

# Math 202B Lecture Notes

Sky Cao

Spring 2015

# Chapter 1

## Integration

### 1.1 Measurable Functions

Let  $X$  be a set,  $S$  a  $\sigma$ -ring of subsets,  $B$  a Banach space. Given the measurable space  $(X, S)$ , we wish to answer the question: what kinds of functions  $f : X \rightarrow B$  can we integrate?

Given  $E \in S$ , let  $\chi_E$  be its characteristic function. Given the Banach space  $B$ , for  $b \in B$ , we can obtain the function  $b\chi_E$ , i.e.  $b\chi_E : X \rightarrow B$  is such that:

$$b\chi_E = \begin{cases} b & x \in E \\ 0 & x \notin E \end{cases}$$

**Definition.** A “simple  $S$ -measurable  $B$ -valued” function is a function  $f : X \rightarrow B$  such that the range of  $f$  is finite, and if  $b \in \text{range}(f)$ ,  $b \neq 0$ , then  $f^{-1}(b) \in S$ .

**Definition.**  $\text{MSF}(X, S, B)$  is the collection of all simple  $S$ -measurable  $B$ -valued functions from  $X$  to  $B$ . It will often be denoted **MSF**.

**Remark.** For all  $f \in \text{MSF}(X, S, B)$ ,  $f = \sum_{j=1}^n b_j \chi_{E_j}$  with the  $b_j$  distinct and not equal to 0.

**Remark.**  $\text{MSF}(X, S, B)$  is a vector space for the pointwise operations.

Let  $\mu$  be a measure on  $S$  (or at least finitely additive, if  $S$  is a ring).

**Definition.** A “simple  $\mu$ -integrable  $B$ -valued function” is a function  $f$  in **MSF** such that  $f = \sum_{j=1}^n b_j \chi_{E_j}$  with the  $E_j$  disjoint and  $\mu(E_j) < \infty$  for all  $j$ . Note we don’t require the  $b_j$  to be distinct or not equal to 0.

**Definition.**  $\text{ISF}(X, S, \mu, B)$  is the collection of all simple  $\mu$ -integrable  $B$ -valued functions. It will often be denoted **ISF**.

**Definition.** The **carrier** of a function  $f : X \rightarrow B$ , denoted  $C(f)$ , is defined as:

$$C(f) = \{x \in X \mid f(x) \neq 0\}$$

We want to define integration in such a way so that if  $f \in \text{ISF}(X, S, \mu, B)$ , we have:

$$\int f d\mu = \sum_j b_j \mu(E_j)$$

But first we need to ask if this expression is well-defined, because  $f \in \text{ISF}(X, S, \mu, B)$  can be represented as a sum in more than one way.

**Proposition 1.1.1.**  $\int f d\mu$  is well-defined.

*Proof.* Let  $f = \sum_j b_j \chi_{E_j} = \sum_k c_k \chi_{F_k}$ , with the  $E_j$ 's disjoint and the  $F_k$ 's disjoint. We can assume all the  $b_j, c_k \neq 0$ , because it is clear that taking them out will not affect the value of  $\int f d\mu$ . We then have:

$$C(f) = \bigoplus_j E_j = \bigoplus_k F_k$$

Note  $\chi_{E_j} \chi_{F_k} = \chi_{E_j \cap F_k}$ . Also note  $\{E_j \cap F_k\}_{j,k}$  is a disjoint collection of sets. We have:

$$E_j = \bigoplus_k E_j \cap F_k \text{ and } F_k = \bigoplus_j E_j \cap F_k$$

We can thus obtain:

$$f = \sum_{j,k} b_j \chi_{E_j \cap F_k} = \sum_{k,j} c_k \chi_{E_j \cap F_k}$$

So if  $E_j \cap F_k \neq \emptyset$  for some  $j, k$ , then  $b_j = c_k$  necessarily. This shows that  $\int f d\mu$  is well-defined, because:

$$\begin{aligned} \sum_j b_j \mu(E_j) &= \sum_{j,k} b_j \mu(E_j \cap F_k) \\ &= \sum_{E_j \cap F_k \neq \emptyset} b_j \mu(E_j \cap F_k) \\ &= \sum_{E_j \cap F_k \neq \emptyset} c_k \mu(E_j \cap F_k) \\ &= \sum_{k,j} c_k \mu(E_j \cap F_k) \\ &= \sum_k c_k \mu(F_k) \end{aligned}$$

□

We are making progress. We now have a concept of integration for some relatively simple functions.

**Definition.** For  $G \subseteq X$ , we say that  $G$  is **locally  $S$ -measurable** if  $G \cap E \in S$  for every  $E \in S$ .

**Definition.**  $M_{\text{loc}}(S)$  is the set of locally  $S$ -measurable subsets. Note it is a  $\sigma$ -ring if  $S$  is a  $\sigma$ -ring.

**Remark.** For any  $G \in M_{\text{loc}}(S)$ , we can define:

$$\int_G f d\mu = \sum_j b_j \mu(E_j \cap G)$$

Note  $\int f d\mu = \int_X f d\mu$ .

**Definition.** The **indefinite integral** of  $f$  for  $\mu$  is defined to be the function from  $M_{\text{loc}}(S)$  to  $B$  that takes  $G \mapsto \int_G f d\mu$ .

**Remark.** The indefinite integral is finitely additive:

$$\begin{aligned} \int_{G \oplus H} f d\mu &= \sum_j b_j \mu(E_j \cap (G \oplus H)) \\ &= \sum_j b_j (\mu(E_j \cap G) + \mu(E_j \cap H)) \\ &= \int_G f d\mu + \int_H f d\mu \end{aligned}$$

**Proposition 1.1.2.** Let  $f, g \in ISF$ . Then  $\int (f + g) d\mu = \int f d\mu + \int g d\mu$ . Also, for  $r \in \mathbb{R}$ , we have  $\int (rf) d\mu = r \int f d\mu$ . Thus  $f \mapsto \int f d\mu$  is a linear operator from  $ISF$  to  $B$ .

*Proof.* The scalar multiplication part is easy enough to see. We will focus on the additivity part. Let  $f = \sum_j b_j \chi_{E_j}$  and  $g = \sum_k c_k \chi_{F_k}$ . By adding terms with 0 coefficients on each side, we can ensure that:

$$\bigoplus_j E_j = \bigoplus_k F_k$$

We can then write:

$$f = \sum_{j,k} b_j \chi_{E_j \cap F_k} \text{ and } g = \sum_{k,j} c_k \chi_{E_j \cap F_k}$$

Thus:

$$f + g = \sum_{j,k} (b_j + c_k) \chi_{E_j \cap F_k}$$

We then have:

$$\begin{aligned} \int (f + g) d\mu &= \sum_{j,k} (b_j + c_k) \mu(E_j \cap F_k) \\ &= \sum_{j,k} b_j \mu(E_j \cap F_k) + \sum_{k,j} c_k \mu(E_j \cap F_k) \\ &= \int f d\mu + \int g d\mu \end{aligned}$$

□

Now assume that  $B$  has a norm. Given any  $B$ -valued function  $f$ , we can set  $\|f\|(x) = \|f(x)\|_B$ .

**Proposition 1.1.3.** *For  $f \in ISF$ :*

- (a)  $\|f\| \in ISF$  ( $\mathbb{R}$ -valued)
- (b)  $\left\| \int f d\mu \right\| \leq \int \|f\| d\mu$
- (c) Set  $\|f\|_1 = \int \|f\| d\mu$ . Then  $f \mapsto \|f\|_1$  is a seminorm on  $ISF$ .

*Proof.* Let  $f = \sum_j b_j \chi_{E_j}$ . Then  $\|f\| = \sum_j \|b_j\|_B \chi_{E_j}$ . This proves (a). For (b), note:

$$\left\| \int f d\mu \right\|_B = \left\| \int \sum_j b_j \mu(E_j) \right\|_B \leq \sum_j \|b_j\|_B \mu(E_j) = \int \|f\| d\mu$$

And finally, to prove that (c) is a seminorm, we need to prove that it satisfies the triangle inequality and that it preserves scalar multiplication. The scalar multiplication part is apparent; we will focus on proving that the triangle inequality is satisfied. Let  $f = \sum_j b_j \chi_{E_j}$ ,  $g = \sum_k c_k \chi_{F_k}$ . Arrange that  $\bigoplus E_j = \bigoplus F_k$ . Then:

$$f = \sum_{j,k} b_j \chi_{E_j \cap F_k} \text{ and } g = \sum_{k,j} c_k \chi_{E_j \cap F_k}$$

and thus:

$$f + g = \sum_{j,k} (b_j + c_k) \chi_{E_j \cap F_k}$$

So:

$$\begin{aligned}
\|f + g\|_1 &= \int \|f + g\| d\mu \\
&= \sum_{j,k} \|b_j + c_k\|_B \mu(\chi_{E_j \cap F_k}) \\
&\leq \sum_{j,k} \|b_j\|_B \mu(\chi_{E_j \cap F_k}) + \sum_{k,j} \|c_k\|_B \mu(\chi_{E_j \cap F_k}) \\
&\leq \|f\|_1 + \|g\|_1
\end{aligned}$$

Note  $\|\cdot\|_1$  is not a norm, because we can define  $f = b\chi_E$ , where  $b \neq 0$ ,  $E \neq \emptyset$ ,  $\mu(E) = 0$ . Then  $\|f\|_1 = b\mu(E) = 0$ , but  $f \neq 0$ .  $\square$

**Definition.** We define  $\text{NISF} = \{f \in \text{ISF} \mid \|f\|_1 = 0\}$ . NISF is a vector subspace of ISF. Note  $f \in \text{NISF}$  if and only if  $f = \sum_j b_j \chi_{E_j}$  where the  $E_j$ 's are disjoint,  $b_j \neq 0$ , and  $\mu(E_j) = 0$ . In other words  $f \in \text{NISF}$  if and only if  $\mu(C(f)) = 0$ .

**Definition.** We say that a property  $P(\cdot)$  that depends on  $x \in X$  holds  $\mu$ -almost everywhere if  $\{x \mid P(x) \text{ is not true}\}$  is a null set for  $\mu$ . We often write this as **a.e.** $\mu$ .

**Remark.** We see that  $f \in \text{NISF}$  if and only if  $f(x) = 0$  a.e. $\mu$ .

We can think of NISF as the set of functions that is keeping  $\|\cdot\|_1$  from being a norm. To turn  $\|\cdot\|_1$  into a norm, we must “zero” out this set. This equates to taking the quotient space  $\text{ISF} / \text{NISF}$ . Formally, we have:

$$\text{ISF} / \text{NISF} = \{f + \text{NISF} \mid f \in \text{ISF}\}$$

Note that the elements of the quotient space are actually equivalence classes defined by the relation  $f \sim g$  if  $f - g \in \text{NISF}$ . This is the same as saying  $f \sim g$  if  $f = g$  a.e. $\mu$ . Thus  $\|\cdot\|_1$  “drops” to a norm on  $\text{ISF} / \text{NISF}$ .

From this norm, we obtain a metric,  $d(f, g) = \|f - g\|_1$ . One question that we can naturally ask is whether the space is complete. The answer is usually no. We then may ask what exactly is the space’s completion? This is where the notion of the abstract completion of a metric space comes in. Given any non-complete metric space  $(M, d)$ , we can form its abstract completion, which consists of equivalence classes of Cauchy sequences. Given two Cauchy sequences  $\{x_n\}, \{y_n\}$ , we say that they are equivalent if the sequence  $\{d(x_n, y_n)\}$  converges to 0.

Denote this set of equivalence classes by  $\tilde{M}$ . We can define a metric on  $\tilde{M}$  by:

$$\tilde{d}(\{x_n\}, \{y_n\}) = \lim_{n \rightarrow \infty} d(x_n, y_n)$$

We can check that with this metric,  $\tilde{M}$  is complete. We can then identify  $M$  as a subset of  $\tilde{M}$  by  $x \mapsto \{x\}$ , the constant sequence  $x_n = x$ . Then  $\tilde{d}|_M = d$ . One can show that  $M$  is dense in  $\tilde{M}$ .

If we have a vector space  $V$  and a norm  $\|\cdot\|$  on  $V$ , we can form its abstract completion  $\tilde{V}$ . On  $\tilde{V}$ , we have a natural sense of addition and scalar multiplication, making  $\tilde{V}$  into a vector space, with  $V$  a vector subspace of  $\tilde{V}$ . The norm on  $V$  extends to a norm on  $\tilde{V}$ .

With this discussion in mind, we can form the abstract completion of ISF/NSF for  $\|\cdot\|_1$ . We wish to represent the elements of this completion by functions, because they are easier to think about and deal with than equivalence classes of Cauchy sequences. The functions we obtain will be the  $\mu$ -integrable functions. We will thus begin to discuss sequences of integrable simple functions. For a while, we really only need  $B$  to be a metric space with the special point 0.

**Definition.** A function  $f : X \rightarrow B$  is  **$S$ -measurable** if there is a sequence  $\{f_n\} \subseteq MSF$  that converges pointwise to  $f$ .

**Definition.**  $\mathcal{M}(X, S, B)$  is defined to be the set of  $S$ -measurable functions.

**Definition.** Given a measure space  $(X, S, \mu)$ , a  $B$ -valued function  $f$  is  **$\mu$ -measurable** if there is a sequence  $\{f_n\} \subseteq MSF$  that converges pointwise to  $f$  a.e. $\mu$ .

**Remark.**  $\mu$ -measurability is weaker than  $S$ -measurability.

**Remark.** The definition of  $\mu$ -measurability allows  $f$  to be undefined on a null set.

Question: If  $\{f_n\}$  is a sequence in  $\mathcal{M}(X, S, B)$ , and if  $f_n \rightarrow f$  pointwise for some  $f : X \rightarrow B$ , is  $f \in \mathcal{M}(X, S, B)$ ? To answer this question takes quite a bit of work, which we carry out here.

## 1.2 Properties of $\mathcal{M}(X, S, B)$

Let  $\{f_n\} \subseteq MSF$ ,  $f_n \rightarrow f$  pointwise. Let  $R_n = \text{range } f_n$ , a finite subset of  $B$ . Then  $\bigcup_n R_n$  is countable. We can show:

$$\text{range } f \subseteq \overline{\bigcup_n R_n}$$

We thus see that  $\text{range } f$  is separable. We can generalize and strengthen this result:

**Proposition 1.2.1.** *Let  $\{f_n\}$  be a sequence of functions from  $X$  to  $B$ , such that  $\text{range } f_n$  is separable for all  $n$ . Let  $f_n \rightarrow f$  pointwise. Then  $\text{range } f$  is separable.*

*Proof.* Let  $M_n$  be a countable dense subset of  $\text{range } f_n$ . Let  $M = \bigcup_n M_n$ . Note  $M$  is countable. Let  $f(x) \in \text{range } f$ . Then  $f(x) = \lim_{n \rightarrow \infty} f_n(x)$ . Note  $\{f_n(x)\} \subseteq M$ . Thus we see that  $f(x)$  is a limit point of  $M$ , and so  $f(x) \in \bar{M}$ . Thus  $\text{range } f \subseteq \bar{M}$ . Therefore  $\text{range } f$  is separable.  $\square$

**Corollary 1.2.2.** *If  $f \in \mathcal{M}(X, S, B)$ , then  $\text{range } f$  is separable.*

**Corollary 1.2.3.** *If  $f$  is the pointwise limit of a sequence in  $\mathcal{M}(X, S, B)$ , then  $\text{range } f$  is separable.*

Note also that if  $f \in MSF$ , then given any open set  $U \subseteq B$ , we have  $f^{-1}(U \setminus \{0\}) \in S$ . This motivates another proposition:

**Proposition 1.2.4.** *Let  $\{f_n\}$  be a sequence of functions from  $X$  to  $B$  having the property that for each  $n$ , if  $U \subseteq B$  is open, then:*

$$f_n^{-1}(U \setminus \{0\}) \in S$$

*If  $\{f_n\}$  converges pointwise to some  $f$  and if  $S$  is a  $\sigma$ -ring, then  $f$  has the property that for all open  $U \subseteq B$ ,  $f^{-1}(U \setminus \{0\}) \in S$ .*

*Proof.* Let  $U$  be an open subset of  $B$ . For simplicity, we can assume  $0 \notin U$ . For each  $m$ , let:

$$U_m = \{b \in U \mid \text{dist}(b, U^c) > 1/m\}$$

Note:

$$U = \bigcup_m U_m$$

Consider the following chain of if and only if's:

$$\begin{aligned} x \in f^{-1}(U) &\iff f(x) \in U \\ &\iff \exists m, \exists K \text{ such that } \forall k \geq K, f_k(x) \in U_m \\ &\iff x \in \bigcup_{m=1}^{\infty} \bigcup_{K=1}^{\infty} \bigcap_{k \geq K} f_k^{-1}(U_m) \end{aligned}$$

This chain implies:

$$f^{-1}(U) = \bigcup_{m=1}^{\infty} \bigcup_{K=1}^{\infty} \bigcap_{k \geq K} f_k^{-1}(U_m)$$

Now since  $S$  is a  $\sigma$ -ring, we have that  $f^{-1}(U) \in S$ .  $\square$

**Corollary 1.2.5.** *If  $f \in \mathcal{M}(X, S, B)$ , then  $f$  has the property.*

**Corollary 1.2.6.** *If  $f$  is the pointwise limit of a sequence of functions in  $\mathcal{M}(X, S, B)$ , then  $f$  also has the property.*



So now we know that if  $f \in \mathcal{M}(X, S, B)$ , then  $\text{range } f$  is separable and for every open  $U \subseteq B$ ,  $f^{-1}(U \setminus \{0\}) \in S$ . Is the converse true? The answer is in fact yes.

**Proposition 1.2.7.** *Given a measurable space  $(X, S)$ , and a Banach space  $B$ , let  $f : X \rightarrow B$  satisfy:*

- (a) *range  $f$  is separable*
- (b) *for any open  $U \subseteq B$ , we have  $f^{-1}(U \setminus \{0\}) \in S$*

*Then  $f \in \mathcal{M}(X, S, B)$ , i.e. it is  $S$ -measurable, i.e. there is a sequence  $\{f_n\} \subseteq MSF$  converging pointwise to  $f$ .*

*Proof.* Before we get into the details, here is a high-level overview of the proof, as the proof itself is a bit technical. Because  $\text{range } f$  is separable, we can use a dense subset to “approximate”  $\text{range } f$ . We can then approximate  $f$  at some point  $x$  by finding an element  $b \in B$  that is say  $\varepsilon$ -close to  $f(x)$ , and then taking the preimage of the  $\varepsilon$ -ball about  $f(x)$ . Call this preimage  $E$ . Then the function  $bE$  is in some sense a good approximation of  $f$  at  $x$ . We can do this for every point in the range of  $f$ , and then add up all the  $bE$ ’s to obtain an approximation of  $f$  on all of  $X$ .

Let  $\{b_n\}$  be a sequence of elements of  $B$  that is dense in  $\text{range } f$ . We can let  $b_1 = 0$ . Define:

$$C_{ji} = f^{-1} \left( \left\{ b \mid \|b - b_i\| < \frac{1}{j}, b \neq 0 \right\} \right)$$

Note by definition  $C_{ji} \subseteq C(f)$  for all  $j, i$ . By hypothesis,  $C_{ji} \in S$ . We can put the lexicographic order on the pairs  $(j, i)$ . We want to “disjointize” the  $C_{ji}$  in some sense, so for all  $n \in \mathbb{N}$ , set:

$$E_{ji}^n = C_{ji} \setminus \bigcup_{(j,i) < (l,k) \leq (n,n)} C_{lk}$$

Now set:

$$f_n = \sum_{j,i \leq n} b_i \chi_{E_{ji}^n}$$

Note  $f_n \in MSF$ . We want to show that  $f_n \rightarrow f$  pointwise. Let  $x \in X$ . First note that if  $x \notin C(f)$ , then since  $C_{ji} \subseteq C(f)$  and  $E_{ji}^n \subseteq C_{ji}$ , we get that  $f_n(x) = 0$  for all  $n$ . So assume  $x \in C(f)$ . Let  $\varepsilon > 0$ . We need to find an  $N$  such that if  $n \geq N$ , then  $\|f(x) - f_n(x)\| < \varepsilon$ . Choose  $j_0$  such that  $1/j_0 < \varepsilon$ . Choose  $i_0$  with  $\|f(x) - b_{i_0}\| < 1/j_0$ . Let  $N = \max\{i_0, j_0\}$ . Let  $n \geq N$ . Let  $(l, k)$  be the largest tuple less than or equal to  $(n, n)$  such that  $\|f(x) - b_k\| < 1/l$ . We know there must exist one, since  $(j_0, i_0) \leq (N, N) \leq (n, n)$  is such a tuple. Note  $(l, k) \geq (j_0, i_0)$ , so that  $l \geq j_0$ . We then have that  $x \in E_{lk}^n$ , so that  $f_n(x) = b_k$ , and thus

$$\|f(x) - f_n(x)\| = \|f(x) - b_k\| < \frac{1}{l} < \frac{1}{j_0} < \varepsilon$$

as needed. □

We can now combine this proposition with the previous corollaries to obtain the following theorem:

**Theorem 1.2.8.** *A function  $f : X \rightarrow B$  is  $S$ -measurable if and only if  $f$  satisfies:*

- (a) *range  $f$  is separable*
- (b) *for any  $U \subseteq B$  open,  $f^{-1}(U \setminus \{0\}) \in S$*

We have also shown that if we have a sequence of functions satisfying these two properties, then their pointwise limit also satisfies these two properties, and that any  $f \in \mathcal{M}(X, S, B)$  satisfies these two properties. We thus obtain:

**Theorem 1.2.9.** *If  $\{f_n\}$  is a sequence in  $\mathcal{M}(X, S, B)$  and if  $f_n \rightarrow f$  pointwise, then  $f \in \mathcal{M}(X, S, B)$ .*

We can extend our results to the case where we have a measure space  $(X, S, \mu)$ , by considering  $f$  everywhere except for on a null set. Here is theorem statement:

**Theorem 1.2.10.** *Given a measure space  $(X, S, \mu)$  and a Banach space  $B$ ,  $f \in \mathcal{M}(X, S, \mu, B)$  (i.e.  $f$  is  $\mu$ -measurable) if and only if there is a null set  $N$  such that:*

- (a) *(range  $f$ )  $\setminus N$  is separable*
- (b) *for any  $U \subseteq B$  open we have  $f^{-1}(U \setminus \{0\}) \setminus N$  is  $\mu$ -measurable, i.e. is in  $M(\mu^*)$ .*

The proof is similar enough to what we have done so far that we will omit it. It involves keeping track of a null set  $N$  throughout.

## 1.3 Convergence

Thus far we have seen two types of convergence: pointwise and uniform. We will see in this section a number of new types of convergence that will help us towards our goal of being able to integrate certain functions.

**Definition.** Given a sequence of  $B$ -valued functions  $\{f_n\}$ , we say that this sequence converges “**almost uniformly**” (a.u.) to  $f$  on  $E$  if for all  $\varepsilon > 0$ , there exists  $F \subseteq E$ ,  $\mu^*(E \setminus F) < \varepsilon$  such that  $\{f_n\}$ , on  $F$ , converges uniformly to  $f$ .

Note in this definition, we don’t need an entire measure space, but rather just an outer measure  $\mu^*$ . This definition is used in the following theorem:

**Theorem 1.3.1. (Egoroff's Theorem)** Let  $\{f_n\}$  be a sequence in  $\mathcal{M}(X, S, \mu, B)$  that converges to  $f$  a.e. Let  $E \in S$ ,  $\mu(E) < \infty$ . Then  $\{f_n\}$  converges almost uniformly to  $f$  on  $E$ .

*Proof.* By removing a null set from  $E$ , we can assume  $f_n \rightarrow f$  pointwise everywhere on  $E$ . For each  $m, n \in \mathbb{N}$ , set:

$$E_n^m = \{x \in E \mid \|f(x) - f_k(x)\| > 1/m \text{ for some } k \geq n\}$$

Note  $E_n^m \in S$ , because  $f - f_k \in \mathcal{M}(X, S, \mu, B)$  for all  $k$ , and  $E_n^m$  can be written as the union of a bunch of  $(f - f_k)^{-1}(U)$ , where  $U$  is the open set that is the complement of the closed ball  $B(0, 1/m)$  for some  $m$ .

For fixed  $m$ , for each  $x$ , there exists  $n$  such that  $x \notin E_n^m$  by pointwise convergence. We thus have:

$$E_n^m \xrightarrow{n \rightarrow \infty} \emptyset$$

for fixed  $m$ . Since  $E_n^m \subseteq E$ , and  $\mu(E) < \infty$ , we have  $\mu(E_n^m) \rightarrow 0$  as  $n \rightarrow \infty$ , by continuity results for measures. Note this is where the essential assumption  $\mu(E) < \infty$  comes into play, because without this, we cannot conclude  $\mu(E_n^m) \rightarrow 0$ .

Given  $\varepsilon > 0$ , choose for each  $m$  an  $n_m$  such that:

$$\mu(E_{n_m}^m) < \frac{\varepsilon}{2^m}$$

Let:

$$F = E \setminus \bigcup_{m=1}^{\infty} E_{n_m}^m$$

Note:

$$\mu\left(\bigcup_{m=1}^{\infty} E_{n_m}^m\right) = \sum_{m=1}^{\infty} \mu(E_{n_m}^m) < \varepsilon$$

and thus  $\mu(E \setminus F) < \varepsilon$ . We claim that  $f_n \rightarrow f$  uniformly on  $F$ . Let  $\delta > 0$  be given. Choose  $m$  so that  $1/m < \delta$ . Let  $N = n_m$ . Then for  $k \geq N$ , and for  $x \in F$ , we have  $x \notin E_{n_m}^m$ , and thus:

$$\|f(x) - f_k(x)\| \leq \frac{1}{m} < \delta$$

as needed. □

**Remark.** Let  $X = [0, 1]$ ,  $\mu$  be the Lebesgue measure. Define a sequence of subsets  $E_n$  by:

$$\{[0, 1], [0, 1/2], [1/2, 1], [0, 1/3], [1/3, 2/3], [2/3, 1], [0, 1/4], \dots\}$$

The pattern should be clear. Note  $\mu(E_n) \rightarrow 0$  as  $n \rightarrow \infty$ . Let  $f_n = \chi_{E_n}$  in ISF. Note  $\|f_n\|_1 \rightarrow 0$  as  $n \rightarrow \infty$ . Thus  $f_n \rightarrow 0$  for  $\|\cdot\|_1$ . So  $\{f_n\}$  is a Cauchy sequence for

$\|\cdot\|_1$ . But notice  $f_n$  does not converge pointwise to 0;  $f_n(0) = 1$  for all  $n$ . We need to deal with this difference between convergence under the norm and pointwise convergence.

Let  $\{f_n\}$  be a sequence of ISF( $X, S, \mu, B$ ) that converges for  $\|\cdot\|_1$  to  $f \in \text{ISF}(X, S, \mu, B)$ . Let  $\varepsilon > 0$ . Let:

$$E_n = \{x \in X \mid \|f(x) - f_n(x)\| \geq \varepsilon\}$$

Note  $f - f_n \in \text{ISF}$ , which implies that  $E_n \in S$ , and that  $\mu(E_n) < \infty$ . Consider  $\varepsilon\chi_{E_n}$ . We have:

$$\varepsilon\chi_{E_n} \leq \|f - f_n\|_B$$

which implies:

$$\varepsilon\mu(E_n) = \int \varepsilon\chi_{E_n} d\mu \leq \int \|f - f_n\|_B d\mu$$

and so:

$$\mu(E_n) \leq \frac{1}{\varepsilon} \|f - f_n\|_1$$

Thus as  $n \rightarrow \infty$ ,  $\mu(E_n) \rightarrow 0$ . This motivates the following definition:

**Definition.** A sequence of  $B$ -valued functions  $\{f_n\}$  **converges to a function  $f$  in measure** if for any  $\varepsilon > 0$ :

$$\mu^*({x \mid \|f(x) - f_n(x)\| \geq \varepsilon}) \rightarrow 0 \text{ as } n \rightarrow \infty$$

Note this definition involves the outer measure  $\mu^*$  because we are not assuming we have a measure space.

Analogously, let  $\{f_n\}$  be a sequence in ISF that is Cauchy for  $\|\cdot\|_1$ . Given  $\varepsilon > 0$ , let:

$$E_{mn} = \{x \mid \|f_m(x) - f_n(x)\| \geq \varepsilon\}$$

we can similarly find that:

$$\lim_{m,n \rightarrow \infty} \mu(E_{mn}) = 0$$

which motivates the following definition:

**Definition.** A sequence of functions  $\{f_n\}$  is **Cauchy in measure** if:

$$\mu^*({x \mid \|f_m(x) - f_n(x)\| \geq \varepsilon}) \rightarrow 0 \text{ as } m, n \rightarrow \infty$$

Let us take a step back right now and think about what our goals are. Our ultimate goal is to find a set of functions that we can integrate. If you recall, we talked about how the vector space ISF / NISF is not complete. We then discussed how given an abstract non-complete vector space, we could form its completion by taking equivalence classes of Cauchy sequences. We then wanted to be able to represent these Cauchy sequences as functions, so that we could better work with them. Our goal is thus to associate each Cauchy sequence with a function in such a way that equivalent Cauchy sequences are associated with the same function.

**Definition.** A sequence  $\{g_n\}$  is **almost uniformly Cauchy (a.u.Cauchy)** on  $E \in S$  if for all  $\varepsilon > 0$ , there exists  $F \in S$  with  $\mu(E \setminus F) < \varepsilon$  such that  $\{g_n\}$  is uniformly Cauchy on  $F$ .

We have seen that a sequence of ISF that is Cauchy for  $\|\cdot\|_1$  is Cauchy in measure. The following lemma shows that being Cauchy in measure has certain implications.

**Lemma 1.3.2. (Riesz, Weyl)** *Let  $\{f_n\}$  be a sequence in  $\mathcal{M}(X, S, \mu, B)$  that is Cauchy in measure. Then there is a subsequence that is almost uniformly Cauchy.*

*Proof.* Let  $\{f_n\}$  be a sequence that is Cauchy in measure. We want a subsequence  $\{f_{n_m}\}$ . Let  $n_1 = 1$ . Define the sequence inductively as follows: if we have  $n_1, \dots, n_m$ , choose  $n_{m+1}$  such that:

- (a)  $n_{m+1} > n_m$
- (b)  $\forall k \geq n_{m+1}$ , we have:

$$\mu\left(\{x \mid |f_{n_{m+1}}(x) - f_k(x)| \geq 2^{-(m+1)}\}\right) < 2^{-(m+1)}$$

Let  $g_m = f_{n_m}$ . We claim that  $\{g_m\}$  is a.u.Cauchy. Let  $\varepsilon > 0$  be given. Choose  $l$  such that:

$$\sum_{m=l}^{\infty} 2^{-m} < \varepsilon$$

Let:

$$G = \bigcup_{m \geq l} \{x \mid |g_{m+1}(x) - g_m(x)| \geq 2^{-m}\}$$

Note  $\mu(G) < \varepsilon$ , by how we selected the  $g_m$ . Let  $E = \bigcup C(g_m) \in S$ . Let  $F = E \setminus G$ , so that  $\mu(E \setminus F) < \varepsilon$ . To show  $\{g_m\}$  is uniformly Cauchy on  $F$ , let  $\delta > 0$  be given. Choose  $M$  such that:

$$\sum_{m=M}^{\infty} 2^{-m} < \delta$$

Then for  $m, n \geq M$ ,  $m < n$ , we have for all  $x \in F$ :

$$\|g_n(x) - g_m(x)\| \leq \|g_n(x) - g_{n-1}(x)\| + \cdots + \|g_{m+1}(x) - g_m(x)\| \leq 2^{-(n-1)} + \cdots + 2^{-m} < \delta$$

as needed.  $\square$

Once we know that a sequence is a.u.Cauchy, we can deduce that it converges almost uniformly to a function.

**Proposition 1.3.3.** *Let  $\{f_n\}$  be a sequence that is a.u.Cauchy on  $E \in S$ . Then there is a function  $f$  such that  $f_n \rightarrow f$  a.u. and a.e. on  $E$ .*

*Proof.* For each  $m$ , choose  $F_m \in S$  such that:

$$\mu(E \setminus F_m) < \frac{1}{m}$$

and  $\{f_n\}$  is uniformly Cauchy on  $F_m$ . Let  $F = \bigcup F_m$ . Note  $E \setminus F \subseteq E \setminus F_m$  for all  $m$ , so  $\mu(E \setminus F) \leq \mu(E \setminus F_m) < \frac{1}{m}$  for all  $m$ . Thus  $\mu(E \setminus F) = 0$ .

On each  $F_m$ ,  $\{f_n\}$  converges pointwise to a function, because  $\{f_n\}$  is uniformly Cauchy, and thus each for each  $x \in F_m$ ,  $\{f_n(x)\}$  is a Cauchy sequence, and because  $B$  is complete,  $\{f_n(x)\}$  converges. Call this function  $f$ . Then  $f_n \rightarrow f$  pointwise on  $F_m$ . Since  $\{f_n\}$  is uniformly Cauchy, for all  $\varepsilon > 0$ , there exists  $N$  such that for  $m, n \geq N$ , we have for all  $x \in F_m$ :

$$\|f_n(x) - f_m(x)\| < \varepsilon$$

Letting  $n \rightarrow \infty$ , we get:

$$\|f(x) - f_m(x)\| \leq \varepsilon$$

and thus we see that  $f_n \rightarrow f$  uniformly on  $F_m$ .

Because  $F = \bigcup F_m$ , and  $\{f_n\}$  converges pointwise on each  $F_m$ , we get that  $\{f_n\}$  converges pointwise on  $F$ , say to  $g$ . Set  $g(x) = 0$  for  $x \notin F$ . Because we also have that  $\{f_n\}$  converges uniformly on every  $F_m$ , we see that  $f_n \rightarrow g$  a.u. on  $F$ , and thus on  $E$ . Note also since  $\mu(E \setminus F) = 0$ , we have that  $f_n \rightarrow g$  a.e. on  $E$ .  $\square$

As the following proposition shows, almost uniform convergence implies convergence in measure.

**Proposition 1.3.4.** *If  $f_n \rightarrow f$  a.u., then  $f_n \rightarrow f$  in measure.*

*Proof.* Given  $\varepsilon > 0$ , we need to consider:

$$\mu(\{x \mid \|f(x) - f_n(x)\| \geq \varepsilon\})$$

Given  $\delta > 0$ , choose  $F \subseteq E$ ,  $\mu(E \setminus F) < \delta$  such that  $f_n \rightarrow f$  uniformly on  $F$ . Choose  $N$  such that for  $n \geq N$ ,  $\|f(x) - f_n(x)\| < \varepsilon$  for all  $x \in F$ . This implies for  $n \geq N$ :

$$\{x \mid \|f(x) - f_n(x)\| \geq \varepsilon\} \subseteq F^c$$

and thus:

$$\mu(\{x \mid \|f(x) - f_n(x)\| \geq \varepsilon\}) < \delta$$

as needed.  $\square$

We can now deduce that if  $\{f_n\}$  is a sequence of ISF that is Cauchy for  $\|\cdot\|_1$ , then  $\{f_n\}$  is Cauchy in measure, which implies that there is a subsequence  $\{f_{n_k}\}$  that is a.u. Cauchy. We then have that  $\{f_{n_k}\}$  converges to some  $f$  a.u., which implies that  $\{f_{n_k}\}$  converges to  $f$  in measure. The natural next step is to show that then  $\{f_n\}$  must also converge to  $f$  in measure.

**Proposition 1.3.5.** *If  $\{f_n\}$  is Cauchy in measure, and if it has a subsequence that converges in measure to some  $f \in \mathcal{M}(X, S, \mu, B)$ , then  $\{f_n\}$  converges to  $f$  in measure.*

*Proof.* Given  $\varepsilon > 0$ , we need to consider:

$$\mu(\{x \mid \|f(x) - f_n(x)\| \geq \varepsilon\})$$

Note:

$$\{x \mid \|f(x) - f_n(x)\| \geq \varepsilon\} \subseteq \left\{x \mid \|f(x) - f_{n_k}(x)\| \geq \frac{\varepsilon}{2}\right\} \cup \left\{x \mid \|f_{n_k}(x) - f_n(x)\| \geq \frac{\varepsilon}{2}\right\}$$

Given  $\delta > 0$ , there exists  $N$  such that for  $n_k > N$ :

$$\mu\left(\left\{x \mid \|f(x) - f_{n_k}(x)\| \geq \frac{\varepsilon}{2}\right\}\right) < \frac{\delta}{2}$$

and for  $n \geq N$ :

$$\mu\left(\left\{x \mid \|f_{n_k}(x) - f_n(x)\| \geq \frac{\varepsilon}{2}\right\}\right) < \frac{\delta}{2}$$

We then have for  $n \geq N$ :

$$\mu(\{x \mid \|f(x) - f_n(x)\| \geq \varepsilon\}) < \frac{\delta}{2} + \frac{\delta}{2} = \delta$$

as needed.  $\square$

We have thus shown that if  $\{f_n\}$  is a sequence of ISF that is Cauchy for  $\|\cdot\|_1$ , then  $f_n \rightarrow f \in \mathcal{M}(X, S, \mu, B)$  in measure.

**Proposition 1.3.6.** *Let  $\{f_n\}$  be a sequence that converges in measure to  $f$  and to  $g$ . Then  $f = g$  a.e.*

*Proof.* For given  $\varepsilon > 0$ , we have:

$$\{x \mid \|f(x) - g(x)\| \geq \varepsilon\} \subseteq \left\{x \mid \|f(x) - f_n(x)\| \geq \frac{\varepsilon}{2}\right\} \cup \left\{x \mid \|g(x) - f_n(x)\| \geq \frac{\varepsilon}{2}\right\}$$

Given  $\delta > 0$ , by convergence in measure, there exists  $N$  such that for  $n \geq N$ , we have that the measures of the two sets on the right are both less than  $\frac{\delta}{2}$ , which implies:

$$\mu(\{x \mid \|f(x) - g(x)\| \geq \varepsilon\}) < \delta$$

Since this holds for arbitrary  $\delta$ , we obtain:

$$\mu(\{x \mid \|f(x) - g(x)\| \geq \varepsilon\}) = 0$$

Note:

$$\{x \mid \|f(x) - g(x)\| \neq 0\} \subseteq \bigcup_{k=1}^{\infty} \left\{x \mid \|f(x) - g(x)\| \geq \frac{1}{k}\right\}$$

We have shown that the measure of every set in the union on the right is 0. The measure of the set on the left must be less than or equal to the sum of the measures of the sets on the right. We thus obtain:

$$\mu(\{x \mid \|f(x) - g(x)\| \neq 0\}) = 0$$

as needed. □

We thus see that if  $f_n \rightarrow f$  in measure, then  $f$  is unique a.e. We are almost done. We need to show that if  $\{f_n\}, \{g_n\}$  are two equivalent Cauchy sequences in the sense that:

$$\|f_n - g_n\|_1 \rightarrow 0 \text{ as } n \rightarrow \infty$$

then if  $f_n \rightarrow f$  in measure, then  $g_n \rightarrow f$  in measure also.

**Proposition 1.3.7.** *Let  $\{f_n\}, \{g_n\}$  be sequences in ISF that are Cauchy for  $\|\cdot\|_1$ . If  $\{f_n\}$  is equivalent to  $\{g_n\}$ , and if  $f_n \rightarrow f$  in measure, then  $g_n \rightarrow f$  in measure.*

*Proof.* Consider the sequence:

$$\{f_1, g_1, f_2, g_2, f_3, g_3, \dots\}$$

which is Cauchy for  $\|\cdot\|_1$ . Then the sequence in Cauchy is measure. It has a subsequence  $\{f_n\}$  that converges to  $f$  in measure. Thus  $\{f_1, g_1, \dots\}$  must also converge to  $f$  in measure. Therefore  $g_n \rightarrow f$  in measure, because any subsequence of a sequence that converges to  $f$  in measure must also converge to  $f$  in measure. □



After all this work, we have finally associated for every Cauchy sequence a function  $f$  in such a way that equivalent Cauchy sequences are assigned the same  $f$ . Note how essential defining convergence in measure and Cauchy in measure were to the process; the notion of “in measure” is exactly what we needed to get the job done. In some sense, pointwise convergence is a much more restrictive condition when working with sequences in ISF.

## 1.4 The Bochner-Lebesgue Integral

**Definition.** A sequence of  $\{f_n\}$  in ISF is said to be **mean-Cauchy** if it is Cauchy for  $\|\cdot\|_1$ .

**Proposition 1.4.1.** *Given a measure space  $(X, S, \mu)$ , a Banach space  $B$ , and a function  $f : X \rightarrow B$ , the following are equivalent:*

- (a) *There is a mean-Cauchy sequence of ISF that converges to  $f$  in measure.*
- (b) *There is a mean-Cauchy sequence of ISF that converges to  $f$  a.u.*
- (c) *There is a mean-Cauchy sequence of ISF that converges to  $f$  a.e.*

*Proof.* (a)  $\rightarrow$  (b). Let  $\{f_n\}$  be a mean-Cauchy sequence that converges to  $f$  in measure. Then it is Cauchy in measure, and by the Riesz-Weyl theorem (Lemma 1.15), there is a subsequence  $\{f_{n_k}\}$  that is a.u.Cauchy. By Proposition 1.16,  $f_{n_k} \rightarrow g$  a.u. for some  $g$ . By Proposition 1.17, we know that a.u. convergence implies convergence in measure, so that we have that  $f_{n_k} \rightarrow g$  in measure. Now because  $\{f_{n_k}\}$  is a subsequence of  $\{f_n\}$ , we must have that  $f = g$  a.e., and therefore we see that  $f_{n_k} \rightarrow f$  a.u.

(b)  $\rightarrow$  (c). Let  $\{f_n\}$  be a mean-Cauchy sequence that converges to  $f$  a.u. We then have that  $\{f_n\}$  is a.u.Cauchy. From the argument in Proposition 1.16, we get that  $f_n \rightarrow g$  a.e. and a.u. for some  $g$ . Because  $f_n \rightarrow g, f$  a.u., we get that by Proposition 1.17,  $f_n \rightarrow g, f$  in measure, which implies by Proposition 1.19 that  $f = g$  a.e. Therefore  $f_n \rightarrow f$  a.e.

(b)  $\rightarrow$  (a). By Proposition 1.17, convergence a.u. implies convergence in measure.

(c)  $\rightarrow$  (b). Let  $\{f_n\}$  be a mean-Cauchy sequence that converges to  $f$  a.e. Since  $\{f_n\}$  is mean-Cauchy, it is Cauchy in measure, and so by Riesz-Weyl, there is a subsequence  $\{f_{n_k}\}$  that is a.u.Cauchy. By proposition 1.16,  $f_{n_k} \rightarrow g$  a.u. and a.e. for some  $g$ . We then have that  $f = g$  a.e., and therefore  $f_{n_k} \rightarrow f$  a.u.

□

**Definition.** If  $f : X \rightarrow B$  satisfies any one of the properties listed in the proposition above, then it is said to be **Bochner-Lebesgue integrable**. The set of Bochner-Lebesgue integral functions will be denoted by  $\mathcal{L}^1(X, S, \mu, B)$ .

Let  $\{f_n\}$  be a mean-Cauchy sequence in ISF. Then:

$$\left\| \int f_n d\mu - \int f_m d\mu \right\|_B = \left\| \int (f_n - f_m) d\mu \right\|_B \leq \int \|f_n - f_m\| d\mu = \|f_n - f_m\|_1$$

Since  $\{f_n\}$  is mean-Cauchy, we get that  $\|f_n - f_m\|_1 \rightarrow 0$  as  $n, m \rightarrow \infty$ . Thus we see that  $\left\{ \int f_n d\mu \right\}$  is a Cauchy sequence in  $B$ . Since  $B$  is complete,  $\left\{ \int f_n d\mu \right\}$  converges to an element of  $B$ . We can finally define:

**Definition.** If  $f \in \mathcal{L}^1(X, S, \mu, B)$ , then we can define the **Bochner-Lebesgue integral** of  $f$ :

$$\int f d\mu = \lim_{n \rightarrow \infty} \int f_n d\mu$$

Before we go any further we need to show that  $\int f d\mu$  is well-defined. Let  $\{f_n\}, \{g_n\}$  be mean-Cauchy sequences that converge to  $f$  in measure. We need that  $\{f_n\}, \{g_n\}$  are equivalent, i.e.  $\|f_n - g_n\|_1 \rightarrow 0$  as  $n \rightarrow \infty$ . By our previous inequality, this will imply that  $\left\| \int f_n d\mu - \int g_n d\mu \right\|_B \rightarrow 0$  as  $n \rightarrow \infty$ , which implies  $\int f d\mu = \int g d\mu$ .

**Proposition 1.4.2.** *Let  $\{f_n\}, \{g_n\}$  be mean-Cauchy sequences in ISF. If they converge in measure to the same function  $f$ , then they are equivalent.*

*Proof.* By Riesz-Weyl and other propositions, there are subsequences of  $\{f_n\}, \{g_n\}$  that converge a.u. to  $f$ . It suffices to show that these subsequences are equivalent. Thus we can assume from the start that  $\{f_n\}, \{g_n\}$  converge a.u. to  $f$ . Let  $h_n = g_n - f_n$ . We want  $\|h_n\|_1 \rightarrow 0$ . We know that  $h_n \in \text{ISF}$ , and  $h_n \rightarrow 0$  a.u. Also,  $\{h_n\}$  is mean-Cauchy.

Let  $\varepsilon > 0$ . Choose  $N$  such that for  $m, n \geq N$ :

$$\|h_n - h_m\|_1 < \varepsilon$$

Let  $E = C(h_N)$ . Since  $h_N \in \text{ISF}$ , we know  $\mu(E) < \infty$ . For any  $n \geq N$ , consider:

$$\int_{E^c} \|h_n\| d\mu = \int \chi_{E^c} \|h_n\| d\mu$$

Note  $E^c$  is locally  $S$ -measurable, because intersecting with  $E^c$  is the same as taking away  $E$ , i.e. set difference. We have:

$$\int_{E^c} \|h_n\| d\mu = \int_{E^c} \|h_n - h_N\| d\mu$$

because  $h_N$  is 0 on  $E^c$ . We have:

$$\int_{E^c} \|h_n - h_N\| d\mu \leq \int_X \|h_n - h_N\| d\mu = \|h_n - h_N\|_1 < \varepsilon$$

Since  $h_N \in \text{ISF}$ , it is bounded. Let  $B = \|h_N\|_\infty$ . Since  $h_n \rightarrow 0$  a.u., there exists  $F \subseteq E$ ,  $\mu(E \setminus F) < \frac{\varepsilon}{B+1}$  such that  $h_n \rightarrow 0$  uniformly on  $F$ . We can choose  $M \geq N$  such that for  $n \geq M$ ,  $\|h_n\|_\infty < \frac{\varepsilon}{\mu(F)}$ . Thus:

$$\int_F \|h_n\| d\mu \leq \int_F \frac{\varepsilon}{\mu(F)} d\mu = \varepsilon$$

We also have:

$$\begin{aligned} \int_{E \setminus F} \|h_n\| d\mu &\leq \int_{E \setminus F} \|h_n - h_N\| d\mu + \int_{E \setminus F} \|h_N\| d\mu \\ &\leq \|h_n - h_N\|_1 + \int_{E \setminus F} \|h_N\|_\infty d\mu \\ &\leq \|h_n - h_N\|_1 + \|h_N\|_\infty \mu(E \setminus F) \\ &\leq \varepsilon + \varepsilon = 2\varepsilon \end{aligned}$$

Putting everything together, we have for  $n \geq M$ :

$$\begin{aligned} \|h_n\|_1 &= \int \|h_n\| d\mu = \int_{E^c} \|h_n\| d\mu + \int_{E \setminus F} \|h_n\| d\mu + \int_F \|h_n\| d\mu \\ &\leq \varepsilon + 2\varepsilon + \varepsilon = 4\varepsilon \end{aligned}$$

And therefore  $h_n \rightarrow 0$ . □

Having finally achieved our goal of integration, let us go over some properties of the set of Bochner-Lebesgue integrable functions.

**Properties of  $\mathcal{L}^1(X, S, \mu, B)$ .**

Let  $E$  be locally  $S$ -measurable. If  $f, g \in \mathcal{L}^1(X, S, \mu, B)$ , then:

1.  $f + g \in \mathcal{L}^1(X, S, \mu, B)$ , and:

$$\int_E (f + g) d\mu = \int_E f d\mu + \int_E g d\mu$$

2. If  $r \in \mathbb{R}$ , then:

$$\int_E (rf) d\mu = r \int_E f d\mu$$

3. If  $f \geq 0$  a.e., then  $\int_E f d\mu \geq 0$ .

4. For  $f \in \mathcal{L}^1(X, S, \mu, B)$ , the function taking  $x \mapsto \|f(x)\|_B$  is in  $\mathcal{L}^1(X, S, \mu, \mathbb{R})$ , and:

$$\left\| \int_E f d\mu \right\|_B \leq \int_E \|f\|_B d\mu$$

From these properties, we see that  $\mathcal{L}^1$  is a vector space, and for a fixed locally  $S$ -measurable set  $E$ ,  $\int_E : \mathcal{L}^1 \rightarrow B$  is linear. For  $f \in \mathcal{L}^1$ , set:

$$\|f\|_1 = \int \|f(x)\|_B d\mu(x)$$

We will call this the  $L^1$  norm. So property (4) is saying:

$$\left\| \int f d\mu \right\|_B \leq \|f\|_1$$

But note that the  $L^1$  norm is not in fact a norm on  $\mathcal{L}^1$ . This is because the norm will have value 0 on any function that is 0 a.e. To deal with this, for  $f \in \mathcal{L}^1(X, S, \mu, B)$ , let  $[f]$  be the set of functions that equal  $f$  a.e. Let  $L^1(X, S, \mu, B)$  be the set of all the  $[f]$  such that  $f \in \mathcal{L}^1$ . We can define  $[f] + [g] \equiv [f + g]$ , and  $r[f] \equiv [rf]$ . Note then  $\int [f] d\mu$  is well-defined, and that  $L^1$  also satisfies properties (1)-(4) listed above. We want to show that the  $L^1$  norm is in fact a norm on  $L^1$ . The hard part is showing that if  $\|[f]\|_1 = 0$  then  $[f] = [0]$ , i.e.  $f = 0$  a.e. To show this, first a couple observations.

**Lemma 1.4.3.** *If  $f \in \mathcal{L}^1$ , then  $C(f)$  is  $\sigma$ -finite.*

*Proof.* There exists a mean-Cauchy sequence of ISF  $\{f_n\}$  that converges to  $f$  a.e. Each  $f_n = \sum_{k=1}^{k_n} b_k^n \chi_{E_k^n}$ , with  $\mu(E_k^n) < \infty$ . We have:

$$C(f) \subseteq \left( \bigcup_{n,k} E_k^n \right) \cup N$$

where  $N$  is a null-set. □

**Lemma 1.4.4.** *If  $f \in \mathcal{L}^1(X, S, \mu, \mathbb{R})$ , and if  $E \in S$ ,  $\chi_E \leq f$  a.e., then  $\mu(E) < \infty$ .*

*Proof.* We have that  $E \subseteq C(f)$  except for on a null-set. Thus  $E$  is  $\sigma$ -finite, i.e.  $E \subseteq \bigcup_{n=1}^{\infty} E_n$ ,  $\mu(E_n) < \infty$ . We can take intersections so that we can write  $E = \bigcup E_n$ . Let  $F_n = \bigcup_{j=1}^n E_j$ . Then  $F_n \subseteq F_{n+1}$ ,  $F_n \subseteq E$ . We also have  $\lim F_n = E$  and  $\mu(F_n) < \infty$ . Thus  $\chi_{F_n} \in \text{ISF}$ , and thus  $\chi_{F_n} \in \mathcal{L}^1$ . We have  $\chi_{F_n} \leq f$  a.e., and so:

$$\mu(F_n) = \int \chi_{F_n} d\mu \leq \int f d\mu < \infty$$

And thus we see that by continuity results for measures:

$$\mu(E) = \lim_{n \rightarrow \infty} \mu(F_n) \leq \int f d\mu < \infty$$

□

Note that this lemma states that if  $\chi_E \leq f$  a.e., then  $\chi_E \in \text{ISF}$ , so that  $\chi_E \in \mathcal{L}^1$ .

**Proposition 1.4.5.** *If  $f \in \mathcal{L}^1(X, S, \mu, \mathbb{R})$ ,  $f \geq 0$ , and  $\int f d\mu = 0$ , then  $f = 0$  a.e.*

*Proof.* For each  $n$ , let  $E_n = \{x \mid f(x) > \frac{1}{n}\}$ . We have  $\chi_{E_n} \leq nf$  for all  $n$ , and so we have:

$$\mu(E_n) \leq n \int f d\mu = 0$$

which implies  $\mu(E_n) = 0$  for all  $n$ . Note:

$$C(f) = \bigcup_{n=1}^{\infty} E_n$$

and thus we have that  $\mu(C(f)) = 0$ . □

**Corollary 1.4.6.** *The  $L^1$  norm is in fact a norm on  $L^1(X, S, \mu, B)$ .*

**Remark.** On ISF, the  $L^1$  norm agrees with the norm on ISF\NISF. So we can write:

$$\text{ISF} \setminus \text{NISF} \subseteq L^1$$

This is in line with our ultimate goal of thinking of  $L^1(X, S, \mu, B)$  as the metric completion of ISF\NISF for  $\|\cdot\|_1$ . To fully validate our goal, we need to show that  $L^1$  is in fact complete.

Now that we have finally developed the Bochner-Lebesgue integral, we want to be able to tell when a given function is Bochner-Lebesgue integrable. To do this, we first examine some properties of Bochner-Lebesgue integrable functions.

**Definition.** For given  $f \in \mathcal{L}^1(X, S, \mu, B)$ , the function  $G \mapsto \int_G f(x) d\mu(x)$  for given  $G \in \mathcal{M}_{\text{loc}}(S)$  is called the **indefinite integral** of  $f$ . It will be denoted it by  $\mu_f$ .

**Proposition 1.4.7.** *Let  $f \in \mathcal{L}^1(X, S, \mu, B)$ . Then  $f$  satisfies:*

(a) *For all  $\varepsilon > 0$ , there exists  $E \in S$ ,  $\mu(E) < \infty$  such that:*

$$\int_{X \setminus E} \|f\| d\mu < \varepsilon$$

(b) For all  $\varepsilon > 0$ , there exists  $\delta > 0$  such that if  $E \in S$ ,  $\mu(E) < \delta$ , then  $\|\mu_f(E)\|_B < \varepsilon$ .

*Proof.* (a) Given  $\varepsilon > 0$ , choose  $g \in \text{ISF}$  with  $\|f - g\|_1 < \varepsilon$ . Let  $E = C(g)$ . Since  $g \in \text{ISF}$ ,  $\mu(C(g)) < \infty$ . We have:

$$\begin{aligned} \left\| \int_{X \setminus E} f d\mu \right\| &= \left\| \int (f - g) d\mu \right\| \\ &\leq \int_{X \setminus E} \|f - g\| d\mu \\ &\leq \|f - g\|_1 < \varepsilon \end{aligned}$$

(b) Given  $\varepsilon > 0$ , let  $g \in \text{ISF}$  such that  $\|f - g\|_1 < \frac{\varepsilon}{2}$ . Note  $\|g\|_\infty < \infty$ . Let  $\delta = \frac{\varepsilon}{2} \|g\|_\infty$ . Suppose  $E \in S$ ,  $\mu(E) < \delta$ . Then:

$$\begin{aligned} \|\mu_f(E)\| &= \left\| \int_E f d\mu \right\| \\ &\leq \left\| \int_E (f - g) d\mu \right\| + \left\| \int_E g d\mu \right\| \\ &\leq \|f - g\|_1 + \|g\|_\infty \mu(E) \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \end{aligned}$$

□

The properties displayed are important enough to make them definitions.

**Definition.** A function  $f$  is **almost supported** on a set of finite measure if for all  $\varepsilon > 0$ , there exists  $E \in S$ ,  $\mu(E) < \infty$  such that:

$$\int_{X \setminus E} \|f\| d\mu < \varepsilon$$

**Definition.** For any measure  $\nu$ , we say that  $\nu$  is **strongly absolutely continuous** with respect to  $\mu$  if for all  $\varepsilon > 0$ , there exists  $\delta > 0$  such that if  $\mu(E) < \delta$ , then  $\|\nu(E)\| < \varepsilon$ .

These notions allow us to formulate and prove the following important theorem:

**Theorem 1.4.8. (Lebesgue Dominated Convergence Theorem)** Let  $\{f_n\}$  be a sequence in  $\mathcal{L}^1(X, S, \mu, B)$  that converges a.e. to  $f$ . If there is a function  $g \in \mathcal{L}^1(X, S, \mu, \mathbb{R})$  such that for all  $n$ ,  $\|f_n(x)\| \leq g(x)$  a.e., then  $\{f_n\}$  is mean-Cauchy, and  $f \in \mathcal{L}^1(X, S, \mu, B)$ .

*Proof.* Let  $\varepsilon > 0$ . Since  $g \in \mathcal{L}^1(X, S, \mu, \mathbb{R})$ , it is almost supported on a set of finite measure, so there exists  $E \in S$ ,  $\mu(E) < \infty$  such that:

$$\int_{X \setminus E} g d\mu < \frac{\varepsilon}{6}$$

Thus:

$$\begin{aligned} \int_{X \setminus E} \|f_n - f_m\| d\mu &\leq \int_{X \setminus E} \|f_n\| d\mu + \int_{X \setminus E} \|f_m\| d\mu \\ &\leq 2 \int_{X \setminus E} g d\mu < \frac{\varepsilon}{3} \end{aligned}$$

We also have that  $f_n \rightarrow f$  a.e., and  $\mu(E) < \infty$ . Thus by Egoroff's theorem,  $f_n \rightarrow f$  a.u. on  $E$ . Since  $\mu_g$  is strongly absolutely continuous with respect to  $\mu$ , we can choose  $\delta > 0$  such that if  $\mu(G) < \delta$ , then  $\mu_g(G) = \int_G g d\mu < \frac{\varepsilon}{6}$ . Because of a.u. convergence, there exists  $F \subseteq E$  such that  $\mu(E \setminus F) < \delta$  and  $f_n \rightarrow f$  uniformly on  $F$ . Thus:

$$\int_{E \setminus F} \|f_n - f_m\| d\mu \leq 2 \int_{E \setminus F} g d\mu < \frac{\varepsilon}{3}$$

Finally, since uniform convergence implies uniformly Cauchy, we can choose  $N$  such that for  $m, n \geq N$ , we have:

$$\|f_n(x) - f_m(x)\| < \frac{\varepsilon}{3} \mu(F)$$

for all  $x \in F$ . Then:

$$\int_F \|f_n - f_m\| d\mu < \int_F \frac{\varepsilon}{3} \mu(F) = \frac{\varepsilon}{3}$$

Putting everything together, we have for  $m, n \geq N$ :

$$\begin{aligned} \|f_n - f_m\|_1 &= \int_{X \setminus E} \|f_n - f_m\| d\mu + \int_{E \setminus F} \|f_n - f_m\| d\mu + \int_F \|f_n - f_m\| d\mu \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon \end{aligned}$$

And therefore  $\{f_n\}$  is mean-Cauchy.

To deduce that  $f \in \mathcal{L}^1(X, S, \mu, B)$ , note that for all  $\frac{1}{n}$ , we can find  $g_n \in \text{ISF}$  such that  $\|f_n - g_n\|_1 < \frac{1}{n}$ . We can see that  $\{g_n\}$  is a mean-Cauchy sequence. Furthermore, by the inequality:

$$\|f - g_n\|_1 \leq \|f - f_n\|_1 + \|f_n - g_n\|_1$$

we get that  $g_n \rightarrow f$  for  $\|\cdot\|_1$ . We have seen that convergence for  $\|\cdot\|_1$  implies convergence in measure (in fact, this is what motivated the definition of convergence in measure). Therefore  $f \in \mathcal{L}^1(X, S, \mu, B)$ .  $\square$

**Corollary 1.4.9.** *If  $f \in \mathcal{M}(X, S, \mu, B)$ , and if there is  $g \in \mathcal{L}^1(X, S, \mu, \mathbb{R})$  with  $\|f(x)\| \leq g(x)$  a.e., then  $f \in \mathcal{L}^1(X, S, \mu, B)$ .*

*Proof.* Let  $\{f_n\}$  be a sequence of MSF that converges to  $f$  a.e. Set:

$$h_n(x) = \begin{cases} f_n(x) & \text{if } \|f_n(x)\| \leq 2g(x) \\ 0 & \text{if } \|f_n(x)\| > 2g(x) \end{cases}$$

Note this new sequence  $\{h_n\}$  is in MSF, and  $h_n \rightarrow f$  a.e., and  $\|h_n(x)\| \leq 2g(x)$ . We have that the last inequality implies  $C(h_n) \subseteq C(g)$  for all  $n$ . As in the proof of Lemma 1.24, we can obtain a sequence of subsets  $\{F_n\}$  of finite measure that increases up to  $C(h_n)$ . This implies that the sequence  $\{h_n F_n\}$  is in ISF. Furthermore, the sequence converges pointwise to  $h_n$ . We therefore have that  $h_n$  is integrable for all  $n$ . Therefore by the Lebesgue Dominated Convergence Theorem,  $f \in \mathcal{L}^1(X, S, \mu, B)$ .  $\square$

**Corollary 1.4.10.** *If  $x \mapsto \|f(x)\|_B$  is in  $\mathcal{L}^1$ , then  $f \in \mathcal{L}^1$ .*

**Theorem 1.4.11. (Monotone Convergence Theorem)** *Let  $\{f_n\}$  be a sequence in  $\mathcal{L}^1(X, S, \mu, \mathbb{R})$ . If  $f_n \geq f_m$  a.e. for  $n \geq m$ , and if there is  $c \in \mathbb{R}$  such that  $\int f_n d\mu \leq c$  for all  $n$ , then  $\{f_n\}$  is a mean-Cauchy sequence, and there is  $f \in \mathcal{L}^1$  such that  $f_n \rightarrow f$  a.e., and  $\int f d\mu = \lim_{n \rightarrow \infty} \int f_n d\mu$ .*

*Proof.* For  $n \geq m$ , we have  $f_n \geq f_m$  a.e., so that  $c \geq \int f_n d\mu \geq \int f_m d\mu$ . We thus see that  $\{\int f_n d\mu\}$  is a bounded monotone sequence of real numbers, and so it has a least upper bound, to which it converges. Thus  $\{\int f_n d\mu\}$  is a Cauchy sequence. For  $n \geq m$ :

$$\|f_n - f_m\|_1 = \int |f_n - f_m| d\mu = \int f_n d\mu - \int f_m d\mu$$

It follows that  $\{f_n\}$  is mean-Cauchy. We therefore have a subsequence that converges a.e. to some  $f$ . Because  $\{f_n\}$  is monotone, the sequence itself must also converge to  $f$  a.e. We then have:

$$\lim_{n \rightarrow \infty} \int f_n d\mu = \int f d\mu$$

$\square$

**Remark.** If  $\{f_n\}$  is a monotone sequence as above, and if  $f_n \rightarrow f$  a.e., but there does not exist an upper bound on  $\int f_n d\mu$ , then we say that  $\int f d\mu = \infty$ .



**Definition.** The **liminf** of a sequence of functions  $\{f_n\}$ ,  $f_n \geq 0$  a.e., is defined by the following procedure. First define:

$$g_{nm} = f_n \wedge f_{n+1} \wedge \cdots \wedge f_m$$

Note  $g_{nm}$  is measurable if  $\{f_n\}$  is measurable. For fixed  $n$ ,  $g_{nm} \downarrow$  as  $m \rightarrow \infty$ . But  $g_{nm} \geq 0$  a.e. So for fixed  $n$ ,  $g_{nm} \downarrow$  a.e. to some function, call it  $g_n$ . Think of  $g_n$  as  $f_n \wedge f_{n+1} \wedge \cdots$ . Now as  $n \rightarrow \infty$ , we have  $g_n \uparrow$ , and since the  $g_n$  are bounded above (possibly by a function that is  $\infty$  a.e.), we get that  $\{g_n\}$  converges a.e. to some function, call this  $\liminf\{f_n\}$ .

**Lemma 1.4.12. (Fatou's Lemma)** Let  $\{f_n\}$  be a sequence in  $\mathcal{L}^1(X, S, \mu, B)$ . Let  $f_n \geq 0$  a.e. for all  $n$ . Then:

$$\int \liminf\{f_n\} d\mu \leq \liminf \left\{ \int f_n d\mu \right\}$$

*Proof.* If  $\liminf \left\{ \int f_n d\mu \right\} = \infty$ , then the inequality is true. So assume  $\liminf \left\{ \int f_n d\mu \right\} < \infty$ . For any  $m \geq n$ , we have  $g_n \leq f_m$ , so that  $\int g_n d\mu \leq \int f_m d\mu$ . This then implies  $\int g_n d\mu \leq \liminf \left\{ \int f_m d\mu \right\}$ . We also have by definition that  $g_n \uparrow \liminf f_n$ . Thus by the Monotone Convergence theorem, we have:

$$\int \liminf f_n d\mu \leq \liminf \int f_n d\mu$$

□

Here we see a nice application of Fatou's Lemma.

**Proposition 1.4.13.** Let  $\{f_n\}$  be a sequence in  $\mathcal{L}^1(X, S, \mu, B)$ , and let  $f_n \rightarrow f$  a.e., and suppose there is a constant  $c$  such that  $\|f_n\|_1 \leq c$  for all  $n$ . Then  $f \in \mathcal{L}^1$ , and  $\|f\|_1 \leq c$ .

*Proof.* Note because of convergence:

$$\|f(\cdot)\|_B = \liminf\{\|f_n(\cdot)\|_B\}$$

So by Fatou's Lemma:

$$\int \|f(x)\| d\mu(x) \leq \liminf \left\{ \int \|f_n(x)\| d\mu(x) \right\} \leq c$$

Thus we see that  $f$  is Lebesgue-integrable, because  $\int \|f\| d\mu$  is finite.

□

## 1.5 Product Measures

Given measure spaces  $(X, S, \mu), (Y, T, \nu)$ ,  $B$  a Banach space, we want to define a “product measure” on  $(X \times Y, S \times T)$  that has certain nice properties.

As a concrete example, suppose we form  $L^1(Y, T, \nu, B)$ , and then form:

$$L^1(X, S, \mu, L^1(Y, T, \nu, B))$$

We can also swap the order, i.e. form:

$$L^1(Y, T, \nu, L^1(X, S, \mu, B))$$

If  $E \in S, \mu(E) < \infty, F \in T, \nu(F) < \infty, b \in B$ , consider:

$$(x, y) \mapsto b\chi_E(x)\chi_F(y)$$

Considering the map as a function of  $x$  gives us an element in the first iterated  $L^1$  space, while considering the map as a function of  $y$  gives an element in the second iterated  $L^1$  space. The norm of the map in both spaces is  $\|b\|\mu(E)\nu(F)$ . We in some sense expect these two iterated  $L^1$  spaces to be basically the same. Note we can write:

$$\int_X \left( \int_Y f(x, y) d\nu \right) d\mu = b\mu(E)\nu(F) = \int_Y \left( \int_X f(x, y) d\mu \right) d\nu$$

We want the product measure  $\mu \times \nu$  that we define to be such that  $L^1(X \times Y, S \times T, \mu \times \nu, B)$  is naturally isomorphic to the two iterated  $L^1$  spaces.

**Definition.** Let  $(X, S, \mu), (Y, T, \nu)$  be measure spaces. Let  $E \in S, F \in T$ . Then  $E \times F \subseteq X \times Y$  is called a **measurable rectangle**. Let:

$$S \times T = \sigma\text{-ring generated by the measurable rectangles}$$

We also define:

$$P = \{E \times F \mid E \in S, F \in T\}$$

We can show that  $P$  is a semiring.

**Definition.** Given  $A \subseteq X \times Y, x \in X$ , let:

$$A_x = \{y \in Y \mid (x, y) \in A\}$$

Similarly, for  $y \in Y$ :

$$A^y = \{x \in X \mid (x, y) \in A\}$$

We call  $A_x$  and  $A^y$  **slices**.

As expected, slices of measurable sets are measurable, as the next proposition shows.

**Proposition 1.5.1.** *If  $A \in S \times T$ ,  $x \in X, y \in Y$ , then  $A_x \in T, A^y \in S$ .*

*Proof.* Let:

$$M = \left\{ A \in X \times Y \mid A_x \in S \ \forall x, A^y \in T \ \forall y \right\}$$

Note all measurable rectangles are in  $M$ . We claim that  $M$  is closed under countable unions and intersections, and under taking differences.

Let  $\{A_n\} \subseteq M$ . We have:

$$\left( \bigcup A_n \right)_x = \bigcup (A_n)_x$$

Similarly:

$$\left( \bigcap A_n \right)_x = \bigcap (A_n)_x$$

Finally:

$$(A \setminus B)_x = A_x \setminus B_x$$

Note the case for  $y$ -sections is exactly the same. We therefore see that  $M$  is a  $\sigma$ -ring, and it contains  $P$ . Now  $S \times T$  is defined as the smallest  $\sigma$ -ring containing  $P$ , and so we therefore have:

$$M \supseteq S \times T$$

□

**Definition.** Let  $B$  be a Banach space, and let  $f : X \times Y \rightarrow B$ . We define  $f_x : Y \rightarrow B$  by  $f_x(y) = f(x, y)$ , and  $f^y : X \rightarrow B$  by  $f^y(x) = f(x, y)$ .

**Proposition 1.5.2.** *If  $f$  is  $S \times T$ -measurable, then for any  $x \in X, y \in Y$ , we have that  $f_x$  is  $T$ -measurable,  $f^y$  is  $S$ -measurable.*

*Proof.*

□

## Chapter 2

# Functional Analysis

At this time we will jump to an unrelated topic and go over the basics of functional analysis. We will use the theory developed in this chapter in future chapters.

### 2.1 Linear Operators

Let  $V, W$  be normed vector spaces. As normed vector spaces, we can put on  $V, W$  the topologies induced by the respective norms. Let  $T : V \rightarrow W$  be a linear operator. We can ask whether  $T$  is continuous, say at 0. We have that  $T$  is continuous at 0 implies:

$$T^{-1}(\overline{B_W(0, 1)}) \supseteq \overline{B_V(0, r)}$$

for some  $r > 0$ . In other words, this says if  $v \in V$  and  $\|v\| \leq r$ , then  $\|Tv\| \leq 1$ . We can divide by  $r$  and say if  $\|v\| \leq 1$ , then  $\|Tv\| \leq \frac{1}{r}$ .

**Definition.** The norm of a linear operator is defined as:

$$\|T\| = \sup \left\{ \|Tv\| \mid \|v\| \leq 1 \right\}$$

We call this the “**operator norm**”. We will show soon that it is in fact a norm.

So if  $T$  is continuous at 0, then  $\|T\| < \infty$ . Now for  $v_1, v_2 \in V$ :

$$\|Tv_1 - Tv_2\| = \|T(v_1 - v_2)\| \leq \|T\| \cdot \|v_1 - v_2\|$$

Thus if  $\|T\| < \infty$ , we see that actually  $T$  is Lipschitz with Lipschitz constant  $\|T\|$ . So we get that  $T$  is uniformly continuous. We thus get the following theorem:

**Theorem 2.1.1.** *Let  $T : V \rightarrow W$  be a linear operator from normed vector space  $V, W$ . Then  $T$  is continuous at 0 if and only if  $\|T\| < \infty$ .*

**Definition.** Given  $T : V \rightarrow W$ , we say that  $T$  is “**bounded**” if  $\|T\| < \infty$ .

**Definition.** For normed vector spaces  $V, W$ , we define  $\mathcal{B}(V, W)$  as the **set of bounded linear operators** from  $V$  to  $W$ .

**Proposition 2.1.2.** *The set  $\mathcal{B}(V, W)$  is a normed vector space for the pointwise operations.*

*Proof.* Let  $S, T \in \mathcal{B}(V, W)$ . Let  $v \in V$ . We have:

$$\|(S + T)(v)\| = \|Sv + Tv\| \leq \|Sv\| + \|Tv\| \leq \|S\| \cdot \|v\| + \|T\| \cdot \|v\| = (\|S\| + \|T\|)\|v\|$$

Thus  $\|S + T\| \leq \|S\| + \|T\|$ , so that  $S + T \in \mathcal{B}(V, W)$ . Note we also see that the operator norm satisfies the triangle inequality. Given  $r$ , we have:

$$\|(rT)(v)\| = \|r(Tv)\| = |r| \cdot \|Tv\| \leq |r| \cdot \|T\| \cdot \|v\|$$

So  $\|rT\| \leq |r| \cdot \|T\|$ . Thus  $rT \in \mathcal{B}(V, W)$ . It also isn't too hard to show that in fact  $\|rT\| = |r| \cdot \|T\|$ .

Now if  $\|T\| = 0$ , then  $\|Tv\| \leq 0 \cdot \|v\|$  for all  $v \in V$ , and thus  $Tv = 0$  for all  $v \in V$ , and so  $T = 0$ . We thus see that the operator norm is in fact a norm.  $\square$

**Proposition 2.1.3.** *Let  $V, W$  be normed vector spaces. If  $W$  is a Banach space (i.e. it is complete), then  $\mathcal{B}(V, W)$  is Banach space (i.e. it is also complete).*

*Proof.* Let  $\{T_n\}$  be a Cauchy sequence in  $\mathcal{B}(V, W)$ . For any  $v \in V$ , we have:

$$\|T_nv - T_mv\| = \|(T_n - T_m)v\| \leq \|T_n - T_m\| \cdot \|v\|$$

We thus see that  $\{T_nv\}$  is a Cauchy sequence in  $W$  for all  $v \in V$ . Since  $W$  is complete,  $\{T_nv\}$  converges. Define a function  $T : V \rightarrow W$ , by  $Tv = \lim T_nv$ . We have:

$$T(v_1 + v_2) = \lim T_n(v_1 + v_2) = \lim T_nv_1 + \lim T_nv_2 = Tv_1 + Tv_2$$

And for a given scalar  $r$ , we have:

$$T(rv) = \lim T_n(rv) = \lim r(T_nv) = r \lim T_nv = r(Tv)$$

Thus  $T$  is linear.

Since  $\{T_n\}$  is Cauchy, there is a constant  $K$  such that  $\|T_n\| \leq K$  for all  $n$ . We then have:

$$\|Tv\| \leq \|Tv - T_nv\| + \|T_nv\| \leq \varepsilon + K\|v\|$$

for all  $n$  large enough. Thus  $\|T\| \leq K$ , and thus  $T \in \mathcal{B}(V, W)$ .

The last thing we need to show is that  $T_n \rightarrow T$ . Let  $\varepsilon > 0$ . There exists  $N$  such that for all  $m, n \geq N$ , we have:

$$\|T_n - T_m\| < \frac{\varepsilon}{2}$$

and for each  $v \in V$ , there exists  $m_v \geq N$  such that:

$$\|T v - T_{m_v} v\| < \frac{\varepsilon}{2}$$

We then have that for all  $n \geq N$ , for all  $\|v\| \leq 1$ :

$$\|T v - T_n v\| \leq \|T v - T_{m_v} v\| + \|T_{m_v} v - T_n v\| < \frac{\varepsilon}{2} + \|T_{m_v} - T_n\| \cdot \|v\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

□

**Example.** Consider  $L^1(X, S, \mu, B)$ . Define  $I : L^1 \rightarrow B$  by:

$$I f = \int f d\mu$$

We have:

$$\left\| \int f d\mu \right\|_B \leq \|f\|_1$$

Thus  $\|I\| \leq 1$  (actually, equality holds).

**Example.** Let  $B, C$  be Banach spaces, and let  $T \in \mathcal{B}(B, C)$ . For  $f \in L^1(X, S, \mu, B)$ , define  $T f$  by:

$$(T f)(x) = T(f(x))$$

So  $T f : X \rightarrow C$ . Note  $f \in \mathcal{L}^1$ , so there is a mean Cauchy sequence of ISF such that  $f_n \rightarrow f$  a.e. We can then show that  $\{T f_n\}$  is a sequence of MSF, and that  $T f$  is measurable. Also:

$$\|(T f)(x)\| \leq \|T\| \cdot \|f(x)\|$$

Thus by the corollary to the LDCT, we get that  $T f \in L^1(X, S, \mu, C)$ . Also:

$$\|T f\|_1 \leq \|T\| \cdot \|f\|_1$$

We can then define  $\tilde{T} : L^1(X, S, \mu, B) \rightarrow L^1(X, S, \mu, C)$  by:

$$\tilde{T}(f) = T f$$

We have that  $\tilde{T} \in \mathcal{B}(L^1(X, B), L^1(X, C))$

**Example.** Let  $V, W, Z$  be normed vector spaces, and let  $Z$  be a Banach space. Let  $T \in \mathcal{B}(V, W), S \in \mathcal{B}(W, Z)$ . Define  $ST : V \rightarrow Z$  by  $(ST)(v) = S(Tv)$ . We have:

$$\|(ST)(v)\| = \|S(Tv)\| \leq \|S\| \cdot \|Tv\| \leq \|S\| \cdot \|T\| \cdot \|v\|$$

So  $ST \in \mathcal{B}(V, Z)$ , and  $\|ST\| \leq \|S\| \cdot \|T\|$  (not necessarily equal).

In particular, consider  $\mathcal{B}(V, V) = \mathcal{B}(V)$ . If  $S, T \in \mathcal{B}(V)$ , then  $ST \in \mathcal{B}(V)$ . So  $\mathcal{B}(V)$  is an algebra with an identity (the identity operator). If  $V$  is complete, so is  $\mathcal{B}(V)$ .

**Definition.** A **Banach algebra** is an algebra that is a Banach space over  $\mathbb{R}$  or  $\mathbb{C}$ , such that  $\|ab\| \leq \|a\| \cdot \|b\|$  for all  $a, b$  in the algebra.

We thus see that  $\mathcal{B}(V)$  is a Banach algebra.

**Definition.** Given a normed vector space  $V$  over  $\mathbb{R}$  or  $\mathbb{C}$  (denote it as  $F$ ), consider  $\mathcal{B}(V, F)$  – the continuous linear functions on  $V$ . We call this space the **dual** of  $V$ , and denote it as  $V^*$ . Note since  $F$  is complete, we have that  $V^*$  is a Banach space.

We will see later many natural examples of dual spaces.

## 2.2 Quotient Spaces

Let  $V$  be a normed vector space. Let  $W$  be a subspace of  $V$ . We can form the quotient vector space:

$$V/W = \{v + W \mid v \in V\}$$

Define the quotient map  $\pi : V \rightarrow V/W$  by:

$$\pi(v) = v + W$$

Define a seminorm on  $V/W$  by:

$$\|\pi(v)\| = \inf \left\{ \|v - w\| \mid w \in W \right\}$$

We can think of  $\|\pi(v)\|$  as the distance from  $v$  to the subspace  $W$ .

**Proposition 2.2.1.** *The function we defined on  $V/W$  is in fact a seminorm.*

*Proof.* Given  $\varepsilon > 0, v_1, v_2 \in V$ , we can find  $w_1, w_2 \in W$  such that:

$$\|v_j - w_j\| < \|\pi(v_j)\| + \varepsilon$$

We thus have:

$$\|\pi(v_1 + v_2)\| \leq \|(v_1 + v_2) - (w_1 + w_2)\| \leq \|v_1 - w_1\| + \|v_2 - w_2\| \leq \|\pi(v_1)\| + \|\pi(v_2)\| + 2\varepsilon$$

We thus have that the triangle inequality is satisfied.

Given  $r \neq 0$  a scalar, we have:

$$\begin{aligned} \|\pi(rv)\| &= \inf \left\{ \|rv - w\| \mid w \in W \right\} \\ &= \inf \left\{ r\|v - w\| \mid w \in W \right\} \\ &= r \inf \left\{ \|v - w\| \mid w \in W \right\} \\ &= r\|\pi(v)\| \end{aligned}$$

And if  $r = 0$ , then  $\|\pi(rv)\| = 0 = r\|\pi(v)\|$ . □

**Proposition 2.2.2.**  $\|\cdot\|_{V/W}$  is a norm if and only if  $W$  is closed.

*Proof.* Suppose  $W$  is closed. If  $\|\pi(v)\| = 0$ , then there exists  $\{w_n\} \subseteq W$  such that  $\|v - w_n\| \rightarrow 0$ , and thus  $v \in \bar{W} = W$ , and thus  $\pi(v) = 0$ .

If  $W$  is not closed, then there exists a limit point  $v$  of  $W$ ,  $v \notin W$ . Note then that  $\pi(v) \neq 0$ . However, we can obtain a sequence  $\{w_n\} \subseteq W$  such that  $\|v - w_n\| \rightarrow 0$ , which means that  $\|\pi(v)\| = 0$ . □

**Remark.** Note that from the definition of  $\|\pi(v)\|$ , we have that  $\|\pi(v)\| \leq \|v\|$  for all  $v \in V$ . We thus see that  $\pi$  is a bounded linear operator, and that  $\|\pi\| \leq 1$ .

**Remark.** If  $z \in V/W$ , then:

$$\|z\|_{V/W} = \inf \left\{ \|v\| \mid \pi(v) = z \right\}$$

This is because if  $\pi(v) = z$ , then  $\pi(v - w) = z$  for all  $w \in W$ .

**Lemma 2.2.3.**  $\pi : V \rightarrow V/W$  is an open function. In fact, for any  $r > 0$ :

$$\pi(B_V(0, r)) = B_{V/W}(0, r)$$

*Proof.* First, if  $\|v\| < r$ , then  $\|\pi(v)\| \leq \|v\| < r$ . Thus:

$$\pi(B_V(0, r)) \subseteq B_{V/W}(0, r)$$

Conversely, if  $z \in V/W$ ,  $\|z\| < r$ , choose  $\varepsilon > 0$  such that  $\|z\| + \varepsilon < r$ . Choose  $v \in V$ ,  $\pi(v) = z$ ,  $\|v\| \leq \|z\| + \varepsilon < r$ . So  $v \in B_V(0, r)$ . Thus:

$$B_{V/W}(0, r) \subseteq \pi(B_V(0, r))$$

□



We now may ask under what conditions is  $V/W$  a Banach space? It turns out that if  $V$  is a Banach space and  $W$  is a closed subspace, then  $V/W$  is a Banach space. To show this, we need a lemma from real analysis.

**Definition.** Let  $(M, d)$  be a metric space. A sequence  $\{x_n\}$  is “**rapidly Cauchy**” if:

$$\sum_{n=1}^{\infty} d(x_n, x_{n+1}) < \infty$$

Equivalently,  $\{x_n\}$  is rapidly Cauchy if there is a sequence  $\{\varepsilon_n\} \subseteq \mathbb{R}^+$  with:

$$\sum \varepsilon_n < \infty$$

and  $d(x_n, x_{n+1}) < \varepsilon_n$  for all  $n$ .

**Lemma 2.2.4.** *Every Cauchy sequence has a subsequence that is rapidly Cauchy.*

*Proof.* Let  $\{x_n\}$  be a Cauchy sequence. Choose  $\varepsilon_n$  so that  $\sum \varepsilon_n < \infty$ . There exists  $N_1$  such that for all  $m, n \geq N_1$ , we have  $d(x_m, x_n) < \varepsilon_1$ . Let  $y_1 = x_{N_1}$ . There exists  $N_2 > N_1$  such that for all  $m, n \geq N_2$ , we have  $d(x_m, x_n) < \varepsilon_2$ . Let  $y_2 = x_{N_2}$ . Suppose we have  $N_1 < \dots < N_k$  and  $y_1, \dots, y_k$ . There exists  $N_{k+1} > N_k$  such that for all  $m, n \geq N_{k+1}$ , we have  $d(x_m, x_n) < \varepsilon_{k+1}$ . Let  $y_{k+1} = x_{N_{k+1}}$ . We have that  $\{y_k\}$  is a rapidly Cauchy subsequence.  $\square$

**Theorem 2.2.5.** *If  $V$  is a Banach space and  $W$  a closed subspace, then  $V/W$  is a Banach space.*

*Proof.* It suffices to show that every rapidly Cauchy sequence converges, because every Cauchy sequence has a rapidly Cauchy subsequence, and if a subsequence of a Cauchy sequence converges, then the Cauchy sequence itself converges to the same limit.

So let  $\{z_n\} \subseteq V/W$  be rapidly Cauchy. Then:

$$\|z_{n+1} - z_n\| < \varepsilon_n$$

for all  $n$ , for some  $\{\varepsilon_n\}$  with  $\sum \varepsilon_n < \infty$ . Choose  $v_1$  such that  $\pi(v_1) = z_1$ . Choose for  $n \geq 1$ ,  $u_n \in V$  such that  $\pi(u_n) = z_{n+1} - z_n$ , and  $\|u_n\| < \varepsilon_n$ . Let  $v_n, n > 1$  be:

$$v_n = u_1 + u_2 + \dots + u_n$$

Then for  $n > 1$ , we have  $\|v_{n+1} - v_n\| = \|u_{n+1}\| < \varepsilon_{n+1}$ . So  $\{v_n\}$  is rapidly Cauchy. Because  $V$  is complete, we have that  $v_n \rightarrow v \in V$ . We have:

$$\pi(v_n) = z_1 + (z_2 - z_1) + \dots + (z_{n+1} - z_n) = z_{n+1}$$

Note  $\pi$  is bounded, and thus continuous. We have:

$$\lim z_{n+1} = \lim \pi(v_n) = \pi(\lim v_n) = \pi(v)$$

and therefore we see that  $z_n \rightarrow \pi(v)$ .  $\square$

## 2.3 Baire Category Theorem and Consequences

We start off with the central theorem of this section.

**Theorem 2.3.1. (Baire Category Theorem)** *Let  $(M, d)$  be a complete metric space, and let  $\{D_n\}_{n=1}^{\infty}$  be a sequence of dense open sets in  $M$ . Then  $\bigcap_{n=1}^{\infty} D_n$  is dense in  $M$ .*

*Proof.* We must show that for all nonempty open sets  $U$  of  $M$ , we have:

$$\left(\bigcap_{n=1}^{\infty} D_n\right) \cap U \neq \emptyset$$

Since  $D_1$  is dense, we have that  $D_1 \cap U \neq \emptyset$ , and since both sets are open, we can obtain  $x_1 \in M$ ,  $0 < r_1 < 2^{-1}$  such that  $B(x_1, r_1) \subseteq D_1 \cap U$ . Now consider  $D_2 \cap B(x_1, r_1)$ . This intersection is also nonempty and open, so that we can pick  $x_2 \in M$ ,  $0 < r_2 < 2^{-2}$  such that  $\bar{B}(x_2, r_2) \subseteq D_1 \cap B(x_1, r_1)$ . Now suppose we have picked  $x_1, \dots, x_k, r_1, \dots, r_k$  in the given pattern. We then have that  $D_{k+1} \cap B(x_k, r_k) \neq \emptyset$ , so that as before, we can pick  $x_{k+1} \in M$ ,  $0 < r_{k+1} < 2^{-(k+1)}$  such that:

$$\bar{B}(x_{k+1}, r_{k+1}) \subseteq D_{k+1} \cap B(x_k, r_k)$$

Now consider the sequence  $\{x_n\}$  that we have picked. For any given  $N$ , we have  $x_n, x_m \in B(x_N, r_N)$  for  $n, m \geq N$ . Now since  $r_N < 2^{-N}$ , we see that  $\{x_n\}$  is Cauchy. Because  $M$  is complete, we get that  $\{x_n\}$  converges to some point  $x \in M$ . Now for  $N > 1$ , we have  $x_n \in B(x_N, r_N)$  for all  $n \geq N$ . This implies  $x \in \bar{B}(x_N, r_N)$  for all  $N > 1$ . We then have:

$$x \in \bar{B}(x_N, r_N) \subseteq D_{N-1} \cap B(x_{N-1}, r_{N-1}) \subseteq D_{N-1} \cap B(x_1, r_1) \subseteq D_{N-1} \cap U$$

for all  $N > 1$ . We can rewrite this as  $x \in D_N \cap U$  for all  $N$ . We therefore have:

$$x \in \left(\bigcap_{n=1}^{\infty} D_n\right) \cap U \neq \emptyset$$

□

**Corollary 2.3.2.** *Let  $(M, d)$  be a complete metric space. Then  $M$  is not the countable union of nowhere dense sets.*

*Proof.* Let  $\{E_n\}_{n=1}^{\infty}$  be a sequence of nowhere dense sets in  $M$ . Consider  $\{\bar{E}_n^c\}_{n=1}^{\infty}$ . Because each  $E_n$  is nowhere dense, we have that by definition,  $\bar{E}_n^c$  has empty

interior, which implies that the closure of  $\bar{E}_n^c$  is all of  $M$ . I.e.,  $\bar{E}_n^c$  is dense in  $M$  for all  $n$ . By the Baire Category Theorem, we have that:

$$\bigcap_{n=1}^{\infty} \bar{E}_n^c$$

is dense in  $M$ , which implies that the set is nonempty. This implies:

$$X \neq \left( \bigcap_{n=1}^{\infty} \bar{E}_n^c \right)^c = \bigcup_{n=1}^{\infty} \bar{E}_n$$

Finally, note:

$$\bigcup_{n=1}^{\infty} E_n \subseteq \bigcup_{n=1}^{\infty} \bar{E}_n$$

□

**Remark.** The Baire Category Theorem is a topological statement. Therefore any topological space that is homeomorphic to a complete metric space will satisfy the conditions given in the theorem.

Inspired by this theorem, we make a definition.

**Definition.** Let  $X$  be a topological space. We say that a set  $E \subseteq X$  is **of the first category** if it is the countable union of nowhere dense sets. Otherwise, we say that  $E$  is **of the second category**. Another name used to denote a set of the first category is **meager**.

We now shift our attentions to three seemingly unrelated theorems. We will see how the Baire Category Theorem is used in each of the proofs.

**Theorem 2.3.3. (Open Mapping Theorem)** *Let  $B, C$  be Banach spaces, and let  $T \in \mathcal{B}(B, C)$ . If  $T$  is surjective, then  $T$  is open.*

*Proof.* To prove a map  $f$  between metric spaces  $X, Y$  is open, it suffices to show that for any open ball  $B(x, r) \subseteq X$ , there exists an open ball  $B(f(x), q) \subseteq Y$  such that:

$$T(B(x, r)) \supseteq B(f(x), q)$$

In our present situation, we are working with normed linear spaces. Thus we can scale and shift, which means it suffices to show that for the open ball  $B(0, 1) \subseteq B$ , there exists  $B(0, r) \subseteq C$  such that:

$$T(B(0, 1)) \supseteq B(0, r)$$

Let  $B_t = B(0, t) \subseteq B$ . Note:

$$B = \bigcup_{n=1}^{\infty} B_n$$

Since  $T$  is surjective, we have:

$$C = T(B) = \bigcup_{n=1}^{\infty} T(B_n)$$

Note  $C$  is complete. The map  $y \mapsto ny$  is a homeomorphism from  $C$  to itself, which maps  $T(B_1)$  to  $T(B_n)$ . Therefore by the Baire Category Theorem, we have that  $T(B_1)$  cannot be nowhere dense. Thus there exists  $B(y_0, 4r) \subseteq \overline{T(B_1)}$ . Choose  $y_1 = Tx_1 \in T(B_1)$  such that  $\|y_1 - y_0\| < 2r$ , then  $B(y_1, 2r) \subseteq \overline{T(B_1)}$ . Now if  $\|y\| < 2r$ , we have:

$$y = Tx_1 + (y - y_1)$$

Note  $y_1 - (y - y_1) = y$ , so that  $y - y_1 \in B(y_1, 2r)$ . Thus:

$$y \in Tx_1 + B(y_1, 2r) \subseteq Tx_1 + \overline{T(B_1)} \subseteq \overline{T(x_1 + B_1)} \subseteq \overline{T(B_2)}$$

We can divide by 2 to obtain if  $\|y\| < r$ , then  $y \in \overline{T(B_1)}$ .

Our goal now is to replace the  $\overline{T(B_1)}$  by  $T(B_1)$ , because then we would be done. First note that for  $\|y\| < r2^{-n}$ , we have  $y \in \overline{T(B_{2^{-n}})}$ . So suppose  $\|y\| < r/2$ , we then have  $x_1 \in B_{1/2}$  such that  $\|y - Tx_1\| < r/4$ . Note then that  $y - Tx_1 \in \overline{T(B_{2^{-2}})}$ . We can then find  $x_2 \in B_{2^{-2}}$  such that  $\|y - Tx_1 - Tx_2\| < r2^{-3}$ . Proceeding as so, we can find a sequence  $\{x_n\}$  such that for all  $n$ , we have:

$$\|y - Tx_1 - \dots - Tx_n\| < r2^{-(n+1)}$$

We thus have:

$$y = \sum_{n=1}^{\infty} Tx_n = T\left(\sum_{n=1}^{\infty} x_n\right)$$

Note since  $B$  is complete, the series  $\sum_{n=1}^{\infty} x_n$  converges to a point  $x \in B$ . We thus have  $y = Tx$ . Furthermore, we have:

$$\|x\| \leq \sum_{n=1}^{\infty} \|x_n\| < \sum_{n=1}^{\infty} 2^{-n} = 1$$

Therefore for all  $\|y\| < r/2$ , we have  $y \in T(B_1)$ , as needed.  $\square$

**Corollary 2.3.4.** *Let  $B, C$  be Banach spaces and let  $T \in \mathcal{B}(B, C)$ . If  $T$  is bijective, then  $T$  is an isomorphism.*

*Proof.* Since  $T$  is surjective, it is open. Since  $T$  is bijective,  $T$  is open implies that  $T^{-1}$  is continuous.  $\square$

For our next theorem, we first need to make a few definitions.

**Definition.** Given normed vector spaces  $X, Y$  and  $T : X \rightarrow Y$  a linear operator, we define the **graph** of  $T$  as so:

$$\Gamma(T) = \{(x, y) \in X \times Y \mid y = Tx\}$$

Note  $\Gamma(T)$  is a subspace of  $X \times Y$ .

**Definition.** A linear operator  $T : X \rightarrow Y$  between normed linear spaces is closed if  $\Gamma(T)$  is a closed subspace of  $X \times Y$ .

We can see that if  $T$  is continuous, then  $T$  is closed, because if  $\{(x_n, y_n)\}$  is a sequence in  $\Gamma(T)$  that converges to  $(x, y)$ , then  $\{x_n\}$  is a sequence in  $X$  that converges to  $x$  and  $\{y_n\}$  is a sequence in  $Y$  that converges to  $y$ . By the continuity of  $T$ , we have:

$$Tx = T(\lim x_n) = \lim Tx_n = \lim y_n = y$$

The next theorem shows that if  $X, Y$  are complete, then the converse is true.

**Theorem 2.3.5. (Closed Graph Theorem)** *Let  $X, Y$  be Banach spaces and  $T : X \rightarrow Y$  a linear operator. If  $T$  is closed, then  $T$  is bounded.*

*Proof.* Let  $\pi_1, \pi_2$  be the projections from  $\Gamma(T)$  to  $X, Y$  respectively. I.e.,  $\pi_1(x, Tx) = x$  and  $\pi_2(x, Tx) = Tx$ . Note  $\pi_1 \in \mathcal{B}(\Gamma(T), X)$  and  $\pi_2 \in \mathcal{B}(\Gamma(T), Y)$ . Note because  $X, Y$  are complete, we have that  $X \times Y$  is complete, and thus since  $T$  is closed, we get that  $\Gamma(T)$  is complete. Note  $\pi_1$  is a bijection from  $\Gamma(T)$  to  $X$ . By the corollary to the Open Mapping Theorem, we get that  $\pi_1^{-1}$  is bounded. Note that  $T = \pi_2 \circ \pi_1^{-1}$ . Therefore  $T$  is the composition of bounded maps, and thus  $T$  itself is bounded.  $\square$

Moving on, let  $\Lambda$  be an index set. For each  $\lambda \in \Lambda$ , let  $Y_\lambda$  be a Banach space (over  $\mathbb{R}$ ). Form (notice the prime):

$$Y = \prod'_{\lambda \in \Lambda} Y_\lambda = \left\{ y : \Lambda \rightarrow \bigcup Y_\lambda \mid y(\lambda) \in Y_\lambda \forall \lambda, \sup \{\|y(\lambda)\| \mid \lambda \in \Lambda\} < \infty \right\}$$

For  $y \in Y$ , set  $\|y\| = \sup \{\|y(\lambda)\| \mid \lambda \in \Lambda\}$ . This gives a norm on the vector space  $Y$ . Since each  $Y_\lambda$  is complete, we can show that  $Y$  is complete. We use this construction in the next theorem.

**Theorem 2.3.6.** (*Principle of Uniform Boundedness*) Let  $V, W$  be Banach spaces. For some index set  $\Lambda$ , for each  $\lambda \in \Lambda$ , let  $T_\lambda \in \mathcal{B}(V, W)$ . Suppose for each  $v \in V$ , there is a constant  $c_v$  such that:

$$\|T_\lambda v\| \leq c_v$$

for all  $\lambda \in \Lambda$ . Then there is a constant  $c$  such that:

$$\|T_\lambda\| \leq c$$

for all  $\lambda \in \Lambda$ .

*Proof.* For each  $\lambda$ , let  $Y_\lambda = W$ , and form  $Y = \prod' Y_\lambda$ . Define  $T : V \rightarrow Y$  by:

$$Tv = (T_\lambda v)_{\lambda \in \Lambda}$$

Note by the boundedness condition in the statement of the theorem,  $T$  does map into  $Y$ . For all  $\lambda \in \Lambda$ , and for all  $v \in V$ , we have:

$$\|T_\lambda v\| \leq \sup \{\|T_\lambda v\| \mid \lambda \in \Lambda\} = \|Tv\|$$

This implies that  $\|T_\lambda\| \leq \|T\|$  for all  $\lambda \in \Lambda$ . Therefore to prove the theorem, it suffices to show that  $T$  is bounded.

Since  $V$  and  $Y$  are both complete, by the Closed Graph Theorem, it suffices to show that  $T$  is closed. Let  $\{(v_n, T v_n)\}$  be a sequence in  $\Gamma(T)$  that converges to  $(v, y) \in V \times Y$ . Note then  $v_n \rightarrow v$  and  $T v_n \rightarrow y$ . We need to show that  $y = Tv$ . First note that for all  $\lambda$ , we have  $(T v_n)_\lambda \rightarrow y_\lambda$ . We also have that  $T_\lambda$  is bounded for each  $\lambda$ . This implies  $T_\lambda v_n \rightarrow T_\lambda v$ . Now note that  $(T v_n)_\lambda = T_\lambda v_n$ . Therefore we have that both  $T_\lambda v_n \rightarrow T_\lambda v$ , and  $T_\lambda v_n \rightarrow y_\lambda$ . This implies  $T_\lambda v = y_\lambda$  for all  $\lambda$ , and thus  $Tv = y$ .  $\square$

## 2.4 Linear Functionals

We now switch gears and consider the dual space of a vector space –  $\mathcal{B}(V, F)$ , where  $F = \mathbb{R}$  or  $F = \mathbb{C}$ . As mentioned before, elements in this dual are also called linear functionals. The main goal of this section is to show that there are in fact nontrivial continuous linear functionals on a given vector space  $V$ . This result may seem self evident at first, but it does take some effort to prove.

**Definition.** Let  $V$  be a vector space over  $\mathbb{R}$ . By a **Minkowski functional** on  $V$ , we mean a function  $p : V \rightarrow \mathbb{R}$  such that:

$$(a) \quad p(v + w) \leq p(v) + p(w) \quad \forall v, w \in V$$

$$(b) \ p(rv) = rp(v) \ \forall v \in V, \forall r \in \mathbb{R}, r > 0$$

This definition of a Minkowski functional may seem arbitrary, but in reality it is just a generalization of the properties of seminorms and norms. As we will see, many of the Minkowski functionals that we will be using are in fact seminorms.

We now formulate and prove the key lemma needed to show that there are many continuous linear functionals on a given vector space.

**Lemma 2.4.1.** *Let  $V$  be a vector space over  $\mathbb{R}$ , and let  $p$  be a Minkowski functional on  $V$ . Let  $W$  be a subspace of  $V$ , and suppose  $\varphi$  is a linear functional on  $W$  that satisfies  $\varphi(w) \leq p(w)$  for all  $w \in W$ . Then for any  $v \in V$ ,  $v \notin W$ ,  $\varphi$  has an extension  $\tilde{\varphi}$  to  $W \oplus \mathbb{R}v$ , such that  $\tilde{\varphi}(z) \leq p(z)$  for all  $z \in W \oplus \mathbb{R}v$ .*

*Proof.* We want to find  $\alpha \in \mathbb{R}$  such that if we set  $\tilde{\varphi}(w + sv) = \varphi(w) + s\alpha$ , then  $\tilde{\varphi}$  satisfies the condition. So, for  $s > 0$ , we want:

$$\varphi(w) + s\alpha \leq p(w + sv) \leq p(w) + sp(v)$$

rewriting:

$$s\varphi(w/s) + s\alpha \leq p(s(w/s) + sv) \leq p(s(w/s)) + sp(v)$$

$$s\varphi(w/s) + s\alpha \leq sp(w/s + v) \leq sp(w/s) + sp(v)$$

Dividing by  $s$  on both sides, we see that we just need:

$$\varphi(w) + \alpha \leq p(w + v)$$

for all  $w \in W$ . And if  $s < 0$ , we do a similar calculation, but now we divide by  $|s|$ , instead of  $s$  itself, to make sure the inequality doesn't flip. We get:

$$\varphi(w) - \alpha \leq p(w - v)$$

Putting these two inequalities together, we want an  $\alpha$  such that:

$$\varphi(w) - p(w - v) \leq \alpha \leq p(w + v) - \varphi(w)$$

for all  $w \in W$ . Let us proceed to find this  $\alpha$ . Let  $w, w_1 \in W$ . Then:

$$\begin{aligned} \varphi(w) + \varphi(w_1) &= \varphi(w + w_1) \leq p(w + w_1) \\ &\leq p((w + v) + (w_1 - v)) \\ &\leq p(w + v) + p(w_1 - v) \end{aligned}$$

We thus have:

$$\varphi(w_1) - p(w_1 - v) \leq p(w + v) - \varphi(w)$$

for all  $w, w_1 \in W$ . We thus have for all  $w \in W$ :

$$\sup_{w_1 \in W} \{\varphi(w_1) - p(w_1 - v)\} \leq p(w + v) - \varphi(w)$$

and since this is true for all  $w \in W$ , we have:

$$\sup_{w_1 \in W} \{\varphi(w_1) - p(w_1 - v)\} \leq \inf_{w \in W} \{p(w + v) - \varphi(w)\}$$

So we can choose any  $\alpha$  between the sup and the inf. This  $\alpha$  satisfies:

$$\varphi(w) - p(w - v) \leq \alpha \leq p(w + v) - \varphi(w)$$

for all  $w \in W$ , as needed.  $\square$

The next theorem is essential in proving the existence of continuous linear functionals.

**Theorem 2.4.2. (Hahn-Banach Theorem (Real Case))** *Let  $V$  be a vector space over  $\mathbb{R}$ , and let  $p$  be a Minkowski functional on  $V$ . Let  $W$  be a subspace of  $V$ . Let  $\varphi$  be a linear functional on  $W$  such that  $\varphi(w) \leq p(w)$  for all  $w \in W$ . Then there is an extension  $\tilde{\varphi}$  of  $\varphi$  to all of  $V$  such that  $\tilde{\varphi}$  is linear, and  $\tilde{\varphi}(v) \leq p(v)$  for all  $v \in V$ .*

*Proof.* Let  $U \subseteq_S V$  be taken to mean that  $U$  is a linear subspace of  $V$ . Let:

$$\mathcal{E} = \left\{ (U, \varphi_U) \mid W \subseteq U \subseteq_S V, \varphi_U \text{ is a linear extension of } \varphi \text{ to } U, \varphi_U \leq p \text{ on } U \right\}$$

Put on  $\mathcal{E}$  the partial order:

$$(U_1, \varphi_{U_1}) \leq (U_2, \varphi_{U_2})$$

if  $U_1 \leq U_2$  and  $\varphi_{U_2}$  is an extension of  $\varphi_{U_1}$ . Let  $\mathcal{T} \subseteq \mathcal{E}$  be totally ordered. Let:

$$X = \bigcup_{(U, \varphi_U) \in \mathcal{T}} U$$

Since the  $U$ 's are nested,  $X$  is actually a subspace of  $V$ . Furthermore, we can define  $\varphi_X : X \rightarrow F$  by  $\varphi_X(v) = \varphi_U(v)$ , where  $U$  is some subspace such that  $v \in U$ , and  $(U, \varphi_U) \in \mathcal{T}$ . The reason that this  $\varphi_X$  well-defined is because  $\mathcal{T}$  is totally ordered, so given  $(U_1, \varphi_{U_1}), (U_2, \varphi_{U_2}) \in \mathcal{T}$ , either  $\varphi_{U_1}$  extends  $\varphi_{U_2}$  or vice versa. We can also see that  $\varphi_X$  is linear, is an extension of  $\varphi_U$  for all  $(U, \varphi_U) \in \mathcal{T}$ , and satisfies  $\varphi_X \leq p$  on  $X$ . Therefore  $(X, \varphi_X) \in \mathcal{E}$ , and  $(X, \varphi_X)$  is an upper bound on  $\mathcal{T}$ .

We thus see that  $\mathcal{E}$  is inductively ordered. By Zorn's Lemma,  $\mathcal{E}$  contains a maximal element, call it  $(Y, \varphi_Y)$ . So there does not exist  $(U, \varphi_U) \in \mathcal{E}$  such that  $(Y, \varphi_Y) \leq (U, \varphi_U)$ ,  $(Y, \varphi_Y) \neq (U, \varphi_U)$ . We then must have  $V = Y$ , otherwise, we would have  $v \in V, v \notin Y$ . By the lemma, we would then be able to extend  $\varphi_Y$  onto  $Y \oplus \mathbb{R}v$ , and obtain an element strictly bigger than  $(Y, \varphi_Y)$ . Therefore  $\varphi_Y$  is an extension of  $\varphi$  onto all of  $V$ .  $\square$



**Corollary 2.4.3.** *Let  $V$  be a normed vector space over  $\mathbb{R}$ . Let  $W$  be a subspace, and let  $\varphi$  be a bounded linear functional on  $W$ . So  $|\varphi(w)| \leq \|\varphi\| \cdot \|w\|$  for all  $w \in W$ . Note the function  $v \mapsto \|\varphi\| \cdot \|v\|$  is a Minkowski functional. By Hahn-Banach, there is an extension  $\tilde{\varphi}$  of  $\varphi$  to  $V$ . Note this  $\tilde{\varphi}$  satisfies  $\|\tilde{\varphi}\| = \|\varphi\|$ .*

**Example.** One may ask whether the linear extension whose existence is asserted in the Hahn-Banach Theorem is unique. The answer is no, the extension need not be unique. For example, consider  $V = \ell^1$ , the space of absolutely summable sequences, with the  $\|\cdot\|_1$  norm. For fixed  $n_0 \in \mathbb{N}$ , let  $W = \mathbb{R}\delta_{n_0}$ , where  $\delta_{n_0}$  is the indicator function on  $n_0$ . Define  $\varphi$  on  $W$  by  $\varphi(\delta_{n_0}) = 1$ . Then  $\|\varphi\| = 1$ . We want to extend  $\varphi$  to  $\ell^1$ . Choose any  $g \in \ell^\infty$ , i.e. the space of bounded sequences, with  $\|g\|_\infty = 1$ . Then  $g$  determines a linear functional  $\psi_g$  by:

$$\psi_g(f) = \sum f(n)g(n)$$

Note  $\|\psi_g\| = 1$ . If  $g(n_0) = 1$ , then  $\psi_g(\delta_{n_0}) = 1$ . Thus if  $g(n_0) = 1$ , then  $\psi_g$  extends  $\varphi$ . We therefore see that extensions of  $\varphi$  are decidedly not unique.

**Example.** Let  $c_0$  be the space of sequences converging to 0, and consider it as a subspace of  $\ell^\infty$ , the space of bounded sequences. Let  $W = c_0 \oplus \mathbb{R}1$ , where 1 is the constant sequence. So  $W$  is the space of convergent sequences. On  $W$ , define  $\varphi$  by  $\varphi(g) = \lim g(n)$ . Note  $\|\varphi\| = 1$ . So the corollary to the Hahn-Banach Theorem says that  $\varphi$  has an extension  $\tilde{\varphi}$  to  $\ell^\infty$  with  $\|\tilde{\varphi}\| = \|\varphi\| = 1$ . Note  $\tilde{\varphi}(c_0) = 0$ . It turns out that this extension  $\tilde{\varphi}$  is not constructible. It is not surprising that such an example like this exists, given that our proof of Hahn-Banach relies on the Axiom of Choice.

We now prove the Hahn-Banach Theorem for complex normed vector spaces.

**Theorem 2.4.4. (Hahn-Banach Theorem (Complex, Normed Case))** *Let  $V$  be a normed vector space over  $\mathbb{C}$ . Let  $W \subseteq V$  be a subspace. Let  $\varphi$  be a bounded linear functional on  $W$ . There is an extension  $\tilde{\varphi}$  of  $\varphi$  to  $V$  with  $\|\tilde{\varphi}\| = \|\varphi\|$ .*

*Proof.* For any  $z \in \mathbb{C}$ , we can write  $z = x + iy$ ,  $x, y \in \mathbb{R}$ . Note:

$$-\operatorname{Re}(iz) = \operatorname{Im}(z)$$

Let  $\psi(w) = \operatorname{Re}(\varphi(w))$ . Then  $\psi$  is  $\mathbb{R}$ -linear. Also:

$$|\psi(w)| = |\operatorname{Re}(\varphi(w))| \leq |\varphi(w)| \leq \|\varphi\| \cdot \|w\|$$

By the Real Hahn-Banach Theorem, we obtain an extension  $\tilde{\psi}$  of  $\psi$  to  $V$ , such that  $\tilde{\psi}$  is  $\mathbb{R}$ -linear, and  $|\tilde{\psi}(v)| \leq \|\tilde{\psi}\| \cdot \|v\| \leq \|\varphi\| \cdot \|v\|$ . Set:

$$\tilde{\varphi}(v) = \tilde{\psi}(v) - i\tilde{\psi}(iv)$$

Since  $\tilde{\psi}$  is  $\mathbb{R}$ -linear, we have that  $\tilde{\varphi}$  is  $\mathbb{R}$ -linear. Also:

$$\begin{aligned}\tilde{\varphi}(iv) &= \tilde{\psi}(iv) - i\tilde{\psi}(-v) \\ &= \tilde{\psi}(iv) + i\tilde{\psi}(v) \\ &= i(\tilde{\psi}(v) - i\tilde{\psi}(iv)) \\ &= i\tilde{\varphi}(v)\end{aligned}$$

This shows that  $\tilde{\varphi}$  is also  $\mathbb{C}$ -linear. We also have that  $\tilde{\varphi}$  extends  $\varphi$ , since  $\tilde{\psi}$  extends  $\psi$ , and the relation  $-\operatorname{Re}(iz) = \operatorname{Im}(z)$  for all  $z \in \mathbb{C}$ . Given  $v \in V$ , we can find  $\alpha \in \mathbb{C}$ ,  $|\alpha| = 1$ , such that  $\tilde{\varphi}(\alpha v) \in \mathbb{R}$  and  $\tilde{\varphi}(\alpha v) \geq 0$ . This implies:

$$|\tilde{\varphi}(\alpha v)| = |\tilde{\psi}(\alpha v)| \leq \|\varphi\| \cdot \|\alpha v\| = \|\varphi\| \cdot \|v\|$$

We also have:

$$|\tilde{\varphi}(v)| = |\alpha| \cdot |\tilde{\varphi}(v)| = |\alpha \tilde{\varphi}(v)| = |\tilde{\varphi}(\alpha v)|$$

We thus have that  $\|\tilde{\varphi}\| \leq \|\varphi\|$ . Since  $\tilde{\varphi}$  extends  $\varphi$ , we get that actually  $\|\tilde{\varphi}\| = \|\varphi\|$ .  $\square$

**Corollary 2.4.5.** *Let  $V$  be a normed vector space. Let  $v \in V$ . Then there exists  $\varphi \in V^*$  with  $\|\varphi\| = 1$ ,  $\varphi(v) = \|v\|$ .*

*Proof.* Let  $W = \mathbb{R}v$ , and let  $\varphi(v) = \|v\|$ . Extend  $\varphi$  to  $V$ .  $\square$

**Corollary 2.4.6.** *For all  $v \in V$ , we have:*

$$\|v\| = \sup \left\{ |\varphi(v)| \mid \varphi \in V^*, \|\varphi\| \leq 1 \right\}$$

**Corollary 2.4.7.** *Given  $v \in V$ ,  $v = 0$  if and only if for all  $\varphi \in V^*$ , we have  $\varphi(v) = 0$ .*

Given a normed vector space  $V$ , consider  $V^{**} = (V^*)^*$ . We have a canonical map  $j: V \rightarrow V^{**}$ :

$$v \mapsto \left( \varphi \mapsto \varphi(v) \right)$$

Note:

$$|(j(v))(\varphi)| = |\varphi(v)| \leq \|\varphi\| \cdot \|v\|$$

Also, given  $v$ , we know there exists  $\varphi_v$  such that  $\|\varphi_v\| = 1$  and  $\varphi_v(v) = \|v\|$ . Then:

$$(j(v))(\varphi_v) = \varphi_v(v) = \|v\|$$

We therefore see that  $\|j(v)\| = \|v\|$ . Therefore the map  $j$  is an isometry from  $V$  into its double dual. If  $V$  is complete, i.e. a Banach space, then  $j(V)$  is a closed subspace of  $V^{**}$ .

**Example.** Let  $V = c_0$ , the space of sequences converging to 0. Then  $V^* = \ell^1$ , the space of absolutely summable sequence, and  $V^{**} = \ell^\infty$ , the space of bounded sequences. The function  $j$  is the evident inclusion of  $c_0$  in  $\ell^\infty$ . Note  $c_0$  is separable, but  $\ell^\infty$  is not separable. Therefore  $j$  is not surjective.

The example shows that  $j$  need not be surjective. When it is, we have a special name for  $V$ .

**Definition.** Let  $V$  be a Banach space. We say that  $V$  is **reflexive** if  $j : V \rightarrow V^{**}$  is surjective, and thus an isomorphism.

Let  $V, W$  be normed vector spaces. Let  $T \in \mathcal{B}(V, W)$ . Define  $T^* : W^* \rightarrow V^*$  by:

$$(T^*(\psi))(v) = \psi(Tv)$$

We have:

$$|(T^*(\psi))(v)| = |\psi(Tv)| \leq \|\psi\| \cdot \|Tv\| \leq \|\psi\| \cdot \|T\| \cdot \|v\|$$

thus:

$$\|T^*(\psi)\| \leq \|\psi\| \cdot \|T\|$$

and thus:

$$\|T^*\| \leq \|T\|$$

Now note for any  $\varepsilon > 0$ , there exists  $v \in V$ ,  $\|v\| \leq 1$ ,  $\|Tv\| \geq \|T\| - \varepsilon$ . We can then find  $\psi \in W^*$  such that  $\|\psi\| = 1$  and  $\psi(Tv) = \|Tv\|$ . Then:

$$|(T^*(\psi))(v)| = |\psi(Tv)| = \|Tv\| \geq \|T\| - \varepsilon$$

Therefore  $\|T^*(\psi)\| \geq \|T\| - \varepsilon$ , and since  $\|\psi\| = 1$ , we get  $\|T^*\| \geq \|T\| - \varepsilon$ . Since  $\varepsilon$  was arbitrary, we get  $\|T^*\| \geq \|T\|$ , and so we can conclude  $\|T^*\| = \|T\|$ .

Note if we have  $T : V \rightarrow W$  and  $S : W \rightarrow Z$ , then we can obtain  $S^* : Z^* \rightarrow W^*$  and  $T^* : W^* \rightarrow V^*$ . We can obtain  $(TS)^* = S^*T^*$ . Side note – the  $*$  operator is a functor from the category of Banach spaces into itself (whatever that means).

## 2.5 Weak and Weak-\* Topologies

We first review some topological concepts. Let  $X$  be a set,  $\{Y_\lambda\}_{\lambda \in \Lambda}$  be a collection of topological spaces,  $\{f_\lambda\}_{\lambda \in \Lambda}$  a collection of functions such that  $f_\lambda : X \rightarrow Y_\lambda$  for all  $\lambda \in \Lambda$ . We can put on  $X$  the corresponding **initial topology** – i.e. the weakest topology on  $X$  making all the  $f_\lambda$ 's continuous. The topology has the subbase:

$$\mathcal{S} = \left\{ f_\lambda^{-1}(O) \mid \lambda \in \Lambda, O \in \mathcal{T}_\lambda \right\}$$

A base for this topology consists of finite intersections of elements of  $\mathcal{S}$ .

Now let  $V$  be a vector space, and  $\mathcal{S} = \{p_\lambda\}_{\lambda \in \Lambda}$  a collection of seminorms on  $V$ . Note each  $p_\lambda$  maps into  $F$ , a field. We can put on  $V$  the initial topology for  $\mathcal{S}$ .

**Example.** Let  $\mathcal{S}$  be the collection of linear functionals on  $V$ . For each  $\varphi \in \mathcal{S}$ , define  $p_\varphi$  by:

$$p_\varphi(v) = |\varphi(v)|$$

The initial topology here is the same as the initial topology for  $\mathcal{S}$ .

**Example.** Let  $V = C^\infty([0, 1])$ , the set of infinitely differentiable functions on the unit interval. We can set  $\|f\|_n = \|f^{(n)}\|_\infty$ , and obtain the initial topology on  $V$  given by  $\{\|\cdot\|_n\}$ .

**Remark.** If  $p$  is a seminorm on a vector space  $V$ , then the set:

$$p^{-1}(\bar{B}(0, r)) = \{v \in V \mid p(v) \leq r\}$$

is convex, because  $p$  satisfies the triangle inequality.

**Definition.** A **locally convex topological vector space (LCTVS)** is a vector space equipped with a family of seminorms. Put on  $V$  the weakest topology making all the seminorms continuous. The seminorms give a subbase for the neighborhood system at 0. Shifting gives a neighborhood system at any other point of  $V$ .

**Remark.** If the seminorms come from linear functionals on  $V$ , i.e.  $p(v) = |\varphi(v)|$  for some  $\varphi$ , for every  $p$ , then we often use “weak” in the name of the topology.

**Definition.** Let  $V$  be a normed vector space. Consider  $V^*$  and  $V^{**}$ . The initial topology on  $V^*$  coming from  $V^{**}$  is called the **weak topology**. Note also we have the canonical map  $j : V \rightarrow V^{**}$ . The initial topology on  $V^*$  given by  $j(V)$  is called the **weak-\* topology**.

We thus see that there are many different topologies that we can put on  $V^*$ . We can use the topology coming from the norm on  $V^*$ , or the weak topology, or the weak-\* topology. In general, these topologies may all be different. Note if  $V$  is reflexive, i.e.  $j$  is surjective, then the weak and weak-\* topologies coincide.

The following theorem explains why it is important to consider different topologies on  $V^*$ .

**Theorem 2.5.1. (Alaoglu’s Theorem)** *The unit ball of  $V^*$  is compact for the weak-\* topology.*

*Proof.* The proof is an exercise in using Tychonoff's theorem. Let  $B$  be the unit ball in  $V^*$ . For each  $v \in V$ , let  $D_v \subseteq \mathbb{R}$  (or  $\mathbb{C}$ ) be the closed ball of radius  $\|v\|$ . So  $D_v$  is compact. Form  $P = \prod_{v \in V} D_v$ . Give  $P$  the product topology. By Tychonoff's theorem,  $P$  is compact. Define  $i : B \rightarrow P$  by:

$$(i(\varphi))_v = \varphi(v)$$

We have:

$$|(i(\varphi))_v| = |\varphi(v)| \leq \|\varphi\| \cdot \|v\| \leq \|v\|$$

since  $\varphi \in B$ , so that  $\|\varphi\| \leq 1$ . Thus  $(i(\varphi))_v \in D_v$  for all  $v$ , and thus  $i$  does in fact map into  $P$ . We want to show that  $i$  is actually a homeomorphism from  $B$  to  $i(B)$ .

Note if  $\varphi \neq \varphi'$ , then there exists  $v$  such that  $\varphi(v) \neq \varphi'(v)$ , so that  $(i(\varphi))_v \neq (i(\varphi'))_v$ . Thus  $i$  is injective. The product topology is the smallest topology making all the projections  $\pi_v : P \rightarrow D_v$ ,  $\pi_v(f) = f_v$  continuous. The subbase for the product topology is given by:

$$\left\{ \pi_v^{-1}(B(x, r)) \mid v \in V, x \in \mathbb{R}, r > 0 \right\}$$

Now let  $\pi_v^{-1}(B(x, r))$  be an element in the subbase. Consider:

$$\begin{aligned} i^{-1}(\pi_v^{-1}(B(x, r))) &= \left\{ \varphi \in B \mid i(\varphi) \in \pi_v^{-1}(B(x, r)) \right\} \\ &= \left\{ \varphi \in B \mid (i(\varphi))_v \in B(x, r) \right\} \\ &= \left\{ \varphi \in B \mid \varphi(v) \in B(x, r) \right\} \end{aligned}$$

Now let  $p_v \in j(V)$  be defined by  $p_v(\varphi) = \varphi(v)$ . We then have:

$$\begin{aligned} &= \left\{ \varphi \in B \mid p_v(\varphi) \in B(x, r) \right\} \\ &= p_v^{-1}(B(x, r)) \cap B \end{aligned}$$

Note that  $p_v^{-1}(B(x, r)) \cap B$  is a subbasic open set in the relative weak-\* topology on  $B$ . Thus  $B$  is continuous. Now consider any subbasic open set  $p_v^{-1}(B(x, r)) \cap B$  in the relative weak-\* topology on  $B$ . We have:

$$\begin{aligned} i(p_v^{-1}(B(x, r)) \cap B) &= i\left(i^{-1}(\pi_v^{-1}(B(x, r)))\right) \\ &= (\pi_v^{-1}(B(x, r))) \cap i(B) \end{aligned}$$

Note  $(\pi_v^{-1}(B(x, r))) \cap i(B)$  is a subbasic open set in the relative product topology on  $i(B)$ . We therefore see that  $i$  is a homeomorphism.

To show that  $B$  is compact, it suffices to show that  $i(B)$  is closed, because  $i(B) \subseteq P$ , and  $P$  is compact, and closed subsets of compact sets are compact. To show that  $i(B)$  is closed, we can show that the every limit of every net in  $i(B)$  belongs to  $i(B)$ . So suppose we have a net  $\{i(\varphi_\lambda)\}$  in  $i(B)$ , and suppose that this net converges to  $f \in P$ . We need to show that  $f \in i(B)$ . Define  $\varphi : V \rightarrow \mathbb{R}$  by  $\varphi(v) = f_v$ . Let  $v_1, v_2 \in V$ . We have:

$$\begin{aligned} \varphi(v_1 + v_2) &= f_{v_1+v_2} = \lim_{\lambda} (i(\varphi_\lambda))_{v_1+v_2} \\ &= \lim_{\lambda} \varphi_\lambda(v_1 + v_2) \\ &= \lim_{\lambda} \varphi_\lambda(v_1) + \lim_{\lambda} \varphi_\lambda(v_2) \\ &= \lim_{\lambda} (i(\varphi_\lambda))_{v_1} + \lim_{\lambda} (i(\varphi_\lambda))_{v_2} \\ &= f_{v_1} + f_{v_2} \\ &= \varphi(v_1) + \varphi(v_2) \end{aligned}$$

We can do the same for scalar multiplication, and thus we see that  $\varphi$  is linear, i.e.  $\varphi \in V^*$ . Now since  $f \in P$ , we have  $\varphi(v) = f_v \in D_v$ , i.e.  $|\varphi(v)| \leq \|v\|$  for all  $v$ . Therefore  $\|\varphi\| \leq 1$ . Thus  $\varphi \in B$ . Note by how  $\varphi$  is defined, we have  $i(\varphi) = f$ . Thus  $f \in i(B)$ , as needed.  $\square$

**Corollary 2.5.2.** *Any Banach space can be realized as a closed subspace of some  $C(K)$ , the space of continuous functions on  $K$ , where  $K$  is compact.*

*Proof.* Let  $V$  be a Banach space, and let  $B$  be the unit ball in  $V^*$ , which by Alaoglu's Theorem, is weak-\* compact. It is also Hausdorff. Consider  $C(B)$ . Define a map  $V \rightarrow C(B)$  by sending  $v \mapsto f_v$ , where  $f_v(\varphi) = \varphi(v)$ . Note:

$$\|f_v\|_{\infty} = \sup \{|\varphi(v)| \mid \varphi \in B\} \leq \|v\|$$

since  $\|\varphi\| \leq 1$ . But by Hahn-Banach, we can find  $\varphi$  such that  $\|\varphi\| = 1$  and  $\varphi(v) = \|v\|$ . So we see that  $\|f_v\|_{\infty} = \|v\|$ . Thus  $v \mapsto f_v$  is an isometry. Now since  $V$  is a Banach space, its image is complete, and therefore its image is closed in  $C(B)$ .  $\square$

## 2.6 Convex Sets

In this section we explore properties of convex sets and how they relate to linear functionals. We start off with an important lemma.

**Lemma 2.6.1.** *Let  $V$  be a locally convex topological vector space, and let  $U$  be a convex neighborhood of  $0_V$  in  $V$ . For any  $v \in V$ , set:*

$$m_U(v) = \inf \{s \mid s > 0, s^{-1}v \in U\}$$

Then  $m_U(v)$  is a Minkowski functional, with the following property:

$$U \supseteq \{v \mid m_U(v) < 1\}$$

*Proof.* Let  $v \in V$ . Consider the map  $\alpha \mapsto \alpha v$  for  $\alpha \in \mathbb{R}$ . Call this map  $s_v$ . Because  $V$  is a topological vector space, this map is continuous. Note  $0 \in s_v^{-1}(U)$ , since  $0 \cdot v = 0_V \in U$ . By continuity, we have that  $s_v^{-1}(U)$  is a neighborhood of 0. There thus exists an open ball  $B(0, r) \subseteq s_v^{-1}(U)$ . Thus for all  $n \in \mathbb{N}$  such that  $\frac{1}{n} < r$ , we have  $\frac{1}{n}v \in U$ . Therefore  $m_U(v) < \infty$ . Also note  $m_U(0_V) = 0$ .

Let  $t > 0$ . We have:

$$\begin{aligned} m_U(tv) &= \inf \{s \mid s > 0, s^{-1}(tv) \in U\} \\ &= \inf \{ts' \mid s' > 0, s'^{-1}v \in U\} \\ &= t \inf \{s' \mid s' > 0, s'^{-1}v \in U\} \\ &= tm_U(v) \end{aligned}$$

Now let  $v, w \in V$ ,  $s, t > 0$  such that  $s^{-1}v, t^{-1}w \in U$ . We have that  $\frac{s}{s+t}, \frac{t}{s+t} \in [0, 1]$ , and that they sum to 1. Thus since  $U$  is convex, we have:

$$\frac{s}{s+t}(s^{-1}v) + \frac{t}{s+t}(t^{-1}w) \in U$$

which can be rewritten:

$$(s+t)^{-1}(v+w) \in U$$

We therefore see that:

$$m_U(v+w) \leq s+t$$

Since this holds for all  $s, t$  such that  $s^{-1}v, t^{-1}w \in U$ , we have:

$$m_U(v+w) \leq m_U(v) + m_U(w)$$

We thus see that  $m_U$  is indeed a Minkowski functional.

Now let  $v$  be such that  $m_U(v) < 1$ . There exists  $m_U(v) \leq r < 1$  such that  $r^{-1}v \in U$ . Now since  $0_V \in U$ , and  $U$  is convex, for all  $t > r$ , we have that  $t^{-1}v = t^{-1}v + (r^{-1} - t^{-1}) \cdot 0_V \in U$ . Thus since  $1 > r$ , we get that  $v \in U$ . Therefore:

$$U \supseteq \{v \mid m_U(v) < 1\}$$

as needed. □

In order to prove the theorem we want to prove, we also need a small lemma concerning continuous linear functionals.

**Lemma 2.6.2.** *Let  $\varphi : V \rightarrow \mathbb{R}$  be a linear functional, and let  $V$  be a topological vector space. If  $\varphi$  is bounded on a neighborhood of  $0_V$ , then  $\varphi$  is continuous.*

*Proof.* Let  $U$  be a neighborhood of  $0_V$  such that  $\varphi$  is bounded on  $U$ . Then there exists  $M$  such that for all  $v \in U$ , we have  $\varphi(v) < M$ . If we then consider the neighborhood  $W = (\varepsilon/M)V$ , we have that  $\varphi(v) < \varepsilon$  for all  $v \in W$ . Thus  $\varphi$  is continuous.  $\square$

With these two lemmas in hand, we can now prove the following theorem:

**Theorem 2.6.3. (Hahn-Banach Separation Theorem)** *Let  $V$  be a locally convex topological vector space (i.e. its topology is given by a family of seminorms). Let  $O$  be an open convex set in  $V$ , and let  $C$  be a convex set in  $V$  such that  $O \cap C = \emptyset$ . Then there exists  $\varphi \in V^*$  and  $t_0 \in \mathbb{R}$  such that  $\varphi(O) < t_0 \leq \varphi(C)$ .*

*Proof.* Let:

$$O - C = \bigcup_{c \in C} O - c$$

Note that each of the translates  $O - c$  is open, and thus  $O - C$  itself is open. Given  $o_1 - c_1, o_2 - c_2 \in O - C$ ,  $s, t \in [0, 1]$ ,  $s + t = 1$ , we have:

$$s(o_1 - c_1) + t(o_2 - c_2) = so_1 + to_2 - (sc_1 + tc_2)$$

Note  $O, C$  are convex, so that  $so_1 + to_2 \in O$ ,  $sc_1 + tc_2 \in C$ . Therefore we see that  $O - C$  is also convex. Note that since  $O \cap C = \emptyset$ , we have that  $0_V \notin O - C$ . Pick any  $v_0 \in O - C$ , and consider the set:

$$U = O - C - v_0$$

Note  $U$  is both open and convex, and  $0_V \in U$ . Thus  $U$  is a convex neighborhood of  $0_V$ . Consider the minkowski functional  $m_U$  as defined in the lemma. Since  $O \cap C = \emptyset$ , we have that  $-v_0 \notin U$ . Define  $\varphi_0$  on  $\mathbb{R}(-v_0)$  by  $\varphi_0(-v_0) = 1$ . Then for  $s > 0$ , we have  $\varphi_0(s(-v_0)) = s$ , and:

$$m_U(s(-v_0)) = sm_U(-v_0) \geq s$$

Since as proved in the lemma, if  $m_U(v) < 1$ , then  $v \in U$ , so that if  $v \notin U$ , we have  $m_U(v) \geq 1$ . Now if  $s < 0$ , we have  $\varphi_0(s(-v_0)) \leq 0$ , while  $m_U(s(-v_0)) \geq 0$ . We therefore have that  $\varphi_0 \leq m_U$  on  $\mathbb{R}(-v_0)$ . By the Hahn-Banach Theorem,  $\varphi_0$  extends to  $\varphi$  on  $V$ , with  $\varphi \leq m_U$  on  $V$ .

Now note that  $U \cap (-U)$  is an open neighborhood of  $0_V$ . If  $v \in U \cap (-U)$ , then:

$$\varphi(v) \leq m_U(v) \leq 1$$



(if  $v \in U$ , then from the definition of  $m_U$ , we have  $m_U(v) \leq 1$ ). Also:

$$\varphi(-v) \leq m_U(-v) \leq 1$$

thus:

$$|\varphi(v)| \leq 1$$

for all  $v \in U \cap (-U)$ . We thus see that  $\varphi$  is bounded on a neighborhood of  $0_V$ , and thus by our lemma, we have that  $\varphi$  is continuous, i.e.  $\varphi \in V^*$ .

The last step is to show that this  $\varphi$  we have obtained separates our convex sets  $O$  and  $C$ . Given  $w \in O, c \in C$ , we have  $w - c - v_0 \in U$ . We have:

$$\varphi(w - c - v_0) \leq m_U(w - c - v_0) \leq 1$$

and thus:

$$\varphi(w) - \varphi(c) - \varphi(v_0) \leq 1$$

But note  $\varphi(-v_0) = 1$ , and thus:

$$\varphi(w) \leq \varphi(c)$$

Note this holds for all  $w \in O$ , and thus:

$$\sup \{ \varphi(w) \mid w \in O \} \leq \varphi(c)$$

and since this holds for all  $c \in C$ , we have:

$$\sup \{ \varphi(w) \mid w \in O \} \leq \inf \{ \varphi(c) \mid c \in C \}$$

Now choose some  $t_0$  between these two. We then have:

$$\varphi(w) \leq t_0 \leq \varphi(c)$$

for all  $w \in O, c \in C$ . Finally, note that  $O$  is open, so we actually must have:

$$\varphi(w) < t_0 \leq \varphi(c)$$

for all  $w \in O, c \in C$ . □

**Corollary 2.6.4.** *For a locally convex (Hausdorff) topological vector space  $V$ ,  $V^*$  separates points.*

*Proof.* It suffices to show that  $V^*$  separates points from 0. Let  $v \in V, v \neq 0$ . Set  $C = \{0\}$ , and let  $O$  be some open convex set such that  $0_V \in O, v \notin O$  (we can do this since  $V$  is Hausdorff). Then apply the Hahn-Banach Separation Theorem to obtain our  $\varphi$ . □

Let us now transition to another topic regarding convex sets. Our goal is to prove a certain theorem called the Krein-Milman Theorem. But first, we need to set the stage for it.

**Definition.** Let  $V$  be a vector space,  $C \subseteq V$  a convex subset. A nonempty, convex subset  $F \subseteq C$  is a **face** of  $C$  if whenever  $v \in F$ ,  $v = tw_1 + (1 - t)w_2$ ,  $w_1, w_2 \in C$ ,  $0 < t < 1$ , then  $w_1, w_2 \in F$ .

An **extreme point** of  $C$  is a point  $v \in C$  such that  $\{v\}$  is a face of  $C$ .

**Proposition 2.6.5.** *Let  $F$  be a face of  $C$ , and  $G$  a face of  $F$ . Then  $G$  is a face of  $C$ .*

*Proof.* Let  $v \in G$ ,  $v = tw_1 + (1 - t)w_2$ ,  $w_1, w_2 \in C$ ,  $0 < t < 1$ . Since  $G \subseteq F$ , we have  $v \in F$ . Since  $F$  is a face of  $C$ , we have  $w_1, w_2 \in F$ . Since  $G$  is a face of  $F$ , we have  $w_1, w_2 \in G$ . Thus we see that  $G$  is a face of  $C$ .  $\square$

Now that we have defined faces and extreme points, we may wonder, how do we actually obtain examples of the two? This is a good point, because it is not immediately obvious how to go about constructing faces and extreme points.

Let  $V$  be a locally convex (Hausdorff) topological vector space, and let  $C \subseteq V$  be convex and compact. Let  $\varphi \in V^*$ . Then  $\varphi|_C$  is a continuous function on a compact set, and so it takes on its supremum and infimum. Let  $F$  be the subset of  $C$  where  $\varphi$  takes on its supremum on  $C$ . Then  $F$  is a face of  $C$ , for if:

$$v = tw_1 + (1 - t)w_2 \in F$$

where  $w_1, w_2 \in C$ ,  $0 < t < 1$ , then:

$$\sup_{x \in C} \varphi(x) = \varphi(v) = t\varphi(w_1) + (1 - t)\varphi(w_2)$$

Clearly,  $\varphi$  must take on its supremum at both  $w_1, w_2$  in order for this equality to be satisfied. Thus  $w_1, w_2 \in F$ . So compactness provides lots of faces.

Suppose  $C$  has more than one point. We then have  $w_1, w_2 \in C$ ,  $w_1 \neq w_2$ . Consider a convex neighborhood  $N$  of  $w_1$  such that  $w_2 \notin N$ . Then the Hahn-Banach Separation Theorem gives us a  $\varphi \in V^*$  such that  $\varphi(w_1) \neq \varphi(w_2)$ . So at least one of  $w_1, w_2$  is where  $\varphi$  does not take its max. Thus  $F$  as defined before is a proper face of  $C$ . Thus if  $C$  has more than one point, then it has a proper face.

We have shown that compact sets contain many faces. But what about extreme points? The next theorem answers that question.

**Theorem 2.6.6. (Krein-Milman Theorem)** *Let  $V$  be a locally convex (Hausdorff) topological vector space. Let  $C$  be a compact, convex set in  $V$ . Let  $E$  be the set of extreme points of  $C$ . Then:*

$$C = \bar{c}(E)$$

where  $\bar{c}(E)$  is the closure of the convex hull of  $E$ . Note  $c(E)$  is defined as:

$$c(E) = \left\{ t_1 e_1 + \cdots + t_n e_n \mid t_j \geq 0, \sum t_j = 1, e_j \in E \right\}$$

*Proof.* Let  $F_0$  be any closed face of  $C$  (it could even be  $C$  itself). Let:

$$\mathcal{F} = \{\text{set of closed faces of } F_0\}$$

Order  $\mathcal{F}$  by reverse inclusion, i.e.  $F_1 \geq F_2$  if  $F_1 \subseteq F_2$ . We claim that  $\mathcal{F}$  is inductively ordered. Let  $\mathcal{C} \subseteq \mathcal{F}$  be a chain. Let:

$$F_* = \bigcap \{F \in \mathcal{C}\}$$

Now note that  $F_0$  is compact, and  $\mathcal{C}$  satisfies the finite intersection property, so that the intersection of the elements of  $\mathcal{C}$  is nonempty, i.e.  $F_* \neq \emptyset$ . Thus  $F_*$  is an upper bound of  $\mathcal{C}$ .

So by Zorn's Lemma,  $\mathcal{F}$  has a maximal element,  $F_\#$ . Note  $F_\#$  can only have one point, or else it would contain a proper face and thus not be maximal. Thus every closed face of  $C$  contains an extreme point of  $C$ .

Let  $D = \bar{c}(E)$ . Note  $D \subseteq C$ , since  $C$  is both convex and compact, and  $V$  is Hausdorff, so that  $C$  is closed. If  $D \neq C$ , let  $w \in C, w \notin D$ . We can find an open convex neighborhood  $U$  of  $w$  that does not meet  $D$ . By Hahn-Banach separation, we can find  $\varphi \in V^*$ ,  $\varphi(w) > \varphi(v)$  for all  $v \in D$ . Then the closed face where  $\varphi$  takes its maximum value will contain an extreme point that does not belong to  $E$ , since  $D \supseteq E$ . This is a contradiction. Therefore  $D = C$ .  $\square$

**Remark.** As a side note, Alaoglu's Theorem and the Krein-Milman Theorem are often used in tandem to prove that certain Banach spaces cannot be the dual of any space. The general procedure is to suppose a given Banach space  $V$  is the dual of some space. Then the unit ball of  $V$  is weak-\* compact, and so by the Krein-Milman Theorem, the unit ball of  $V$  should have many extreme points. We then show that it turns out that  $V$  does not have enough extreme points, i.e. the closure of the convex hull of the set of extreme points of the unit ball cannot be equal to the unit ball. This contradiction then shows that  $V$  cannot be the dual of some space.

## 2.7 Hilbert Spaces

In this section, we talk about some of the “nicest” examples of Banach spaces – Hilbert spaces. They can be thought of as generalizations of Euclidean space.

**Definition.** A Hilbert space  $\mathcal{H}$  is a vector space over  $\mathbb{R}$  or  $\mathbb{C}$  equipped with an inner product  $\langle \cdot, \cdot \rangle$  that is linear in the first argument and conjugate linear in the second. Also,  $\mathcal{H}$  must be complete with respect to the metric coming from the norm defined by:

$$\|\xi\| = \langle \xi, \xi \rangle^{1/2}$$

In short, a Hilbert space is a Banach space with inner product, with the norm coming from the inner product.

**Remark.** For any vector space with inner product, we have:

$$\langle \xi, \eta \rangle = \frac{1}{4} \sum_{n=0}^3 i^n \langle \xi + i^n \eta, \xi + i^n \eta \rangle$$

This is called **polarization**. The function  $\xi \mapsto \langle \xi, \xi \rangle$  is called the corresponding **quadratic form**. The equation says that the inner product is characterized by its quadratic form. Note also in inner product spaces, we have the **parallelogram law**:

$$\|\xi + \eta\|^2 + \|\xi - \eta\|^2 = 2(\|\xi\|^2 + \|\eta\|^2)$$

**Proposition 2.7.1.** Let  $\mathcal{H}$  be a Hilbert space, and let  $C$  be a closed convex subset. Let  $\xi_0 \in \mathcal{H}$ ,  $\xi_0 \notin C$ . Then there exists a unique point of  $C$  closest to  $\xi_0$ .

*Proof.* We can shift both  $C$  and  $\xi_0$  by  $\xi_0$ , and so it suffices to prove for the case  $\xi_0 = 0$ ,  $0 \notin C$ . Let:

$$M = \inf\{\|h\| \mid h \in C\}$$

We can find a sequence  $\{h_n\}$  in  $C$  such that  $\|h_n\| \rightarrow M$ . Rearranging the parallelogram law, we can write:

$$\frac{1}{4}\|h_m - h_n\|^2 = \frac{1}{2}(\|h_m\|^2 + \|h_n\|^2) - \|(h_m + h_n)/2\|^2$$

Note since  $C$  is convex, we have  $(h_m + h_n)/2 \in C$ , and thus:

$$\|(h_m + h_n)/2\|^2 \geq M^2$$

Also note that  $\|h_m\|^2, \|h_n\|^2 \rightarrow M^2$ . These two facts combine to show that

$$\|h_m - h_n\| \rightarrow 0$$

as  $m, n \rightarrow \infty$ . Thus  $\{h_n\}$  is Cauchy. Since  $\mathcal{H}$  is complete, we have that  $h_n \rightarrow h_0$  for some  $h_0 \in \mathcal{H}$ . Since  $C$  is closed, we have that  $h_0 \in C$ . We then have:

$$\|h_0\| = \lim_{n \rightarrow \infty} \|h_n\| = M$$

So that  $h_0$  is indeed a closest point. Now if we also have  $h_* \in C$ ,  $\|h_*\| = M$ , then:

$$\frac{1}{4}\|h_0 - h_*\|^2 = \frac{1}{2}(\|h_0\|^2 + \|h_*\|^2) - \|(h_0 + h_*)/2\|^2 \leq \frac{1}{2}(M^2 + M^2) - M^2 = 0$$

and thus  $\|h_0 - h_*\| = 0$ , and thus  $h_0 = h_*$ . Therefore  $h_0$  is unique.  $\square$

This theorem is one example why Hilbert spaces are “nice” Banach spaces. Let us uncover a few others.

**Definition.** For  $\xi, \eta \in \mathcal{H}$ , we say that they are **orthogonal** if  $\langle \xi, \eta \rangle = 0$ . We denote this by  $\xi \perp \eta$ .

**Definition.** Given a subspace  $V$  of a Hilbert space  $\mathcal{H}$ , we denote:

$$V^\perp = \{\xi \in \mathcal{H} \mid \xi \perp \eta \ \forall \eta \in V\}$$

We call  $V^\perp$  the **orthogonal complement** of  $V$ .

The next theorem shows that Hilbert spaces can be decomposed into closed subspaces and their orthogonal complements.

**Proposition 2.7.2.** *Let  $V$  be a closed subspace of  $\mathcal{H}$ . We then have:*

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

*Proof.* Let  $\xi_0 \in \mathcal{H}$ ,  $\xi_0 \notin V$  (if  $V = \mathcal{H}$ , then  $\mathcal{H} = V \oplus \{0\}$ ). A closed subspace is a closed convex subset. Thus by the previous proposition, we obtain a closest point  $\eta_0 \in V$  to  $\xi_0$ . We have:

$$\|\xi_0 - \eta_0\|^2 \leq \|\xi_0 - (\eta_0 - \eta)\|^2$$

for all  $\eta \in V$ . This can be rewritten:

$$\|\xi_0 - \eta_0\|^2 \leq \|(\xi_0 - \eta_0) - \eta\|^2$$

We have:

$$\|(\xi_0 - \eta_0) + \eta\|^2 = \|\xi_0 - \eta_0\|^2 + \langle \xi_0 - \eta_0, \eta \rangle + \langle \eta, \xi_0 - \eta_0 \rangle + \|\eta\|^2$$

Thus:

$$0 \leq 2 \operatorname{Re} \langle \xi_0 - \eta_0, \eta \rangle + \|\eta\|^2$$

for all  $\eta \in V$ . Now for all  $t > 0$ , we can replace  $\eta$  by  $t\eta$ :

$$0 \leq 2 \operatorname{Re} \langle \xi_0 - \eta_0, t\eta \rangle + \|t\eta\|^2$$

Dividing by  $t$ , we obtain:

$$0 \leq 2 \operatorname{Re} \langle \xi_0 - \eta_0, \eta \rangle + t \|\eta\|^2$$

By letting  $t \rightarrow 0$ , we see that:

$$\operatorname{Re} \langle \xi_0 - \eta_0, \eta \rangle \geq 0$$

for all  $\eta \in V$ . Now we can replace  $\eta$  by  $\alpha\eta$  for all  $\alpha \in \mathbb{C}$ . Specifically, if we choose any  $|\alpha| = 1$ , we are rotating the complex number  $\langle \xi_0 - \eta_0, \eta \rangle$ , through some angle. The only way that the real part of this complex number is nonnegative for all rotations is if in fact  $\langle \xi_0 - \eta_0, \eta \rangle = 0$ . We thus see that  $\xi_0 - \eta_0 \perp \eta$  for all  $\eta \in V$ . Thus  $\xi_0 - \eta_0 \in V^\perp$ . Thus any element not in  $V$  is the sum of an element in  $V$  with an element in  $V^\perp$ . Therefore we have:

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

□

Note how this result is a generalization for what happens in, say,  $\mathbb{R}^2$ . For example, if we have a line in  $\mathbb{R}^2$ , and a point  $(x, y)$  not on the line, then the closest point on the line to  $(x, y)$  is obtained by dropping a perpendicular from  $(x, y)$  to the line, and finding the intersection. Facts about Hilbert spaces can be visualized in this sense.

We have seen nice properties of Hilbert spaces. But perhaps the nicest property of them all is the fact that the duals of Hilbert spaces can be very easily described – in fact, the dual of a Hilbert space  $\mathcal{H}$  is itself!

**Theorem 2.7.3.** *The dual of a Hilbert space is itself.*

*Proof.* For each  $\eta$ , let  $\varphi_\eta$  be the function  $\varphi_\eta(\xi) = \langle \xi, \eta \rangle$ . Note:

$$|\varphi_\eta(\xi)| = |\langle \xi, \eta \rangle| \leq \|\xi\| \cdot \|\eta\|$$

So  $\varphi_\eta \in \mathcal{H}^*$ ,  $\|\varphi_\eta\| \leq \|\eta\|$ . Note that  $\varphi_\eta(\eta) = \|\eta\|^2$ , so that in fact  $\|\varphi_\eta\| = \|\eta\|$ . Therefore we see that the map  $\eta \mapsto \varphi_\eta$  is an isometric inclusion of  $\mathcal{H}$  into  $\mathcal{H}^*$ . Because of the conjugate-linearity of the second argument of  $\langle, \rangle$ , the map is actually anti-linear.

Now to show that  $\mathcal{H} = \mathcal{H}^*$ , we need to show that the map we have defined is onto. So let  $\psi \in \mathcal{H}^*$ ,  $\psi \neq 0$ . Let  $\mathcal{K} = \ker \psi$ . So  $\mathcal{K}$  is a closed (proper) subspace of  $\mathcal{H}$ . We can then write  $\mathcal{H} = \mathcal{K} \oplus \mathcal{K}^\perp$ . Thus there exists  $\eta_1 \in \mathcal{K}^\perp$ ,  $\psi(\eta_1) = 1$ . Then for any  $\xi \in \mathcal{H}$ , we have:

$$\psi(\xi - \psi(\xi)\eta_1) = 0$$

Thus  $\xi - \psi(\xi)\eta_1 \in \mathcal{H}$  for all  $\xi \in \mathcal{H}$ . Since  $\eta_1 \in \mathcal{H}^\perp$ , we have:

$$\begin{aligned}\langle \xi - \psi(\xi)\eta_1, \eta_1 \rangle &= 0 \\ \langle \xi, \eta_1 \rangle &= \psi(\xi)\langle \eta_1, \eta_1 \rangle\end{aligned}$$

We thus have:

$$\psi(\xi) = \left\langle \xi, \frac{\eta_1}{\|\eta_1\|^2} \right\rangle$$

and therefore  $\psi = \varphi_{\eta_1/\|\eta_1\|^2}$ .  $\square$

The theorem shows that  $\mathcal{H} \cong \mathcal{H}^*$ , where the  $\cong$  in this case denotes anti isomorphism. Now if we take the double dual, we get that  $\mathcal{H} \cong \mathcal{H}^{**}$ , where now  $\cong$  denotes isomorphism. Therefore  $\mathcal{H}$  is reflexive. Note this implies (by Alaoglu's Theorem) that the unit ball in  $\mathcal{H}$  is weakly compact.

Let  $\mathcal{H}, \mathcal{K}$  be Hilbert spaces, and let  $T \in \mathcal{B}(\mathcal{H}, \mathcal{K})$ . For any  $\eta \in \mathcal{K}$ , define:

$$\psi_\eta(\xi) = \langle T\xi, \eta \rangle$$

This is a bounded linear functional on  $\mathcal{H}$ . Since  $\mathcal{H}$  is self-dual, there exists  $\xi_\eta \in \mathcal{H}$  such that:

$$\langle T\xi, \eta \rangle = \langle \xi, \xi_\eta \rangle$$

We thus see that the map  $\eta \mapsto \xi_\eta$  is linear, and plugging in  $\xi = \xi_\eta$ , we get:

$$\langle T\xi_\eta, \eta \rangle = \langle \xi_\eta, \xi_\eta \rangle = \|\xi_\eta\|^2$$

By Cauchy-Schwarz, we get:

$$\|\xi_\eta\|^2 \leq \|T\xi_\eta\| \cdot \|\eta\| = \|T\| \cdot \|\xi_\eta\| \cdot \|\eta\|$$

and therefore we can obtain:

$$\|\xi_\eta\| \leq \|T\| \cdot \|\eta\|$$

Denote the map  $\eta \mapsto \xi_\eta$  by  $T^*$ . We have:

$$\langle T\xi, \eta \rangle = \langle \xi, T^*\eta \rangle$$

for all  $\xi \in \mathcal{H}, \eta \in \mathcal{K}$ . Also, since  $\|\xi_\eta\| \leq \|T\| \cdot \|\eta\|$ , we get  $\|T^*\| \leq \|T\|$ . It turns out that actually,  $\|T^*\| = \|T\|$ . Furthermore,  $T \mapsto T^*$  is anti-linear,  $(T^*)^* = T$ , and  $(ST)^* = T^*S^*$ .

**Definition.** The map  $T^*$  we have defined is called the **adjoint** of  $T$ .

**Definition.** Let  $T \in \mathcal{B}(\mathcal{H}, \mathcal{H})$ . We say that  $T$  is **self-adjoint** if  $T = T^*$ . We say that  $T$  is **unitary** if  $T^* = T^{-1}$ .

There are a few more facts about Hilbert that we will quickly go over. For a more in depth treatment, please consult a textbook.

**Definition.** We say that a set  $E \subseteq \mathcal{H}$  is orthonormal if for all  $\xi \in E$ , we have  $\|\xi\| = 1$ , and for all  $\xi, \eta \in E$ ,  $\xi \neq \eta$ , we have  $\xi \perp \eta$ .

Let  $\{e_j\} \subseteq \mathcal{H}$  be orthonormal. Given  $\xi \in \mathcal{H}$ , let:

$$\eta_n = \sum_{j=1}^n \langle \xi, e_j \rangle e_j$$

Note if  $\xi \perp \eta$ , then  $\|\xi + \eta\|^2 = \|\xi\|^2 + \|\eta\|^2$ . Applying this to our present situation, we get:

$$\|\eta_n\|^2 = \sum_{j=1}^n |\langle \xi, e_j \rangle|^2$$

We also have  $\xi - \eta_n \perp \eta_n$ , so that:

$$\|\xi\|^2 = \|\xi - \eta_n\|^2 + \|\eta_n\|^2$$

therefore  $\|\eta_n\| \leq \|\xi\|$  for all  $n$ . Writing this out:

$$\sum_{j=1}^n |\langle \xi, e_j \rangle|^2 \leq \|\xi\|^2$$

This is called **Bessel's Inequality**.

**Definition.** We say that  $\{e_j\} \subseteq \mathcal{H}$  is an **orthonormal basis** if  $\{e_j\}$  is orthonormal and the span of  $\{e_j\}$  is dense in  $\mathcal{H}$ .

It can be shown that if  $\{e_j\}$  is an orthonormal basis, then for all  $\xi \in \mathcal{H}$ , we have:

$$\xi = \sum_{j=1}^{\infty} \langle \xi, e_j \rangle e_j$$

We can then show:

$$\|\xi\|^2 = \sum_{j=1}^{\infty} |\langle \xi, e_j \rangle|^2$$

This is called **Parseval's Equality**.



# Chapter 3

## $L^p$ Spaces

We now return to integration and consider a more general class of “integrable” functions, the  $L^p$  spaces, for  $0 < p < \infty$ . We will also use the theory of Functional Analysis to obtain some important results.

### 3.1 $L^p$ as a Banach Space

**Definition.** We define:

$$\mathcal{L}^p(X, S, \mu, B) = \left\{ f \in \mathcal{M}(X, S, \mu, B) \mid (x \mapsto \|f(x)\|^p) \in \mathcal{L}^1(X, S, \mu, R) \right\}$$

**Proposition 3.1.1.**  $\mathcal{L}^p$  is a vector space for the pointwise operations.

*Proof.* Let  $f, g \in \mathcal{L}^p$ . Then:

$$\begin{aligned} \|f(x) + g(x)\|^p &\leq (\|f(x)\| + \|g(x)\|)^p \\ &\leq (2(\|f(x)\| \vee \|g(x)\|))^p \\ &\leq 2^p (\|f(x)\| \vee \|g(x)\|)^p \\ &\leq 2^p (\|f(x)\|^p + \|g(x)\|^p) \end{aligned}$$

Note  $x \mapsto \|f(x)\|^p, x \mapsto \|g(x)\|^p \in \mathcal{L}^1$ , and so  $x \mapsto \|f(x) + g(x)\|^p$  is in  $\mathcal{L}^1$ , by the corollary to the Lebesgue Dominated Convergence (LDC) theorem. Scalar multiplication is easier:

$$\|rf(x)\|^p = |r|^p \|f(x)\|^p$$

and  $x \mapsto |r|^p \|f(x)\|^p \in \mathcal{L}^1$ . □

**Definition.** For  $f \in \mathcal{L}^p$ , we set its  $p$ -norm to be:

$$\|f\|_p = \left( \int \|f(x)\|^p d\mu(x) \right)^{\frac{1}{p}}$$

**Remark.** Note then that  $\|rf\|_p = |r| \cdot \|f\|_p$ .

However, for  $0 < p < 1$ ,  $\|\cdot\|_p$  is not a norm (the triangle inequality fails). In fact, the closed convex hull of:

$$\{f \mid \|f\|_p \leq 1\}$$

is all of  $\mathcal{L}^p$  (if  $\|\cdot\|_p$  is a norm, then it should be the unit ball). Looking ahead, this means that there are no linear functionals on  $\mathcal{L}^p$ ,  $0 < p < 1$ , that are continuous.

We can also note that  $\|f\|_p = 0$  implies  $f = 0$  a.e. Thus we can obtain  $L^p$ , the a.e. equivalence classes of  $\mathcal{L}^p$ . To prove that  $\|\cdot\|_p$  is a norm on  $L^p$  for  $1 \leq p < \infty$ , we need a Cauchy-Scharz type inequality.

**Lemma 3.1.2.** Let  $1 < p < \infty$ . Let  $q$  be defined by:

$$\frac{1}{p} + \frac{1}{q} = 1$$

Note  $q$  is uniquely determined. Then for  $r, s \in \mathbb{R}^+$ :

$$rs \leq \frac{r^p}{p} + \frac{s^q}{q}$$

*Proof.* Let  $a = r^p$ ,  $b = s^q$ . We need:

$$a^{1/p} b^{1/q} \leq \frac{a}{p} + \frac{b}{q}$$

Dividing by  $b$ :

$$a^{1/p} b^{-1/p} \leq \frac{a/b}{p} + 1/q$$

$$\left(\frac{a}{b}\right)^{1/p} \leq \frac{a/b}{p} + \frac{1}{q}$$

So we need for  $t = \frac{a}{b} \in \mathbb{R}$ ,  $0 < t < \infty$ :

$$t^{1/p} \leq \frac{t}{p} + \frac{1}{q} = \frac{t}{p} + 1 - \frac{1}{p}$$

Let  $f(t) = \frac{t}{p} + 1 - \frac{1}{p} - t^{1/p}$ . We want  $f \geq 0$ . This is a calculus problem. Note  $f(1) = 0$ , and:

$$f'(t) = \frac{1}{p} - \frac{1}{p} t^{(1/p)-1}$$

Note:

$$\frac{1}{p} - 1 = \frac{1}{q} < 0$$

For  $t > 1$ ,  $t^{(1/p)-1} < 1$ , so  $f'(t) > 0$ . For  $0 < t < 1$ , we can see that  $f'(t) < 0$ . Thus we get that  $f \geq 0$ .  $\square$

**Lemma 3.1.3. (Holder's Inequality)** For  $1 < p < \infty$ ,  $\frac{1}{p} + \frac{1}{q} = 1$ ,  $f \in \mathcal{L}^p(X, S, \mu, B)$ ,  $g \in \mathcal{L}^q(X, S, \mu, B)$ , we have:

$$(x \mapsto \|f(x)\| \cdot \|g(x)\|) \in \mathcal{L}^1(X, S, \mu, \mathbb{R})$$

and:

$$\int \|f\| \cdot \|g\| d\mu \leq \|f\|_p \|g\|_q$$

*Proof.* We have:

$$\|f(x)\| \cdot \|g(x)\| \leq \frac{\|f(x)\|^p}{p} + \frac{\|g(x)\|^q}{q}$$

for all  $x \in X$ . Note by definition,  $x \mapsto \frac{\|f(x)\|^p}{p}$ ,  $x \mapsto \frac{\|g(x)\|^q}{q} \in \mathcal{L}^1$ . We thus see that  $x \mapsto \|f(x)\| \cdot \|g(x)\|$  is dominated by a function in  $\mathcal{L}^1$ , and thus by the corollary to LDC, we get that  $x \mapsto \|f(x)\| \cdot \|g(x)\| \in \mathcal{L}^1$ . Even more:

$$\int \|f(x)\| \cdot \|g(x)\| d\mu(x) \leq \frac{1}{p} \int \|f(x)\|^p d\mu(x) + \frac{1}{q} \int \|g(x)\|^q d\mu(x)$$

Applying this inequality to the functions  $\frac{\|f(x)\|}{\|f\|_p}$ ,  $\frac{\|g(x)\|}{\|g\|_q}$ , we get:

$$\frac{1}{\|f\|_p \|g\|_q} \int \|f(x)\| \cdot \|g(x)\| d\mu(x) \leq \frac{1}{p} \int \frac{\|f(x)\|^p}{\|f\|_p^p} d\mu(x) + \frac{1}{q} \int \frac{\|g(x)\|^q}{\|g\|_q^q} d\mu(x)$$

By the definition of  $\|f\|_p$ ,  $\|g\|_q$ , we obtain:

$$\leq \frac{1}{p} + \frac{1}{q} = 1$$

and therefore:

$$\int \|f(x)\| \cdot \|g(x)\| d\mu(x) \leq \|f\|_p \|g\|_q$$

□

**Lemma 3.1.4. (Minkowski's Inequality)** For  $p > 1$ ,  $f, g \in \mathcal{L}^p$ , we have:

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p$$

*Proof.* Let  $q$  be defined by  $\frac{1}{p} + \frac{1}{q} = 1$ . Then for  $x \in X$ :

$$\begin{aligned} \|f(x) + g(x)\|^p &= \|f(x) + g(x)\| \cdot \|f(x) + g(x)\|^{p-1} \\ &= \|f(x) + g(x)\| \cdot \|f(x) + g(x)\|^{p/q} \\ &\leq \|f(x)\| \cdot \|f(x) + g(x)\|^{p/q} + \|g(x)\| \cdot \|f(x) + g(x)\|^{p/q} \end{aligned}$$

Note that the terms on the right are integrable. We have:

$$\int \|f + g\|^p d\mu \leq \int \|f\| \cdot \|f + g\|^{p/q} d\mu + \int \|g\| \cdot \|f + g\|^{p/q} d\mu(x)$$

Note  $f + g \in \mathcal{L}^p$ , and thus  $(f + g)^{p/q} \in \mathcal{L}^q$ . Thus by Holder's inequality, the first term on the right satisfies:

$$\begin{aligned} \int \|f\| \cdot \|f + g\|^{p/q} d\mu &\leq \|f\|_p \cdot \|f + g\|_q^{p/q} \\ &\leq \|f\|_p \left( \int (f + g)^p d\mu \right)^{\frac{1}{q}} \\ &\leq \|f\|_p \left( \int (f + g)^p d\mu \right)^{\frac{1}{p} \frac{p}{q}} \\ &\leq \|f\|_p \cdot \|f + g\|_p^{p/q} \end{aligned}$$

Similarly, the second term on the right is  $\leq \|g\|_p \cdot \|f + g\|_p^{p/q}$ . We can therefore obtain:

$$\|f + g\|_p^p \leq \|f\|_p \cdot \|f + g\|_p^{p/q} + \|g\|_p \cdot \|f + g\|_p^{p/q}$$

Assuming  $\|f + g\|_p \neq 0$ , we get:

$$\begin{aligned} \|f + g\|_p^{p - \frac{p}{q}} &\leq \|f\|_p + \|g\|_p \\ \|f + g\|_p &\leq \|f\|_p + \|g\|_p \end{aligned}$$

□

**Theorem 3.1.5.** For  $1 \leq p < \infty$ ,  $\|\cdot\|_p$  is a norm on  $L^p$ .

*Proof.* We have already proven this for  $p = 1$ . For  $1 < p < \infty$ , we have shown that  $\|\cdot\|_p$  satisfies the triangle inequality. The other two conditions are simple enough to see. □

So we have now turned  $L^p$  ( $1 \leq p < \infty$ ) into a vector space. One natural question to ask is whether it is in fact a Banach space; i.e., is it complete?

**Lemma 3.1.6.** Convergence in “ $p$ -mean” implies convergence in measure.

*Proof.* Let  $\varepsilon > 0$ ,  $f \in \mathcal{L}^p$ ,  $E_\varepsilon = \{x \mid \|f(x)\| \geq \varepsilon\}$ . We have:

$$\chi_{E_\varepsilon}(x) \leq \frac{\|f(x)\|}{\varepsilon} \leq \frac{\|f(x)\|^p}{\varepsilon^p} \in \mathcal{L}^1$$

And so by a previous lemma, we have that  $\mu(E_\varepsilon) < \infty$ , and actually:

$$\mu(E_\varepsilon) \leq \int \frac{\|f(x)\|^p}{\varepsilon^p} d\mu \leq \frac{1}{\varepsilon^p} \|f\|_p^p$$

Applying this to  $f - f_n$ , let  $E_\varepsilon = \{x \mid \|f(x) - f_n(x)\| \geq \varepsilon\}$ . We have:

$$\mu(E_\varepsilon) \leq \frac{1}{\varepsilon^p} \|f - f_n\|_p^p$$

We thus see that if  $\{f_n\}$  is a sequence in  $\mathcal{L}^p$ , and if  $\|f - f_n\|_p \rightarrow 0$  ( $f \in \mathcal{L}^p$ ), then  $\mu(E_\varepsilon) \rightarrow 0$ , so that  $f_n \rightarrow f$  in measure.  $\square$

**Remark.** In the same way, we see that when we apply this proof to functions of the form  $f_m - f_n$ , we get that a sequence that is  $p$ -mean Cauchy is Cauchy in measure.

**Theorem 3.1.7.**  $L^p$  is complete.

*Proof.* Let  $\{f_n\}$  be a  $p$ -mean Cauchy sequence in  $\mathcal{L}^p$ . It is Cauchy in measure, so there is a subsequence that converges a.u. (and so a.e.) to some function  $f$  that is measurable. We know  $f$  is unique a.e., and  $f_n \rightarrow f$  in measure. We want to show that if  $\{f_n\}$  is  $p$ -mean Cauchy, and if  $f_n \rightarrow f$  a.e., then  $f \in \mathcal{L}^p$  and  $f_n \rightarrow f$  in  $p$ -mean. I.e.:

$$\int \|f - f_n\|^p d\mu \rightarrow 0$$

We have for fixed  $n$ :

$$\|f_m(x) - f_n(x)\|^p \rightarrow \|f(x) - f_n(x)\|^p \quad (m \rightarrow \infty)$$

a.e. Thus we can apply Fatou's lemma:

$$\int \|f - f_n\|^p d\mu \leq \liminf \int \|f_m - f_n\|^p d\mu$$

Given  $\varepsilon > 0$ , there exists  $N$  such that for  $m, n \geq N$ :

$$\|f_m - f_n\|_p < \varepsilon$$

i.e.:

$$\int \|f_m - f_n\|^p d\mu < \varepsilon^p$$

In particular, for fixed  $n \geq N$ , and all  $m \geq N$ , we have for these conditions, that:

$$\liminf \left\{ \int \|f_m - f_n\|^p d\mu \right\} < \varepsilon^p$$

Thus for  $n \geq N$ :

$$\int \|f - f_n\|^p d\mu < \varepsilon^p$$

Since this integral is finite, we get that  $f - f_n \in \mathcal{L}^p$ , and thus  $f \in \mathcal{L}^p$ . We can also see that  $\|f - f_n\|_p \rightarrow 0$ .  $\square$