f - g_n\|^2 + \|f - g_m\|^2\right) - 4\left\|f - \frac{g_n + g_m}{2}\right\|^2 \\
&\leq 2\left(\|f - g_n\|^2 + \|f - g_m\|^2\right) - 4d^2
\end{aligned}$$

By construction, we have that $\|f - g_n\|$ and $\|f - g_m\|$ tend to d, hence the right hand side tends to 0. We can thereby conclude that the sequence is indeed Cauchy.

We have therefore shown that there exists some $g_0 \in \mathcal{S}$ such that $\|f - g_0\| = \inf_{g \in \mathcal{S}} \|f - g\|$. We now conclude the proof by proving that $f - g_0 \perp \mathcal{S}$. Let $g \in \mathcal{S}$, we show that $\Re(\langle f - g_0, g \rangle) = 0$, the imaginary part can be done similarly. For any $\varepsilon \in \mathbb{R}$, we have

$$\|f - (g_0 - \varepsilon g)\|^2 \geq \|f - g_0\|^2$$

Then after expanding the left hand side, we obtain

$$2\varepsilon \Re(\langle f - g_0, g \rangle) + \varepsilon^2 \|g\|^2 \geq 0$$

Suppose $\Re(\langle f - g_0, g \rangle) > 0$, then we can find $\varepsilon < 0$ such that the above equation is negative. Similarly, if $\Re(\langle f - g_0, g \rangle) < 0$, then we can find $\varepsilon > 0$ such that the above equation is negative. We can therefore conclude that $\Re(\langle f - g_0, g \rangle) = 0$. For the imaginary part, taking $\imath\varepsilon$ instead of ε yields the desired result. □

Perhaps the most important result of the above result is that we may decompose a Hilbert space $\mathcal{H}$ into closed subspaces. Certainly, let $\mathcal{H}$ be a Hilbert space over a field $\mathbb{K}$ and suppose $\mathcal{S}$ is a closed subspace of $\mathcal{H}$. We shall now define:

$$\mathcal{S}^{\perp} := \{f \in \mathcal{H} : \langle f, g \rangle = 0, \forall g \in \mathcal{S}\}$$

to be the orthogonal component in $\mathcal{H}$ to $\mathcal{S}$. We now claim that $\mathcal{S}^{\perp}$ is a closed subspace of $\mathcal{H}$ whenever $\mathcal{S}$ is. Certainly, let (f_n) be a subsequence in $\mathcal{S}^{\perp}$ converging to $f \in \mathcal{H}$. We must show $f \perp g$ for all $g \in \mathcal{S}$. Indeed, for all n large consider:

$$\langle f_n, g \rangle = 0$$

therefore, in the limit we find that for each n:

$$|\langle f, g \rangle| = |\langle f, g \rangle - \langle f_n, g \rangle| = \langle f - f_n, g \rangle \leq \|g\| \, \|f - f_n\|$$

which of course tends to 0 as $n \to \infty$. Therefore, we have found that $\mathcal{S}^{\perp} \leq \mathcal{H}$. The truly powerful result is the following:

Theorem 1.6. *Let $\mathcal{S} \leq \mathcal{H}$, where $\mathcal{H}$ is a Hilbert space over $\mathbb{K}$. Then, $\mathcal{H} = \mathcal{S} \oplus \mathcal{S}^{\perp}$.*

PROOF. Let $f \in \mathcal{H}$ be given. By Theorem 1.5 there exists some $g \in \mathcal{S}$ such that $(f - g) \perp \mathcal{S}$. Therefore, we may of course write out $f = (f - g) + g$, which is of the form $\mathcal{S}^{\perp} + \mathcal{S}$. Therefore, each vector in $\mathcal{H}$ has a representation as the sum of elements in $\mathcal{S}$ and $S^{\perp}$. Now, it remains only to prove that this representation is unique.

Suppose that $g_0 + h_0 = g_1 + h_1$ where $g_j \in \mathcal{S}$ and $h_j \in \mathcal{S}^{\perp}$ for $j = 0, 1$. It follows that $(g_0 - g_1) = (h_1 - h_0)$. The proof will now be complete if we can show that $\mathcal{S} \cap \mathcal{S}^{\perp} = \{0\}$ for then $g_0 - g_1 = 0$ and $h_1 - h_0 = 0$. To see this, note that as vector subspaces we immediately have $\mathcal{S} \cap \mathcal{S}^{\perp} \supseteq \{0\}$ since 0 is an element of each and every vector space. To see the reverse inclusion, let $f \in \mathcal{S} \cap \mathcal{S}^{\perp}$. Then, since $f \in \mathcal{S}^{\perp}$ and $f \in \mathcal{S}$ one has trivially

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

whence $f \perp f$ so that $f = 0$. This proves the theorem. We then know $\mathcal{H} = \mathcal{S} \oplus \mathcal{S}^{\perp}$. □

We shall later see that $\mathcal{S}$ is also a Hilbert space in its own right, provided it is a closed subspace of $\mathcal{H}$. Completeness is more or less immediate, but its separability is an issue. Of course, if (f_n) is a Cauchy sequence in $\mathcal{S}$ it is again Cauchy in $\mathcal{H}$, where it must converge to $f \in \mathcal{H}$.[5] Since $\mathcal{S}$ is topologically closed it follows that $f \in \mathcal{S}$ and hence $\mathcal{S}$ is complete.

[5] Since $\mathcal{H}$ is complete.

The issue with separability lies in the following. Let $\mathcal{H}$ be a Hilbert space over a field $\mathbb{K}$ and suppose $\mathcal{S} \leq \mathcal{H}$ is a closed Hilbert subspace of $\mathcal{H}$. However, if $\mathcal{Q}$ is a countable dense subset of $\mathcal{H}$ there is no-guarantee that $\mathcal{Q} \cap \mathcal{S} \neq \varnothing$. We shall return to this problem once we have defined the notion of a projection mapping onto a closed subspace $\mathcal{S} \leq \mathcal{H}$.

1.3 Linear Operators and Functionals

Although we had introduced linear operators in the previous section, in this section we shall distinguish them from the linear transformations you have all seen in linear algebra. In this section we look at these operators from a perspective that is truly analytic.

We shall say that a linear operator $\mathrm{L} : \mathcal{H}_1 \to \mathcal{H}_2$ is bounded provided there exists some $M > 0$, independent of $f \in \mathcal{H}_1$, such that $\|\mathrm{L}(f)\|_{\mathcal{H}_2} \leq M \|f\|_{\mathcal{H}_1}$. In the future we shall omit these subscripts on the norm, when the context and space cannot cause confusion.

We shall call a linear operator $\mathrm{L} : \mathcal{H}_1 \to \mathcal{H}_2$ continuous provided it satisfies the *sequential criterion for continuity*[6]: for each sequence in $\mathcal{H}$ $(f_n) \to f$, with $f \in \mathcal{H}$ one has

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

Note that these definitions did not use the properties of $\mathcal{H}_1$ or $\mathcal{H}_2$ that distinguish it from traditional normed vector spaces, say, Banach spaces.Thence we see that these definitions remain equally valid over Banach spaces. The more surprising result is that these concepts are equivalent.

Theorem 1.7. *Let $\mathcal{B}_1, \mathcal{B}_2$ be Banach spaces over a field $\mathbb{F}$,*[7] *a linear operator $\mathrm{L} : \mathcal{B}_1 \to \mathcal{B}_2$ is continuous if and only if it is bounded.*

PROOF. Suppose that L is continuous, but not bounded. In the negation of our definition of a bounded linear operator we find that for each $n \in \mathbb{N}$ there is some $f_n \in \mathcal{B}_1$ such that $\|\mathrm{L}(f_n)\| > n \|f_n\|$. Observe that none of the f_n are zero, and it makes sense to define an auxiliary sequence of vectors in $\mathcal{B}_1$ by setting for $n \in \mathbb{N}$:

$$g_n := \frac{f_n}{n \|f_n\|}$$

By a calculation we subsequently discover that

$$\|\mathrm{L}(g_n)\| = \frac{1}{n \|f_n\|} \cdot \|\mathrm{L}(f_n)\| > \frac{1}{n \|f_n\|} \cdot n \|f_n\| = 1$$

[6] This corresponds to the $\varepsilon - \delta$ definition in metric spaces, but is easier to work with.
[7] Here we do not require that $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$

therefore

$$\lim_{n\to\infty} \mathrm{L}(g_n) \neq 0$$

On the other-hand, it is clear that $g_n \to 0$ as $n \to \infty$, since all these g_n have norm $\frac{1}{n}$. This contradicts the continuity of L at 0.

Conversely, suppose that L is a bounded linear operator. By definition, L then satisfies a "Lipschitz-like" condition, and must be continuous. Indeed, let (f_n) be a sequence in $\mathcal{B}_1$ converging to $f \in \mathcal{B}_1$. We then write:

$$\|\mathrm{L}(f_n) - \mathrm{L}(f)\| = \|L(f_n - f)\| \leq M \|f_n - f\| \xrightarrow{n\to\infty} 0$$

which proves this theorem. □

The above theorem is often called the *characterization of continuity theorem*. We now return mostly to intrinsic properties of Hilbert spaces. Again, we consider Hilbert spaces $\mathcal{H}$ over $\mathbb{K}$, either the real or complex numbers. We recall the notion of a dual space. If $\mathcal{H}$ is a Hilbert space it is necessarily a vector space, and therefore has a dual space: $\mathcal{H}^*$ consisting of all *linear functionals* over $\mathcal{H}$. That is,

$$\mathcal{H}^* := \{\ell : \mathcal{H} \to \mathbb{K},\ \ell \text{ a linear map}\} \tag{5}$$

These ℓ are called **linear functionals**. In other-terms, ℓ is said to be a linear functional over $\mathcal{H}$ if it is a linear operator:

$$\ell : \mathcal{H} \to \mathbb{K}$$

Some examples are in order. Clearly, if we fix a vector $g \in \mathcal{H}$ then one can define a linear functional $\ell(\cdot) := \langle \cdot, g \rangle$. By Cauchy-Schwarz, this is a bounded linear functional and hence a continuous linear operator.[8] The truly astounding result is the *Riesz Representation Theorem*, which states that there are no other bounded linear functionals:

Theorem 1.8 (Riesz Representation Theorem)**.** *Let $\mathcal{H}$ be a Hilbert space over $\mathbb{K}$ and suppose $\ell : \mathcal{H} \to \mathbb{K}$ is a continuous (or bounded) linear functional. There exists a unique vector $g \in \mathcal{H}$ such that*

$$\ell(f) = \langle f, g \rangle, \quad \forall f \in \mathcal{H} \tag{6}$$

Proof. This is a proof by construction. We consider now the *null-space* of ℓ, denoted below by

$$\mathcal{N} := \{g \in \mathcal{H} : \ell(g) = 0\}$$

[8]A linear functional ℓ is clear a linear operator in its own right.

We claim that $\mathcal{N} \leq \mathcal{H}$. Clearly, by linearity of ℓ it is clear that $\mathcal{N} < \mathcal{H}$, i.e. $\mathcal{N}$ is a vector subspace of $\mathcal{H}$. To see now that $\mathcal{N}$ is topologically closed as well, pick a convergent sequence (f_n) in $\mathcal{N}$ with some limit point in $\mathcal{H}$, but not known to be in $\mathcal{N}$. Then, for all n one has

$$\ell(f_n) = 0, \quad \ell(f_n) \xrightarrow{n\to\infty} \ell(f)$$

whence $\ell(f) = 0$; by continuity of ℓ. Therefore, $\mathcal{N} \leq \mathcal{H}$ as was required. Now, there must exist an orthogonal complement $\mathcal{N}^\perp$ such that

$$\mathcal{H} = \mathcal{N} \oplus \mathcal{N}^\perp$$

If $\ell \equiv 0$, we set $g = 0$ and we are done. Else, $\mathcal{N}^\perp$ we may select a unit vector $h \in \mathcal{N}^\perp$. Then, define $g := \overline{\ell(h)}h \neq 0$. Now, define for each $f \in \mathcal{H}$ some associated $u := \ell(h)f - h\ell(f)$; clearly $u \in \mathcal{N}$. Therefore, $u \perp h$ and therefore we find that

$$\begin{aligned} 0 = \langle u, h\rangle = \langle f\ell(h) - \ell(f)h, h\rangle &= \ell(h)\langle f, h\rangle - \ell(f)\langle h, h\rangle \\ &= \ell(h)\langle f, h\rangle - \ell(f) \qquad (h \text{ is a unit vector}) \end{aligned}$$

whence, $\ell(f) = \ell(h)\langle f, h\rangle = \left\langle f, \overline{\ell(h)}h \right\rangle = \langle f, g\rangle$. It now only remains to show uniqueness of this g. Suppose both $g, h \in \mathcal{H}$ satisfy the claim: i.e. $\ell(f) = \langle f, g\rangle = \langle f, h\rangle$ for all $f \in \mathcal{H}$. In particular, one has $\langle f, g - h\rangle = 0$ for all $f \in \mathcal{H}$, and thus

$$\overline{\langle g - h, f\rangle} = 0, \quad \forall f \in \mathcal{H}$$

let $\{e_j\}_{j\in\mathbb{N}}$ be an orthonormal basis for $\mathcal{H}$, since we have $\overline{\langle g - h, e_j\rangle} = 0$ for each j it follows from Theorem 1.3 that g is identical to h. □

We are now ready to study special analogues to operators over Hilbert spaces. First, we should like to extend our notion of norm to the set of all linear operators between two Hilbert spaces. For our purposes, it will suffice to extend $\|\cdot\|$ to endomorphisms.[9]

Definition. If $\mathrm{L} : \mathcal{H} \to \mathcal{H}$ be a bounded endomorphism of a Hilbert space $\mathcal{H}$ over $\mathbb{K}$ we may define the norm of L, denoted $\|\mathrm{L}\|$, to be the infimum of all $M > 0$ such that $\|\mathrm{L}(f)\| \leq M\|f\|$ for all $f \in \mathcal{H}$. In symbolic terms:

$$\|\mathrm{L}\| := \inf\{M > 0 : \|\mathrm{L}(f)\| \leq M\|f\|, \forall f \in \mathcal{H}\} \tag{7}$$

This is in practice not a useful definition, but is by far the easiest to consider

[9]Linear operators from a Hilbert space into itself, i.e. $\mathrm{L} : \mathcal{H} \to \mathcal{H}$.

and motivate. We shall give an alternative characterization of this quantity below, and it shall be of greater use but it will become self evident why it makes for a rather poor definition. We remark that the definition above makes sense whenever L is a bounded linear operator, for the set of all such M is non-empty.

Proposition 1.9. *Let* L *be a bounded endomorphism of a Hilbert space* $\mathcal{H}$*. Then,*

$$\|\mathrm{L}\| = \sup_{\substack{f,g\in\mathcal{H} \\ \|f\|=\|g\|=1}} |\langle \mathrm{L}(f), g\rangle| \tag{8}$$

Proof. Note that since L is a bounded linear operator this supremum, say, S, exists and is finite. We shall show that S is admissible as a "bound" on L as in the definition and also that $S \leq M$ for all admissible M.

Let $M > 0$ be admissible, i.e. assume $\|\mathrm{L}(f)\| \leq M \|f\|$ for all $f \in \mathcal{H}$ and let $f, g \in \mathcal{H}$ be unit vectors. Then, by Cauchy-Schwarz we obtain the following:

$$|\langle \mathrm{L}(f), g\rangle| \leq \|\mathrm{L}(f)\| \leq M \|f\| = M$$

Taking the infimum over all such M we obtain that $|\langle \mathrm{L}(f), g\rangle| \leq \|\mathrm{L}\|$, and taking the supremum over all such pairs of unit vectors (f, g) one obtains the first inequality: $S \leq \|L\|$. We must now show that S is a bound on L. Let $f \in \mathcal{H}$ be given, if $f = 0$ or $\mathrm{L}(f) = 0$ then we trivially have $\|\mathrm{L}(f)\| \leq S \|f\|$ since linear transformations must preserve the zero-vector. Else, we safely define:

$$\widetilde{f} := \frac{f}{\|f\|}, \quad \widetilde{g} := \frac{\mathrm{L}(f)}{\|\mathrm{L}(f)\|}$$

Since g is then a unit vector, we know $|\langle \mathrm{L}(g), g\rangle| \leq S$. More precisely,

$$\left|\left\langle \frac{\mathrm{L}(f)}{\|f\|}, \frac{\mathrm{L}(f)}{\|\mathrm{L}(f)\|}\right\rangle\right| \leq S$$

noting that

$$\left\langle \frac{\mathrm{L}(f)}{\|f\|}, \frac{\mathrm{L}(f)}{\|\mathrm{L}(f)\|}\right\rangle = \frac{\|\mathrm{L}(f)\|}{\|f\|}$$

we find that $\|\mathrm{L}\| \leq S$, which proves the proposition. □

For the remainder of this section let us fix a bounded endomorphism $\mathrm{L} : \mathcal{H} \to \mathcal{H}$. We wish to construct its **adjoint operator**. Namely, we claim there is a unique associated endomorphism $\mathrm{L}^* : \mathcal{H} \to \mathcal{H}$, also bounded that satisfies

$$\langle \mathrm{L}(f), g\rangle = \langle f, \mathrm{L}^*(g)\rangle \tag{9}$$

for all $f, g \in \mathcal{H}$. Certainly, such a linear operator exists by the Riesz representation theorem we have given previously. Fix $g \in \mathcal{H}$ and let $f \in \mathcal{H}$ vary. We note that one can view this inner product as a mapping

$$\ell : \mathcal{H} \to \mathcal{H}, \quad f \overset{\ell}{\mapsto} \langle \mathrm{L}(f), g \rangle$$

Since g is fixed, this is a bounded linear functional, and is hence continuous. There must therefore exist some unique vector $h \in \mathcal{H}$ such that $\ell(f) = \langle f, h \rangle$ for all $f \in \mathcal{H}$. Hence, we may define a mapping:

$$\mathrm{L}^* : \mathcal{H} \to \mathcal{H}, \quad g \overset{\mathrm{L}^*}{\mapsto} h$$

We now see clearly that under this definition we have satisfied equation (9). Since the vector given to us by the Riesz representation theorem is unique, this is a well-defined linear transformation/operator. To see that this mapping is unique, let T be any endomoprhism satisfying (9). Let $\{e_j\}_{j\in\mathbb{N}}$ be an orthonormal basis for $\mathcal{H}$. Fix $g \in \mathcal{H}$, it then follows from the linearity of the inner-product that

$$\langle e_j, \mathrm{L}^*(g) - T(g) \rangle = 0, \quad \forall j \in \mathbb{N}$$

whence by Theorem 1.3 we once again observe that $\mathrm{L}^*(g) = T(g)$, and therefore $\mathrm{L}^* \equiv T$ since g was arbitrary.

We end this section with some definitions that allow us to construct special classes of endomorphisms on a Hilbert space $\mathcal{H}$ over $\mathbb{K}$. We shall call a linear operator $\mathrm{L} : \mathcal{H} \to \mathcal{H}$ **hermitian**[10] provided $\mathrm{L}^* = \mathrm{L}$.

A linear operator $\mathrm{L} : \mathcal{H} \to \mathcal{H}$ is said to be **compact** if and only if for each bounded sequence (f_n) in $\mathcal{H}$ there is a subsequence (f_{n_k}) such that $\mathrm{L}(f_{n_k})$ converges. We shall always assume that these L are bounded (and therefore continuous).

1.4 The Spectral Theorem for Compact Operators

In this section we shall attempt to generalize to infinite dimensional spaces the results from linear algebra regarding hermitian matrices. Namely, we would like to answer the following question: if $\mathrm{L} : \mathcal{H} \to \mathcal{H}$ is a bounded, compact, hermitian operator does there exist an orthonormal basis $\{\varphi_j\}_j$ of $\mathcal{H}$ and associated $\lambda_j \in \mathbb{K}$ such that

$$\mathrm{L}(\varphi_j) = \lambda_j \varphi_j$$

As we shall see by the end of this section, the answer to this question is **yes**.

[10]Often this is called symmetric or self-adjoint.

This may be summarized in the following theorem:

Theorem 1.10 (Spectral Theorem). *Let $\mathcal{H}$ be a Hilbert space over $\mathbb{K}$ and $\mathrm{L} : \mathcal{H} \to \mathcal{H}$ a bounded, hermitian compact operator. There exists an orthonormal basis $\{\varphi_j\}_j$ and $\lambda_j \in \mathbb{K}$ such that*

$$\mathrm{L}(\varphi_j) = \lambda_j \varphi_j, \quad \forall j \tag{10}$$

Moreover, $\Im \lambda_j = 0$ and if $\dim(\mathcal{H}) = \infty$ the $\lambda_j \to 0$ as $j \to \infty$.

In the above the vectors φ are called *eigenvectors* and the λ are *eigenvalues*. Note that by definition, since they form an orthonormal basis, no eigenvector can be zero.

We do not yet have the tools to prove this theorem, and we shall build up results in this section to prove this theorem. In fact, we shall prove this theorem in several steps instead. For the remainder of this section we shall assume that L is a bounded, compact and hermitian endomorphism on $\mathcal{H}$.

Lemma 1.11. *If λ is an eigenvalue for some eigenvector φ, then $\lambda \in \mathbb{R}$. Moreover, if $\lambda_1 \neq \lambda_2$ are two eigenvalues with eigenvectors φ_1, φ_2 respectively, then $\varphi_1 \perp \varphi_2$.*

PROOF. For the first claim one needs only write:

$$\begin{aligned} \lambda \langle \varphi, \varphi \rangle = \langle \lambda\varphi, \varphi \rangle = \langle \mathrm{L}\varphi, \varphi \rangle = \langle \varphi, L\varphi \rangle &= \langle \varphi, \lambda\varphi \rangle \\ &= \overline{\lambda} \langle \varphi, \varphi \rangle \end{aligned}$$

This implies that $\lambda \|\varphi\|^2 = \overline{\lambda} \|\varphi\|^2$, and since $\varphi \neq 0$ we get $\lambda = \overline{\lambda}$ whence $\lambda \in \mathbb{R}$.

For the second claim, we wish to show that $\langle \varphi_1, \varphi_2 \rangle = 0$. To see this, let us calculate:

$$\begin{aligned} \lambda_1 \langle \varphi_1, \varphi_2 \rangle = \langle \lambda_1\varphi_1, \varphi_2 \rangle = \langle \mathrm{L}(\varphi_1), \varphi_2 \rangle &= \langle \varphi_1, \lambda_2\varphi_2 \rangle \\ &= \lambda_2 \langle \varphi_1, \varphi_2 \rangle \qquad (\lambda_2 \in \mathbb{R}) \end{aligned}$$

Now, as $\lambda_1 \neq \lambda_2$ we find that $\langle \varphi_1, \varphi_2 \rangle = 0$ whence $\varphi_1 \perp \varphi_2$. □

Note that in the above we did not require boundedness nor compactness: only that L is hermitian. Our next lemma is more intricate, and actually relies on continuity and compactness of the endomorphism L.

Lemma 1.12. *For $\mathrm{L} : \mathcal{H} \to \mathcal{H}$ compact, hermitian and continuous one has:*

(1) For each $\lambda \neq 0$ the null space of $\mathrm{L} - \lambda I$ has finite dimension.

(2) If $(\lambda_n)_{n\in\mathbb{N}}$ is a sequence of pairwise distinct eigenvalues then

$$\lim_{n\to\infty} \lambda_n = 0 \tag{11}$$

(3) There are at-most countably many eigenvalues for L.

PROOF.

(1) Let $\lambda \neq 0$ be given, and real. Define $T_\lambda := \mathrm{L} - \lambda I$. Since L, I are bounded linear operators, it follows that T_λ is bounded as well, and therefore by the argument used in the Riesz representation theorem we know that $\ker T_\lambda$ is a closed subspace of $\mathcal{H}$. It therefore inherits a Hilbert space structure. Suppose by way of contradiction that $\dim(\ker T_\lambda) = \infty$. Then, we may select an orthogonal set $\{\varphi_n\}_{n\in\mathbb{N}}$ in $\ker T_\lambda$.

For each such n one has $T_\lambda(\varphi_n) = 0$ and therefore $\mathrm{L}(\varphi_n) = \lambda\varphi_n$. Since these are orthogonal, we may assume without harm that they are unit vectors as well. Therefore, we have found a bounded sequence. Since L is compact, there is a subsequence φ_k such that $\mathrm{L}(\varphi_k)$ is convergent. In particular, the sequence $(\mathrm{L}(\varphi_k))_{k\in\mathbb{N}}$ is Cauchy in $\mathcal{H}$. On the other-hand, for all $k \neq m$ in $\mathbb{N}$ one has

$$\|\mathrm{L}(\varphi_k) - \mathrm{L}(\varphi_m)\|^2 = \|\lambda\varphi_k - \lambda\varphi_m\|^2 = \lambda^2 \|\varphi_k - \varphi_m\|^2 = \lambda^2 \qquad \text{(Prop. 1)}$$

which contradicts that $(\mathrm{L}(\varphi_k))_{k\in\mathbb{N}}$ is Cauchy since $\lambda \neq 0$

(2) Assume now that we have collected infinitely many eigenvalues. We claim that for each $\varepsilon > 0$ there are finitely many λ_n with $|\lambda_n| \geq \varepsilon$. Of course, we argue by contradiction. Otherwise, there exists $\varepsilon_0 > 0$ and a countably infinite family of *distinct* eigenvalues $\{\lambda_k\}_k$ such that $|\lambda_k| \geq \varepsilon_0$. Now, let φ_k be an eigenvector associated to λ_k for each $k \in \mathbb{N}$. By our previous lemma, these φ_k must be orthogonal and therefore we may assume without harm that they are orthonormal. Indeed, this follows from the fact that for $C \neq 0$:

$$\mathrm{L}\left(\frac{\varphi}{C}\right) = \frac{1}{C}\mathrm{L}(\varphi) = \lambda \cdot \frac{\varphi}{C}$$

Therefore, a non-zero constant multiple of an eigenvector is again an eigenvector (for the same eigenvalue). Hence, we again have a bounded sequence in $\mathcal{H}$ whence by compactness of our endomorphism L we may assume $\mathrm{L}(\varphi_k)$ is convergent; otherwise passing to a subsequence. However, it follows that $(\mathrm{L}(\varphi_k))_k$ must be Cauchy which would contradict:

$$\|\mathrm{L}(\varphi_k) - \mathrm{L}(\varphi_m)\|^2 = \lambda_k^2 + \lambda_m^2 \geq 2\varepsilon_0^2$$

for all $k \neq m$.

(3) Let $\bigwedge$ be the set of all eigenvalues for L, save possibly $\lambda = 0$ (if 0 is an eigenvalue of L). Clearly, there are at most-countably many eigenvalues for

L if and only if $\bigwedge$ is countable. Observe now that one may write:

$$\bigwedge := \bigcup_{n\in\mathbb{N}} \left\{\lambda : \lambda \geq \frac{1}{n} \text{ is an eigenvalue}\right\}$$

By (2) each of these sets is finite, and therefore $\bigwedge$ is countable.

□

Now comes the time to prove that the point of all this is not moot; namely that an operator $\mathrm{L} : \mathcal{H} \to \mathcal{H}$ that is compact, hermitian and bounded *does* have eigenvalues and eigenvectors. First consider the case $\mathrm{L} = \mathbf{0}$. Clearly, this is hermitian, compact and bounded. We observe that 0 is an eigenvalue for L with eigenvector 1. Now, for the more general case:

Lemma 1.13. *Let* $\mathrm{L} \neq 0$ *be a compact, hermitian and bounded endomorphism on* $\mathcal{H}$*. Then,* $\|\mathrm{L}\|$ *is an eigenvalue of* L *with at least one eigenvector.*

PROOF. Note first that $\|\mathrm{L}\| > 0$, for otherwise one would have $\mathrm{L} \equiv 0$. Observe that it follows from Proposition 1.9 together with the hermitian property: $\mathrm{L} = \mathrm{L}^*$ that we have[11]

$$\|\mathrm{L}\| = \sup_{f \text{ unit}} |\langle \mathrm{L}(f), f\rangle|$$

In particular, we have exactly one of the two cases below:

$$\|\mathrm{L}\| = \sup_{f \text{ unit}} \langle \mathrm{L}(f), f\rangle \qquad \text{or} \qquad \|\mathrm{L}\| = -\inf_{f \text{ unit}} \langle \mathrm{L}(f), f\rangle \tag{12}$$

We handle only the first case here, the second follows easily by a symmetric argument, or arguing for $-\mathrm{L}$ instead of L. In the first case, observe that by definition of the supremum we may select a sequence of unit vectors (f_n) in $\mathcal{H}$ such that

$$\lim_{n\to\infty} \langle \mathrm{L}(f_n), f_n\rangle = \|\mathrm{L}\|$$

By compactness of the endomorphism L we may presume without harm to the proof that $\mathrm{L}(f_n) \to g \in \mathcal{H}$; passing to a subsequence in the case where this does not hold. Ultimately, we claim that $\lambda := \|\mathrm{L}\|$ is an eigenvalue for the vector g; i.e. that $\mathrm{L}(g) = \lambda g$ and $g \neq 0$.

[11] This is not immediate. For a short proof, we refer the reader to [1, page 184]

First, write out:

$$\begin{aligned} \|\mathrm{L}(f_n) - \lambda f_n\|^2 &= \langle \mathrm{L}(f_n) - \lambda f_n, \mathrm{L}(f_n) - \lambda f_n \rangle \\ &= \|\mathrm{L}(f_n)\|^2 + \lambda^2 \|f_n\|^2 - \lambda \langle \mathrm{L}(f_n), f_n \rangle - \lambda \langle f_n), \mathrm{L}(f_n)) \rangle \\ &= \|\mathrm{L}(f_n)\|^2 + \lambda^2 - 2\lambda \langle \mathrm{L}(f_n), f_n \rangle \end{aligned}$$

where in this last step we have used that L is hermitian and the f_n's are unit vectors. Hence,

$$\begin{aligned} \|\mathrm{L}(f_n) - \lambda f_n\|^2 &\leq \|L\|^2 \|f_n\|^2 + \lambda^2 - 2\lambda \langle \mathrm{L}(f_n), f_n \rangle \\ &= 2\lambda^2 - 2\lambda \langle \mathrm{L}(f_n), f_n \rangle \end{aligned}$$

where this last term vanishes as $n \to \infty$ since $\lambda \langle \mathrm{L}(f_n), f_n \rangle \to \lambda$. Now, this tells us mostly that $\lambda f_n \to g$ as $n \to \infty$. Now, by continuity we obtain then:

$$\mathrm{L}(g) = \lim_{n\to\infty} \mathrm{L}(\lambda f_n) = \lambda \lim_{n\to\infty} \mathrm{L}(f_n) = \lambda g$$

The proof will then be complete if we can show that $g \neq 0$. Suppose for a contradiction that $g = 0$, then by continuity we have:

$$\lim_{n\to\infty} \mathrm{L}(f_n) = 0$$

However, we have shown that $\|\mathrm{L}(f_n) - \lambda f_n\|$ tends to 0 as $n \to \infty$, with $\lambda \neq 0$ by hypothesis. But, for all $n \in \mathbb{N}$ one has $\|\lambda f_n\| = |\lambda| > 0$ which is a contradiction. □

Having established all of these lemmata we are now ready to prove the *Spectral Theorem*:

PROOF OF THEOREM 1.10. Consider the pre-space $\mathfrak{S}$ consisting of the vectors in $\mathcal{H}$ spanned by the eigenvectors of the endomorphism L. Note that by the previous lemma together with the remark preceding it, we know $\mathfrak{S}$ is non-empty. The main idea here is to let $\mathcal{S}$ be the topological closure of $\mathfrak{S}$ in $\mathcal{H}$.

By construction, $\mathcal{S} \leq \mathcal{H}$. Note by construction that the collection of eigenvectors, say, $\{\varphi_n\}_n$ are orthogonal by a previous lemma, and therefore we may "normalize" these vectors. By construction, this family $\{\varphi_n\}_n$ is hence an orthonormal basis for $\mathcal{S}$.[12] We may therefore define an orthogonal complement to $\mathcal{S}$:

$$\mathcal{H} = \mathcal{S} \oplus \mathcal{S}^\perp$$

[12] Any element in the closure of a metric subspace may be approximated by elements in the original set.

The goal here is to show that $\mathcal{S}^\perp = \{0\}$, whence $\mathcal{H} = \mathcal{S}$. First we need to show that L maps $\mathcal{S}$ back into itself and $\mathcal{S}^\perp$ back into itself. Certainly, let $f \in \mathcal{S}$, then we may write:

$$f = \sum_n \zeta_n \varphi_n, \quad \zeta_n \in \mathbb{K}$$

where this sum is countable by a previous lemma; here we have used the fact that $\{\varphi_n\}$ is an orthonormal basis for $\mathcal{S}$. If this family of φ_n is finite, then clearly $\mathrm{L}(f) = \sum_n \zeta_n \mathrm{L}(\varphi_n)$ by linearity. If $\{\varphi_n\}_n$ is countably infinite we know that $\sum_{n \leq N} \zeta_n \varphi_n \to f$ as $N \to \infty$ and therefore by continuity of L one finds:

$$\mathrm{L}(f) = \lim_{N\to\infty} \mathrm{L}\left(\sum\nolimits_{n \leq N} \zeta_n \varphi_n\right) = \lim_{N\to\infty} \sum_{n \leq N} \zeta_n \lambda_n \varphi_n = \sum_n \xi_n \varphi_n$$

where we have set $\xi_n := \zeta_n \lambda_n$. Hence, $\mathrm{L}(S) \subseteq \mathcal{S}$. We also claimed that $\mathrm{L}\left(\mathcal{S}^\perp\right) \subseteq \mathcal{S}^\perp$. If $\mathcal{S}^\perp = \{0\}$ there is nothing to show, otherwise pick $g \in \mathcal{S}^\perp$. We claim $\mathrm{L}(g) \perp f$ for all $f \in \mathcal{S}$. This follows mostly from the hermitian property that L carries:

$$\langle f, \mathrm{L}g \rangle = \langle \mathrm{L}(f), g \rangle = 0$$

since $\mathrm{L}(f) \in \mathcal{S}$. Therefore, we may consider restrictions of L to both $\mathcal{S}$ and $\mathcal{S}^\perp$, since L is an endomorphism on each of these closed subspaces: and thus Hilbert spaces. Moreover, it is obvious that on these subspaces L is compact, continuous and hermitian. Suppose by way of contradiction that there is a non-zero vector $g \in \mathcal{S}^\perp$. If the restriction of L on $\mathcal{S}^\perp$ is identically zero, take $\lambda = 0$, then g is an eigenvalue of L. If instead $\|\mathrm{L}\| > 0$ on $\mathcal{S}^\perp$, we have by the previous lemma the existence of an eigenvector with eigenvalue $\mathcal{S}^\perp$. In any case we have reached a contradiction since $\mathcal{S}$ contains all eigenvectors of L.

The spectral theorem is now proven. □

2 Fourier Analysis

In this chapter we study the Fourier transform from a rigorous point of view. Let us recall from the measure theory text the space of *square integrable functions*:

$$L^2(\mathbb{R}^d) := \left\{ f \in \mathcal{M}(\mathbb{R}^d) : \int_{\mathbb{R}^d} |f(x)|^2 \, \mathrm{d}x < \infty \right\} \tag{13}$$

where $\mathcal{M}(\mathbb{R}^d)$ is the space of all measurable functions $f : \mathbb{R}^d \to \mathbb{C}$. Of all L^p spaces, L^2 is the only one that is a Hilbert space, where the inner-product on

$L^2(\mathbb{R}^d)$ is defined by:

$$\langle f, g\rangle := \int_{\mathbb{R}^d} f(x)\overline{g(x)}\,\mathrm{d}x \tag{14}$$

The purpose of this chapter is to develop a Fourier transform that takes advantage of the generality intrinsic to the Lebesgue integral. The strategy is to develop the Fourier transform for a "smaller" space of "nicer" functions, which we shall show to be dense in $L^2(\mathbb{R}^d)$ and extend this uniquely to all of $L^2(\mathbb{R}^d)$.

2.1 Schwartz Space and the Fourier Transform

In this section we define and study basic properties of the Fourier transform on a smaller space of functions called *Schwartz* functions. Without harm we shall consider the case where $d = 1$, for simplicity. The same arguments carry over nicely to higher dimensions, with the aid of Fubini's theorem. The Schwartz class of functions is defined below:

Definition (Schwartz Space)**.** The Schwartz space, denoted $\mathcal{S} = \mathcal{S}(\mathbb{R})$ is defined as the subspace of $L^p(\mathbb{R})$ for all $p \geq 1$:

$$\mathcal{S} := \left\{ f \in C^\infty(\mathbb{R}) : \forall \alpha, \beta \in \mathbb{N}_0, \exists C_{\alpha,\beta} \geq 0 \text{ with } \left\| x^\alpha \left(\frac{\mathrm{d}}{\mathrm{d}x}\right)^\beta f \right\|_\infty \leq C_{\alpha,\beta} \right\}^{13}.$$

Loosely speaking, $\mathcal{S}$ consists of all smooth functions decaying faster at infinity, along with its derivatives, than any rational function. In the hopes of avoiding confusion, despite having used $\mathcal{S}$ to denote the Schwartz space, the reader should not assume that $\mathcal{S}$ is a closed subspace of L^p. Certainly, note that, given p, by choosing α appropriately with $\beta = 0$ we find $\mathcal{S}(\mathbb{R}) \subsetneq L^p(\mathbb{R})$. To see that this inclusion is strict, consider the Dirichlet function:

$$\mathfrak{d} : \mathbb{R} \to \mathbb{R}, \quad x \mapsto \mathbf{1}_{\mathbb{Q}}(x).$$

Since $\mu(\mathbb{Q}) = 0$ we note that $\mathfrak{d} = 0$ almost everywhere and therefore $\int_{\mathbb{R}} |\mathfrak{d}(x)|^p \ \mathrm{d}x = 0$ for all $p \geq 1$. However, $\mathfrak{d}$ is discontinuous everywhere and is certainly not an element of $\mathcal{S}$.

Moreover we have that $\mathcal{S}$ cannot be a closed subspace of L^p. Indeed, we shall later see that $\mathcal{S}$ is dense in $L^p(\mathbb{R}^d)$ for all $p \geq 1$, and then it follows from the following proposition that $\mathcal{S}$ cannot be a closed subspace of $L^p(\mathbb{R})$:

[13] We write here $\mathbb{N}_0 := \{0, 1, \ldots\}$.

Proposition 2.1. *Let (X, d) be a metric space and $Q \subseteq X$ a closed dense subspace of X. Then, $Q = X$. Namely, the only closed dense subspace of X is X itself.*

PROOF. Given $x \in X$ choose a sequence (q_n) approximating x, i.e. $(q_n) \to x$ as $n \to \infty$. Since Q is closed, $x \in Q$ and we have $Q \supseteq X$. □

With these preliminary concerns out of the way, we are prepared to consider the Fourier transform.

Definition (Fourier Transform). For any $f \in \mathcal{S} = \mathcal{S}(\mathbb{R})$ we define the Fourier transform of f, denoted $\widehat{f}$ or $\mathcal{F}[f]$, to be the function $\widehat{f} : \mathbb{R} \to \mathbb{C}$ defined by the integral:

$$\mathcal{F}[f](\xi) = \hat{f}(\xi) := \int_{\mathbb{R}} f(x) e^{-2\pi i x \xi} \, \mathrm{d}x \tag{15}$$

We shall alternate between the notations above, when one is more convenient than the other. We claim that the above is well defined, i.e. that for all ξ the integrand $f(x)e^{-2\pi i x \xi}$ is integrable. Certainly, by our earlier observations $f \in \mathcal{S}$ implies $f \in L^1(\mathbb{R})$, and since $\left| f(x) e^{-2\pi i x \xi} \right| = |f(x)|$ we have $f(x)e^{-2\pi i x \xi} \in L^1(\mathbb{R})$.

Before we proceed further, we require a further result on the decay of such integrands.

Theorem 2.2 (Riemann-Lebesgue Lemma). *Let $f \in L^1(\mathbb{R})$. Then,*

$$\lim_{|\xi| \to \infty} \int_{\mathbb{R}} f(x) e^{-2\pi i x \xi} \, \mathrm{d}x = 0 \tag{16}$$

PROOF. We begin by arguing for the characteristic functions of intervals. Consider now $f(x) = \mathbf{1}_{[a,b]}(x)$ for some compact interval $[a, b]$ in $\mathbb{R}$. Then, it is easy to verify now by direct calculation that for all $\xi \in \mathbb{R}$:

$$\begin{aligned}
\left| \int_{\mathbb{R}} f(x) e^{-2\pi i x \xi} \, \mathrm{d}x \right| = \left| \int_a^b e^{-2\pi i x \xi} \, \mathrm{d}x \right| &= \left| -\frac{e^{-2\pi i x \xi}}{2\pi i \xi} \Big|_a^b \right| \\
&= \frac{1}{2\pi |\xi|} \left| e^{-2\pi i b \xi} - e^{-2\pi i a \xi} \right| \\
&\leq \frac{1}{2\pi |\xi|}
\end{aligned}$$

vanishing in the limit as $|\xi| \to \infty$.

Now, for the general case we recall that step functions are dense in $L^1(\mathbb{R})$. Let $f \in L^1(\mathbb{R})$ be given and fix $\varepsilon > 0$. There exists a step function $\varphi =$

$\sum_{n=1}^{N} \alpha_n \mathbf{1}_{I_n}(x)$, where each I_n is an interval of the form $[a_n, b_n]$ with $\|f-\varphi\|_{L^1} \leq \varepsilon$. Now, we write:

$$\begin{aligned}\left|\int_{\mathbb{R}} f(x)e^{-2\pi i x\xi}\,\mathrm{d}x\right| &= \left|\int_{\mathbb{R}} (f(x)-\varphi(x))\,e^{-2\pi i x\xi}\,\mathrm{d}x + \int_{\mathbb{R}} \varphi(x)e^{-2\pi i x\xi}\,\mathrm{d}x\right| \\ &\leq \|f-\varphi\|_{L^1} + \sum_{n=1}^{N} |\alpha_n| \left|\int_{a_n}^{b_n} e^{-2\pi i x\xi}\,\mathrm{d}x\right|\end{aligned}$$

By the argument above, letting $|\xi|$ be very large we can guarantee the inequality

$$\sum_{n=1}^{N} |\alpha_n| \left|\int_{a_n}^{b_n} e^{-2\pi i x\xi}\,\mathrm{d}x\right| \leq \varepsilon.$$

Putting all this together, we find

$$\lim_{|\xi|\to\infty} \left|\int_{\mathbb{R}} f(x)e^{-2\pi i x\xi}\,\mathrm{d}x\right| \leq 2\varepsilon$$

since $\varepsilon > 0$ was arbitrary, the proof is complete. □

We shall now prove one more lemma, which will be of great use for calculation purposes:

Lemma 2.3. *Let $f \in \mathcal{S}(\mathbb{R})$ be given. Then,*

(1) $\dfrac{\mathrm{d}}{\mathrm{d}\xi}\widehat{f}(\xi) = (-2\pi i)\widehat{xf}(\xi)$.

(2) $\mathcal{F}[f'](\xi) = 2\pi i\xi \widehat{f}(\xi)$.

PROOF.

(1) Since $f \in \mathcal{S}$ implies that f is smooth we may pass to limits of Riemann integrals, where we may "differentiate under the integral sign". Therefore, we may calculate:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\xi}\widehat{f}(\xi) = \frac{\mathrm{d}}{\mathrm{d}\xi}\int_{\mathbb{R}} f(x)e^{-2\pi i x\xi}\,\mathrm{d}x &= \int_{\mathbb{R}} f(x)\cdot\frac{\partial}{\partial\xi}e^{-2\pi i x\xi}\,\mathrm{d}x \\ &= \int_{\mathbb{R}} f(x)\cdot(-2\pi i x)e^{-2\pi i x\xi}\,\mathrm{d}x \\ &= (-2\pi i)\int_{\mathbb{R}} f(x)e^{-2\pi i x\xi}\,\mathrm{d}x\end{aligned}$$

whence we have proven the first claim.

(2) For the second, we pass to a similar argument together with *integration by parts* to discover that $\mathcal{F}\left[f'(x)\right](\xi)$ is given by

$$\int_{\mathbb{R}} f'(x)e^{-2\pi ix\xi}\ \mathrm{d}x = f(x)e^{-2\pi ix\xi}\big|_{-\infty}^{\infty} - \int_{\mathbb{R}} f(x)(-2\pi i\xi)e^{-2\pi ix\xi}\ \mathrm{d}x$$

Since $f \in \mathcal{S}$ it vanishes quickly at infinity, and we are left with $(2\pi i\xi)\int_{\mathbb{R}} f(x)e^{-2\pi ix\xi}\ \mathrm{d}x$ which is precisely the quantity

$$2\pi i\xi\widehat{f}(\xi)$$

□

This implies the following:

Corollary 2.4. *The linear operator $\mathcal{F}[\cdot]$ is an endomorphism of $\mathcal{S}$.*

PROOF. We must show that given $f \in \mathcal{S}$ the image point $\widehat{f}$ is in $\mathcal{S}(\mathbb{R})$. Let $\alpha \geq 0$ be an integer. We note that the previous lemma together with simple induction implies:

$$\xi^\alpha\widehat{f}(\xi) = \frac{1}{(2\pi i)^\alpha}\widehat{f^{(\alpha)}}(\xi) = \frac{1}{(2\pi i)^\alpha}\int_{\mathbb{R}} f^{(\alpha)}(x)e^{-2\pi ix\xi}\ \mathrm{d}x$$

since f decays fast at infinity we may invoke the Riemann Lebesgue lemma as $|\xi| \to \infty$ to deduce that $\xi^\alpha\widehat{f}(\xi)$ vanishes at infinity. Since this function is continuous in ξ, it follows that we may bound $\xi^\alpha\widehat{f}(\xi)$ uniformly, as is required.

Now assume $\beta \geq 0$ is an integer. Again, by induction on the previous lemma we note that

$$\left(\frac{\mathrm{d}}{\mathrm{d}\xi}\right)^\beta \widehat{f}(\xi) = (-2\pi i)^\beta\mathcal{F}\left[x^\beta f\right](\xi) = (-2\pi i)^\beta\int_{\mathbb{R}} x^\beta f(x)e^{-2\pi ix\xi}\,\mathrm{d}\xi$$

By definition of $\mathcal{S}$, it follows that $x^\beta f(x)$ is again in $\mathcal{S}$ and therefore is of class L^1, an application of Riemann-Lebesgue proves the corollary. □

We now show that $\mathcal{F}[\cdot]$ behaves somewhat like the convolution operator, in the sense that we may "exchange" the the transform under the integral.

Proposition 2.5. *Let $f, g \in \mathcal{S}$, then*

$$\int_{\mathbb{R}} \hat{f}(x)g(x)\ \mathrm{d}x = \int_{\mathbb{R}} f(y)\hat{g}(y)\ \mathrm{d}y \tag{17}$$

PROOF. We pass unto Fubini-Tonelli:

$$\begin{aligned}\int_{\mathbb{R}} \hat{f}(x)g(x)\ \mathrm{d}x &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(y)e^{-2\pi iyx}\ \mathrm{d}y \right) g(x)\ \mathrm{d}x \\ &= \int_{\mathbb{R}\times\mathbb{R}} f(y)g(x)e^{-2\pi iyx}\ \mathrm{d}x\,\mathrm{d}y \\ &= \int_{\mathbb{R}} f(y) \left(\int_{\mathbb{R}} g(x)e^{-2\pi ixy}\ \mathrm{d}x \right) \mathrm{d}y \\ &= \int_{\mathbb{R}} f(y)\hat{g}(y)\ \mathrm{d}y\end{aligned}$$

as was to be shown. □

There are other algebraic identities that are of particular use in the practice. For $a \in \mathbb{R}$ let us define $f_a(x) := f(x+a)$. Then,

$$\widehat{f_a}(\xi) = \int_{\mathbb{R}} f(x+a)e^{-2\pi ix\xi}\ \mathrm{d}x \xrightarrow{u:=x+a} \int_{\mathbb{R}} f(u)e^{-2\pi i(u-a)\xi}\ \mathrm{d}u$$

whence we find

$$\boxed{\widehat{f_a}(\xi) = e^{2\pi ia\xi}\widehat{f}(\xi)}, \quad \text{where } f_a(x) = f(x+a) \tag{18}$$

In a similar vein, for $f_a(x) := f\left(\frac{x}{a}\right)$, where $a \neq 0$, we calculate:

$$\widehat{f_a}(\xi) = \int_{\mathbb{R}} f\left(\frac{x}{a}\right) e^{-2\pi ix\xi}\ \mathrm{d}x \xrightarrow{u:=x/a} a\int_{\mathbb{R}} f(u)e^{-2\pi i(ax)\xi}\ \mathrm{d}x$$

therefore, for $f_a(x) = f(x/a)$ we obtain

$$\boxed{\widehat{f_a}(\xi) = a\widehat{f}(a\xi)}, \quad \text{where } f_a(x) := f\left(\frac{x}{a}\right) \tag{19}$$

2.2 The Fourier Inversion Theorem

In this section we consider the *recovery problem*, i.e. how to recover f from $\widehat{f}$. First, we shall need to calculate a particular Fourier transform. Thankfully, this provides us with an example of a Fourier transform being used in practice!

Fourier Transform of Gaussian

Consider the function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) := e^{-\pi x^2}$. We wish to calculate $\mathcal{F}[f](\xi)$. First, we note that by direct calculation one has

$$f'(x) = -2\pi x f(x)$$

implying that $\mathcal{F}[f'](\xi) = -2\pi \mathcal{F}[xf](\xi) = \frac{1}{i}\partial_\xi \mathcal{F}[f](\xi)$ by Lemma 2.3. In a similar vein, by this very same lemma one easily finds:

$$\mathcal{F}[f'](\xi) = 2\pi i \xi \mathcal{F}[f](\xi)$$

Moreover, by direct calculation we have:

$$\mathcal{F}[f](0) = \int_{\mathbb{R}} e^{-\pi x^2} \, \mathrm{d}x = 1$$

Putting all of this data togther, we obtain the following IVP:

$$\begin{cases} \partial_\xi \mathcal{F}[f] = -2\pi\xi\mathcal{F}[f](\xi) & \text{in } \mathbb{R} \\ \mathcal{F}[f](0) = 1 \end{cases} \tag{20}$$

This differential equation has solution $\mathcal{F}[f](\xi) = Ce^{-\pi\xi^2}$, for some $C \in \mathbb{R}$. Using that $\mathcal{F}[f](0) = 1$ we find $C = 1$ and therefore;

$$\mathcal{F}[f](\xi) = e^{-\pi\xi^2}$$

Proof of The Fourier Inversion Formula

Theorem 2.6 (Fourier Inversion Formula). *Let $f \in \mathcal{S}$, we know $\widehat{f} \in \mathcal{S}$. One can recover f by calcuating:*

$$\int_{\mathbb{R}} \widehat{f}(\xi) e^{2\pi i x \xi} \, \mathrm{d}\xi = f(x) \tag{21}$$

PROOF. Consider the smooth map for $a \gg 0$ defined by $g(\xi) := e^{-\pi \frac{\xi^2}{a^2}}$. We have by equation (17) that

$$\int_{\mathbb{R}} \widehat{f}(\xi) g(\xi) \, \mathrm{d}\xi = \int_{\mathbb{R}} f(x) \widehat{g}(x) \, \mathrm{d}x.$$

By our calculations in the previous subsection for the Fourier transform of $e^{-\pi y^2}$, together with (19) we note that this implies

$$\int_{\mathbb{R}} f(\xi)\widehat{g}(\xi)\ \mathrm{d}\xi = \int_{\mathbb{R}} f(x) a e^{-\pi(xa)^2}\ \mathrm{d}x.$$

Letting now $u := ax$ we find that:

$$\int_{\mathbb{R}} f(\xi)\widehat{g}(\xi)\ \mathrm{d}\xi = \int_{\mathbb{R}} f\left(\frac{u}{a}\right) e^{-\pi u^2}\ \mathrm{d}u.$$

On one hand, the left hand side as a tends to infinity is equivalent to:

$$\lim_{a\to\infty}\int_{\mathbb{R}} \widehat{f}(\xi)g(\xi)\ \mathrm{d}\xi = \lim_{a\to\infty}\int_{\mathbb{R}} \widehat{f}(\xi)e^{-\pi\frac{\xi^2}{a^2}}\ \mathrm{d}\xi = \int_{\mathbb{R}} \widehat{f}(\xi)\ \mathrm{d}\xi$$

by dominated convergence. However, in the right-hand side we find instead:

$$\lim_{a\to\infty}\int_{\mathbb{R}} f\left(\frac{u}{a}\right) e^{-\pi u^2}\ \mathrm{d}u = f(0)\int_{\mathbb{R}} e^{-\pi u^2}\ \mathrm{d}u = f(0)$$

by dominated convergence. Whence,

$$f(0) = \int_{\mathbb{R}} \widehat{f(\xi)}\ \mathrm{d}\xi$$

Now we are almost done. For any $z \in \mathbb{R}$ define a function $h(x) := f(x+z) = f_z(x)$. We may repeat the above procedure to discover in this case that:

$$f(z) = h(0) = \int_{\mathbb{R}} \widehat{h}(\xi)\ \mathrm{d}\xi = \int_{\mathbb{R}} e^{2\pi i z\xi}\widehat{f}(\xi)\ \mathrm{d}\xi \qquad \text{(eqn (18))}$$

which proves the theorem. □

The formula is "nice", but what we really have been seeking is the subsequent corollary:

Corollary 2.7. *The Fourier transform $\mathcal{F}[\cdot]$ is an automorphism of $\mathcal{S}$.*

3 The Plancherel Identities

In this section we develop some useful algebraic identities using the Fourier transform. First consider any $f \in \mathcal{S}$; we have seen that $\widehat{f} \in \mathcal{S}$ as well. We begin with the calculation:

$$\overline{f(x)} = \overline{\int_{\mathbb{R}} e^{2\pi i x\xi} \widehat{f}(\xi)\ \mathrm{d}\xi} = \int_{\mathbb{R}} e^{-2\pi i x\xi} \overline{\mathcal{F}[f](\xi)}\ \mathrm{d}\xi = \mathcal{F}\left[\overline{\mathcal{F}[f]}\right](x)$$

We are now prepared to evaluate:

$$\begin{aligned}\|f\|_{L^2}^2 = \int_{\mathbb{R}} f(x)\overline{f(x)}\ \mathrm{d}x &= \int_{\mathbb{R}} f(x)\mathcal{F}\left[\overline{\mathcal{F}[f]}\right](x)]\ \mathrm{d}x \\ &= \int_{\mathbb{R}} \mathcal{F}[f](x)\overline{\mathcal{F}[f](x)}\ \mathrm{d}x = \|\mathcal{F}[f]\|_{L^2}^2\end{aligned}$$

The above is called **Plancherel's First Identity**. As for the second, let δ denote the Dirac-δ measure. Then, we claim for $f, g \in \mathcal{S}$:

$$\begin{aligned}\int_{\mathbb{R}} \widehat{f}(\xi)\overline{\widehat{g}(\xi)}\ \mathrm{d}\xi &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x)e^{-2\pi i x\xi}\ \mathrm{d}x\right) \overline{\left(\int_{\mathbb{R}} g(y)e^{-2\pi i y\xi}\ \mathrm{d}y\right)}\ \mathrm{d}\xi \\ &= \iiint_{\mathbb{R}^3} f(x)\overline{g(y)}e^{-2\pi i(x-y)\xi}\ \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}\xi \\ &= \iiint_{\mathbb{R}^3} f(x)\overline{g(y)}e^{-2\pi i(x-y)\xi}\ \mathrm{d}\xi\,\mathrm{d}y\,\mathrm{d}x \\ &= \iint_{\mathbb{R}^2} f(x)\overline{g(y)}\delta(x-y)\ \mathrm{d}y\,\mathrm{d}x \\ &= \int_{\mathbb{R}} f(x)\overline{g(x)}\ \mathrm{d}x.\end{aligned}$$

The above marks **Plancherel's Second Identity**.

3.1 Extending $\mathcal{F}$ to $L^2(\mathbb{R})$

Here we shall accept without proof a fact from mollifier theory; although it is easy to believe. We already know that continuous functions of compact support are dense in L^p for $p \geq 1$. In a similar vein, it can be shown that $C_0^\infty(\mathbb{R})$, the smooth functions of compact support are dense in $L^p(\mathbb{R})$. This is typically achieved by "smoothing-out" these functions. In any case, this implies that $\mathcal{S}$ is a *dense* subspace of $L^2(\mathbb{R})$.

We now show that the bounded linear operator on a dense subspace of a Banach space $\mathcal{B}$ may be extended to the entire space. First, we claim that $\mathcal{F}$ is a bounded operator in $\mathcal{S}$. Indeed, let $f \in \mathcal{S}$; we compute:

$$\|\mathcal{F}[f]\|_{L^2} = \|f\|_{L^2}$$

by Plancherel's first identity. Therefore, it suffices to prove the following theorem:

Theorem 3.1. *Let $\mathcal{B}_1, \mathcal{B}_2$ be Banach spaces and $\mathcal{S}$ a dense subspace of $\mathcal{B}_1$. If $\mathrm{L} : \mathcal{S} \to \mathcal{B}_2$ is a bounded linear operator there exists a unique bounded extension of L to $\mathcal{B}_1$.*

PROOF. Our first job is to define $\mathrm{L}(f)$ for a general vector $f \in \mathcal{B}_1$. Let (f_n) be a sequence in $\mathcal{S}$ such that $(f_n) \to f$ as $n \to \infty$. Especially, (f_n) is Cauchy in $\mathcal{B}_1$ and as a bounded linear operator we note that $(\mathrm{L}(f_n))$ is Cauchy in $\mathcal{B}_2$, which we know to be complete. There exists then a limit point g such that

$$\lim_{n\to\infty} \mathrm{L}(f_n) = g.$$

We then define $\mathrm{L}(f) := g$. We claim that this is well defined and independent of choice of sequence (f_n). Indeed, if (g_n) is another sequence approximating f then we have especially that

$$\|f_n - g_n\| \xrightarrow{n\to\infty} 0.$$

Whence, $\|\mathrm{L}(f_n) - \mathrm{L}(g_n)\| = \|\mathrm{L}(f_n - g_n)\| \to 0$ as $n \to \infty$ by continuity of L at 0.[14] We claim that this resulting operator is bounded. On $\mathcal{S}$ we know that $\|\mathrm{L}(\cdot)\| \le M \|\cdot\|$. For each $f \in \mathcal{B}_1$ let (f_n) be a sequence in $\mathcal{S}$ converging to f. Then, by definition:

$$\mathrm{L}(f) = \lim_{n\to\infty} \mathrm{L}(f_n) \le \lim_{n\to\infty} M \|f_n\| = M \|f\|$$

by continuity of the norm. Now, to show uniqueness assume there are two extensions, say, L, T. Then, they must agree on $\mathcal{S}$. For any $f \in \mathcal{B}_1$ pick a sequence (f_n) in $\mathcal{S}$ approximating f. Then, by continuity we have both:

$$\mathrm{L}(f) = \lim_{n\to\infty} \mathrm{L}(f_n), \quad T(f) = \lim_{n\to\infty} T(f_n).$$

Since $\mathrm{L}(f_n) = T(f_n)$ for each n we have $T \equiv \mathrm{L}$. □

We apply the above procedure to $\mathcal{F}[\cdot]$, whence we obtain a Fourier transform $\mathcal{F}[\cdot]$ on $L^2(\mathbb{R})$. By this we mean, that if $(f_n) \subseteq \mathcal{S}$ is a sequence approximating $f \in L^2$ we define:

$$\lim_{n\to\infty} \mathcal{F}[f_n] =: \mathcal{F}[f]$$

Note that the above limit is defined in the L^2 sense. We claim now that Plancherel's first identity survives this extension. Surely, let $f \in L^2(\mathbb{R})$ and

[14] Note that the zero vector lies in any subspace of $\mathcal{B}_1$.

pick a sequence (f_n) in $\mathcal{S}$ such that $\|f_n - f\|_{L^2} \to 0$ as $n \to \infty$. Then, since any norm $\|\cdot\|$ is continuous

$$\|\mathcal{F}[f]\| = \lim_{n\to\infty} \|\mathcal{F}[f_n]\| = \lim_{n\to\infty} \|f_n\| = \|f\|$$

We may repeat the above process to extend the inverse transform: $\mathcal{F}^{-1}$ to L^2; we wish to show that $\mathcal{F} \circ \mathcal{F}^{-1} \equiv \mathcal{F}^{-1} \circ \mathcal{F} \equiv \mathbf{1}$. This is clear from the fact that the composition of two bounded operators is again bounded (and hence continuous). Thence, if $f \in L^2(\mathbb{R})$ and (f_n) is a sequence in $\mathcal{S}$ with $(f_n) \to f$ in $L^2(\mathbb{R})$ we find by continuity of composition

$$\left(\mathcal{F} \circ \mathcal{F}^{-1}\right)(f) = \lim_{n\to\infty} \mathcal{F}\left(\mathcal{F}^{-1}(f_n)\right) = \lim_{n\to\infty} f_n = f$$

and an identical argument proves the case for $\mathcal{F}^{-1} \circ \mathcal{F}$. Therefore, this extension is an automorphism of $L^2(\mathbb{R})$, and moreover, is a unitary mapping.

4 Solved Problems

These problems are taken from [1] Chapters 1-2. We shall omit portions of the problems involving Cantor sets $\mathcal{C}$.

Problem 1. Let $E \subseteq \mathbb{R}^d$ be a measurable set and define

$$\mathcal{O}_n := \left\{ x \in \mathbb{R}^d : d(x, E) < \frac{1}{n} \right\}.$$

Prove that:

(1) If E is compact then $\mu(E) = \lim_{n\to\infty} \mu(\mathcal{O}_n)$.

(2) If E is merely closed but unbounded then (1) may not hold true.

Solution.

(1) To prove this first statement we observe that since $d(x, E) = 0$ for all $x \in E$ one has immediately $E \subseteq \bigcap_{n\in\mathbb{N}} \mathcal{O}_n$. Now, we observe that $\mathcal{O}_{n+1} \subseteq \mathcal{O}_n$ for all $n \in \mathbb{N}$. Indeed, clearly $d(x, E) < \frac{1}{n+1} < \frac{1}{n}$. Therefore, if E is compact then the $\mathcal{O}_n$ are uniformly bounded whence

$$\mu\left(\bigcap_{n\in\mathbb{N}} \mathcal{O}_n\right) = \lim_{n\to\infty} \mu(\mathcal{O}_n)$$

It therefore suffices to show that $E = \bigcap_{n\in\mathbb{N}} \mathcal{O}_n$. To see the reverse inclusion, let $x \in \bigcap_{n\in\mathbb{N}} \mathcal{O}_n$, then $d(x,E) < 1/n$ for all n implies that $d(x,E) = 0$. Therefore, by definition of $d(x,E)$ there exists a sequence (y_n) in E such that

$$0 \leq |y_n - x| \leq \frac{1}{n}$$

passing to the limit we observe that $y_n \to x$; since E is closed (it is compact) $x \in E$ and we are done.

(2) A simple counter example is $\mathbb{N}$. This is a closed subspace of $\mathbb{R}$, since one may write:

$$\mathbb{N}^c = (-\infty, 1) \cup \bigcup_{k=1}^{\infty} (k, k+1)$$

which is clearly open as the union of open sets. However, $\mathbb{N}$ is unbounded and since it is discrete we have $\mu(\mathbb{N}) = 0$. On the other-hand, for each $n \in \mathbb{N}$ one has:

$$\mathcal{O}_n = \bigcup_{k\in\mathbb{N}} \left(k - \frac{1}{n}, k + \frac{1}{n}\right)$$

whence, $\mu(\mathcal{O}_n) = \sum_{k\in\mathbb{N}} \frac{2}{n} = \infty$ for all n; clearly not identical to $\mu(\mathbb{N})$.

□

Problem 2. Let E be a measurable subset of $\mathbb{R}^d$ with finite measure. Then, if $\delta \in \mathbb{R}^d$ is a vector with positive components then the set δE is also measurable and carries Lebesgue measure $\mu(\delta E) = |\delta|\, \mu(E)$.

SOLUTION. We first show δE is measurable. Let $\varepsilon > 0$ be given, since $E \in \mathcal{M}(\mathbb{R}^d)$ we may find an open set $O \supseteq E$ with $\mu(O \setminus E) \leq \varepsilon$. It is clear that the dilation of an open ball is again an open ball, and therefore it follows that δO is an open set containing δE. Now, we may select a cubic covering $\{Q_j\}_{j\in\mathbb{N}}$ for $O \setminus E$ such that

$$\sum_{j\in\mathbb{N}} |Q_j| \leq \mu(O \setminus \varepsilon) + \varepsilon$$

by definition of the infimum. Clearly, $\{\delta Q_j\}_j$ is then a family of cubes covering $\delta O \setminus \delta E$. Whence, by monotonicity it follows that

$$\begin{aligned}\mu(\delta O \setminus \delta E) \leq \sum_{j\in\mathbb{N}} |\delta Q_j| = \delta \sum_{j\in\mathbb{N}} |Q_j| &\leq \delta\mu(O \setminus E) + \delta\varepsilon \\ &\leq \varepsilon\delta + \varepsilon\delta = 2\varepsilon\delta\end{aligned}$$

Proving that $\delta O \in \mathcal{M}(\mathbb{R}^d)$.

Now, let $\{Q_j\}_{j\in\mathbb{N}}$ be a collection of cubes covering E. It is clear that $\{\delta Q_j\}_j$ covers δE and therefore;

$$\mu(\delta E) \leq \sum_{j\in\mathbb{N}} |\delta| \, |Q_j| = |\delta| \sum_{j\in\mathbb{N}} |Q_j|$$

Taking the infimum over all such cubes we find $\mu(\delta E) \leq |\delta| \, \mu(E)$. This argument is symmetric with respect to $1/|\delta|$ and therefore we achieve the desired equality: $|\delta| \, \mu(E) = \mu(\delta E)$. □

Problem 3. Prove the Borel-Cantelli Lemma:

"Let $\{E_k\}_{k\in\mathbb{N}}$ be a sequence of measurable sets, with $\sum_{k\in\mathbb{N}} \mu(E_k) < \infty$. Let $\mathcal{E}$ denote the set of all points lying in infinitely many of the E_k. Then, $\mathcal{E}$ is a null set; i.e. $\mu(\mathcal{E}) = 0$."

PROOF. The key is to observe that:

$$\mathcal{E} = \bigcap_{n\in\mathbb{N}} \bigcup_{k\geq n} E_k$$

Clearly, $\mathcal{E} \subseteq \bigcap_{n\in\mathbb{N}} \bigcup_{k\geq n} E_k$. Conversely, if $x \in \bigcap_{n\in\mathbb{N}} \bigcup_{k\geq n} E_k$ then for each $n \in \mathbb{N}$ there is at least one E_k with $k \geq n$ and $x \in E_k$; i.e. x belongs to infinitely many E_k.

Now, since $\sum_{k\in\mathbb{N}} \mu(E_k) < \infty$ the series converges absolutely and it follows that the tail of this sequence converges to zero. Let $\varepsilon > 0$, then there is some $N \in \mathbb{N}$ such that $\sum_{k\geq N} \mu(E_k) \leq \varepsilon$. Now, by monotonicity of the Lebsegue measure on $\mathbb{R}^d$ one has

$$\mathcal{E} \subseteq \bigcap_{k\geq N} E_k$$

whence

$$\mu(\mathcal{E}) \leq \sum_{k\geq N} \mu(E_k) \leq \varepsilon$$

since $\varepsilon > 0$ was arbitrary we find $\mu(\mathcal{E}) = 0$. □

Problem 4. Let $\{f_n\}$ be a sequence of measurable functions on $[0, 1]$ with $|f_n(x)| < \infty$ for almost all x and all $n \in \mathbb{N}$. There exists a sequence of $c_n > 0$ such that

$$\lim_{n\to\infty} \frac{f_n(x)}{c_n}$$

for almost all $x \in [0, 1]$.

SOLUTION. We show that for each $n \in \mathbb{N}$ there is some $c_n > 0$ such that

$$\mu\left(\left\{x : \frac{|f_n(x)|}{c_n} > \frac{1}{n}\right\}\right) < 2^{-n}$$

Indeed, for $m \in \mathbb{N}$ let us define:

$$E_m := \left\{x : \frac{|f_n(x)|}{m} > \frac{1}{n}\right\} \in \mathcal{M}(\mathbb{R})$$

We claim that $\bigcap_{m=1}^{\infty} E_m$ is a null set. Indeed, note that if $x \in \bigcap_{m=1}^{\infty} E_m$ then:

$$\lim_{m\to\infty} \frac{|f_n(x)|}{m} \geq \frac{1}{n}$$

implying that $f_n(x) = \pm\infty$. Therefore, it follows that $\mu\left(\bigcap_{m\in\mathbb{N}} E_m\right) = 0$. Furthermore, note that $E_{m+1} \subseteq E_m$ since

$$\frac{|f_n(x)|}{m+1} \leq \frac{|f_n(x)|}{m}, \quad \forall m \in \mathbb{N}$$

Therefore, $\lim_{m\to\infty} \mu(E_m) = 0$ whence there is $c_n \in \mathbb{N}$ such that:

$$\mu(E_m) = \mu\left(\left\{x : \frac{|f_n(x)|}{c_n} > \frac{1}{n}\right\}\right) < 2^{-n}$$

Therefore, taking $\{F_n\}$ to be collection of such E_m we observe that $\sum_{n\in\mathbb{N}} \mu(F_n) < \infty$. Now we note that if

$$\frac{f_n(x)}{c_n} \not\to 0 \text{ as } n \to \infty$$

one must have $x \in \bigcap_{n\in\mathbb{N}} \bigcup_{k\geq n} F_k$ for otherwise $\frac{f_n(x)}{c_n} \leq \frac{1}{n}$ for all sufficiently large n. Therefore, by Borel-Cantelli the set of all such x has measure 0. □

Problem 5. There is no continuous function $\phi : \mathbb{R} \to \mathbb{R}$ equal to $\mathbf{1}_{[a,b]}(x)$ almost everywhere.

SOLUTION. By way of contradiction assume $\phi : \mathbb{R} \to \mathbb{R}$ is continuous almost and equal to $\mathbf{1}_{[a,b]}(x)$ almost everywhere. We note that $\phi(a) = 1$ for otherwise we have either $\phi(a) < 1$ or $\phi(a) > 1$. In either case, there is an interval about a, say, $(a-\delta, a+\delta)$ where $\phi > 1$ or $\phi < 1$, respectively. In either-case, there is a set of positive measure where ϕ disagrees with $\mathbf{1}_{[a,b]}(x)$.

We claim that $\phi(a) = 0$ as well. Otherwise, by continuity there is a neighbourhood $(a-\varepsilon, a+\varepsilon)$ where $\phi \neq 0$. However, then $\phi \neq \mathbf{1}_{[a,b]}$ on $(a-\varepsilon, a]$, which has positive measure. □

Problem 6. Let $A, B \subset \mathbb{R}$ be measurable sets with $\mu(B) < \infty$ and $\mu(A) = \mu(B)$. If $A \subseteq E \subseteq B$ then E is measurable and $\mu(E) = \mu(B)$.

SOLUTION. We begin by proving that E is measurable. We may write $B = A \sqcup (B \setminus A)$. That is, we have $\mu(B) = \mu(A) + \mu(B \setminus A)$ whence $\mu(B \setminus A)$. By monotonicity together with the fact that $E \setminus A \subseteq B \setminus A$ we conclude that $E \setminus A$ is a null-set and is therefore measurable. Now, we may write $E = A \sqcup (E \setminus A)$, whence E is measurable and the result follows from additivity of μ over $\mathcal{M}(\mathbb{R})$. □

Problem 7. Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function, the curve of f in the plane $\mathbb{R}^2$ is defined by

$$\Gamma := \{(x, f(x)) : x \in \mathbb{R}\}$$

Show that $\mu(\Gamma) = 0$ in $\mathbb{R}^2$.

SOLUTION. By monotonicity together with the fact that $\mathbb{R}$ may be covered with countably many compact intervals of the form $[a, b]$. Hence, we would rather show that the portion of the curve lying in the image of $[a, b]$ has measure 0 in $\mathbb{R}^2$. Since f is continuous and $[a, b]$ is compact it follows that f is uniformly continuous on $[a, b]$.

Fix $\varepsilon > 0$, there is some $\delta > 0$ such that for all pairs $(x, y) \in [a, b]^2$ with $|x - y| < \delta$ one has $|f(x) - f(y)| < \varepsilon$. We may cover this interval $[a, b]$ by $3\left(\frac{b-a}{\delta}\right)$ intervals of length less than δ. Now, take an open rectangle of side δ and height, say, 2ε. We claim first that this collection covers $\Gamma \cap [a, b]$. Indeed, for $x \in [a, b]$ such that $(x, f(x)) \in \Gamma$ the point x must lie in $(x_0 - \delta/2, x_0 + \delta/2)$ for some x_0. Then,

$$|f(x) - f(x_0)| < \varepsilon$$

proving that $(x, f(x))$ lies in the cube. Now, summing over the area of all such cubes gives us

$$\mu(\Gamma \cap [a, b]) \leq \frac{3(b-a)}{\delta} \cdot \delta\varepsilon = 3(b-a)\varepsilon$$

Proving the result. □

Problem 8. The statement of this problem is quite long, and we refer the reader to [1] for details.

SOLUTION.

(a) Let $\varepsilon > 0$ be given, since $\delta(y) = \inf_{z \in F} |y - z|$ there is some $z \in F$ such that $\delta(y) \leq |z - y| < \delta(y) + \varepsilon$. By the triangle inequality we then have

$\delta(x) \leq |x-z| \leq |x-y| + |y-z| < \delta(y) + \varepsilon + |x-y|$ whence taking $\varepsilon \to 0$ we obtain:

$$\delta(x) \leq \delta(y) + |x-y| \iff \delta(x) - \delta(y) \leq |x-y|$$

Similarly, for such $\varepsilon > 0$ there is $w \in F$ such that $\delta(x) \leq |x-w| < \delta(x) + \varepsilon$. Therefore,

$$\delta(y) \leq |y-w| \leq |x-y| + |x-w| \leq |x-y| + \delta(x) + \varepsilon$$

again, we find then that $\delta(y) - \delta(x) \leq |x-y| + \varepsilon$ and taking $\varepsilon \to 0$ yields

$$\delta(x) - \delta(y) \geq -|x-y|$$

which concludes this part.

(b) Define $I(x) = \int_{\mathbb{R}} \frac{\delta(y)}{|x-y|^2}\, \mathrm{d}y$. We claim that if $x \notin F$ then $I(x) = \infty$. This proof will be filled in later (I dislike this problem).

□

Problem 9. Let $f : \mathbb{R} \to \mathbb{R}$ be uniformly continuous. If f is integrable, then $f(x)$ vanishes at infinity.

SOLUTION. We argue by contradiction. Suppose not, that is suppose

$$\neg(\forall \varepsilon > 0)(\exists M > 0)(|x| \geq M \implies |f(x)| < \varepsilon)$$

The above is logically equivalent the existence of some $\varepsilon_0 > 0$ such that for all $M > 0$ there is some $|x| \geq M$ with $|f(x)| \geq 2\varepsilon_0$. For such an ε_0 there is some $\delta > 0$ such that for all x, y with $|x-y| < \delta$ we are guaranteed $|f(x) - f(y)| < \varepsilon_0$. There must then exist a sequence of $(x_n) \in \mathbb{R}$ with $|x_n| \to \infty$ with $|f(x_n)| \geq 2\varepsilon_0$ for all indices n. Since we are guaranteed infinitely many such x_n going to infinity in norm, we are free to take them such that the intervals $(x_n - \delta, x_n + \delta)$ are disjoint for distinct n.

Note that for all $x \in (x_n - \delta, x_n + \delta)$ we have $|x - x_n| < 2\delta$ and so $|f(x_n) - f(x)| < \varepsilon_0$ and thus

$$|f(x_n)| - \varepsilon_0 \leq |f(x)|$$

implying $|f(x)| \geq \varepsilon_0$. Thus,

$$\|f\|_{L^1} \geq \int_{\bigcup_n (x_n - \delta, x_n + \delta)} |f(x)|\ \mathrm{d}x \geq \varepsilon_0 \int_{\bigcup_n (x_n - \delta, x_n + \delta)} \mathrm{d}x = \infty$$

which is a contradiction. □

Problem 10. Let $f \in L^1(\mathbb{R})$. Then, $F(x) := \int_{-x}^{\infty} f(t)\ \mathrm{d}t$ is uniformly continuous.

SOLUTION. Recall that for each $\varepsilon > 0$ there is a $\delta > 0$ so that whenever $\mu(E) < \delta$ one has $\int_E |f| < \varepsilon$. Let now $|x - y| < \delta$ but suppose without loss of generality that $y < x$. In the difference we find:

$$|F(x) - F(y)| \leq \int_y^x |f(t)|\ \mathrm{d}t < \varepsilon$$

□

Problem 11 (Chebyshev's Inequality)**.** If $f \geq 0$ is integrable. Define for each $\alpha > 0$ a set $E_\alpha := \{x \in \mathbb{R}^d : f(x) > \alpha\}$. Then,

$$\mu(E_\alpha) \leq \frac{1}{\alpha} \int_{\mathbb{R}} f(x)\ \mathrm{d}x$$

SOLUTION. This follows essentially from monotonicity:

$$\alpha\mu(E_\alpha) = \int_{E_\alpha} \alpha\ \mathrm{d}x \leq \int_{E_\alpha} f(x)\ \mathrm{d}x \leq \|f\|_{L^1(\mathbb{R}^d)}$$

□

Problem 12. Let $f \in L^1(\mathbb{R})$ and suppose that for each $E \in \mathcal{M}(\mathbb{R}^d)$ one has $\int_E f(x)\ \mathrm{d}x \geq 0$. Then $f \geq 0$ almost everywhere. Similarly, if $\int_E f(x)\ \mathrm{d}x = 0$ for all $E \in \mathcal{M}(\mathbb{R}^d)$ then $f = 0$ almost everywhere.

SOLUTION. Assume first $\int_E f(x)\ \mathrm{d}x \geq 0$ for all $E \in \mathcal{M}(\mathbb{R}^d)$. Define for $n \in \mathbb{N}$:

$$A_n := \left\{x \in \mathbb{R}^d : f(x) < -\frac{1}{n}\right\}$$

Clearly, $A_n \in \mathcal{M}(\mathbb{R}^d)$ and $\{x \in \mathbb{R}^d : f(x) < 0\} = \bigcup_{n \in \mathbb{N}} A_n$. We claim $\mu(A_n) = 0$ for all n, indeed write

$$0 \leq \int_{A_n} f(x)\ \mathrm{d}x < -\frac{1}{n}\mu(A_n)$$

since $\mu(A_n) \geq 0$ we conclude $\mu(A_n) = 0$, as was required. For the second part, if $\int_E f(x)\ \mathrm{d}x = 0$ for all $E \in \mathcal{M}(\mathbb{R}^d)$ by the above argument we have $f \geq 0$ almost everywhere. Let

$$B_n := \left\{x \in \mathbb{R}^d : f(x) > \frac{1}{n}\right\}$$

so that $\left\{x \in \mathbb{R}^d : f(x) > 0\right\} = \bigcup_{n\in\mathbb{N}}$. Now, to see that $\mu(B_n) = 0$ for all n note that

$$0 = \int_{B_n} f(x)\ \mathrm{d}x \geq \frac{1}{n}\mu(B_n)$$

Now, the result follows from sub-additivity. □

Problem 13. Define a function $f : \mathbb{R} \to \mathbb{R}$ by letting:

$$f(x) := \begin{cases} x^{-1/2} & 0 < x < 1, \\ 0 & \text{else.} \end{cases}$$

Now fix an enumeration $\{r_n\}_{n\in\mathbb{N}}$ of $\mathbb{Q}$. Define now:

$$F(x) := \sum_{n=1}^{\infty} 2^{-n} f(x - r_n)$$

Then, F is integrable but F is unbounded on any interval.[15]

SOLUTION. We first show that $\int_{\mathbb{R}} f(x)\ \mathrm{d}x < \infty$. For $k \in \mathbb{N}$ let us define a function:

$$f_k(x) := \begin{cases} x^{-1/2} & \frac{1}{k+1} < x < 1 \\ 0 & \text{else} \end{cases}$$

Clearly, $f_k(x) \to f(x)$ for almost every x as $k \to \infty$ and the f_k's are Riemann integrable. We may then apply the fundamental theorem of calculus to deduce:

$$\int_{\mathbb{R}} f_k(x)\ \mathrm{d}x = \int_{\frac{1}{k+1}}^{1} \frac{1}{\sqrt{x}}\ \mathrm{d}x = 2 - \frac{2}{\sqrt{k+1}}$$

By monotone convergence,

$$\int_{\mathbb{R}} f(x)\ \mathrm{d}x = \lim_{k\to\infty} \int_{\mathbb{R}} f_k(x)\ \mathrm{d}x = \lim_{k\to\infty} \left(2 - \frac{2}{\sqrt{k+1}}\right) = 2$$

Hence, f is integrable. Now, since all terms in the series are non-negative we

[15] A singleton in not considered an interval here.

write:

$$\begin{aligned}
\int_{\mathbb{R}} \left[\sum_{n\in\mathbb{N}} 2^{-n} f(x - r_n) \right] \, \mathrm{d}x &= \sum_{n\in\mathbb{N}} 2^{-n} \int_{\mathbb{R}} f(x - r_n) \, \mathrm{d}x \\
&= \sum_{n\in\mathbb{N}} 2^{-n} \int_{\mathbb{R}} f(x) \, \mathrm{d}x \qquad \text{(translation invariance)} \\
&= 2 \sum_{n\in\mathbb{N}} 2^{-n} < \infty
\end{aligned}$$

Now, to see that f is unbounded on any interval note that it suffices to show this for any open interval $(a, b) \subset \mathbb{R}$. By the density of $\mathbb{Q}$ in $\mathbb{R}$ we may extract some $r_N \in (a, b)$. Now, for any $M > 0$ there exists $\eta > 0$ small so that whenever $0 < x < \eta$ one has

$$\frac{1}{\sqrt{x}} > M$$

because $\lim_{x\to 0} x^{-1/2} = \infty$. Therefore, by approaching the point $x = r_N$ the function $f(x)$ can be made arbitrarily large. □

Problem 14. Let $f : [0, 1] \to \mathbb{C}$ be measurable and suppose $|f(x)| < \infty$ for all $x \in [0, 1]$. If $|f(x) - f(y)|$ is integrable on $[0, 1] \times [0, 1]$ then f is integrable on $[0, 1]$.

SOLUTION. This is mostly an application of Fubini's theorem. We write out:

$$\begin{aligned}
\int_0^1 |f(x)| \, \mathrm{d}x &\le \int_0^1 |f(x) - f(y)| \, \mathrm{d}x + \int_0^1 |f(y)| \, \mathrm{d}x \\
&= \int_0^1 |f(x) - f(y)| \, \mathrm{d}x + f(y)
\end{aligned}$$

Now, $|f(y)| < \infty$ and by Fubini $\int_0^1 |f(x) - f(y)| < \infty$ for almost all y. Taking y such that this holds true we have that f is integrable on $[0, 1]$. □

These are problem(s) from assignments that did not appear in [1] or [2].

Problem 15. Let $f : \mathbb{R} \to \mathbb{R}$ be a monotone function, then f is measurable.

SOLUTION. Without loss of generality assume that f is increasing, replacing f with $-f$ otherwise. Let $\alpha \in \mathbb{R}$ be given, we consider the pre-image $\{f > \alpha\}$. If this set is empty we are done. Otherwise, it is non-empty, and set $a = \inf_{x\in\mathbb{R}}\{f(x) > \alpha\}$. If $a = -\infty$ then $\{f > \alpha\} = \mathbb{R}$ by monotonicity, else $a > -\infty$ and the pre-image $\{f > \alpha\}$ is either of the form (a, ∞) or $[a, \infty)$. □

References

[1] E. Stein, R. Shakarchi. *Measure Theory, Integration and Hilbert Spaces*, Princeton Lectures in Analysis, 3rd edition, 2007.

[2] E. Stein, R. Shakarchi. *Complex Analysis*, Princeton Lectures in Analysis, 3rd edition, 2007.

[3] Partial Differential Equations, Lawrence C. Evans. volume 19, Graduate Studies in Mathematics.

[4] Partial Differential Equations, Jurgen Jost. Graduate Texts in Mathematics. ISBN: 978-1-4614-4809-9; Springer.