1. L^2-spaces

In this section we establish that the space of all square integrable functions form a Hilbert space. To start, consider all continuous functions on some interval I which may be the half line or the whole real line and define

$$L^2(I) = \{f : \int_I |f(x)|^2 dx < \infty\} \ .$$

It is quite easy to verify that $L^2(I)$ is a linear space with inner product

$$(f,g) = \int_I \overline{f(x)} g(x) dx \ .$$

Unfortunately, this spaces is not complete. Consider $I = [-1,1]$ and the sequence of functions $f^{(k)}(x) = -1$ for $-1 \le x \le -\frac{1}{k}$, $f^{(k)}(x) = kx$ for $-\frac{1}{k} \le x \le \frac{1}{k}$ and $f^{(k)}(x) = 1$ for $\frac{1}{k} \le x \le 1$. Clearly these functions are continuous for each $k = 1, 2, \dots$. If $\ell \ge k$ we have that

$$\int_I |f^{(\ell)}(x) - f^{(k)}(x)|^2 dx = \int_{-\frac{1}{k}}^{\frac{1}{k}} |f^{(\ell)}(x) - f^{(k)}(x)|^2 dx \le 4 \times \frac{2}{k}$$

from which we see that $f^{(k)}$ is a Cauchy sequence. The limit of this sequence, however, is not a continuous function and the limit is not in our linear space. There is a process of completing this space at the price that the integral has to be interpreted according to Lebesgue.

The idea is the following. Consider a positive function f and we want to give a definition of

$$\int_I f(x) dx \ .$$

We interpret this integral as the area underneath the graph of f over I. One way of approximating this area is according to Riemann which you have learned in your analysis course. Another one is to look at the **length** of the **level sets** of the function f which is given by

$$\{x \in I : f(x) > t\} \ .$$

If we denote by $|\{x \in I : f(x) > t\}|$ the length of these level sets we can think of the area as

$$\int_0^M |\{x \in I : f(x) > t\}| dt \ , \tag{1}$$

where M is the maximal value of f. Note that $|\{x \in I : f(x) > t\}|$ is a decreasing function of t and hence it is Riemann integrable.

Now observe, that the level sets can be quite crazy sets that do not necessarily have a length. So the first step is to state properties that such sets must have in order to have a chance of making sense out of this integral. We call such sets **measurable** and require the following:

a) If $A \subset I$ is measurable, so is its complement A^c.
b) I is measurable.
c) If $A_1, A_2, \dots$ is a countable family of measurable sets, then their union is also measurable.

Any collection of sets that have the above properties we call a **sigma algebra**.

In a further step we now define what we mean by the volume of such sets,i.e., the **measure** of such sets. A **measure** μ is a function from a sigma algebra Σ into the positive real numbers that has the following properties

a) $\mu(A) \le \mu(B)$ if $A \subset B$ and $A, B \in \Sigma$.

b) Let $A_1, A_2, \ldots$ be a **countable** collection of **disjoint** sets in Σ. Then

$$\mu(\cup_{j=1}^{\infty} A_j) = \sum_{j=1}^{\infty} \mu(A_j) \ .$$

This last property is called **countable additivity** of the the measure μ. This property is the key in establishing completeness of spaces of integrable functions.

A consequence of the countable additivity are the following two statements:

a) If $A_1 \subset A_2 \subset \cdots$ is an increasingly nested sequence of sets in Σ, then

$$\lim_{N\to\infty} \mu(\cup_{j=1}^{N} A_j) = \mu(\cup_{j=1}^{\infty} A_j)$$

and

b) If $A_1 \supset A_2 \supset \cdots$ is a decreasingly nested sequence of set in Σ, then

$$\lim_{N\to\infty} \mu(\cap_{j=1}^{N} A_j) = \mu(\cap_{j=1}^{\infty} A_j) \ .$$

Now we close in on our definition of the integral. A function $f : I \to \mathbb{R}_+$ is **measurable** if the sets $\{x \in I : f(x) > t\}$ are measurable for all $t \in \mathbb{R}$.

Given a non-negative measurable function f and a measure μ we say that the function is **summable** or **integrable** if

$$\int_I f(x)\mu(dx) := \int \mu(\{x \in I : f(x) > t\})dt < \infty$$

where, as before, the last integral is a Riemann integral, since the function $t \to \mu(\{x \in I : f(x) > t\})$ is decreasing.

Remark 1.1. *There could be sets that have zero measure. Thus modifying the function on a set of zero measure would not affect the integral. We say that a certain property holds* **almost everywhere with respect to** μ *if the set where the property does not hold has zero* μ *measure.*

There are two important theorems that follow from these definitions.

Theorem 1.2. Monotone convergence *Let $f^{(k)}$ be a sequence of summable functions and assume $f^{(k)}(x) \le f^{(k+1)}(x)$ for almost all $x \in I$. Then the limit*

$$\lim_{k\to\infty} f^{(k)}(x) := f(x)$$

exists for almost every x and is measurable. Moreover,

$$\lim_{k\to\infty} \int_I f^{(k)}(x)\mu(dx) = \int_I f(x)\mu(dx) \ ,$$

in the sense that if one of the quantities is infinite so is the other.

The other important theorem is

Theorem 1.3. Dominated convergence *Let $f^{(k)}(x)$ be a sequence of summable functions that converges almost everywhere with respect to μ to a function f. If there exists a summable function $G(x)$ such that*

$$|f^{(k)}(x)| \le G(x)$$

for all $k = 1, 2, 3, \dots$, then $f(x)$ is measurable, summable and

$$\lim_{k\to\infty} \int_I f^{(k)}(x)\mu(dx) = \int_I f(x)\mu(dx) \ .$$

The big questions is whether such σ algebras and measures exist. This is the hard part of the theory and you are referred to the books on measure theory. on the real line there exists a unique translation invariant measure $\mathcal{L}$, the Lebesgue measure. Translation invariant means that $\mathcal{L}(B) = \mathcal{L}(A)$ whenever B is a translate of A.

The beauty of all this is that it works in great generality. We can replace the interval I by any set Ω and μ be any measure on a sigma algebra of subsets of Ω. The theorems stated above continue to hold in this case too.

Definition 1.4. **$L^2(\Omega, \mu)$-space** *This space consists of all square summable functions $f : \Omega \to \mathbb{C}$.*

We are ready to state the important

Theorem 1.5. **Riesz-Fischer** *The space $L^2(\Omega, \mu)$ endowed with the inner product*

$$(f, g) = \int_\Omega \overline{f(x)} g(x)\mu(dx)$$

is a Hilbert space.

Proof. Let $f^{(k)}$ be a Cauchy sequence in $L^2(\Omega, \mu)$. This means that for any $\varepsilon > 0$ there exists N so that for all $k, \ell > N$

$$\|f^{(k)} - f^{(\ell)}\| < \varepsilon \ .$$

Hence, there exists k_1 so that for all $\ell > k_1$

$$\|f^{(k_1)} - f^{(\ell)}\| < \frac{1}{2} \ .$$

Likewise, there exists $k_2 > k_1$ so that for all $\ell > k_2$

$$\|f^{(k_2)} - f^{(\ell)}\| < \frac{1}{2^2} \ .$$

Continuing this way we find a sequence $k_1, k_2, k_3, \dots$ such that for all $j = 1, 2, \dots$

$$\|f^{(k_j)} - f^{(k_{j+1})}\| < \frac{1}{2^j} \ .$$

Now consider the sequence $f^{(k_j)}$ and write

$$f^{(k_j)} = f^{(k_1)} + [f^{(k_2)} - f^{(k_1)}] + [f^{(k_3)} - f^{(k_2)}] + \cdots + [f^{(k_j)} - f^{(k_{j-1})}] \ .$$

If we set

$$F^{(j)} = |f^{(k_1)}| + |f^{(k_2)} - f^{(k_1)}| + |f^{(k_3)} - f^{(k_2)}| + \cdots + |f^{(k_j)} - f^{(k_{j-1})}| \ ,$$

we obviously have that

$$|f^{(k_j)}| \le F^{(j)} \ .$$

The sequence $F^{(j)}(x)$ is a monotone increasing sequence and hence converges to a function $F(x)$. This implies that the sequence $f^{(k_j)}(x)$ converges to some function $f(x)$ since the partial sums converge absolutely. Further,

$$\|F^{(j)}\| \le \|f^{(k_1)}\| + \|f^{(k_2)} - f^{(k_1)}\| + \|f^{(k_3)} - f^{(k_2)}\| + \cdots + \|f^{(k_j)} - f^{(k_{j-1})}\|$$

which is bounded above by

$$\|f^{(k_1)}\| + \frac{1}{2} + \frac{1}{2^2} + \cdots + \frac{1}{2^j} < \|f^{(k_1)}\| + 1 \ .$$

Hence by the monotone convergence theorem we find that F is square summable and

$$|f^{(k_j)}(x) - f(x)| \le |f^{(k_j)}(x)| + |f(x)| \le 2F(x) \ .$$

Since $f^{(k_j)}(x)$ converges to $f(x)$ for every x we have by the dominated convergence theorem that

$$\lim_{j\to\infty} \int_\Omega |f^{(k_j)}(x) - f(x)|^2 \mu(dx) = 0 \ .$$

In other words we have for the *subsequence* k_j that

$$\lim_{j\to\infty} \|f^{(k_j)} - f\| = 0 \ .$$

We have to show that the whole sequence converges. For this, fix and $\varepsilon > 0$ and pick N such that for $k_j > N$, $\|f^{(k_j)} - f\| < \varepsilon/2$ and for $\ell > N$, $\|f^{(k_j)} - f^{(\ell)}\| < \varepsilon/2$. Then for all $\ell > N$

$$\|f^{(\ell)} - f\| \le \|f^{(k_j)} - f^{(\ell)}\| + \|f^{(k_j)} - f\| < \varepsilon.$$

□