Chapter 2

Normed and Hilbert Spaces

2.1 Topics to be covered

- Normed spaces

  ℓ^p spaces, Holder inequality, Minkowski inequality, Riesz-Fischer theorem

  The space $C(X)$

  Quotients and conditions for completeness, the 2/3's theorem

  Finite dimensional normed spaces, equivalence of norms

  Convexity, absolute convexity, the bipolar theorem

  Consequences of Baire's theorem:

  Principle of Uniform Boundedness, Resonance Principle

  Open mapping, closed graph and bounded inverse theorems

  Hahn-Banach theorem

  Krein-Milman theorem

  Dual spaces and adjoints

  The double dual

  Weak topologies, weak convergences

- Hilbert spaces

  Cauchy-Schwarz inequality

  Polarization identity, Parallelogram Law

  Jordan-von Neumann theorem

Orthonormal bases and Parseval identities

Direct sums

Bilinear maps and tensor products of Banach and Hilbert spaces

Infinite tensor products and quantum spin chains

2.2 Banach Spaces

All vector spaces will be over the field $\mathbb{R}$ or $\mathbb{C}$, when we wish to consider a notion that applies to both fields we shall write $\mathbb{F}$.

Definition 2.1. Let X be a vector space over $\mathbb{F}$. Then a **norm** on X is a function, $\|\cdot\| : X \to \mathbb{R}^+$ satisfying:

1. $\|x\| = 0 \iff x = 0$,
2. $\|\lambda x\| = |\lambda| \cdot \|x\|$, $\forall \lambda \in \mathbb{F}$, $\forall x \in X$,
3. (Triangle Inequality) $\|x + y\| \leq \|x\| + \|y\|$, $\forall x, y \in X$.

We call the pair $(X, \|\cdot\|)$ a **normed linear space** or **n.l.s.**, for short.

Proposition 2.2. *Let $(X, \|\cdot\|)$ be a n.l.s.. Then $d(x, y) = \|x-y\|$ is a metric on X (called the "induced metric") that satisfies $d(x + z, y + z) = d(x, y)$ ("translation invariance") and $d(\lambda x, \lambda y) = |\lambda| d(x, y)$ ("scaled").*

Conversely, if X is a vector space and d is a metric on X that is scaled and translation invariant, then $\|x\| = d(0, x)$ is a norm on X.

Definition 2.3. A n.l.s. $(X, \|\cdot\|)$ is a **Banach space** if and only if it is complete in the induced metric, $d(x, y) = \|x - y\|$.

Some examples are in order.

Example 2.4. Let $X = C_{\mathbb{R}}([0, 1])$ denote the vector space of continuous real-valued functions on the unit interval. Given $f \in X$ we set

$$\|f\|_\infty = \sup\{|f(t)| : 0 \leq t \leq 1\}.$$

We leave it to the reader to verify that this is a norm. We claim that $(X, ||\cdot||_\infty)$ is a Banach space. Note that a sequence of functions $\{f_n\} \subseteq X$ is Cauchy if and only if for each $\epsilon > 0$ there is N so that for $n, m > N$,

$$d_\infty(f_n, f_m) = \sup\{|f_n(t) - f_m(t)| : 0 \leq t \leq 1\} < \epsilon.$$

Now we recall some arguments from undergraduate analysis. For each $0 \leq t_0 \leq 1$, $|f_n(t_0) - f_m(t_0)| \leq \|f_n - f_m\|_\infty$ and so $\{f_n(t_0)\}$ is a Cauchy sequence of real numbers. Hence, $\lim f_n(t_0)$ exists and we defined $f(t_0)$ to be this value. Next one shows that

$$\lim_n \sup\{|f(t) - f_n(t)| : 0 \leq t \leq 1\} = 0,$$

so that the sequence $\{f_n\}$ converges *uniformly* to f and recall from undergraduate analysis that this implies that f is continuous. Thus, $f \in X$ and the last equation shows that

$$\|f - f_n\|_\infty \to 0.$$

Thus, $(X, \|\cdot\|_\infty)$ is a Banach space.

The next example is a vector space that is not a Banach space.

Example 2.5. Let $X = C_{\mathbb{R}}([0,1])$ as before, but now set

$$\|f\|_1 = \int_0^1 |f(t)|dt.$$

Note that if $f \in X$ and $f \neq 0$ then, since f is continuous, there is a $\delta > 0$ and a small interval $[a, b]$ on which $|f(t)| \geq \delta$. Hence, $\|f\|_1 \geq \delta(b - a)$. This proves that $\|f\|_1 = 0$ if and only if $f = 0$, and we have verified one property of a norm. We leave it to the reader to verify that the remaining properties of a norm are met.

We claim that $(X, \|\cdot\|_1)$ is not complete. To see this consider the following sequence of functions. For each $n \geq 3$ set

$$f_n(t) = \begin{cases} 0, & 0 \leq t \leq \frac{1}{2} - \frac{1}{n}, \\ \frac{2nt-n+2}{4}, & \frac{1}{2} - \frac{1}{n} \leq t \leq \frac{1}{2} + \frac{1}{n}, \\ 1, & \frac{1}{2} + \frac{1}{n} \leq t \leq 1. \end{cases}$$

For $n, m > N$ the functions f_n and f_m are equal except on a subset of the interval $[\frac{1}{2} - \frac{1}{N}, \frac{1}{2} + \frac{1}{N}]$, and on this subinterval they can differ by at most 1. Hence,

$$\|f_n - f_m\|_1 \leq \frac{2}{N},$$

and so the sequence is Cauchy. Now check that no continuous function can be the limit of these functions in $\|\cdot\|_1$.

Example 2.6. For $1 \leq p < +\infty$ and $n \in \mathbb{N}$, and $x = (x_1, ..., x_n) \in \mathbb{R}^n$, set

$$\|(x_1, \ldots, x_n)\|_p = \big(\sum_{j=1}^{n} |x_j|^p\big)^{1/p},$$

then this is a norm called the *p-norm* and we write ℓ_n^p to denote $\mathbb{R}^n$ endowed with this norm. It is not so easy to see that this satisfies the triangle inequality, this depends on some results below. For $p = +\infty$ we set

$$\|(x_1, \ldots, x_n)\| = \max\{|x_1|, \ldots, |x_n|\}.$$

Example 2.7. More generally, for $1 \leq p < +\infty$, we let

$$\ell^p = \{(x_1, x_2, \ldots) : \sum_{j=1}^{+\infty} |x_j|^p < +\infty\},$$

and for $x \in \ell^p$ we set $\|x\|_p = \big(\sum_{j=1}^{+\infty} |x_j|^p\big)^{1/p}$.

For $p = +\infty$, we set

$$\ell^\infty = \{(x_1, x_2, \ldots) : \sup_j |x_j| < +\infty\},$$

and define $\|x\|_\infty = \sup_j |x_j|$.

Then for $1 \leq p \leq +\infty$, $(\ell^p, \|\cdot\|_p)$ are all Banach spaces. This fact relies on a number of theorems that we will state below.

Problem 2.8. Prove that $(\ell^1, \|\cdot\|_1)$ and $(\ell^\infty, \|\cdot\|_\infty)$ are Banach spaces.

Problem 2.9. Let $c_0 = \{(x_1, x_2, ...) : \lim_n x_n = 0\}$. Prove that c_0 is a vector subspace of ℓ^∞ and that it is closed in the metric induced by the $\|\cdot\|_\infty$ norm.

Lemma 2.10 (Young's Inequality)**.** *Let $1 < p, q < +\infty$ with $\frac{1}{p} + \frac{1}{q} = 1$. Then for any $a, b \geq 0$ we have $ab \leq \frac{a^p}{p} + \frac{b^q}{q}$.*

Problem 2.11. Prove this inequality.

Proposition 2.12 (Holder's Inequality)**.** *Let $1 < p, q < +\infty$ with $\frac{1}{p} + \frac{1}{q} = 1$, and let $x = (x_1, \ldots) \in \ell^p$, $y = (y_1, \ldots) \in \ell^q$, then $(x_1 y_1, \ldots) \in \ell^1$ and*

$$\sum_n |x_n y_n| \leq \|x\|_p \|y\|_q.$$

Problem 2.13. Prove Holder's inequality.

Theorem 2.14 (Minkowski's Inequality)**.** *Let $1 \le p \le +\infty$, and let $x, y \in \ell^p$, then $x + y \in \ell^p$ and $\|x + y\|_p \le \|x + y\|_p$.*

Proof. We shall only prove the case that $1 < p < +\infty$ for succinctness. First note that

$$|x_i + y_i| \le \begin{cases} 2|x_i| & \text{when } |x_i| \ge |y_i| \\ 2|y_i| & \text{when } |x_i| \le |y_i| \end{cases}.$$

Hence, $|x_i + y_i|^p \le \max\{2^p|x_i|^p,\, 2^p|y_i|^p\}$, from which it follows that $\sum_i |x_i + y_i|^p < +\infty$.

Next notice that

$$|x_i + y_i|^p \le |x_i + y_i|^{p-1}(|x_i| + |y_i|).$$

Also, $p + q = pq$ so that $(p-1)q = p$. From this it follows that the sequence $|x_i + y_i|^{p-1}$ is q-summable. By Holder's inequality,

$$\begin{aligned} \sum_i |x_i + y_i|^p &\le \big(\sum_i (|x_i + y_i|^{p-1})^q\big)^{1/q}\big((\sum_i |x_i|^p)^{1/p} + (\sum_i |y_i|^p)^{1/p}\big) \\ &= \big(\sum_i |x_i + y_i|^p\big)^{1/q}\big((\sum_i |x_i|^p)^{1/p} + (\sum_i |y_i|^p)^{1/p}\big). \end{aligned}$$

Cancelling the common term from each side and using $\frac{1}{q} = 1 - \frac{1}{p}$ yields the result. $\square$

Theorem 2.15 (Riesz-Fischer)**.** *The spaces $(\ell^p, \|\cdot\|_p)$ are Banach spaces.*

Given a normed space $(X, \|\cdot\|)$, a sequence of vectors $\{x_n\}$ and a vector x, we write

$$x = \sum_{n=1}^{\infty} x_n,$$

to mean that $\lim_N \|x - \sum_{n=1}^{N} x_n\| = 0$, i.e., that the partial sums converge in norm. Similarly, we write $x = \sum_{n \in \mathbb{N}} x_n$ to mean that the net of finite sums $s_F = \sum_{n \in F} x_n$ converges in the metric induced by the norm, i.e., that given $\epsilon > 0$, there is a finite set F_0 such that whenever F is a finite set with $F_0 \subseteq F$, then $\|x - s_F\| < \epsilon$.

We shall let e_i, also denoted $|i\rangle$, be the vector that is 1 in the i-th entry and 0 in every other entry. Given $x = (x_1, x_2, ...) \in \ell^p$, we can write

$$x = \sum_{i=1}^{\infty} x_i e_i.$$

For $1 \le p < +\infty$ this notation makes sense as the series does converge. In fact, one can show the even stronger statement that $x = \sum_{i\in\mathbb{N}} x_i e_i$. However, this notation is somewhat misleading for $p = \infty$. For example, if we set $x_i = 1$ for all i, then $x + (1, 1, ...) \in \ell^\infty$ and $\|x\|_\infty = 1$. But

$$\|x - \sum_{i=1}^{N} x_i e_i\|_\infty = 1$$

so the series does not converge to x. In fact, any two partial sums of this series are distance 1 apart, so the partial sums are not even Cauchy. Later we will see that there is a different topology on ℓ^∞ for which this series does converge to x in that topology.

Problem 2.16. Prove that for $x = (x_1, ...) \in \ell^p, 1 < +\infty$, we have $x = \sum_{i=1}^\infty x_i e_i = \sum_{i\in\mathbb{N}} x_i e_i$. Thus verifying the above claim.

Problem 2.17. Let $x = (x_1, ...) \in \ell^\infty$. Prove that the series $\sum_{i=1}^\infty x_i e_i$ converges to x (in norm) if and only if $x \in c_0$. What about $\sum_{i\in\mathbb{N}} x_i e_i$?

Another useful test for determining if a normed space is a Banach space is given in terms of convergent series.

Theorem 2.18. *A normed space $(X, \|\cdot\|)$ is a Banach space if and only if whenever $\{x_n\} \subseteq X$ is a sequence such that $\sum_{n=1}^\infty \|x_n\| < +\infty$ then there is an element $x \in X$ such that $\sum_{n=1}^{+\infty} x_n = x$.*

This theorem is sometimes stated as a normed space is a Banach space if and only if every *absolutely convergent series* is convergent. Here a series $\sum_{n=1}^\infty x_n$ is called absolutely convergent provided that $\sum_{n=1}^\infty \|x_n\| < +\infty$.

One final example of a family of Banach spaces.

Example 2.19. Let $(K, \mathcal{T})$ be a compact Hausdorff space and let $C(K)$ denote the space of continuous real or complex valued functions on K. This is clearly a vector space. If we set $\|f\|_\infty = \sup\{|f(x)|; x \in K\}$, then this defines a norm and $(C(K), \|\cdot\|_\infty)$ is a Banach space. Completeness follows from the fact that convergence in this norm is uniform convergence and the fact that uniformly convergent sequences of continuous functions converge to a continuous function.

2.2.1 Bounded and Continuous

Proposition 2.20. *Let $(X, \|\cdot\|_1)$ and $(Y, \|\cdot\|_2)$ be normed spaces and let $T : X \to Y$ be a linear map. Then the following are equivalent:*

1. *T is continuous,*

2. *T is continuous at 0,*

3. *there is a constant M such that $\|Tx\|_2 \leq M\|x\|_1$, $\forall x \in X$*

Definition 2.21. Let X and Y be normed spaces and let $T : X \to Y$ be a linear map. Then T is **bounded** if there exists a constant M such that $\|Tx\| \leq M\|x\|, \forall x \in X$. We set

$$\|T\| := \sup\{\|Tx\| : \|x\| \leq 1\} < \infty$$

and this is called the **operator norm** of T.

Some other formulas for this value include

$$\|T\| = \sup\{\|Tx\| : \|x\| = 1\} = \inf\{M : \|Tx\| \leq M\|x\|, \forall x \in X\}.$$

2.2.2 Equivalence of Norms

Suppose that $\|\cdot\|_i, i = 1, 2$ are norms on X and that $d_i, i = 1, 2$ are the metrics that they induce. Recall that we say the metrics are equivalent if and only if the maps $id : (X, d_1) \to (X, d_2)$ and $id : (X, d_2) \to (X, d_1)$ are both continuous. But since these are normed spaces this is the same as requiring that both maps be bounded, which is the same as both metrics being uniformly equivalent. Thus, the metrics that come from norms are equivalent if and only if they are uniformly equivalent.

This leads to the following definition.

Definition 2.22. Let $(X, \|\cdot\|_i), i = 1, 2$ be two norms on X. Then we say that these norms are **equivalent** provided that there exists constants, $A, B > 0$ such that

$$A\|x\|_1 \leq \|x\|_2 \leq B\|x\|_1, \forall x \in X.$$

Note that these last inequalities are the same as requiring that there are constants $C, D > 0$ such that $\|x\|_2 \leq C\|x\|_1$ and $\|x\|_1 \leq D\|x\|_2$. If we think geometrically then this is the same as requiring that $A \cdot \mathcal{B}_1 \subseteq \mathcal{B}_2 \subseteq B \cdot \mathcal{B}_1$, where $\mathcal{B}_i$ denotes the unit balls in each norm.

Finally, note that "equivalence of norms" really is an equivalence relation on the set of all norms on a space X, that is, it is a symmetric ($\|\cdot\|_1$ equivalent to $\|\cdot\|_2$ if and only if $\|\cdot\|_2$ is equivalent to $\|\cdot\|_1$) and transitive ($\|\cdot\|_1$ equivalent to $\|\cdot\|_2$ and $\|\cdot\|_2$ equivalent to $\|\cdot\|_3$ implies that $\|\cdot\|_1$ is equivalent to $\|\cdot\|_3$) relation.

Proposition 2.23 (Reverse Triangle Inequality)**.** *Let $(X, \|\cdot\|)$ be a normed space. Then*

$$|\|x\| - \|y\|| \leq \|x - y\| \quad \forall x, y \in X.$$

Theorem 2.24. *All norms on $\mathbb{R}^n$ are equivalent.*

Proof. It will be enough to show that an arbitrary norm $\|\cdot\|$ on $\mathbb{R}^n$ is equivalent to the Euclidean norm $\|\cdot\|_2$. Let $e_i = |i\rangle$ denote the standard basis for $\mathbb{R}^n$, so that $x = (x_1, ..., x_n) = x_1 e_1 + \cdots + x_n e_n$. Then,

$$\|x\| \leq |x_1|\|e_1\| + \cdots + |x_n|\|e_n\| \leq \big(\sum_i |x_i|^2\big)^{1/2}\big(\sum_i \|e_i\|^2\big)^{1/2} \leq C\|x\|_2,$$

where $C = \big(\sum_i \|e_i\|^2\big)^{1/2}$.

Let $f : \mathbb{R}^n \to \mathbb{R}$ be defined by $f(x) = \|x\|$. By the reverse triangle inequality we have that,

$$|f(x) - f(y)| \leq \|x - y\| \leq C\|x - y\|_2,$$

which shows that f is continuous from $\mathbb{R}^n$ to $\mathbb{R}$ when $\mathbb{R}^n$ is given the usual Euclidean topology.

Now $K = \{x : \|x\|_2 = 1\}$ is a compact set and so for some $x_0 \in X$,

$$A = \inf\{f(x) : x \in K\} = f(x_0) = \|x_0\| \neq 0.$$

Now for any non-zero $x \in \mathbb{R}^n$, $\frac{x}{\|x\|_2} \in K$ and hence,

$$A \leq \|\frac{x}{\|x\|_2}\| = \frac{\|x\|}{\|x\|_2},$$

so that $A\|x\|_2 \leq \|x\|$ and we are done. □

2.3 Consequences of Baire's Theorem

Many results are true for Banach spaces that are not true for general normed spaces. In particular, this is because Banach spaces are complete metric spaces and so we have Baire's theorem as a tool. In this section we present several facts about Banach spaces that rely on Baire's theorem. The version of Baire's theorem that we shall use most often is the version that says that in a complete metric space, a countable union of nowhere dense sets is still nowhere dense.

2.3.1 Principle of Uniform Boundedness

We begin with an example to show why the following result is perhaps surprising. For $n \geq 2$, let $f_n : [0,1] \to \mathbb{R}$ be defined by

$$f_n(t) = \begin{cases} n^2 t, & 0 \leq t \leq 1/n \\ 2n - n^2 t, & 1/n \leq t \leq 2/n \\ 0, & 2/n \leq t \leq 1 \end{cases},$$

so that each f_n is continuous.

We have $\|f_n\|_\infty = n$ and so each function is bounded. Also for each x we have that for n large enough, $f_n(x) = 0$. Hence, we also have that for each x there is a constant M_x so that $|f_n(x)| \leq M_x$, $\forall n$. However, clearly $\sup\{\|f_n\|_\infty\} = +\infty$.

Note that in this case the domain is a complete metric space.

Theorem 2.25 (Principle of Uniform Boundedness)**.** *Let X be a Banach space, Y a normed space, and let $\mathcal{F}$ be a set of linear mappings from X to Y. If for each $A \in \mathcal{F}$ there is a constant C_A such that $\|Ax\| \leq C_A\|x\|$(i.e., each A is bounded) and for each $x \in X$ there is a constant M_x so that $\|Ax\| \leq M_x$ (i.e., the collection $\mathcal{F}$ is bounded at each point), then*

$$\sup\{\|A\| : A \in \mathcal{F}\} < \infty.$$

Proof. Let $E_n = \{x : \|Ax\| \leq n, \forall A \in \mathcal{F}\}$. Since each A is continuous, it is not hard to see that each E_n is closed. By the hypothesis of pointwise boundedness, $X = \cup_n E_n$. Hence, one of these sets is not nowhere dense, say for n_0. Since E_{n_0} is already closed, there must exist x_0 and a constant $r > 0$ so that $\{y : \|y - x_0\| < r\} \subseteq E_{n_0}$.

Given any $0 \neq z \in X$, we have that $\|\frac{rz}{2\|z\|}\| < r$ so that $x_0 + \frac{rz}{2\|z\|} \in E_{n_0}$. Now given any $A \in \mathcal{F}$ we have that

$$\|A(\frac{rz}{2\|z\|})\| = \|A(x_0 + \frac{rz}{2\|z\|}) - A(x_0)\| \leq n_0 + n_0.$$

Hence, for any $0 \neq z$,

$$\|Az\| \leq \frac{4n_0\|z\|}{r} \text{ and } \|A\| \leq \frac{2n_0}{r}, \forall A \in \mathcal{F}.$$

□

The following is just the contrapositive statement of the above theorem, but it is often useful.

Theorem 2.26 (Resonance Principle)**.** *Let X be a Banach space and let Y be a normed space. If $\mathcal{F}$ is a set of bounded linear maps from X to Y with* $\sup\{\|A\|; A \in \mathcal{F}\} = +\infty$, *then there is a vector $x \in X$ such that* $\sup\{\|Ax\| : A \in \mathcal{F}\} = +\infty$.

The vector given by the above result is often called a *resonant vector* for the set $\mathcal{F}$.

Here is an example to show what can go wrong if the domain X is not a Banach space.

Definition 2.27. Let $C_{00} = \text{span}\,\{e_n : n \in \mathbb{N}\} \subseteq \ell^\infty$

Thus, $x = (x_1, x_2, ...) \in C_{00}$ if and only if there is an N_x so that $x_n = 0$ for all $n > N_x$. The number N_x varies from vector to vector.

Example 2.28. Define $A_n : C_{00} \to C_{00}$ by

$$A_n(x) = (x_1, 2x_2, 3x_3, ..., nx_n, 0, 0, ...).$$

Clearly, each A_n is a bounded linear map with $\|A_n\| = n$. Also for each $x \in C_{00}$ we have an N_x as above and so $\sup_n \|A_n x\| \leq N_x < +\infty$.

Problem 2.29. Let X and Y be normed spaces with X Banach, and let $A_n : X \to Y$ be a sequence of bounded linear maps such that for each $x \in X$ the sequence $A_n(x)$ converges in norm to an element of Y. Prove that if we set $A(x) = \lim_n A_n(x)$ then A defines a bounded linear map. (The sequence $\{A_n\}$ is said to converge *strongly* to A and we write $A_n \xrightarrow{S} A$**.**

2.3.2 Open Mapping, Closed Graph, and More!

The following results are also all consequences of Baire's theorem applied to Banach spaces. In fact, given any one as the starting point, it is possible to deduce the others as consequences. These require that the domain and range both be Banach spaces.

Lemma 2.30. *Let X and Y be Banach spaces, and let $A : X \to Y$ be a bounded linear map that is onto. Then there exists $\delta > 0$ such that*

$$\mathcal{B}_Y(0; \delta) := \{y \in Y : \|y\| < \delta\} \subseteq A\big(\mathcal{B}_X(0; 1)\big),$$

where $\mathcal{B}_X(0; 1) = \{x \in X : \|x\| < 1\}$.

Theorem 2.31 (Bounded Solutions). *Let X and Y be Banach spaces and let $A : X \to Y$ be a bounded linear map. If for each $y \in Y$ there is $x \in X$ solving $Ax = y$, then there is a constant $C > 0$ such that for each $y \in Y$ there is a solution to $Ax = y$ with $\|x\| \leq C\|y\|$.*

Theorem 2.32 (Bounded Inverse Theorem). *Let X and Y be Banach spaces and let $A : X \to Y$ be a bounded linear map. If A is one-to-one and onto then the inverse map $A^{-1} : Y \to X$ is bounded.*

Theorem 2.33 (Equivalence of Norms). *Let X be a space with two norms $\|\cdot\|_1$ and $\|\cdot\|_2$, such that X is a Banach space in both norms. If there is a constant such that $\|x||_1 \leq C\|x\|_2$, $\forall x \in X$, then the norms are equivalent.*

As an application of the equivalence of norms theorem and Baire's theorem, one can prove the following.

Theorem 2.34 (Open Mapping Theorem). *Let X and Y be Banach spaces, and let $A : X \to Y$ be a bounded linear map. If A is onto, then for every open set $U \subseteq X$ the set $A(U)$ is open in Y.*

If X and Y are vector spaces, then their Cartesian product is also a vector space with operations, $(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$ and $\lambda(x, y) = (\lambda x, \lambda y)$. This space is sometimes denoted $X \oplus Y$. There are many equivalent norms that can be put on $X \oplus Y$, we will use $\|(x, y)\|_1 := \|x\| + \|y\|$. It is fairly easy to see that a subset $C \subseteq X \oplus Y$ is closed in the $\|\cdot\|_1$ norm if and only if $\|x_n - x\| \to 0$ and $\|y_n - y\| \to 0$ implies that $(x, y) \in C$.

Given a linear map, $A : X \to Y$ its *graph* is the subset $G_A := \{(x, Ax) : x \in X\}$.

Theorem 2.35 (Closed Graph Theorem). *Let X and Y be Banach spaces, and let $A : X \to Y$ be a linear map. If $\|x_n - x\| \to 0$ and $\|Ax_n - y\| \to 0$ implies that $Ax = y$ (this is equivalent to the statement that G_A is a closed subset), then A is bounded.*

In general, a linear map $A : X \to Y$ is called *closed* provided that G_A is closed. When X and Y are Banach spaces, then closed implies bounded by the above. But if X is not Banach then the map need not be bounded. Here is an example.

Example 2.36. Consider the following function space,

$$C^1([0, 1]) = \{f : [0, 1] \to \mathbb{R} : f' \text{ exists and is continuous }\}.$$

Set $\|f\| = \|f\|\_\infty$ and let $Y = C([0,1])$, also with $\|\cdot\|_\infty$. Let $D : C^1([0,1]) \to C([0,1])$ by $D(f) = f'$. By Rudin, Theorem 7.1.7, if $\|f_n - f\|_\infty \to 0$ and $\|f_n' - g\|_\infty \to 0$, then f is differentiable and $g = f'$. This result is equivalent to saying that the graph of D is closed. But D is not bounded since $\|D(t^n)\|_\infty = n\|t^n\|_\infty$. What goes wrong is that $C^1([0,1])$ is not complete in $\|\cdot\|_\infty$. This last fact can be seen directly, or it can be deduced as an application of the closed graph theorem.

Problem 2.37. Let X be a Banach space and let Y and Z be two closed subspaces of X such that $Y \cap Z = (0)$ and $Y + Z := \{y + z : y \in Y, z \in Z\} = X$. Thus, each $x \in X$ has a unique representation as $x = y + z$. Define $\|x\|_1 = \|y\| + \|z\|$. Prove that $(X, \|\cdot\|_1)$ is a Banach space and that $\|\cdot\|$ and $\|\cdot\|_1$ are equivalent norms on X.

2.4 Hahn-Banach Theory

The Hahn-Banach theorem is usually stated for norm extensions, but it is really more general with the proofs being no harder, so we state it in full generality.

Definition 2.38. Let X be a vector space. Then a function $p : X \to \mathbb{R}$ is **sublinear** if:

1. $p(x + y) \leq p(x) + p(y)$, $\forall x, y \in X$,
2. for all $t > 0$, $p(tx) = tp(x)$.

Note that, unlike a norm, it is not required that $p(-x) = p(x)$.

Lemma 2.39 (One-Step Extension)**.** *Let X be a real vector space, $Y \subseteq X$ a subspace, $p : X \to \mathbb{R}$ a sublinear functional and $f : Y \to \mathbb{R}$ a linear functional satisfying $f(y) \leq p(y)$, $\forall y \in Y$. If $x \in X \backslash Y$ and $Z = span\{Y, x\}$, then there exists a linear functional $g : Z \to \mathbb{R}$ such that $g(y) = f(y), \forall y \in Y$ (i.e., g is an "extension" of f) and $g(z) \leq p(z)$, $\forall z \in Z$.*

Proof. Since every vector in Z is of the form $z = rx + y$ for some $r \in \mathbb{R}$ every possible extension of f is of the form $g_\alpha(rx + y) = r\alpha + f(y)$. So we need to show that for some choice of α we will have that $r\alpha + f(y) \leq p(rx + y)$, $\forall r, y$.

First note that if $y_1, y_2 \in Y$, then

$$f(y_1 + y_2) \leq p(y_1 + y_2) \leq p(y_1 + x) + p(y_2 - x).$$

Hence,

$$f(y_2) - p(y_2 - x) \leq p(y_1 + x) - f(y_1), \forall y_1, y_2 \in Y.$$

Thus,

$$\alpha_0 := \sup\{f(y) - p(y - x) : y \in Y\} \leq \inf\{p(y + x) - f(y) : y \in Y\} =: \alpha_1.$$

The proof is completed by showing that for any α with $\alpha_0 \leq \alpha \leq \alpha_1$ we will have that g_α is an extension with $g_\alpha(rx + y) \leq p(rx + y)$. □

Theorem 2.40 (Hahn-Banach Extension Theorem for Sublinear Functions)**.** *Let X be a real vector space, $p : X \to \mathbb{R}$ a sublinear functional, $Y \subseteq X$ a subspace and $f : Y \to \mathbb{R}$ a real linear functional satisfying $f(y) \leq p(y)$, $\forall y \in Y$. Then there exist a linear functional $g : X \to \mathbb{R}$ that extends f and satisfies, $g(x) \leq p(x) < \forall x \in X$.*

The idea of the proof is to use the one-step extension lemma and then apply Zorn's lemma to show that an extension exists whose domain is maximal among all extensions and argue that necessarily this maximal extension has domain equal to all of X. Zorn's lemma is equivalent to the *Axiom of Choice*, so this is one theorem that requires this additional axiom, which almost every functional analyst is comfortable with assuming.

Next we look at complex versions. The following result is often useful for passing between real linear maps and complex linear maps.

Proposition 2.41. *Let X be a complex vector space.*

1. *If $f : X \to \mathbb{R}$ is a real linear functional, then setting*

$$g(x) = f(x) - if(ix)$$

   defines a complex linear functional.

2. *If $f : X \to \mathbb{C}$ is a complex linear functional, then $Ref(x) = Re(f(x))$, is a real linear functional and*

$$f(x) = Ref(x) - iRef(ix).$$

Theorem 2.42. *Let X be a complex vector space, $p : X \to \mathbb{R}$ be a function that satisfies, $p(x+y) \leq p(x)+p(y)$ and $p(\lambda x) = |\lambda|p(x)$, $\forall \lambda \in \mathbb{C}$. If $Y \subseteq X$ is a complex subspace and $f : Y \to \mathbb{C}$ is a complex liner functional satisfying $|f(y)| \leq p(y)$, $\forall y \in Y$, then there exists a complex linear functional $g : X \to \mathbb{C}$ that extends f and satisfies $|g(x)| \leq p(x)$, $\forall x \in X$.*

Proof. Consider the real linear functional Ref, observe that it satisfies $Ref(y) \leq p(y)$. Apply the previous theorem to extend it to a real linear functional $h : X \to \mathbb{R}$ satisfying $h(x) \leq p(x)$.

Now set $g(x) = h(x) - ih(ix)$ and check that g works. □

Given a bounded linear functional $f : Z \to \mathbb{F}$ we define $\|f\| = \sup\{|f(z)| : \|z\| \leq 1\}$.

Theorem 2.43 (Hahn-Banach Extension Theorem)**.** *Let X be a normed linear space over $\mathbb{F}$, let $Y \subseteq X$ be a subspace, and let $f : Y \to \mathbb{F}$ be a bounded linear functional. Then there exists an extension $g : X \to \mathbb{F}$ that is also a bounded linear functional and satisfies, $\|g\| = \|f\|$.*

Proof. Apply the above result with $p(x) = \|f\| \cdot \|x\|$. □

Corollary 2.44. *Let X be a normed space and let $x_0 \in X$. Then there exists a bounded linear functional $g : X \to \mathbb{F}$ with $\|g\| = 1$ such that $g(x_0) = \|x_0\|$.*

Proof. Let Y be the one dimensional space spanned by x_0 and let $f(\lambda x_0) = \lambda\|x_0\|$. Check that $\|f\| = 1$ and apply the last result. □

2.4.1 Convex Sets

Convex sets play an important role in many settings. We gather a few facts about these sets here. Many of these facts are ultimately consequences of the Hahn-Banach theorem, especially the version for sublinear functions.

Recall that a subset K of a real vector space X is called **convex** if whenever $x, y \in K$, then $tx+(1-t)y \in K$, $\forall 0 \leq t \leq 1$. Writing $tx+(1-t)y = y + t(x - y)$, we see that K is convex if and only if whenever $x, y \in K$, then the line segment from y to x is in K.

Note that if K is convex and $x_1, x_2, x_3 \in K$ then

$$\{s\big(tx_1 + (1-t)x_2\big) + (1-s)x_3 |\, 0 \leq t \leq 1, 0 \leq s \leq 1\} = \{t_1x_1 + t_2x_2 + t_3x_3 |\, t_i \geq 0, \sum_{i=1}^{3} t_i = 1\},$$

is a subset of K. Inductively, one can show that if K is convex, $x_1, ..., x_n \in K$ and $t_i \geq 0, 1 \leq i \leq n$ with $\sum_{i=1}^n t_i = 1$, then $\sum_{i=1}^n t_ix_i \in K$. Such a sum is called a **convex combination** of $x_1, ..., x_n$. It readily follows that K is convex if and only if every convex combination of elements of K is again in K.

Given a non-empty subset S of a real vector space X, the smallest convex set that contains S is called the **convex hull** of S and is denoted by $convh(S)$. Using this last fact it is not hard to see that

$$convh(S) = \{\sum_{i=1}^{n} t_i x_i | t_i \geq 0, x_i \in S,\ 1 \leq i \leq n, \sum_{i=1}^{n} t_i = 1,\ \text{with n arbitrary }\}.$$

For an infinite dimensional vector space the number of terms needed in these sums is, generally, not bounded, but in finite dimensions there is a bound.

Theorem 2.45 (Caratheodory). *Let S be a non-empty subset of $\mathbb{R}^n$, then*

$$convh(S) = \{\sum_{i=1}^{n+1} t_i x_i |\, t_i \geq 0, x_i \in S, 1 \leq i \leq n+1, \sum_{i=1}^{n+1} t_i = 1\}.$$

This has a nice geometric interpretation. Given a non-empty set S in a vector space, we let $\ell^{(1)}(S) = \{tx + (1-t)y |\, x, y \in S,\ 0 \leq t \leq 1\}$, which is the union of all line segments joining points is S, and let $\ell^{(k+1)}(S) = \ell^{(1)}(\ell^{(k)}(S))$. Caratheodory's result says that for $S \subseteq \mathbb{R}^n$, this process terminates with $convh(S) = \ell^{(n)}(S) = \ell^{(m)}(S)$, $\forall m \geq n$.

One corollary of Caratheodory's result which is very useful is the following.

Corollary 2.46. *Let S be a non-empty compact subset of $\mathbb{R}^n$, then $convh(S)$ is also compact.*

Proof. Set

$$K = \{(t_1, ..., t_{n+1}, x_1, ..., x_{n+1}) |\, t_i \in\ bbR, t_i \geq 0, \sum_{i=1}^{n+1} t_i = 1,\ x_i \in S\}.$$

Since $(n+1) + (n+1)n = (n+1)^2$, K is a subset of $\mathbb{R}^{(n+1)^2}$, and is closed and bounded. Define $L : K \to \mathbb{R}^n$ by

$$L((t_1, ..., t_{n+1}, x_1, ..., x_{n+1})) = \sum_{i=1}^{n+1} t_i x_i,$$

which is continuous. By Caratheodory, $conv(S) = L(K)$ and since continuous images of compact sets are compact, the result follows. $\square$

In particular, in finite dimensions, the convex hull of a compact set is automatically closed. This is false in infinite dimensions, the convex hull of the compact set $S = \{0\} \cup \{e_n/n : n \in \mathbb{N}\} \subseteq \ell^1$ is not closed.

Definition 2.47. Given a real normed space X, a bounded linear functional $f : X \to \mathbb{R}$ and $\alpha \in \mathbb{R}$ we set

$$H_{f,\alpha} := \{y \in X : f(y) \leq \alpha\},$$

and we call $H_{f,\alpha}$ a **closed 1/2-space**.

The following result is a consequence of the more general sublinear functionals version of the Hahn-Banach theorem.

Theorem 2.48 (Hahn-Banach Separation Theorem)**.** *Let X be a real normed space and let $K \subseteq X$ be a convex subset.*

1. *If* $\inf\{\|x - y\| : y \in K\} \neq 0$*, then there exists a bounded linear functional $f : X \to \mathbb{R}$ and $\alpha \in \mathbb{R}$ such that $f(y) \leq \alpha, \forall y \in K$, while $\alpha < f(x_0)$.*

2. *Every closed convex set is an intersection of 1/2-spaces.*

Proof. We only sketch the key ideas of the proof of the first statement, since the proof introduces some important concepts. First, by applying a translation one can assume that $0 \in K$. Let us set $d(z, K) = \inf\{\|z - y\| : y \in K\}$. Next one fixes $0 < \epsilon < 1$ with $\epsilon < d(x, K)$ and sets $K_\epsilon = \{z : d(z, K) \leq \epsilon\}$ and shows that this is a closed convex set.

The **Minkowski functional** of the set K_ϵ is defined by

$$p(z) = \inf\{t > 0 : t^{-1}z \in K_\epsilon\}.$$

One proves that this is a sublinear functional with $p(K_\epsilon) \leq 1 < p(x)$. Now define a linear functional on the one-dimensional space spanned by x by $g(rx) = rp(x)$, check that $g(rx) \leq p(rx)$ and extend to obtain a linear functional $f : X \to \mathbb{R}$ with $f(z) \leq p(z)$. One now needs to prove that f is a bounded linear functional and that $f(K) \leq 1 < f(x)$. $\square$

Definition 2.49. Given a convex set K, a point $x \in K$ is called an **extreme point** of K, if whenever there exist $y, z \in K$ and $t, 0 \leq t \leq 1$ such that $x = ty + (1 - t)z$, then necessarily $y = z = x$. We let $Ext(K)$ denote the set of extreme points of K.

For example in $\mathbb{R}^2$, the extreme points of the unit square are its four corners, while the extreme points of the closed unit disk is the unit circle. Note that the open unit disk is convex but has no extreme points. The following theorem tells us that for compact convex sets, there are always plenty of extreme points.

Theorem 2.50 (Krein-Milman, normed version). *Let K be a compact, convex subset of a normed linear space, then K is the closure of the convex hull of its extreme points.*

Even when K is a convex and compact subset of $\mathbb{R}^3$, the set $Ext(K)$ need not be a closed set, but it is always bounded. However, there is a better theorem in fnite dimensions.

Theorem 2.51 (Caratheodory-Minkowski). *Let $K \subseteq \mathbb{R}^n$ be compact and convex. Then every point in K is a convex combination of at most $n+1$ extreme points.*

2.4.2 Separable and Entangled States and Entanglement Witnesses

In this section we give some applications of the above ideas to quantum information theory. For these applications we assume that the reader has some familiarity with matrix theory.

Given vectors $x = (a_1, ..., a_n)$ and $y = (b_1, ..., b_n)$ in $\mathbb{C}^n$, we set their **inner product** equal to

$$\langle x|y\rangle = \sum_{i=1}^{n} a_i \overline{b_i}.$$

Note that this is linear in the first variable and conjugate in the second variable. The (Euclidean) length of a vector is

$$\|x\| = \sqrt{\langle x|x\rangle}.$$

An $n \times n$ matrix $P = (p_{i,j})$ is called **positive semidefinite**(denoted $P \geq 0$), provided

$$\langle x|Px\rangle = \sum_{i,j=1}^{n} p_{i,j} a_j \overline{a_i} \geq 0, \, \forall x \in \mathbb{C}^n.$$

We shall write M_n to denote the set of $n \times n$ matrices and M_n^+ to denote the set of positive semidefinite matrices. Note that if $P, Q \in M_n^+$, then $P + Q \in M_n^+$. Also if $P \in M_n^+$, then P is self-adjoint, i.e., is equal to its own conjugate transpose, and every eigenvalue of P must be non-negative.

For each unit vector $x = (a_1, ..., a_n)$ the matrix $P_x = (a_i \overline{a_j})$ satisfies

$$P_x(y) = \langle y|x\rangle x,$$

and thus is the **orthogonal projection** onto the 1-dimensional space spanned by x.

Finally, the **trace** of a matrix is

$$Tr(P) = \sum_{i=1}^{n} p_{i,i}.$$

Recall that $Tr(P)$ is equal to the sum of the eigenvalues of P.

A matrix P is called a **density matrix** if $P \in M_n^+$ and $Tr(P) = 1$. The states of a quantum system are usually identified with density matrices.

It is readily seen that the set of density matrices is a closed convex set. Hence by Krein-Milman, the set of density matrices will be the convex hull of its extreme points.

Problem 2.52. Use the fact that self-adjoint matrices are diagonalizable to show that the extreme points of the density matrices in M_n are the projections onto 1-dimensional subspaces and that every density matrix is the convex combination of n extreme points.

Since M_n is a $2n^2$-dimensional real vector space, by Caratheodory's theorem we know that we should need at most $2n^2 + 1$ extreme points, but the problem shows that for density matrices we can use considerably fewer.

We now look at the case of *bipartite* quantum systems. States of such systems are typically represented by density matrices in a tensor product. So we need to discuss the tensor product of the space of matrices. One of the simplest ways to view tensor products of matrices is through the concept of **block matrices**. Given n^2 matrices, $A_{i,j}$, $1 \le i, j \le n$ with each $A_{i,j}$ a $k \times k$ matrix, we form an $nk \times nk$ matrix A as follows:

$$A := (A_{i,j}) = \begin{bmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & \ddots & \vdots \\ A_{n,1} & \cdots & A_{n,n} \end{bmatrix}.$$

We write $M_n(M_k)$ to indicate that we are considering $nk \times nk$ matrices as written in such blocks. Clearly every element of M_{nk} can be considered as an element of $M_n(M_k)$ by just introducing parentheses around every $k \times k$ block. Block matrices are also another way to think about the tensor product $M_n \otimes M_k$. If we let $E_{i,j}$, $1 \le i, j \le n$ denote the $n \times n$ matrix that is 1 in the (i, j)-entry and 0 elsewhere, then these form a basis for the vector space M_n. Consequently, every element of the tensor product $M_n \otimes M_k$ has a unique representation as $\sum_{i,j=1}^{n} E_{i,j} \otimes A_{i,j}$ for some set of matrices $A_{i,j} \in M_k$.

In this manner we have identified, $M_{nk} = M_n(M_k) = M_n \otimes M_k$. Note that with this identification, if we start with $B = (b_{i,j}) \in M_n$ and $C \in M_k$

then we have that

$$B \otimes C = (\sum_{i,j=1}^{n} b_{i,j} E_{i,j}) \otimes C = \sum_{i,j=1}^{n} E_{i,j} \otimes (b_{i,j} C) = (b_{i,j} C),$$

where the latter matrix is the block matrix with (i, j)-th block equal to $b_{i,j}C$. This block matrix representation of $B \otimes C$ is often called the **Kronecker tensor**.

Given two densities matrices $P = (p_{i,j}) \in M_n$ and $Q \in M_k$ it is not hard to see that $P \otimes Q = (p_{i,j}Q)$ is a density matrix in $M_n \otimes M_k = M_n(M_k) = M_{nk}$, and hence, the convex hull of all such density matrices is again a set of density matrices.

Definition 2.53. The set of matrices in $M_n \otimes M_k = M_n(M_k) = M_{nk}$ that are in the convex hull of the matrices of the form $P \otimes Q$, with $P \in M_n$, $Q \in M_k$ density matrices, is called the set of **separable density matrices**. Any density matrix in $M_n(M_k)$ that is not separable is called an **entangled density matrix**.

Since the set of density matrices is a compact set, it follows that the set of all matrices of the form $P \otimes Q$ with P and Q density matrices is also a compact set. Hence, by the corollary to Caratheodory's theorem, the set of separable density matrices is a compact subset of the $2n^2k^2$ real dimensional vector space, $M_n(M_k)$.

Problem 2.54. Prove that the extreme points of the separable density matrices are the matrices of the form $P_x \otimes P_y$ where $x \in \mathbb{C}^n$ and $y \in \mathbb{C}^k$ where x and y are unit vectors. Conclude that every separable density matrix is a convex combination of at most $2n^2k^2 + 1$ such matrices.

By the Hahn-Banach separation theorem, the set of separable density matrices is an intersection of closed 1/2-spaces. In particular, if R is an entangled density matrix, then there must be a real linear functional f and $\alpha \in \mathbb{R}$, such that $f(S) \leq \alpha$ for every separable density matrix and $\alpha < f(R)$. In this sense, the pair (f, α) provides the evidence that R is entangled.

Finding such pairs (f, α) and protocols for how to implement them physically are at the heart of the theory of **entanglement witnesses**.

2.5 Dual Spaces

Let X be a normed linear space. We let X^* denote the set of bounded linear functionals on X. It is easy to see that this is a vector space and that

$\|f\| = \sup\{|f(x)| : \|x\| \leq 1\}$ is a norm on this space. The normed space $(X^*, \|\cdot\|)$ is called the *dual space of* X.

Other notations used for the dual space are X', X^d and sometimes $X^\dagger$.

Proposition 2.55. *Let $(X, \|\cdot\|)$ be a normed space, then $(X^*, \|\cdot\|)$ is a Banach space.*

Here are some examples of dual spaces.

Example 2.56 (The dual of ℓ^1)**.** Given $y \in \ell^\infty$ define $f_Y : \ell^1 \to \mathbb{R}$ by $f_y(x) = \sum_n x_n y_n$, i.e., the dot product! It is not hard to see that f_y is bounded and that $\|f_y\| = \|y\|_\infty$. This gives a map from ℓ^∞ into $(\ell^1)^*$.

It is easy to see that this is a linear map, i.e., that $f_{y_1+y_2} = f_{y_1} + f_{y_2}$ and $f_{\lambda y} = \lambda f_y$. Also we have seen that it satisfies $\|f_y\| = \|y\|_\infty$, i.e., that it is isometric and hence must be one-to-one. Finally, we show that it is onto. To see this start with $f : \ell^1 \to \mathbb{R}$ and let $y_n = f(e_n)$ where e_n is the vector that is 1 in the n-th entry and 0 elsewhere. We have that $|y_n| \leq \|f\|\|e_n\|_1 = \|f\|$, so that $y = (y_1, ...) \in \ell^\infty$. Now check that $f_y(x) = f(x)$ for every x so that $f_y = f$ and we have onto.

The above results are often summarized as saying that $(\ell^1)^* = \ell^\infty$, when what they really mean is that there is the above linear, onto isometry between these two spaces.

Example 2.57 (The dual of ℓ^p, $1 < p < +\infty$)**.** Let $\frac{1}{p} + \frac{1}{q} = 1$ and for each $y \in \ell^q$ define

$$f_y : \ell^p \to \mathbb{R} \text{ by setting } f_y(x) = \sum_{n=1}^{+\infty} y_n x_n.$$

By Minkowski's inequality f_y is a bounded linear functional and $\|f_y\| = \|y\|_q$. Moreover, given any $f : \ell^p \to \mathbb{R}$ if we set $y_n = f(e_n)$ where e_n is the vector with 1 in the n-th coordinate and 0's elsewhere, then it follows that $y = (y_n) \in \ell^q$. This proves that the map $L : \ell^q \to (\ell_p)^*$ given by $L(y) = f_y$ is onto. It is clear that it is linear and by the above it is isometric. Thus, briefly $(\ell^p)^* = \ell^q$.

However, this fails when $p = +\infty$, as we shall show.

Example 2.58 (The space C_0)**.** We set $C_0 = \{x = (x_1, x_2, ...) \in \ell^\infty : \lim_n x_n = 0\}$. It is easy to show that this is a closed subspace. In fact C_0 is the closed linear span of $e_n : n \in \mathbb{N}$. Given a bounded linear functional $f : C_0 \to \mathbb{R}$ if we set $y_n = f(e_n)$, then the same calculations as above show that $y \in \ell^1$ and that $f = f_y$. This leads to $C_0^* = \ell^1$.

Example 2.59 ($(\ell^\infty)^* \neq \ell^1$)**.** Now let $f : \ell^\infty \to \mathbb{R}$ be a bounded linear functional. As before let $y_n = f(e_n)$ and show that $y \in \ell^1$. The functional $f - f_y$ has the property that $(f - f_y)(e_n) = 0$ and hence, $(f - f_y)(C_0) = 0$. But this does not guarantee that $f = f_y$.

We now construct an element of ℓ^∞ that is not of the form f_y for any $y \in \ell^1$. Let $e \in \ell^\infty$ be the vector of all 1's. It is easy to check that

$$\inf\{\|e - x\|_\infty : x \in C_0\} = 1.$$

Now let $W = \{re + x : r \in \mathbb{R}, x \in C_0\}$, which is easily seen to be a vector subspace of ℓ^∞. We define a linear functional $g : W \to \mathbb{R}$ by setting $g(re+x) = r$. If $r \neq 0$, then for any $x \in C_0$, $\|re+x\| = |r|\|e+r{-}1x\| \geq |r|$. Hence, $|g(re+x)| = |r| \leq \|re+x\|$, so that g is a bounded linear functional and $\|g\| \leq 1$. Since $1 = g(e) = \|e\|$, we have that $\|g\| = 1$. Now by the Hahn-Banach theorem, we can extend g to a linear functional $f : \ell^\infty \to \mathbb{R}$ with $\|f\| = 1$. However, $f(e_n) = g(e_n) = 0$, so if $f = f_y$ then y would need to be the 0 vector. Since $\|f\| = 1$ we see that $f \neq f_y$ for any $y \in \ell^1$.

Thus, not every bounded linear functional on ℓ^∞ is of the form $f_y, y \in \ell^1$. What is true is that every $f \in \ell^{\infty^*}$ decomposes uniquely as $f = f_y + h$ where $y \in \ell^1$ and $h \in \ell^{\infty^*}$ with $h(x) = 0$ for every $x \in C_0$.

2.5.1 Banach Generalized Limits

It is often useful to be able to take limits even when one doesn't exist! This is one reason that we use $\liminf$ and $\limsup$ but these two operations are not linear. For example if we let x_n be 1 for odd n and 0 for even n and let y_n be one for even n and 0 for odd n, then $\limsup_n x_n = \limsup_n y_n = 1 = \limsup(x_n + y_n)$. Banach generalized limits are one way that we can assign "limits" in a linear fashion.

Proposition 2.60 (Banach Generalized Limits)**.** *There exists a bounded linear functional, $G : \ell^\infty \to \mathbb{R}$ such that:*

- $\|G\| = 1$,
- *when $x = (x_1, x_2, ...)$ and $lim_n x_n$ exists, then $G(x) = \lim_n x_n$,*
- $\forall x \in \ell^\infty$, $\liminf_n x_n \leq G(X) \leq \limsup_n x_n$,
- *given K if we let $y_K = (x_{K+1}, x_{K+2}, ...)$, then $G(y_K) = G(x)$, i.e., the generalized limit doesn't care where we start the sequence.*

The existence of such functional follows from the Hahn-Banach theorem and a Banach generalized limit has the property that $G(C_0) = 0$ so these are examples of functionals that do not come from ℓ^1 sequences.

Example 2.61. For those of you familiar with measure theory. Given any $(X, \mathcal{B}, \mu)$ measure space and $1 \leq p < +\infty$, the set $L^p(X, \mathcal{B}, \mu)$ of equivalence classes of p-integrable functions (*essentially bounded functions* in the case $p = +\infty$) is a Banach space in the p-norm.

The L^p-spaces and their duals behave like the ℓ^p spaces and their duals. In fact, $\ell^p = L^p(\mathbb{N}, 2^{\mathbb{N}}, \mu)$ where μ is counting measure, so some authors treat the ℓ^p spaces as just a special case of the L^p-theory.

A measure space $(X, \mathcal{B}, \mu)$ is called σ-finite, provided that there is a countable collection of sets $B_n \in \mathcal{B}$ with $\mu(B_n) < +\infty$ and $X = \cup_n B_n$.

Theorem 2.62 (Riesz Representation Theorem)**.** *Let $1 < p \leq +\infty$ and $\frac{1}{p} + \frac{1}{q} = 1$ and let $(X, \mathcal{B}, \mu)$ be a measure space. Then every bounded linear functional $L : L^p(X, \mathcal{B}, \mu) \to \mathbb{R}$ is of the form*

$$L(f) = \int_X f\, g \, d\mu,$$

for some $g \in L^q(X, \mathcal{B}, \mu)$. Conversely, every g defines a bounded linear functional via integration. Moreover, $\|L\| = \|g\|_q$, briefly, $L^p(X, \mathcal{B}, \mu)^* = L^q(X, \mathcal{B}, \mu)$. If, in addition, μ is σ-finite, then $L^1(X, \mathcal{B}, \mu)^* = L^\infty(X, \mathcal{B}, \mu)$.

There is one more theorem that has the same name.

Theorem 2.63 (Riesz Representation Theorem)**.** *Let $(K, \mathcal{T})$ be a compact, Hausdorff space, let $C(K)$ denote the space of continuous $\mathbb{R}$-valued functions on K. If $L : C(K) \to \mathbb{R}$ is a bounded linear functional, then there exists a bounded, Borel measure μ such that $L(f) = \int_K f d\mu$ and conversely, every bounded Borel measure defines a bounded linear functional. Moreover, $\|L\| = |\mu|(X)$.*

Example 2.64 (The non-commutative ℓ^p-spaces)**.** These norms on spaces of matrices play an important role in the theory of quantum capacities. We assume that the reader is familiar with the theory of positive semidefinite matrices. We let $M_{m,n}$ denote the vector space of $m \times n$ complex matrices, and for convenience set $M_n = M_{n,n}$. Given $A \in M_{m,n}$ we let $A^* \in M_{n,m}$ denote its conjugate transpose. The matrix $A^*A \in M_n$ is positive semidefinite and so has a complete set of non-negative eigenvalues, $\lambda_1 \geq ... \geq \lambda_n \geq 0$. The numbers $s_1 = \lambda_1^{1/2}, ..., s_n = \lambda_n^{1/2}$ are called the *singular values of A*. An alternative way to get these numbers, in the case that $A \in M_n$ is a square matrix, is to use the *singular valued decompostion of A(SVD)*. That is A can be written as $A = UDV$ where $U^*U = V^*V = I_n$ and D is a diagonal matrix with non-negative entries. In this case, the entries of D arranged in decreasing order are exactly the numbers $s_1 \geq ... \geq s_n$.

Given $A \in M_{m,n}$ and $1 \leq p < +\infty$ we set

$$\|A\|_p = (s_1^p + \cdots + s_n^p)^{1/p},$$

and set $\|A\|_\infty = \max\{s_1, ..., s_n\}$.

Remarkably, these define norms and with these norms $(M_{m,n}, \|\cdot\|_p)$ is called a *non-commutative ℓ^p-space.*

The duality theory for these spaces is similar to the duality theory outlined above. Recall that the *trace, Tr* of a square matrix is the sum of its diagonal entries. Given $B \in M_{n,m}$ we can define a linear map $f_B : M_{m,n} \to \mathbb{C}$ by setting

$$f_B(A) = Tr(BA).$$

Moreover, it is not hard to see that every linear functional on $M_{m,n}$ is of this form for some B. The *non-commutative Holder inequality* says that

$$\sup\{|Tr(BA)| : \|A\|_p \leq 1\} = \|B\|_q,$$

where p and q are any Holder conjugates. Thus, the map from $(M_{n,m}, \|\cdot\|_q) \to (M_{m,n}, \|\cdot\|_p)^*$ given by $B \to f_B$ defines an onto linear isometry and so as in the case of the ℓ^p-spaces we can identify,

$$(M_{m,n}, \|\cdot\|_p)^* = (M_{n,m}, \|\cdot\|_q).$$

We should note that some authors prefer to define $f_B : M_{m,n} \to \mathbb{C}$ by $f_B(A) = Tr(B^tA)$, where B^t is the transpose of B, so that $B \in M_{m,n}$ instead of in $M_{n,m}$. Other authors set $f_B(A) = Tr(B^*A)$, imitating the inner product of vectors. But in this case the map $B \to f_B$ is conjugate linear.

2.5.2 The Double Dual and the Canonical Embedding

Given a normed linear space, X, we let $X^{**} = (X^*)^*$, i.e., the bounded linear functionals on the space of bounded linear functionals. Given $x \in X$ we define a bounded linear functional $\hat{x} : X^* \to \mathbb{R}$ by setting $\hat{x}(f) = f(x)$, i.e., reversing the roles of dependent and independent variables. Thus, $\hat{x} \in X^{**}$.

The map $J : X \to X^{**}$ defined by $J(x) = \hat{x}$ is easily seen to be linear, one-to-one and an isometry. This map is called the *canonical injection* or the *canonical embedding* of X into X^{**}.

Definition 2.65. A normed space X is called **reflexive** provided that $J(X) = X^{**}$, which is often written simply as $X = X^{**}$.

Note that if X is reflexive, then so is X^*. The spaces ℓ^p and $L^p(X, \mathcal{B}, \mu)$ are reflexive for $1 < p < +\infty$. The spaces C_0, ℓ^1, ℓ^∞ are not reflexive. This concept is very important when one discusses the *weak topologies*.

2.5.3 The Weak Topology

Definition 2.66. Give a normed space X, we say that a subset $U \subseteq X$ is **weakly open** provided that whenever $x_0 \in U$ then there exists a finite set $f_1, ..., f_n \in X^*$ and $\epsilon_1, ..., \epsilon_n > 0$ such that

$$\mathcal{B}_{f,\epsilon_i} := \{x \in X : |f_i(x - x_0)| < \epsilon_i, \forall 1 \leq i \leq n\} \subseteq U.$$

The set $\mathcal{T}_w$ of all weakly open subsets of X is a topology on X called the **weak topology.**

We say that a net converges *weakly to* x_0 if it converges to x_0 in the weak topology.

Theorem 2.67. *A net* $\{x_\lambda\}_{\lambda \in D} \subseteq X$ *converges in the weak topology to* $x_0 \in X$ *if and only if*

$$\lim_\lambda f(x_\lambda) = f(x_0), \forall f \in X^*.$$

We write $x_\lambda \xrightarrow{w} x_0$ to indicate that a net converges weakly to x_0.

Example 2.68. Let $1 < p < +\infty$ and let $e_n \in \ell^p$ be the vectors that are 1 in the n-th entry and 0 elsewhere. Then these are all unit vectors, but we claim that $e_n \stackrel{w}{\to} 0$. To see this let $y = (y_1, ...) \in \ell^q$ where q is the Holder conjugate of p. Then $f_y(e_n) = y_n \to 0 = f_y(0)$ since y is q-summable. Since every element of ℓ^{p*} is of the form f_y for some $y \in \ell^q$, $e_n \stackrel{w}{\to} 0$.

In contrast, if we let $p = 1$ so that $\ell^{1*} = \ell^\infty$ and consider $e_n \in \ell^1$, then $y = (1, 1, 1, ...) \in \ell^\infty$ and $f_y(e_n) = 1, \forall n$. Hence, $\{e_n\}$ does not tend to 0 weakly in ℓ^1.

Now consider $e_n \in c_0$ so that $c_0^* = \ell^1$, then again we see that in this space $e_n \stackrel{w}{rightarrow} 0$, since the entries of a vector $y \in \ell^1$ tend to 0.

Finally, it turns out that when we regard $e_n \in \ell^\infty$ then $e_n \stackrel{w}{\to} 0$, but to prove this fact we would first need a description of the dual of ℓ^∞ which we have not given.

2.5.4 The Weak* Topology

Definition 2.69. Given a normed space X, we define a subset $U \subseteq X^*$ to be **weak* open** provided that for every $f_0 \in U$ there exists a finite set $x_1, ..., x_n \in X$ and $\epsilon_1, ..., \epsilon_n > 0$ such that

$$\{f \in X^* : |f(x_i) - f_0(x_i)| < \epsilon_i, \forall 1 \leq i \leq n\} \subseteq U.$$

The set $\mathcal{T}_{w*}$ of all weak* open subsets of X^* is a topology on X^* called the **weak* topology**.

A net of functionals converges *weak* to* f_0 provided that it converges in the weak* topology to f_0.

Theorem 2.70. *A net* $\{f_\lambda\}_{\lambda \in D}$ *converges in the weak* topology to* f_0 *if and only if*

$$\lim_\lambda f_\lambda(x) = f_0(x), \forall x \in X.$$

We write $f_\lambda \xrightarrow{w*} f_0$ to indicate that a net converges in the weak* topology to f_0.

Example 2.71. In parallel with the last example, consider $e_n \in \ell^p$ for $1 < p < +\infty$. Since $\ell^p = (\ell^q)^*$, we see that $e_n \overset{w^*}{\to} 0$.

In fact, we can say a bit more. Since for both the weak topology and the weak* topology on ℓ^p we are pairing with vectors in ℓ^q, then actually the weak and weak*topologies are the same topologies, for $1 < p < +\infty$, so of course the two convergences are the same.

Now consider $e_n \in \ell^1$ and recall that these vectors did not tend to 0 in the weak topology. However, for the weak* topology we use the fact that $\ell^1 = c_0^*$ and so we pair with vectors in c_0. Since the entries of any vector in c_0 tend to 0, we have that $e_n \overset{w*}{\to} 0$.

We cannot discuss the weak* topology on c_0 since we do not have an example of a normed space X for which $c_0 = X^*$. In fact, it is known that it is impossible for c_0 to be the dual of any normed space! So there is in fact no way to define a weak* topology in this case.

Finally, if we consider $e_n \in \ell^\infty = (\ell^1)^*$, then since the entries of every vector in ℓ^1 tend to 0, we have that $e_n \overset{w*}{\to} 0$.

The following is the most important property of the weak* topology.

Theorem 2.72 (Banach-Alaoglu)**.** *Let* X *be a normed space and let* $K = \{f \in X^* : \|f\| \leq 1\}$. *Then* K *is compact in the weak* topology.*

Consequently, any time that we have a bounded net in X^*, we may select a weak* convergent subnet.

Note that when X is reflexive, then regarding $X = (X^*)^*$ we have a weak* topology on X, but this is also the same as the weak topology on X. So weak convergence and weak* convergence mean the same thing.

In particular, in $\ell^p, 1 < p < +\infty$, we see that every bounded net will have a weakly convergent subnet.

2.5.5 Completion of Normed Linear Spaces

We discussed earlier that every metric space (X, ρ) has a canonical complete metric space $(\hat{X}, \hat{\rho})$ that it is isometrically contained in as a dense subset. So if $(X, \|\cdot\|)$ is a normed space then it is contained in a complete metric space, $\hat{X}$. But there is nothing in the earlier theorem that says that $\hat{X}$ can be taken to be a vector space and that the metric $\hat{\rho}$ can be taken to come from a norm. Both of these things are true and can be shown with some care.

However, the following gives us an easier way to see all of this. Given a normed space $(X, \|\cdot\|)$, we have that $X^{**} = (X^*)^*$ and so it is a Banach space. The map $J : X \to X^{**}$ is a linear, isometric embedding. So if we just take the closure of $J(X)$ inside the Banach space X^{**} we will get a closed subspace and that subspace will be a Banach space. Because the completion of a metric space is unique, this shows that the completion of X actually has the structure of a vector space.

2.6 Banach space leftovers (Placeholder Section)

These are currently in no particular order.

It is not hard to see that the induced metric on a n.l.s. satsfies: $\lambda_n \to \lambda \implies \lambda_n x \to \lambda x$ and $x_n \to x, y_n \to y \implies x_n + y_n \to x + y$. If ($X$ is a vector space and ρ is a complete metric on X satsfying these two properties, then (X, ρ) is called a **Frechet space**. The vector space $C(\mathbb{R})$ with the metric ρ corresponding to uniform convergence on compact subsets is an example of a Frechet space that is not a Banach space.

2.6.1 Quotient Spaces

A concept that plays a big role in functional analysis are quotient spaces. Recall from algebra, that if we have a vector space X and a subspace Y then there is a **quotinet space**, denoted X/Y consisting of *cosets*,

$$X/Y := \{x + Y : x \in Y\},$$

and defining, $(x_1 + Y) + (x_2 + Y) = (x_1 + x_2) + Y$, $\lambda(x + Y) = (\lambda x) + Y$ gives well-defined operations that makes this into a vector space. One often used fact is that if $T : X \to Z$ is a linear map and $Y = ker(T)$ then there is a well-defined linear map $\hat{T} : X/Y \to Z$ given by $\hat{T}(x + Y) = T(x)$.

To form quotients of normed spaces, we want one more condition. Let $(X, \|\cdot\|)$ be a normed space and let Y be a closed subspace. On the quotient space X/Y, we define

$$\|x + Y\| := \inf\{\|x + y\| : y \in Y\} = \inf\{\|x - y\| : y \in Y\}.$$

The second formula shows that $\|x + Y\|$ is the distance from x to the set Y. This is the reason that we want Y closed, otherwise there would be a point in the closure that is not in Y and any such point would be distance 0 from Y. Thus, such a vector would be one for which $x + Y \neq 0 + Y$ but $\|x + Y\| = 0$. This quantity is called the **quotient norm**.

Proposition 2.73. *Let $(X, \|\cdot\|)$ be a normed space and let Y be a closed subspace, then the quotient norm is a norm on X/Y.*

The vector space together with the quotient norm is called the **quotient space** or if we want to be very clear the **normed quotient space**.

2.6.2 Conditions for Completeness

Next is one of many *2 out of 3* theorems.

Theorem 2.74 (The 2/3's Theorem)**.** *Let $(X, \|\cdot\|)$ be a normed space, let $Y \subseteq X$ be a closed subspace and let X/Y be the normed quotient space. If any 2 of these normed spaces is a Banach space, then the 3rd space is also a Banach space.*

Proof. We only prove the case, that if Y and X/Y are Banach spaces, then so is X. To this end let $\{x_n\}$ be a Cauchy sequence in X and let $[x_n] = x_n + Y$ be the image of this sequence in X/Y. Since $\|[x_n] - [x_m]\| \leq \|x_n - x_m\|$, this is a Cauchy sequence in X/Y, and hence, converges to some $[x] \in X/Y$. Since $\|[x - x_n]\| \to 0$ we can pick $y_n \in Y$ such that $\|x - x_n - y_n\| \to 0$. Note that

$$\begin{aligned}\|y_n - y_m\| &= \|(x_n + y_n - x) - (x_m + y_m - x) + (x_m - x_n)\| \\ &\leq \|x_n + y_n - x\| + \|x_m + y_m - x\| + \|x_n - x_m\|,\end{aligned}$$

since the first two terms tend to 0 and the 3rd is Cauchy, it is easy to see that $\{y_n\}$ is a Cauchy sequence in Y and hence $y_n \to y$.

This implies that

$$x_n = (x_n + y_n - x) + x - y_n \to x - y,$$

and so X is complete. □

To see how this last result can be used we will show that every finite dimensional normed space is automatically complete(although this is not the shortest way to prove this fact).

Theorem 2.75. *Every finite dimensional normed space is complete.*

Proof. We do the case that the field is $\mathbb{R}$, the case for $\mathbb{C}$ is identical.

This is easy to see when the dimension is one. To see pick any $x \in X$ with $\|x\| = 1$ Then every vector is of the form λx and $\|\lambda_1 x - \lambda_2 x\| = |\lambda_1 - \lambda_2|$. Thus, a sequence $\{\lambda_n x\}$ is Cauchy in X iff $\{\lambda_n\}$ is Cauchy in $\mathbb{R}$.

Now, inductively assume that it is true for spaces of dimension n and let X be $(n+1)$-dimensional take a n-dimensional subspace Y. It will be complete by the inductive hypothesis and hence closed. Also the quotient X/Y is one-dimensional so it is complete. Hence by the 2/3 theorem X is complete. □

2.6.3 Norms and Unit Balls

Let $(X, \|\cdot\|)$ be a normed space, we call $\{x : \|x\| < 1\}$ the **open unit ball** and $\{x : \|x\| \leq 1\}$ the **closed unit ball**. It is not hard to see that these sets are in fact open and closed, respectively and that the closed unit ball is the closure of the open unit ball. We wish to characterize norms in terms of properties of these balls.

Definition 2.76. A subset C of a vector space V is called:

- **convex** if whenever $x, y \in C$ and $0 \leq t \leq 1$ then $tx + (1-t)y \in C$, i.e., provided that the line segment from x to y is in C,
- **absolutely convex** if whenever $x_1, ..., x_N \in C$ and $|\lambda_1| + \cdots + |\lambda_n| = 1$, then $\lambda_1 x_1 + \cdots + \lambda_n x_n \in C$,
- **absorbing** if whenever $y \in V$ then there is a $t > 0$ such that $ty \in C$.

A **ray** is any set of the form $\{ty : t > 0\}$ for some $y \in V$.

Proposition 2.77. *Let $(X, \|\cdot\|)$ be a normed space and let $\mathcal{B}$ be the open(or closed) unit ball, then $\mathcal{B}$ is absolutely convex, absorbing and contains no rays. Moreover,*

$$\|x\| = \inf\{t > 0 : t^{-1}x \in \mathcal{B}\} = \big(\sup\{r > 0 : rx \in \mathcal{B}\}\big)^{-1}.$$

Conversely, if $C \subseteq X$ is any absolutely convex, absorbing set that contains no rays and we define

$$\|x\| = \inf\{t > 0 : t^{-1}x \in C\},$$

then $\|\cdot\|$ is a norm on X and

$$\{x : \|x\| < 1\} \subseteq C \subseteq \{x : \|x\| \leq 1\}.$$

This latter fact gives us a way to generate many norms, by just defining sets C that satisfy the above conditions. Also notice that $\|x\|_1 \leq \|x\|_2$ iff and only if the unit balls satisfy $\mathcal{B}_2 \subseteq \mathcal{B}_1$. Since the p-norms on $\mathbb{R}^n$ satisfy $p \leq p' \implies \|x\|_p \geq \|x\|_{p'}$ we see that $\mathcal{B}_p \subseteq \mathcal{B}_{p'}$.

Suppose that $\mathcal{B}_1$ and $\mathcal{B}_2$ are the unit balls of two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on X. It is easy to see that $\mathcal{B}_1 \cap \mathcal{B}_2$ satisfies the conditions needed to be a unit ball. What is the norm?

Consider $K = \{(x, y) : xy \geq 0 \text{ and } x^2 + y^2 \leq 1\}$. Find the smallest set containing K that is absolutely convex, absorbing find the norm that this set defines.

2.6.4 Polars and the Absolutely Convex Hull

Given any set K the intersection of all absolutely convex sets containing K is the smallest absolutely convex set containing K, this set is called the **absolutely convex hull of K**. Its closure is called the **closed absolutely convex hull of K**.

Given any set $K \subseteq \mathbb{R}^n$ its **polar** is the set

$$K^0 := \{y : |x \cdot y| \leq 1\}.$$

The set $K^{00} := (K^0)^0$ is called the **bipolar**

Theorem 2.78 (The Bipolar Theorem)**.** *Let $K \subseteq \mathbb{R}^n$. Then K^0 is a closed, absolutely convex set. The bipolar K^{00} is the closed absolutely convex hull of K.*

If $\mathcal{B}_i \subseteq \mathbb{R}^n$ are closed unit balls for norms $\|\cdot\|_i$, then $\mathcal{B} = \big(\mathcal{B}_1 \cup \mathcal{B}_2\big)^{00}$ is the closed unit ball for a norm called the **decomposition norm** and it satisfies

$$\|x\| = \inf\{\|y\|_1 + \|z\|_2 : x = y + z\}.$$

2.7 Hilbert Spaces

2.7.1 Hilbert Space Completions

Definition 2.79. Given a vector space V over $\mathbb{F}$ a map, $<|>: V \times V \to \mathbb{F}$ is called an **inner product** provided that it for all vectors $v_1, v_2, w_1, w_2 \in V$ and $\lambda \in \mathbb{F}$, it satisfies:

1. $\langle v_1 + v_2, w_1 \rangle = \langle v_1, w_1 \rangle + \langle v_2, w_1 \rangle$,
2. $\langle v_1, w_1 + w_2 \rangle = \langle v_1, w_1 \rangle + \langle v_1, w_2 \rangle$,
3. $\langle \lambda v_1, w_1 \rangle = \overline{\lambda} \langle v_1, w_1 \rangle$,
4. $\langle v_1, \lambda w_1 \rangle = \lambda \langle v_1, w_1 \rangle$,
5. $\langle v_1, w_1 \rangle = \overline{\langle w_1, v_1 \rangle}$,
6. $\langle v_1, v_1 \rangle \geq 0$,
7. $\langle v_1, v_1 \rangle = 0 \iff v_1 = 0$.

The pair $(V, \langle | \rangle)$ is called an inner product space. If $\langle | \rangle$ only satisifes 1–6, then it is called a **semi-inner product**.

Some examples of inner products:

Example 2.80. The space $\mathbb{F}^n$ with $\langle (x_1, ..., x_n), (y_1, ..., y_n) \rangle = \sum_i \overline{x_i} y_i$

Example 2.81. The vector space $C([a, b])$ of continuous $\mathbb{F}$-valued functions on $[a, b]$ with

$$\langle f, g \rangle = \int_a^b \overline{f(t)} g(t) dt.$$

Example 2.82. The vector space ℓ^2 with $\langle x, y \rangle = \sum_{n=1}^{\infty} \overline{x_n} y_n$.

Proposition 2.83 (Cauchy-Schwarz Inequality)**.** *Let V be an inner product space then*

$$|\langle v, w \rangle| \leq \sqrt{\langle v, v \rangle} \sqrt{\langle w, w \rangle}.$$

Corollary 2.84. *Let V be an inner product space, then $\|x\| = \sqrt{\langle x, x \rangle}$ defines a norm on V.*

This is called the **norm induced by the inner product**.

Definition 2.85. An inner product space is called a **Hilbert space** if it is complete in the norm induced by the inner product.

Examples 8.2 and 8.4 are Hilbert spaces. To see that 8.3 is not, consider the case of [0,1]. It is enough to consider the sequence

$$f_n(t) = \begin{cases} 0, & 0 \le t \le 1/2, \\ n(t-1/2), & 1/2 \le t \le 1/2 + 1/n, \\ 1, & 1/2 + 1/n \le t \le 1 \end{cases},$$

which can be shown to be Cauchy but not converge to any continuous function.

Problem 2.86. Let V be a semi-inner product space. Prove that $\mathcal{N} = \{v : \langle v, v\rangle = 0\}$ is a subspace, that setting $\langle v+\mathcal{N}, w+\mathcal{N}\rangle = \langle v, w\rangle$ is well-defined and gives an inner product on the quotient $V/\mathcal{N}$. (Hint: First show that the Cauchy-Schwarz inequality is still true.)

Proposition 2.87. *Let V be an inner product space and let $\|\cdot\|$ be the norm induced by the inner product. Then:*

Polarization Identity $\langle x, y\rangle = 1/4 \sum_{k=0}^{3} i^k \|x + i^k y\|^2$,

Parallelogram Law $\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2$.

Theorem 2.88 (Jordan-von Neumann)**.** *Let $(X, \|\cdot\|)$ be a Banach space whose norm satisfies the parallelogram law. Then defining $\langle | \rangle$ via the polarization identity, defines an inner product on X and the norm on X is the norm induced by the inner product. Thus, any Banach space satisfying the parallelogram law is a Hilbert space.*

Corollary 2.89. *Let $(V, \langle | \rangle)$ be an inner product space, then there exists a Hilbert space $\mathcal{H}$ and a one-to-one inner product preserving map $J : V \to \mathcal{H}$ such that $J(V)$ is a dens subspace of $\mathcal{H}$.*

The space $\mathcal{H}$ is called the **Hilbert space completion** of V. For example, the Hilbert space completion of 8.3 is the space $L^2([a,b], \mathcal{M}, \lambda)$ of equivalence classes of measurable square integrable functions with respect to Lebesgue measure.

2.8 Bases in Hilbert Space

Definition 2.90. Given an inner product space a set of vectors S is called **orthonormal** if $v \in S$ implies $\|v\| = 1$ and whenever $v, w \in S$ and $v \neq w$ then $\langle v, w\rangle = 0$. We will often write $v \perp w \iff \langle v, w\rangle = 0$.

Proposition 2.91 (Pythagoras). *Let V be an inner product space and let $\{x_n : 1 \leq n \leq N\}$ be an otrhonormal set. Then*

$$\|x\|^2 = \sum_{n=1}^{N} |\langle x_n, x\rangle|^2 + \|x - \sum_{n=1}^{N} \langle x_n, x\rangle x_n\|^2.$$

Proposition 2.92 (Bessel's Inequality). *Let $\{x_a : a \in A\}$ be an orthonormal set. Then $\sum_{a\in A} |\langle x_a, x\rangle|^2$ converges and is less than $\|x\|^2$.*

Definition 2.93. We say that an orthonormal set $\{e_a : a \in A\}$ in a Hilbert space $\mathcal{H}$ is **maximal** if there does not exist an orthonormal set that contains it as a proper subset. A maximal orthonormal set is called an **orthonormal basis(o.n.b.)**.

Maximal orthonormal sets exist by Zorn's Lemma(which is equivalent to the Axiom of Choice). The following explains the reason for calling these bases.

Theorem 2.94 (Parseval's Identities). *Let $\mathcal{H}$ be a Hilbert space and let $\{e_a : a \in A\}$ be an orthonormal basis. Then for any $x, y \in \mathcal{H}$,*

1. $\|x\|^2 = \sum_{a\in A} |\langle e_a, x\rangle|^2$,
2. $x = \sum_{a\in A} \langle e_a, x\rangle e_a$,
3. $\langle x, y\rangle = \sum_{a\in A} \langle x, e_a\rangle\langle e_a, y\rangle$.

Problem 2.95. Prove the 3rd Parseval identity.

Definition 2.96. Two sets S and T are said to have the same **cardinality** if there is a one-to-one onto function between them and we write $\mathbf{card(S) = card(T)}$. When there exists a one-to-one function from S into T we write that $\mathbf{card(S) \leq card(T)}$.

It is a fact that $card(S) = card(T) \iff Card(S) \leq card(T)$ and $card(T) \leq card(S)$. For each natural number $\{0, 1, ...\}$ we have a cardinal number. We set $\aleph_0 := card(\mathbb{N})$ and we set $\aleph_1 := card(\mathbb{R})$. The **continuum hypothesis** asks if there can be a set with $card(\mathbb{N}) < card(S) < card(\mathbb{R})$. This statement is known to be independent of the ZFC= Zermelo-Frankl + Axiom of Choice. So one can add as another axiom the statement that this is true or that it is false to one's axiom system.

We call a set A **countably infinite** if $card(A) = card(\mathbb{N})$. We call a set **countable** if it is either finite or countably infinite.

Cardinal arithmetic is defined as follows: Given disjoint sets A and B we set $card(A) + card(B) := card(A \cup B)$ and we set $card(A) \cdot card(B) := card(A \times B)$, where $A \times B = \{(a, b) : a \in A, b \in B\}$ is the Cartesian product.

Theorem 2.97 (Basis Theorem)**.** *Let $\mathcal{H}$ be a Hilbert space and let $\{e_a : a \in A\}$ and $\{f_b : b \in B\}$ each be an orthonormal basis for $\mathcal{H}$. Then $card(A) = card(B)$.*

Definition 2.98. We call the cardinality of any orthonormal basis for $\mathcal{H}$ the **Hilbert space dimension of** $\mathcal{H}$. We denote it by $dim_{HS}(\mathcal{H})$.

Example 2.99. For every cardinal number, there is a Hilbert space of that dimension. Here is a way to see this fact. Given any non-empty set A set

$$\ell^2_A := \{f : A \to \mathbb{F} : \sum_{a \in A} |f(a)|^2 < +\infty\}.$$

On this vector space we define

$$\langle f, g \rangle = \sum_{a \in A} \overline{f(a)} g(a).$$

The fact that this converges and defines an inner product goes much like the proof for ℓ^2. Now define $e_a \in \ell^2_A$ by

$$e_a(t) = \begin{cases} 1, & t = a \\ 0, & t \neq a \end{cases}.$$

Then the set $\{e_a : a \in A\}$ can be shown to be an o.n.b. for ℓ^2_A. Hence, $dim_{HS}(\ell^2_A) = card(A)$.

Proposition 2.100 (Gram-Schmidt)**.** *Let $\{x_n : 1 \leq n \leq N\}$ be a linearly independent set in a Hilbert space. Then there exists an orthonormal set $\{e_n : 1 \leq n \leq N\}$ such that*

$$span\{x_1, ..., x_k\} = span\{e_1, ..., e_k\}, \ \forall k, 1 \leq k \leq N.$$

Proposition 2.101 (Gram-Schmidt)**.** *Let $\{x_n : n \in \mathbb{N}\}$ be a linearly independent set in a Hilbert space. Then there exists an orthonormal set $\{e_n : n \in \mathbb{N}\}$ such that*

$$span\{x_1, ..., x_k\} = span\{e_1, ..., e_k\}, \ \forall k \in \mathbb{N}.$$

One produces this set by the **Gram-Schmidt orthogonalization process**. In the finite case. For the general case one uses the finite case together with the Principle of Recursive Definition.

Recall that a metric space is called **separable** if it contains a countable dense subset.

Theorem 2.102. *Let $\mathcal{H}$ be a Hilbert space. Then $\mathcal{H}$ is separable(as a metric space) if and only if $dim_{HS}(\mathcal{H}) \leq \aleph_0$, i.e., any o.n.b. for $\mathcal{H}$ is countable.*

2.8.1 Direct Sums

Given two vector spaces V and W we can make a new vector space by taking the Cartesian product

$$\{(v, w) : v \in V, w \in W\},$$

and defining $(v_1, w_1) + (v_2, w_2) = (v_1 + v_2, w_1 + w_2)$ and $\lambda(v, w) = (\lambda v, \lambda w)$. Even though it uses the Cartesian *product*, this space is generally denoted $V \oplus W$. One reason is that in the finite dimensional case if $\{v_1, ..., v_n\}$ is a basis for V and $\{w_1, ..., w_m\}$ is a basis for W, then $\{(v_1, 0), ..., (v_n, 0)\} \cup \{(0, w_1), ..., (0, w_m)\}$ is a basis for $V \oplus W$ so that $dim(V \oplus W) = dim(V) + dim(W)$.

Given Hilbert spaces $(\mathcal{H}_i, \langle, \rangle_i), i = 1, 2$ on their Cartesian product we define

$$\langle (h_1, h_2), (h_1', h_2') \langle = \langle h_1, h_1' \rangle_1 + \langle h_2, h_2' \rangle_2,$$

and it is easy to check that this is an inner product and that the Cartesian product is complete in this inner product. To see this one checks that $(h_{1,n}, h_{2,n})$ is Cauchy if and only if $h_{1,n}$ is Cauchy in $\mathcal{H}_1$ and $h_{2,n}$ is Cauchy in $\mathcal{H}_2$. Hence, there are vectors such that $h_1 = \lim_n h_{1,n}$ and $h_2 = \lim h_{2,n}$. Now one checks that $\|(h_1, h_2) - (h_{1,n}, h_{2,n})\| \to 0$.

This Hilbert space is denoted by $\mathcal{H}_1 \oplus \mathcal{H}_2$. Note that in this space, $(h_1, 0) \perp (0, h_2)$ and that $\|(h_1, 0)\| = \|h_1\|_1$ while $\|(0, h_2)\| = \|h_2\|_2$ so that these subspaces can be identified with the original HIlbert spaces and they are arranged as orthogonal subspaces of $\mathcal{H}_1 \oplus \mathcal{H}_2$.

It is not hard to verify that if $\{e_a : a \in A\}$ is an o.n.b. for $\mathcal{H}_1$ and $\{f_b : b \in B\}$ is an o.n.b. for $\mathcal{H}_2$, then

$$\{(e_a, 0) : a \in A\} \cup \{(0, f_b) : b \in B\},$$

is an o.n.b. for $\mathcal{H}_1 \oplus \mathcal{H}_2$. Thus, $dim_{HS}(\mathcal{H}_1 \oplus \mathcal{H}_2) = dim_{HS}(\mathcal{H}_1) + dim_{HS}(\mathcal{H}_2)$, where the addition is either ordinary in the finite dimensional case or cardinal arithmetic in general.

Note that when we form the direct sum, we can identify $\mathcal{H}_1$ and $\mathcal{H}_2$ canonically as subspaces of the direct sum in the following sense. Namely, if we look at $\widetilde{\mathcal{H}}_1 = \{(h_1, 0) : h_1 \in \mathcal{H}_1\}$ and $\widetilde{\mathcal{H}}_2 = \{(0, h_2) : h_2 \in \mathcal{H}_2\}$, then we have that, as vector spaces these are isomorphic to the original spaces and for any $h_1, h_1' \in \mathcal{H}_1$,

$$\langle h_1 | h_1' \rangle_{\mathcal{H}_1} = \langle (h_1, 0) | (h_1', 0) \rangle_{\mathcal{H}_1 \oplus \mathcal{H}_2},$$

and similarly for any pair of vectors in $\mathcal{H}_2$. We also have that

$$\widetilde{\mathcal{H}_1}^{\perp} = \widetilde{\mathcal{H}_2}.$$

We now generalize this construction to arbitrary collections. Given a family of Hilbert spaces $(\mathcal{H}_a, \langle , \rangle_a)$, $a \in A$. We define a new Hilbert space which is a subset of the Cartesian product via

$$\oplus_{a \in A} \mathcal{H}_a = \{(h_a)_{a \in A} : \sum_{a \in A} \|h_a\|_a^2 < +\infty\},$$

and inner product

$$\langle (h_a), (k_a) \rangle = \sum_{a \in A} \langle h_a, k_a \rangle_a.$$

Proofs similar to the case of ℓ^2_A show that this is indeed a Hilbert space.

2.8.2 Bilinear Maps and Tensor Products

Given vector spaces V, W, Z a map $B : V \times W \to Z$ is called **bilinear** provided that it is linear in each variable, i.e.,

- $B(v_1 + v_2, w) = B(v_1, w) + B(v_2, w)$,
- $B(v, w_1 + w_2) = B(v, w_1) + B(v, w_2)$
- $B(\lambda v, w) = \lambda B(v, w) = B(v, \lambda w)$.

Tensor products were created to linearize bilinear maps. This concept takes some development but, in summary, given two vector spaces their tensor product,

$$V \otimes W = span\{v \otimes w : v \in V, w \in W\},$$

where $v \otimes w$ is called an **elementary tensor** and the rules for adding these mimic the bilinear rules:

- $v_1 \otimes w + v_2 \otimes w = (v_1 + v_2) \otimes w$,
- $v \otimes w_1 + v \otimes w_2 = v \otimes (w_1 + w_2)$,
- $\lambda(v \otimes w) = (\lambda v) \otimes w = v \otimes (\lambda w)$.

If V and W have bases $\{v_a : a \in A\}$ and $\{w_b : b \in B\}$, respectively, then $\{v_a \otimes w_b : a \in A, b \in B\}$ is a basis for $V \otimes W$. Thus, the vector space dimension satisfies

$$dim(V \otimes W) = dim(V) \cdot dim(W),$$

where this is just the product of integers in the finite dimensional case and is true in the sense of cardinal arithmetic in the general case.

Proposition 2.103. *There is a one-to-one correspondence between bilinear maps, $B : V \times W \to Z$ and linear maps $T : V \otimes W \to Z$ given by $T(v \otimes w) = B(v, w)$.*

When V and W are both normed spaces, we would like the tensor product to be a normed space too. There are several "natural" norms to put on the tensor product and on bilinear maps.

Definition 2.104. Let X, Y and Z be normed spaces and let $B : X \times Y$ be a bilinear map. Then we say that B is **bounded** provided that

$$\|B\| = \sup\{||B(x, y)|| : \|x\| \leq 1, \, \|y\| \leq 1\} < +\infty,$$

and we call this quantity the **norm of B**. For $u \in X \otimes Y$ the **projective norm** of u is given by

$$\|u\|_{\vee} = \inf\{\sum \|x_i\| \cdot \|y_i\| : u = \sum_i x_i \otimes y_i\},$$

where the infimum is taken over all ways to express u as a finite sum of elementary tensors. The completion of $X \otimes Y$ in this norm is denoted $X \otimes^{\vee} Y$ and is called the **projective tensor product of X and Y**.

Here is the relationship these two concepts.

Proposition 2.105. *Let X, Y and Z be Banach spaces. Then $B : X \times Y \to Z$ is bounded if and only if the linear map $T_B : X \otimes Y \to Z$ is bounded when $X \otimes Y$ is given the projective norm. In this case $\|B\| = \|T_B\|$ and T_B can be extended to a bounded linear map of the same norm from $X \otimes_{\vee} Y$ into Z.*

The extension of T_B to the completion is, generally, still denoted by T_B.

There is one other tensor norm that is very important. Note that if $f : X \to \mathbb{F}$ and $g : Y \to \mathbb{F}$ are linear, then there is a linear map $f \otimes g : X \otimes Y \to \mathbb{F}$ given by $f \otimes g(x \otimes y) = f(x) \cdot g(y)$. To see this just check that $B(x, y) := f(x)g(y)$ is bilinear.

Definition 2.106. Let X and Y be normed spaces. Given $u \in X \otimes Y$, the **injective norm of u** is given by the formula

$$\|u\|_\wedge := \sup\{|f \otimes g(u)| : f \in X^*,\, g \in Y^*,\, \|f\| \leq 1,\, \|g\| \leq 1\}.$$

This is indeed a norm on $X \otimes Y$ and the completion in this norm is denoted by $X \otimes_\wedge Y$ and is called the **injective tensor product of X and Y**.

Grothendieck was the first to systematically study and classify tensor norms. He defined a **reasonableness condition** for tensor norms and proved that if X and Y are normed spaces, then any reasonable tensor norm on $X \otimes Y$ satisfies, $\|u\|_\wedge \leq \|u\| \leq \|u\|_\vee$.

.

2.8.3 Tensor Products of Hilbert Spaces

All of the above discussion is preliminary to defining a norm on the tensor product of Hilbert spaces that yields a new Hilbert space.

Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces. We wish to define an inner product on $\mathcal{H} \otimes \mathcal{K}$. Given $u = \sum_i h_i \otimes k_i$ and $v = \sum_j v_j \otimes w_j$ two finite sums representing vectors in $\mathcal{H} \otimes \mathcal{K}$, we set

$$\langle u, v\rangle = \sum_{i,j} \langle h_i | v_j\rangle_{\mathcal{H}} \langle k_i | w_j\rangle_{\mathcal{K}}.$$

It is not hard to see that if this formula is well-defined, i.e., does not depend on the particular way that we write u and v as a sum of elementary tensors, then it is sesquilinear.

Proposition 2.107. *The above formula is well-defined and gives an inner product on $\mathcal{H} \otimes \mathcal{K}$.*

Proof. We omit the details of showing that it is well-defined. As noted above, given that it is well-defined it is elementary that it is sesquilinear. So we only prove that it satisfies $u \neq 0 \implies \langle u, u\rangle > 0$.

To this end let $u = \sum_{i=1}^n h_i \otimes k_i$. Let $\mathcal{F} = span\{k_i : 1 \le i \le n\}$ and let $\{e_j : 1 \le j \le m\}$ be an o.n.b. for $\mathcal{F}$ so that $m \le n$. Write each $k_i = \sum_j a_{i,j} e_j$. Then we have that

$$u = \sum_i \sum_j h_i \otimes (a_{i,j} e_j) = \sum_j \sum_i (a_{i,j} h_i) \otimes e_j = \sum_j v_j \otimes e_j,$$

where $v_j = \sum_i a_{i,j} h_i$. Now since $u \neq 0$ we must have that some $v_j \neq 0$. Using the fact that the inner product is independent of the way that we express u as a sum of elementary tensors, we have that

$$\langle u|u\rangle = \sum_{i,j} \langle v_i|v_j\rangle_{\mathcal{H}} \langle e_i|e_j\rangle_{\mathcal{K}} = \sum_i \|v_i\|^2 \neq 0,$$

and we are done. $\square$

Definition 2.108. Given Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$ as above, we let $\mathcal{H}\overline{\otimes}\mathcal{K}$ denote the completion of $\mathcal{H} \otimes \mathcal{K}$ in the above inner product and call it the **Hilbert space tensor product**.

Many authors use $\mathcal{H} \otimes \mathcal{K}$ to denote both the vector space tensor product AND its completion, or they sometimes invent a non-standard notation to denote the vector space tensor product, like $\mathcal{H} \odot \mathcal{K}$, reserving $\mathcal{H} \otimes \mathcal{K}$ for the completion.

Proposition 2.109. *Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces, let $\{e_a : a \in A\}$ be an o.n.b. for $\mathcal{H}$ and $\{f_b : b \in B\}$ be an o.n.b. for $\mathcal{K}$. Then*

$$\{e_a \otimes f_b : a \in A, b \in B\}.$$

is an o.n.b. for $\mathcal{H}\overline{\otimes}\mathcal{K}$. Hence,

$$dim_{HS}(\mathcal{H}\overline{\otimes}\mathcal{K}) = dim_{HS}(\mathcal{H}) \cdot dim_{HS}(\mathcal{K}).$$

In a similar way one can form a Hilbert space tensor product of any finite collection of Hilbert spaces $\mathcal{H}_1, \ldots, \mathcal{H}_n$, denoted $\mathcal{H}_1\overline{\otimes}\cdots\overline{\otimes}\mathcal{H}_n$. The inner product on any two elementary tensors is given by

$$\langle v_1 \otimes \cdots \otimes v_n | w_1 \otimes \cdots \otimes w_n\rangle = \langle v_1|w_1\rangle_{\mathcal{H}_1} \cdots \langle v_n|w_n\rangle_{\mathcal{H}_n},$$

and then is extended linearly to finite sums of elementary tensor products.

Unlike for the direct sum construction, there is no natural way that we can regard $\mathcal{H}$ and $\mathcal{K}$ as subspaces of $\mathcal{H}\overline{\otimes}\mathcal{K}$. In particular, $\{h \otimes 0 : h \in \mathcal{H}\} =$

$(0 \otimes 0)$. One way that we can regard $\mathcal{H}$ as a subspace of the tensor product is to pick a unit vector $k \in \mathcal{K}$. Then we have that $\widetilde{\mathcal{H}} := \{h \otimes k : h \in \mathcal{H}\}$ is isomorphic to $\mathcal{H}$ as a vector space, since $(h_1 + h_2) \otimes k = h_1 \otimes k + h_2 \otimes k$ and $(\lambda h) \otimes k = \lambda(h \otimes k)$ and has the property that for any $h_1, h_2 \in \mathcal{H}$,

$$\langle h_1 | h_2 \rangle_{\mathcal{H}} = \langle h_1 \otimes k | h_2 \otimes k \rangle_{\mathcal{H} \otimes \mathcal{K}}.$$

Similarly, fixing a unit vector in $h \in \mathcal{H}$ allows us to include $\mathcal{K}$ into the tensor product as $\widetilde{\mathcal{K}} := \{h \otimes k : k \in \mathcal{K}\}$.

However, unlike the direct sum these two subspaces are not orthogonal, in fact, their intersection is the one-dimensional subspace spanned by $h \otimes k$.

2.8.4 Infinite Tensor Products

Many physical models call for tensor products of infinitely many Hilbert spaces. This arises for example where one has an infinite lattice of electrons. Each electron has a state space that is a Hilbert space $\mathcal{H}_a$ and the models say that the state space of the whole ensemble should be the tensor product over all electrons.

The first problem that one encounters is that if one has two elementary tensor products of infinitely many vectors, $v_1 \otimes v_2 \otimes \cdots$ and $w_1 \otimes w_2 \otimes \cdots$, then one would like to set

$$\langle v_1 \otimes v_2 \otimes \cdots | w_1 \otimes w_2 \otimes \cdots \rangle = \prod_{i=1}^{\infty} \langle v_i | w_i \rangle,$$

but such an infinite product need not converge. Second, if we had a set of unit vectors $v_1, ...$ and we sat $u = v_1 \otimes (v_2/2) \otimes (v_3/3) \otimes \cdots$, then we would have that

$$\|u\|^2 = \langle u | u \rangle = \prod_{i=1}^{\infty} \frac{1}{i^2} = 0.$$

However, in physics one usually has a lot more. Each electron has a distinguished unit vector in its state space called the *ground state* and it turns out that given Hilbert spaces $\mathcal{H}_a$ and distinguished unit vectors $\psi_a \in \mathcal{H}_a$, then this is enough data to form an infinite tensor product.

The key is, remember that when we have a unit vector then we can "include" $\mathcal{H}$ into $\mathcal{H} \otimes \mathcal{K}$ by sending $h \to h \otimes k$, where k is some fixed unit vector.

We form the infinite tensor product of a collection $(\mathcal{H}_a, \psi_a)$, $a \in A$, where $\psi_a \in \mathcal{H}_a$ is a unit vector as described below.

For each finite subset $F \subseteq A$ there is a well-defined tensor product Hilbert space of these finitely many spaces, which we denote by $\mathcal{H}_F := \otimes_{a\in F}\mathcal{H}_a$(we do not need to complete at this time).

Now given two finite sets $F_1 \subseteq F_2 \subseteq A$, say $F_1 = \{a_1, ..., a_n\}$ and $F_2 = \{a_1, ..., a_n, ...a_m\}$ with $m > n$, we can regard $\mathcal{H}_{F_1}$ as a subspace of $\mathcal{H}_{F_2}$ by identfying $u \in \mathcal{H}_{F_1}$ with $u \otimes \psi_{a_{n+1}} \otimes \cdots \otimes \psi_{a_m} \in \mathcal{H}_{F_2}$. Formally, this means that we have a linear map, $V_{F_1,F_2} : \mathcal{H}_{F_1} \to \mathcal{H}_{F_2}$.

Now take the union over all finite subsets, $\cup_{F\subseteq A}\mathcal{H}_F$, and define a relation on this union by declaring $u_1 \in \mathcal{H}_{F_1}$ **equivalent** to $u_2 \in \mathcal{H}_{F_2}$ provided that there exists a finite set F_3 with $F_1 \subseteq F_3$ and $F_2 \subseteq F_3$ such that $V_{F_1,F_3}(u_1) = V_{F_2,F_3}(u_2)$(we will write $u_1 \sim u_2$ for this relation). It is not hard to show that this is an equivalence relation and we will write $[u]$ for the equivalence class of a vector.

It is not hard to show that $\mathcal{K} = \big(\cup_{F\subseteq A}\mathcal{H}_F\big)/\sim = \{[u] : \exists F, u \in \mathcal{H}_F\}$ is an inner product space with operations defined as follows. Given $\lambda \in \mathbb{F}$ and $[u]$ set $\lambda[u] = [\lambda u]$. Given $[u_1], [u_2]$ with $u_1 \in \mathcal{H}_{F_1}$ and $u_2 \in \mathcal{H}_{F_2}$ pick F_3 with $F_1, F_2 \subseteq F_3$ and set

$$[u_1] + [u_2] = [V_{F_1,F_3}(u_1) + V_{F_2,F_3}(u_2)].$$

Finally, the inner product is defined by

$$\langle [u_1]|[u_2]\rangle_{\mathcal{K}} = \langle V_{F_1,F_3}(u_1)|V_{F_2,F_3}(u_2)\rangle_{\mathcal{H}_{F_3}}.$$

Showing that all of these operations are well-defined, that is, only depend on the equivalence class, not on the particular choices, and that the last formula is an inner product is tedious, but does all work. The completion of the space $\mathcal{K}$ is denoted

$$\overline{\otimes}_{a\in A}(\mathcal{H}_a, \psi_a).$$

In the end this tensor product does not depend on the choice of vectors, in the following sense. If we also choose unit vectors $\phi_a \in \mathcal{H}_a$, then for each Hilbert space there is a unitary map $U_a : \mathcal{H}_a \to \mathcal{H}_a$ such that $U_a(\psi_a) = \phi_a$(we will discuss unitaries in a later section) and these define unitaries $U_F : \mathcal{H}_F \to \mathcal{H}_F$ via $U_F(h_{a_1}\otimes\cdots\otimes h_{a_n}) = U_{a_1}(h_{a_1})\otimes\cdots\otimes U_{a_n}(h_{a_n})$. Together these define a unitary U on $\mathcal{K}$ by for $u \in \mathcal{H}_F$ setting $U([u]) = [U_F(u)]$. Then this unitary satisfies $U([\psi_{a_1} \otimes \cdots \psi_{a_n}]) = [\phi_{a_1} \otimes \cdots \otimes \phi_{a_n}]$, and so induces a unitary on the completions, $U : \overline{\otimes}_{a\in A}(\mathcal{H}_a, \psi_a) \to \overline{\otimes}_{a\in A}(\mathcal{H}_a, \phi_a)$.