Notes in Analysis, AFS-II 2014 Almora

Anant R. Shastri
IIT Bombay

May 13, 2014

Chapter 1

Normed Linear Spaces

1.1 Convex functions and inequalities

(Refer to Rudin's *Real and Complex...*)

Definition 1.1.1 A function $\phi : (a, b) \to \mathbb{R}$ is said to be convex if

$$\phi((1-t)x + ty) \leq (1-t)\phi(x) + t\phi(y) \tag{1.1}$$

for all $x, y \in (a, b)$ and $0 \leq t \leq 1$.

(1.1) is equivalent to say

$$\frac{\phi(z) - \phi(x)}{z - x} \leq \frac{\phi(y) - \phi(z)}{y - z} \tag{1.2}$$

whenever $a < x < z < y < b$. [Put $t = \frac{z-x}{y-x}$ and simplify.]

Theorem 1.1.2
(A) *Let ϕ be differentiable function. Then ϕ is convex iff ϕ' is monotonically increasing.*
(B) *The function* $\exp : \mathbb{R} \to (0, \infty)$ *is a convex function.*
(C) *Any convex function is continuous.*

Proof: (A) In (1.2) let $a < w < x$ to obtain

$$\frac{\phi(x) - \phi(w)}{x - w} \leq \frac{\phi(z) - \phi(x)}{z - x} \leq \frac{\phi(y) - \phi(z)}{y - z}.$$

Now take limit as $w \to x$ and $z \to y$ to see that $\phi'(x) \leq \phi'(y)$. For the converse, use Lagrange Mean Value Theorem.

(C) In (1.2), take $a < u < w < x < z < y < v < b$ and obtain

$$(z - x)\frac{\phi(w) - \phi(u)}{w - u} \leq \phi(z) - \phi(x) \leq (z - x)\frac{\phi(v) - \phi(y)}{v - y}.$$

Put $M = \max\left\{\left|\frac{\phi(w)-\phi(u)}{w-u}\right|, \left|\frac{\phi(v)-\phi(y)}{v-y}\right|\right\}$. Given $\epsilon > 0$ take $\delta = \epsilon/M$. ♠

The following measure theoretic inequality is a far reaching generalization of the (1.1).

Theorem 1.1.3 Jensen's Inequality: *Let (X, μ) be a probability measure space (i.e., $\mu(X) = 1$) and let $f \in L^1(\mu)$ taking values inside (a, b). Then for any convex function ϕ on (a, b), we have*

$$\phi\left(\int_X f\, d\mu\right) \leq \int_X \phi \circ f\, d\mu.$$

Proof: By the average value theorem, it follows that $z := \int_X f\, d\mu \in (a, b)$. Let $\beta = \beta(z)$ denote the supremum of LHS of (1.2) over all x such that $a < x < z$. Then clearly,

$$\frac{\phi(z) - \phi(x)}{z - x} \leq \beta \leq \frac{\phi(y) - \phi(z)}{y - z} \tag{1.3}$$

whenever $a < x < z < y < b$. These two inequalities can be combined to yield

$$\beta(s - z) \leq \phi(s) - \phi(z)$$

for all $a < s < b$. This means that for all $x \in X$

$$\beta(f(x) - z) \leq \phi(f(x) - \phi(z)$$

and hence, upon integration,

$$0 = \beta\left(\int_X f\, d\mu - z\int_X d\mu\right) \leq \int_X \phi \circ f\, d\mu - \phi(z)\int_X d\mu$$

which yields the required result, since $\int_X d\mu = 1$. ♠

Definition 1.1.4 Let p, q be positive real numbers such that $p + q = pq$, equivalently,

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

We then call p, q conjugate pairs of exponents. Often q is denoted by p'.

Clearly, then $1 < p, q < \infty$. Also, if we let $p \to 1$ then we get $q \to \infty$ and vice versa. So, we allow these extreme cases as well, which, ofcourse, require us to pay special attention to them.

Theorem 1.1.5 *Let $1 < p, q < \infty$ be a pair of conjugate exponents. Let (X, μ) be a measure space and $f : X \to [0, \infty]$ be a measurable function. Then*

(A) Hölder Inequality:

$$\int_X fg\, d\mu \leq \left(\int_X f^p\, d\mu\right)^{1/p}\left(\int_X g^q\, d\mu\right)^{1/q}. \tag{1.4}$$

(B) Minkowski Inequality:

$$\left(\int_X (f+g)^p\, d\mu\right)^{1/p} \leq \left(\int_X f^p\, d\mu\right)^{1/p} + \left(\int_X g^p\, d\mu\right)^{1/p} \tag{1.5}$$

(C) Schwarz Inequality:

$$\left(\int_X (f+g)^2\, d\mu\right)^{1/2} \leq \left(\int_X f^2\, d\mu\right)^{1/2} + \left(\int_X g^2\, d\mu\right)^{1/2} \tag{1.6}$$

Proof: (A) Let a and b be the two quantities on the right side of (1.4). If one of them is zero it follows that $fg = 0$ a.e., and hence $\int_X fg\, d\mu = 0$ and hence (1.4) is true. So we may and shall assume that $a \neq 0 \neq b$ and put $F = f/a, G = g/b$. It then follows that

$$\int_X F^p\, d\mu = 1 = \int_X G^q\, d\mu. \tag{1.7}$$

Since $\exp : \mathbb{R} \to (0, \infty)$ is onto, we can choose $s, t \in \mathbb{R}$ such that $e^s = F(x)^p, e^t = G(x)^q$. Since exp is convex, and $1/p + 1/q = 1$, it follows that

$$F(x)G(x) = \exp(s/p + t/q) \leq \frac{1}{p}F(x)^p + \frac{1}{q}G(x)^q, \quad x \in X. \tag{1.8}$$

Integrating over X, we obtain

$$\int_X FG \leq \frac{1}{p} + \frac{1}{q} = 1. \tag{1.9}$$

Now (1.7) yields (1.4).
(B) In order to prove (1.5), we may assume that the quantities on the RHS are finite. By convexity of the power function t^p, it follows that

$$((f + g)/2)^p \leq \frac{1}{2}(f^p + g^p).$$

Integrating this, we obtain that $\int_X (f + g)^p < \infty$.

Write $(f + g)^p = f(f + g)^{p-1} + g(f + g)^{p-1}$ and apply (A) to the two quantities on the right-side to obtain

$$\int_X f(f + g)^{p-1} \leq \left(\int_X f^p\right)^{1/p} \left(\int_X (f + g)^{(p-1)q}\right)^{1/q}$$

$$\int_X g(f + g)^{p-1} \leq \left(\int_X g^p\right)^{1/p} \left(\int_X (f + g)^{(p-1)q}\right)^{1/q}$$

Adding these two and using the fact $(p - 1)q = p$ gives us

$$\int_X (f + g)^p \leq \left[\left(\int_X f^p\right)^{1/p} + \left(\int_X g^p\right)^{1/p}\right] \left(\int_X (f + g)^p\right)^{1/q}.$$

Cancelling out the last factor on the right, we obtain (1.5).
(C) Put $p = 2 = q$ in (B). ♠

Remark 1.1.6 Notice that equality holds in (1.8) iff $s = t$ and equality holds in (1.9) iff equality holds in (1.8), a.e. Therefore, it follows that equality holds in (1.4) iff $F^p = G^q$ a.e. Therefore, we can conclude that equality holds in (1.4) iff there exists constants α, β such that $\alpha f = \beta g$ a.e. Similarly, it follows that equality holds in (1.5) iff there are constants α, β such that $\alpha f^p = \beta g^q$ a.e.

Exercise 1.1.7

1. Let S be a collection of convex functions on (a, b) such that $g(t) = \text{Sup}\,\{f(t) \,:\, f \in S\} < \infty$. Show that g is convex.

2. Let $f_n : (a, b) \to \mathbb{R}$ be a sequence of convex functions. Put $g(t) = \text{limsup}_n f_n(t)$; $h(t) = \text{liminf}_n f_n(t)$. Are g and h convex?

3. Let $\phi : (a, b) \to (c, d)$ be a convex function and $\psi : (c, d) \to \mathbb{R}$ be a non decreasing convex function. Then $\psi \circ \phi$ is convex.

4. Let $f : (a, b) \to \mathbb{R}$ be a function such that for all $x, y \in (a, b)$

$$f(x + y)/2) \leq (f(x) + f(y))/2.$$

(a) Show that f satisfies (1.1) for all $t \in [0, 1] \cap D$ where D denotes the set of all dyadic rationals.
(b) In fact, show that f satisfies (1.1) for all $t \in [0, 1] \cap \mathbb{Q}$.
(c) If f is continuous, show that f is convex.
(d) In fact, if f is bounded on some open interval contained in (a, b) then show that f is convex and hence continuous.
(e) Illustrate the necessity of continuity of f in (c).

1.2 Normed Linear Spaces

Throughout these discussion, $\mathbb{K}$ will denote either $\mathbb{R}$ or $\mathbb{C}$.

Definition 1.2.1 Let X be a vector space over $\mathbb{K}$. By a norm on X we mean a function $\| \ \| : X \to [0, \infty)$ having the following properties:
(i) $\|x\| = 0$ iff $x = 0$.
(ii) $\|\alpha x\| = |\alpha|\|x\|, \ \ \alpha \in \mathbb{K}, x \in X$.
(iii) $\|x + y\| \leq \|x\| + \|y\|, \ \ \ x, y \in X$.
A vector space X together with a norm is called a normed linear space (NLS).

For now on, X will always denote a normed linear space unless mentioned specifically otherwise.

Remark 1.2.2 There is the underlying vector space, the underlying metric space and the underlying topological space. The metric is given by $d(x, y) = \|x - y\|$. with respect to this metric, we write $B_r(x)$ denote the open ball of radius r and centre $x \in X$. The topology gnenerated by the open balles as a basis is the corresponding metric topology. Properties (ii) and (iii) tell you that with respect to this topology, the addition and the scalar multiplication:

$$(x, y) \mapsto x + y, \qquad (\alpha, x) \mapsto \alpha x$$

are continuous. In particular, the maps

$$x \mapsto x + y, y \in X; \qquad x \mapsto kx, k \in \mathbb{K} \setminus \{0\},$$

are homeomorphisms.

Example 1.2.3
(i) $\ell_p^n := \mathbb{K}_p^n$:

Let $1 \le p \le \infty$. On $\mathbb{K}^n$, define $\| \ \|_p$ as follows:

$$\|x\|_p = \begin{cases} \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}, & \text{if } p < \infty; \\ \max\{|x_i| \ : \ 1 \le i \le n\}, & \text{otherwise.} \end{cases}$$

The non trivial thing to verify about these examples is the triangle inequality. However, for $p = 1, \infty$, this is easy and direct.

For $1 < p < \infty$, we appeal to Minkowski's inequality: On the set X of n elements $\{1, 2, \ldots, n\}$ we take the normalized counting measure: $\mu(\{i\}) = 1/n$. Points $x \in \mathbb{K}^n$ give rise to measurable functions $|x| : X \to [0, \infty)$ Then (1.5) translates into

$$\left(\sum_{i=1}^n (|x_i + y_i|)^p\right)^{1/p} \le \left(\sum_{i=1}^n (|x_i| + |y_i|)^p\right)^{1/p} \le \left(\sum_{i=1}^n |x_i|^p\right)^{1/p} + \left(\sum_{i=1}^n |y_i|^p\right)^{1/p}.$$

This is nothing but the triangle inequality for $\| \ \|_p$.

Draw the pictures of unit balls in ℓ_p^2 and see why we need to take $p \ge 1$.

(ii) ℓ_p : We now take X to be an appropriate subset of $\mathbb{K}^\infty$, the set of all sequences $\mathbb{N} \to \mathbb{K}$. The idea is to take infinite sums in the formulae for $\|x\|_p$ and hence the rhs has to make sense. Thus, let ℓ_p denote the set of all those sequences $(x_1, x_2, \ldots,)$ such that

$$\sum_i |x_i|^p < \infty$$

(i.e., absolutely power p-summable sequences. Once again, verification that these are vector spaces and the function $\|-\|_p$ defined as above satisfies conditions other than triangle inequality is easy. The TE itself follows from the same in the finite case upon taking the limit. Finally for the case $p = \infty$, we replace 'max' by 'Sup' and the proofs are immediate consequences of the corresponding results for finite case.

(iii) $L^p(X, \mu)$.

In fact the idea of proving Minkowski inequality for measurable functions is precisely that the space of all power p-integrable functions on a measure space (X, μ) forms a NLS when we define

$$\|f\|_p = \left(\int_X |f|^p \, d\mu\right)^{1/p}.$$

These spaces are also denoted by $L^p(\mu)$ or $L^p(X)$. For the case $p = \infty$ we take $L^\infty(X)$ to be the space of all bounded functions, with

$$\|f\|_\infty = Sup\{|f(x) \ : \ x \in X\}.$$

However, there is a catch!

For $1 \le p < \infty$, $\|f\| = 0$ does not necessarily imply that $f = 0$. It only implies $f = 0$ a.e. But that is quite satisfactory and is the thing that one expects, when doing measure theory any way. Therefore, we **redefine** $L^p(X)$ to be the equivalence classes of p-summable functions: $f \sim g$ iff $f - g = 0$ a.e. And then the positive definiteness of $\| \ \|_p$ is restored.

In view of this, we **redefine** $L^\infty(X)$ also as follows: A measurable function $f : X \to \mathbb{K}$ is said to be *essentially bounded* if there exists $\alpha > 0$ such that $\mu(x \in X \ : \ |f(x)| > \alpha) = 0$. The

equivalence classes of all essentially bounded measurable functions is denoted by $L^\infty(X)$ (the equivalence is defined as before). For $f \in L^\infty(X)$ we then define
$\|f\|_\infty =$ the infimum of α such that $\mu(x \in X \ : \ |f(x)| > \alpha) = 0$.
It is routine to verify that with this norm $L^\infty(X)$ becomes a NLS.
(iv) Let X be any set and $B(X)$ be the space of all bounded $\mathbb{K}$-valued functions on X. For $f \in B(X)$ we define $\|f\|_\infty = \text{Sup}\ \{|f(x)| \ : \ x \in X\}$. Then $B(X)$ becomes a NLS with this **sup norm.** In case X is a metric space, we can consider various subspaces of $B(X)$:
$C(X) = \{f \in B(X) \ : \ f \text{ in continuous}\}$;
$C_0(X) = \{f \in C(X) \ : \ \forall\ \epsilon > 0, \exists \text{ a compact set } K \subset X \text{ with } |f(x)| < \epsilon, x \in X \setminus K\}$.
$C_c(X) = \{f \in C(X) \ : \ \exists \text{ a compact subset } K \subset X \ : \ f(x) = 0 \text{ for all } x \notin K\}$.
Elements of $C_0(X)$ (resp $C_c(X)$) are called continuous functions vanishing at ∞ (continuous functions with compact support). The norm on all these is the sup norm.
(v) We are also interested in spaces $C^m(U;\mathbb{K})$ of m-times differentiable functions on nice subspaces $U \subset \mathbb{K}^n$. For instance, when U is an open or a closed interval, then we can give $C^m(U)$ the norm defined by

$$\|f\| = \sum_{i=0}^{m} \|f^{(i)}\|_\infty$$

where $f^{(i)}$ denotes the i^{th} derivative. One can also consider smooth functions and holomorphic functions and so on.

1.3 Subspaces, Quotients and Products

Definition 1.3.1 Let X be a NLS. Then by a subspace of X we mean a vector subspace Y together with the norm defined for elements of Y as if they are in X. It follows easily that Y itself becomes a NLS and a metric subspace as well as a topological subspace.

Remark 1.3.2 We shall meet lots of subspaces. There is a slightly weaker notion of a subspace which is quite useful. At this stage we shall illustrate this with an example rather than making a definition. Suppose $1 \le p < p' \le \infty$. Let X be a measure space with finite volume. Then we would like to think of $L^{p'}(X)$ as a subspace of $L^p(X)$. For this to hold, we should, first of all identify $L^{p'}$ as a vector-subspace of L^p.

For $p' = \infty$, note that any essentially bounded function is p-integrable, because we have assumed $\mu(X) < \infty$. Thus, $L^\infty(X) \subset L^p(X)$. Moreover, a sequence $\{f_n\}$ converges to 0 in the sup norm implies that it converges to 0 in p-norm. This implies that the inclusion map is continuous as well.

However, there can be sequences $\{f_n\}$ in L^∞ which converge to 0 in L^p but not in L^∞. For instance, take $X = [0,1]$ witn the standard Lebesgue measure, $f_n = 1$ for $t \le 1/n$ and 0 otherwise. It is in this broader sense that we can now treat $L^\infty(X)$ as a subspace of $L^p(X)$, though the norms are different and even the topologies different.

The same comments hold for the case $p' < \infty$, viz., $\|f\|_p \le \|f\|_{p'}$. For this, we have to use the following formula, which is an easy consequence of Hölder's inequality, applied to $|f|^p$ is place of f and constant function 1 in place of g and the conjugate exponents being p'/p and q such that $p/p' + 1/q = 1$:

$$\|f\|_p^p = \int_X |f|^p \, d\mu \leq \left(\int_X (|f|^p)^{p'/p} \, d\mu \right)^{p/p'} \left(\int_X (1)^q d\mu \right)^{1/q} = (\|f\|_{p'})^p.$$

For another concept of subspace, wait till theorem 1.4.7.

Defining quotient spaces is a little more involved.

Definition 1.3.3 Let X be NLS and Y be a closed subspace. On the quotient vector space X/Y, we define 'the induced norm' as follows: for $x \in X$, let

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

Verification that $|||\ \ |||$ defines a norm on X/Y is straight forward. We call this quotient space of X by Y.

The beauty of this construction and the justification for the name comes from the fact that the underlying topological space of X/Y with this norm is equal to the quotient topology coming from the underlying topological space of X. In fact:

Theorem 1.3.4 *The function $\eta : X \to X/Y$ viz., $\eta(x) = x + Y$ is a continuous, open surjection. A sequence $\{x_n + Y\}$ in X/Y converges in X/Y iff there is a sequence $y_n \in Y$ such that $x_n + y_n$ is convergent in X.*

Proof: Suppose $z_n \to z \in X$. Then

$$|||(z_n + Y) - (z + Y)||| = |||(z_n - z) + Y||| \leq \|z_n - z\|$$

shows that $z_n + Y \to z + Y$. This proves the 'if' part of the second claim as well as the continuity of η. Conversely suppose $x_n + Y \to x + Y$. Then there are integers $n_1 < n_2 < \cdots$ such that for all $n > n_k$, $|||(x_n + Y) - (x + Y)||| < 1/k$. This in turn implies that for $n_k \leq n \leq n_{k+1}$, there exist $y_n \in Y$ such that

$$\|x_n - x + y_n\| < 1/k.$$

This then implies that $(x_n + y_n) \to x$.

Since η is clearly surjective, it remains to show that η is open. Let $U \subset X$ be open and $x + Y \in \eta(U)$ with $x \in U$. There exists $r > 0$ such that the open ball $B_r(x) \subset U$. We claim that the open ball $B_r(x + Y) \subset \eta(U)$. So, let $|||(z + Y) - (x + Y)||| < r$. This means that there exist $y \in Y$ such that $\|z - x + y\| < r$ which implies that $z + y \in B_r(x) \subset U$. But then $z + Y = \eta(z + y) \in \eta(U)$. ♠

Remark 1.3.5 This result will come handy later to you in proving open mapping theorem.

Theorem 1.3.6 *Let $1 \leq p \leq \infty$. Let $(X_i, \|\ \|_{(i)}), i = 1, 2, \ldots, k$ be NLSs. Then on the product vector space $X = X_1 \times X_2 \times \cdots \times X_k$, define:*

$$\|x\|_p = \begin{cases} \left(\sum_{i=1}^k \|x_i\|_{(i)}^p \right)^{1/p}; & p < \infty; \\ max \, \{\|x_i\|_{(i)} \;:\; 1 \leq i \leq k\}, & p = \infty. \end{cases}$$

Then $(X, \| \|_p)$ *is a NLS. Moreover, a sequence* $\{x^{(n)}\}$ *in* X *converges iff each of the coordinate sequences* $\{x_i^{(n)}$ *converges in* X_i*. Moreover, the projection maps* $P_i : X \to X_i$ *can be identified with the quotient maps by the subspaces* $Y_i = \{x \ : \ x_i = 0\}$*.*

Proof: Exercise.

Exercise 1.3.7

1. Let $K \subset X$ be compact and $F \subset X$ be a closed subset disjoint from K. Then there exists $r > 0$ such that $(K + B_r(0)) \cap F = \emptyset$.

2. Let $\| \|_{(i)}, i = 1, \ldots, k$ be norms of X and $r_1, \ldots, r_k$ be positive reals. Then for any $1 \leq p \leq \infty$ if we define

$$\|x\| = \begin{cases} \left(\sum_{i=1}^k r_i \|x\|_{(i)}^p\right)^{1/p}; & p < \infty; \\ \text{Sup } \{r_i \|x\|_{(i)} \ : \ i = 1, \ldots, k\}, & p = \infty \end{cases}$$

   Then $\| \|$ is a norm on X. Moreover, a sequence $x_n \in X$ converges to x wrt $\| \|$ iff it is so wrt each $\| \|_{(i)}$.

3. Let $0 < p < 1$ and $n \geq 2$. For $x \in \mathbb{R}^n$, define $\|x\| := (\sum_{i=1}^n |x_i|^p)^{1/p}$. Then show that $\| \|$ is not a norm.

1.4 Linear Functions: continuity

Theorem 1.4.1 *Let* $f : X \to Y$ *be a linear function. Then the following are equivalent:*
(i) f *is continuous on* X*.*
(ii) f *is continuous at* $0 \in X$*.*
(iii) f *is bounded on some closed ball of positive radius around* 0*.*
(iv) f *is uniformly continuous on* X*.*
(v) *There exist* $\alpha > 0$ *such that* $\|f(x)\| \leq \alpha \|x\|$*, for all* $x \in X$*. (Hint: Take* $\alpha = \epsilon/\delta$*.)*

Corollary 1.4.2 A linear functional $f : X \to \mathbb{K}$ is continuous iff the hyperspace $Z(f) = \{x \in X \ : \ f(x) = 0\}$ is a closed set in X.

Proof: If $f \equiv 0$ there is nothing to prove. Otherwise choose $x \in X$ such that $f(x) = 1$. If f is not continuous at 0 then there is a sequence $x_n \in X$ such that $x_n \to 0$ and $|f(x_n)| > 1$. Consider the sequence $z_n = x - x_n/f(x_n)$ which is inside $Z(f)$ but its limit is x which is outside $Z(f)$. That contradicts the closedness of $Z(f)$.

The converse is easy, since $\{0\}$ is closed in $\mathbb{K}$. ♠

Theorem 1.4.3 *A subspace* $Y \subset X$ *is equal to* X *iff its topological interior* Y° *in* X *is non empty. In particular, a hypersurface* Y *is closed iff it is nowhere dense.*

Proof: If $Y = X$ then $Y^\circ = X$ and hence non empty. On the other hand, if $Y^\circ \neq \emptyset$, then there exists an open ball $B_r(a) \subset Y$. Given any $x \in X$ check that $a + \frac{r}{2\|x\|}x \in B_r(a) \subset Y$. That implies $x \in Y$.

The second part follows from the fact that a subset is nowhere dense iff its closure has empty interior. ♠

Definition 1.4.4 Following this important theorem, we shall introduce the notation $B(X, Y)$ for the set of all continuous linear maps from X to Y. It follows easily that $B(X, Y)$ is a vector subspace of all functions from $X \to Y$.

For $f \in B(X, Y)$, it is not difficult to see that

$$\inf \{\alpha \ : \ \|f(x)\| \leq \alpha \|x\|, \text{ for all } x \in X\} = \text{Sup } \{\|f(x)\| \ : \ \|x\| = 1\}.$$

We denote this common value by $\|f\|$. It turns out that this makes $B(X, Y)$ into a NLS (exercise). You shall study this space in more detail later.

An important special case is when $Y = \mathbb{K}$ and then we use the notation

$$X' := B(X, \mathbb{K}).$$

Pay attention to the notation which is not to be confused with X^* which denotes the space of all linear functionals on X. These two spaces co-incide with X is finite dimensional.

Definition 1.4.5 A linear function $f : X \to Y$ of normed linear spaces is said to be norm preserving , if $\|f(x)\| = \|x\|$ for all $x \in X$. If further, it is onto also then we call it an isometry.

Remark 1.4.6 Note that a norm preserving linear function is automatically continuous and injective. If it is an isometry, then its inverse exists, and is automatically linear and norm preserving. Therefore, it is a homeomorphism.

Theorem 1.4.7 *Let $f : X \to Y$ be a linear function. Then f is a homeomorphism onto the subspace $f(X) \subset Y$, iff there exist positive real numbers α, β such that*

$$\alpha \|x\| \leq \|f(x)\| \leq \beta \|x\|, x \in X. \tag{1.10}$$

Proof: The first inequality implies that f is injective. Therefore f is a bijection and its inverse is a linear map. The second inequality now tells us that f is continuous and first one implies that f^{-1} is continuous.

Conversely, Suppose $f : X \to Y$ is a homeomorphism. Then clearly f is onto. We take $\beta = \|f\|$ and $\alpha = \|f^{-1}\|$ and verify the inequality. ♠

Remark 1.4.8 This result leads us to another concept of a subspace that we were waiting for. Under the situation of the above theorem, X can be effectively identified with the subspace $f(X)$ of Y. Even though the norms on X and that on $f(X)$ may differ, the two topologies are the same: a sequence $\{x_n\}$ in X converges iff the sequence $\{f(x_n)\}$ converges in $f(X)$.

Example 1.4.9 Two NLSs are said to be equivalent if there is a linear homeomorphism between them. All ℓ_p^n for $1 \leq p \leq \infty$ are equivalent. This is best illustrated by a picture of the unit balls in $\mathbb{R}^2$ with respect to various norms. However, on infinite dimensional spaces, these norms are not equivalent.

Theorem 1.4.10 *Let $f : \ell_2^n \to X$ be an injective linear map into a NLS X. Then $f : \ell_2^n \to f(\ell_2^n) = Y$ is a linear homeomorphism onto a closed subspace Y.*

Proof: The closedness of Y follows from the completeness of Y, once we establish that f is homeomorphism.

Now put $y_i = f(e_i), i = 1, \ldots, n$. Then $\{y_i\}$ form a basis for Y and hence every element $y \in Y$ has a unique linear combination

$$t_1 y_1 + t_2 y_2 + \cdots + t_n y_n.$$

Also, we have $f(t_1, \ldots, t_n) = t_1 y_1 + \ldots + t_n y_n$. By continuity of the vector space operations, it follows that f is continuous. We need to show the continuity of f^{-1}. This will follow if we show that each $t_i : Y \to \mathbb{K}$ is continuous. This in turn follows if we show $Z(t_i)$ is a closed subspace of Y.

This is done by induction on n.

For $n = 1$, we can take $\alpha = \beta = \|f(e_1)\|$ and verify the inequality (1.10). Suppose the result holds for $n - 1$. This in particular implies that every subspace of X of dimension $n - 1$ is closed and hence in particular, all $Z(t_i)$ are closed. ♠

We now give another useful result which is due to Riesz, especially while dealing with NLS which are not necessarily inner product spaces.

Theorem 1.4.11 Riesz: *Let X be a NLS and Y be a proper closed subspace. Let $0 \leq r < 1$. Then there exists $x = x_r \in X$ such that $\|x\| = 1$ and $r \leq d(x, Y) \leq 1$.*

Proof: Start with any $x' \in X \setminus Y$ and put $d = d(x', Y) > 0$. If $r = 0$ there is nothing to prove. Otherwise, since $d/r > d$, it follows that there exists $y_0 \in Y$ such that $\|x' - y_0\| < d/r$. Put $x = (x' - y_0)/\|x' - y_0\|$ so that $\|x| = 1$. Now for any $y \in Y$,
$\|y - x\| = \frac{\|(\|x'-y_0\|y+y_0)-x'\|}{\|x'-y_0\|} \geq \frac{d}{\|x'-y_0\|} \geq r.$ ♠

Corollary 1.4.12 A NLS is finite dimensional iff any and hence every closed ball of positive radius in X is compact.

Proof: The 'only if' part is the so called Heine- Borel theorem in $\mathbb{K}^n$. To see the 'if' part, assume that $\overline{B_r(x)}$ is compact for some $x \in X$ and hence $\overline{B_1(0)}$ is compact. Cover it with a finitely many open balls $B_{1/2}(x_i), i = 1, 2, \ldots, k$ say. We claim that Span $\{x_1, \ldots, x_k\}$ is the whole of X. If Y is this span, being finite dimensional, it is a closed subspace of X. If $Y \neq X$, then by the above theorem, there exists $x \in X \setminus Y$ such that $\|x\| = 1$ and $d(x, Y) \geq 1/2$. This means that $x \notin B_{1/2}(x_i)$ for any i but $x \in \overline{B_1(0)}$, which is absurd. ♠

Remark 1.4.13 You may wonder the role of putting the restriction $r < 1$ in the above theorem. Of course, in the proof, we have immediately used this condition. That does not automatically mean that this condition is necessary. The problem (1.4.14.7) illustrates the necessity of this condition, which may initially go against your intuition.

Exercise 1.4.14

1. Let X be any infinite dimensional NLS. Show that there exists
   (a) a linear one-to-one function $F : X \to X$ which is not continuous.
   (b) a linear functional $f : X \to \mathbb{K}$ which is not continuous.
   (c) a subspace $Y \subset X$ which is not closed.

2. Let X be finite dimensional. Show that every linear functional $f : X \to \mathbb{K}$ is continuous.

3. Let $F : X \to Y$ be a surjective linear map where Y is finite dimensional. Then F is an open mapping. Also F is continuous iff $Z(F)$ is closed.

4. Let Y_1, Y_2 be subspaces of X. If Y_1 is closed and Y_2 is finite dimensional, then $Y_1 + Y_2$ is closed.

5. Given $a < b \in \mathbb{R}$ show that there is a family of smooth functions $f_{a,b} : \mathbb{R} \to [0, 1]$ satisfing the following
   (i) $f'(t) \geq 0, t \in \mathbb{R}$.
   (ii) $f(t) = 0, t \leq a$.
   (iii) $f(t) = 1, t \geq b$.

6. Show that given $0 \leq r < 1$ there exists a continuous (smooth) function $h : \mathbb{R} \to [0, 1]$ such that $h(0) = 0$ $\int_0^1 h(t)\, dt = r$ and Sup $\{|h(t)\| t \in \mathbb{R}\} < 1$.

7. Consider $X = \{f \in C[0, 1] : f(0) = 0\}$ and $Y = \{f \in X : \int_0^1 f(t)dt = 0\}$, where $C[0, 1]$ is taken with the sup norm. Then Y is a closed subspace of X. However, there are no elements $g \in X$ such that $\|g\| = 1$ and $d(g, Y) = 1$. (Note: In a Hilbert space such elements would exist! (Compare this with theorem 1.4.11.)
   Hint: Use the previous two exercises.

8. (**Hilbert Cube**) Let $1 \leq p < \infty$. Let

$$E = \{x \in \ell_p \ : \ |x(j)|^p \leq 1/j^2, \text{ for all } j \geq 1\}.$$

   Then E is a compact convex subset of ℓ_p and E is not contained in any finite dimensional subspace of ℓ_p. (For $p = 2$, E is called the Hilbert cube. However, note that the corresponding set in ℓ_∞ viz.,

$$\{x \in \ell_\infty \ : \ |x(j)| \leq 1, j \geq 1\}$$

   is not compact.

Chapter 2

Hahn-Banach Theorems

2.1 A Separation Theorem

Theorem 2.1.1 *Let X be a vector space over $\mathbb{C}$ and $X_{\mathbb{R}}$ denote the underlying real vector space.*
(A) *If u is the real part of a complex linear functional on X then $u : X_{\mathbb{R}} \to \mathbb{R}$ is real linear and*

$$f(x) = u(x) - \imath u(\imath x), \text{ for all } x \in X. \tag{2.1}$$

(B) *If $u : X_{\mathbb{R}} \to \mathbb{R}$ is linear then (2.1) defines f as a complex linear functional on X.*
(C) *If X is a NLS and f, u are related as in (2.1), then $\|f\| = \|u\|$.*

Proof: (A) (Note that the real part u of a complex linear functional is a real linear functional.) Since for any complex number z, we have

$$z = \Re(z) - \imath\Re(\imath z)$$

(where $\Re(z)$ denotes the real part of z), we have, in (2.1)

$$LHS = \Re(f(x)) - \imath\Re(\imath f(x)) = u(x) - \imath\Re(f(\imath x) = RHS.$$

(B) It is clear that for a real linear functional u, f defined by (2.1) is also real linear. Moreover,

$$f(\imath x) = u(\imath x) - \imath u(-x) = \imath(u(x) - \imath u(\imath x)) = \imath f(x)$$

which implies that f is complex linear.
(C) Clearly

$$\|f\| = \text{ Sup } \{|f(x)| \ : \ \|x\| \leq 1\} \geq \text{Sup } \{|\Re(f(x))| \ : \ \|x\| \leq 1\} = \|u\|.$$

On the other hand, for any x we can choose $\lambda \in \mathbb{S}^1$ such that $f(\lambda x) = \lambda f(x) = |f(x)|$. It then follows that $u(\lambda x) = f(\lambda x)$. Also if $\|x\| \leq 1$ then $\|\lambda x\| \leq 1$. It follows that $\|f\| \leq \|u\|$ as well. ♠

Lemma 2.1.2 Let X be NLS over $\mathbb{R}$ and $E \neq \emptyset$ be a convex subset of X. Let Y be a subspace of X such that $Y \cap E = \emptyset$. If Y is not a hyperspace then there exists $x \in X \setminus Y$ such that Span $\{Y, x\} \cap E = \emptyset$.

Proof: Consider

$$S = Y + \cup\{rE \ : \ r > 0\}.$$

Since E is open it follows that S is open. We claim $S \cap -S = \emptyset$. For, if $x \in S \cap -S$, then

$$x = y_1 + r_1 e_1 = -(y_2 + r_2 e_2),$$

for some $y_i \in Y$, $e_i \in E$ and $r_i > 0$. This means that

$$-\frac{y_1 + y_2}{r_1 + r_2} = \frac{r_1}{r_1 + r_2} e_1 + \frac{r_2}{r_1 + r_2} e_2$$

which, in turn implies that $Y \cap E \neq \emptyset$ a contradiction.

Also, observe that $Y \cap S = \emptyset = Y \cap -S$. Suppose we show there exists $b \in X \setminus (Y \cup S \cup -S)$. Then this b will satisfy our requirement: For Span $\{Y, b\} \cap E \neq \emptyset$ implies there are elements $y \in Y, e \in E$ and $r \in \mathbb{R}$ such that $y + rb = e$. Clearly, $r \neq 0$. Therefore $b = e/r - y/r$. If $r > 0$ this implies $b \in S$ and if $r < 0$ this implies $b \in -S$ a contradiction in either case.

How do you prove $X \neq Y \cap S \cup -S$. This will differ from author to author. We use the following topological lemma.

Lemma 2.1.3 Let Y be a subspace of X of codimension at least 2. Then $X \setminus Y$ is path connected.

The proof of this lemma is left to the reader as an exercise. Granting that, we see that $S, -S$ are disjoint open subsets of $X \setminus Y$ which is path connected. Therefore, $X \setminus (Y \cup S \cup -S) = (X \setminus Y) \setminus (S \cup -S) \neq \emptyset$. ♠

Theorem 2.1.4 (Banach Separation Theorem) *Let X be NLS over $\mathbb{R}$ and $E \subset X$ be a non empty open convex subset. Given any subspace $Y \subset X$ such that $Y \cap E = \emptyset$, there exists a hyperspace H of X such that $H \cap E = \emptyset$ and $Y \subset H$. In other words, there is $f \in X'$ such that $f|_Y = 0$ and $f(x) \neq 0$ for every $x \in E$.*

Proof: Apply Zorn's lemma to the previous lemma. ♠

Theorem 2.1.5 Hahn-Banach Separation theorem: *Let X be a NLS and E_1, E_2 be two non empty, disjoint, convex subsets of X and let E_1 be open. Then there exist $f \in X'$ and $\alpha \in \mathbb{R}$ such that*

$$\Re f(x_1) < \alpha \leq \Re f(x_2), \quad \textit{for all } x_1 \in E_1, x_2 \in E_2. \tag{2.2}$$

Proof: In view of theorem 2.1.1, it is enough to consider X as a real vector space and find $u \in X'_{\mathbb{R}}$ satisfying the above inequality in place of $\Re f$. We can then take $f(x) = u(x) - \imath u(\imath x)$ and go home.

Since $E_1 \cap E_2 = \emptyset$, $0 \notin E_1 - E_2$. Clearly $E_1 - E_2$ is non empty, open and convex. Therefore by the previous theorem, there exists $f \in X'$ such that $0 \notin f(E_1 - E_2)$. This implies that $f(E_1) \cap f(E_2) = \emptyset$. Both are convex sets and hence are disjoint intervals. By multiplying f by -1, if necessary, we may assume that $f(E_1)$ lies to the left of $f(E_2)$. Finally, since $f(E_1)$ is open, we can take α to be the right-end point of the interval $f(E_1)$. Inequality (2.2) follows. ♠

Corollary 2.1.6 Let X be a NLS and E be a convex subset with a non empty interior. Given $b \notin E^\circ$, there exists $f \in X'$ such that for all $x \in E$ we have

$$\Re f(x) \leq \Re f(b).$$

Proof: Apply the theorem with $E_1 = E^\circ$ and $E_2 = \{b\}$ to obtain $f \in X'$ such that $\Re f(x) < \Re f(b)$ for all $x \in E^\circ$. The result follows by continuity of f. ♠

Remark 2.1.7 The hyperplane defined by $\{x \in X \ : \ \Re f(x) = \Re f(b)\}$ is called a supporting hyperplane for E. When we take $b \in \bar{E} \setminus E^\circ$, this is called a proper supporting hyperplane.

2.2 Extension Theorems

Theorem 2.2.1 *Let X be a NLS and Y be a subspace. For every $g \in Y'$ there exists $f \in X'$ such that $f|_Y = g$ and $\|f\| = \|g\|$.*

Proof: In view of theorem 2.1.1, it is enough to prove the statement for the case when $\mathbb{K} = \mathbb{R}$ and that is what we are going to do.

Again, if $g \equiv 0$ then we can take $f \equiv 0$ and go home. So assume $\|g\| \neq 0$ and by dividing out by $\|g\|$ we may assume $\|g\| = 1$. Let $a \in Y$ be such that $g(a) = 1$. Put $E = B(a, 1)$ the open ball of radius 1 and centre a.

If $y \in E \cap Y$ then

$$\|g(y) - 1\| = \|g(y) - g(a)\| = \|g(y - a)\| \leq \|g\|\|y - a\| < 1$$

and hence $g(y) \neq 0$. This implies that $E \cap Z(g) = \emptyset$. By the Banach separation theorem 2.1.4, there exits $f \in X'$ such that $f \equiv 0$ on $Z(g)$ and is never zero on E. By rescaling we may assume that $f(a) = 1$. But then $f(a) = g(a) = 1$ and $Z(g) \subset Z(f|_Y)$. Therefore $f|_Y = g$. Clearly $\|f\| \geq \|g\| = 1$. On the other hand, for any $x \notin Z(f)$ we have $a - x/f(x) \in Z(f)$. Therefore $a - x/f(x) \notin E$ which is the same as saying $\|a - x/f(x) - a\| \geq 1$, i.e., $\|x\| \geq \|f(x)\|$. Therefore $\|f\| \leq 1$ also. ♠

Corollary 2.2.2 Let X be a NLS and $0 \neq a \in X$. Then there exists $f \in X'$ such that $f(a) = \|a\|$ and $\|f\| = 1$.

Proof: Apply the above theorem with $Y = \text{Span}\ \{a\}$ and $g : Y \to \mathbb{K}$ given by $g(ka) = k\|a\|$.♠

Corollary 2.2.3 Let X be a NLS.
(A) Given a subspace $Y \subset X$ and $a \in X$, $a \in \bar{Y}$ iff for every $f \in X'$ such that $f(Y) = \{0\}$, we have, $f(a) = 0$.
(B) Given linearly independent elements $x_1, \ldots, x_n \in X$, there exist $x'_1, \ldots, x'_n \in X'$ such that $x'_i(x_j) = \delta_{ij}$.

Proof: Exercise.

Example 2.2.4 As an illustration of the application of the above corollary in approximation theory let us prove the following: Let $\{r_1, \ldots, r_n, \ldots\}$ be a countable dense subset of $[0,1]$ and let

$$x_j(t) = \begin{cases} 1, & \text{if } 0 \le t \le r_j \\ 0, & \text{otherwise.} \end{cases}$$

Then every $x \in L^2[0,1]$ can be approximated in the mean square arbitrarily closely by a linear combination of $\{x_1, x_2, \ldots\}$.

Note that proving this statement is the same as showing that the closure of $\text{Span}\{x_1, x_2 \ldots\} = Y$ is the whole of L^2. So, it is enough to show that for every linear functional f on L^2, which vanishes on all x_j, is identically zero. Now we know that there exist $y \in L^2[0,1] = X$ such that

$$f(x) = \int_0^1 xy\, dm, \quad x \in X$$

For every $s \in [0,1]$, put $z(s) = \int_0^s y\, dm$. Then $z \in C[0,1]$ and $z(r_j) = f(x_j) = 0$ for all j. This implies $z \equiv 0$. This in turn implies that $y = 0$ as an element of $L^2[0,1]$ which means $f \equiv 0$.

We shall now consider the uniqueness aspect of the extension. Even in the case of finite dimensional situation, there is no guarantee of uniqueness. Our first step toward this leads to a new concept:

Definition 2.2.5 A NLS is said to be strictly convex if $\|x + y\| < 2$ wherever, $\|x\| = \|y\| = 1$ and $x \neq y$.

Remark 2.2.6 It is an easy consequence of parallelogram law that every inner product space is strictly convex. (Later we shall see that an inner product satisfies even stronger convexity property, viz., uniform convexity.) Thus, ℓ_p^n is strictly convex for $1 < p < \infty$. On the other hand, for $p = 1, \infty$, this is not the case as illustrated by the picture of the unit ball in these spaces.

Theorem 2.2.7 Taylor-Foguel *Let X be a NLS. Then every $f \in Y'$ for every subspace $Y \subset X$ has a unique norm preserving extension to X iff X' is strictly convex.*

Proof: Assume that X' is SC. Let Y be a subspace of X and $f \in Y'$. Suppose $f_1, f_2 \in X'$ are such that $f_j|Y = f$ and $\|f_j\| = \|f\|$. On Y, we have $(f_1 + f_2)/2 = f$ and therefore

$$\|f\| = \frac{1}{2}(\|f_1\| + \|f_2\|) \ge \|(f_1 + f_2)/2\| \ge \|f\|.$$

By strict convexity of X', this implies $f_1 = f_2$.

Conversely, X' is not SC, i.e., there are $f_1, f_2 \in X'$ such that $f_1 \neq f_2$ and such that $1 = \|f_1\| = \|f_2\| = \|(f_1 + f_2)/2\|$. Put $Y = \{x \in X \;:\; f_1(x) = f_2(x)\}$. Clearly, Y is a closed subspace. We shall show that $\|f_j|_Y\| = 1$ for $j = 1, 2$. It then follows that X does not have unique extension property.

Clearly $\|f_j|_Y\| \le \|f_j\| \le 1$, it is enough to show that there is a sequence $y_n \in Y$ such that $\|y_n\| = 1$ and $f_1(y_n) \to 1$.

There exists $a \in X \setminus Y$ such that $f_1(a) = 1 \neq f_2(a)$. Since $\|f_1 + f_2\| = 2$, we can find a sequence x_n of unit norms in X such that $(f_1 + f_2)(x_n) \to 2$. Passing onto subsequences in two stages, we may assume that $\{f_1(x_n)\}$ and $\{f_2(x_n)\}$ converge to α, β respectively. It follows that $|\alpha| \leq 1$ and $|\beta| \leq 1$. But $|\alpha + \beta| = 2$. Therefore, it is necessary that $\alpha = \beta$ and $|\alpha| = 1$. Now by rescaling, we may assume that the sequence x_n of unit norms in X is such that $f_1(x_n) \to 1$ and $f_2(x_n) \to 1$.

Put $k_n = \frac{f_1(x_n) - f_2(x_n)}{1 - f_2(a)}$. Then $k_n \to 0$ and hence there exists n_0 such that $\|x_n - k_n a\| \neq 0$ for $n \geq n_0$. For these n, put $z_n = x_n - k_n a$. Then

$$f_1(z_n) = f_1(x_n) - k_n = \frac{f_2(x_n) - f_2(a) f_1(x_n)}{1 - f_2(a)} = f_2(x_n) - k_n f_2(a) = f_2(z_n).$$

Therefore $z_n \in Y$. Taking $y_n = z_n / \|z_n\|$, we get a sequence as required. ♠

2.3 Completeness of a Norm

Recall that a NLS X is a Banach space if the underlying metric space is complete. We shall now give a criterion for completeness which is quite useful.

Definition 2.3.1 Let $\{x_n\}$ be a sequence in a NLS X. Put $s_n = \sum_1^n x_i$. Then $\{s_n\}$ is called the sequence of partial sums of $\{x_m\}$. We say $\{x_m\}$ is summable or equivalently $\sum_1^\infty x_n$ is convergent if the sequence $\{s_n\}$ is convergent in X to say s and then we write $s = \sum_{i=1}^\infty x_i$. We say $\{x_n\}$ is absolutely summable (summable in norm) if $\sum_{i=1}^\infty \|x_i\| < \infty$. It is not at all clear why absolute summability should imply summability as is the case with sequences in $\mathbb{K}$.

Theorem 2.3.2 *A NLS is a Banach space iff every absolutely summable sequence is summable.*

Proof: Let X be a Banach space and $\sum_0^\infty \|x_i\| < \infty$. Then $\|s_{m+k} - s_m\| \leq \sum_{i=m+1}^{m+k} \|x_i\| < \infty$. This implies $\{s_m\}$ is a Cauchy sequence and hence is convergent. Therefore $\{x_i\}$ is summable.

Conversely, assume that X is NLS in which every absolutely summable sequence is summable. Let $\{s_m\}$ be a Cauchy sequence in X. In order to show that $\{s_n\}$ is convergent, it suffices to display a subsequence which is convergent. Let n_1 be such that for all $n \geq n_1$, we have $\|s_n - s_{n_1}\| \leq 1$. Inductively, having chosen n_k, let $n_{k+1} > n_k$ be such that for all $n \geq n_{k+1}$, $\|s_n - s_{n_{k+1}}\| \leq 1/2^k$. Put $x_k = s_{n_{k+1}} - s_{n_k}$. Then clearly $\sum_1^\infty \|x_k\| \leq \sum_k \frac{1}{2^k} < \infty$. By the hypothesis, this implies that the sequence $\{x_k\}$ is summable, i.e., $\sum_k x_k$ is convergent. This merely implies that the sequence $\{s_{n_k}\}$ is convergent as required. ♠

Theorem 2.3.3 *Let X be a NLS and Y be a closed subspace. Then Y with the induced norm and X/Y with the quotient norm are Banach spaces iff X is.*

Proof: Let X be a Banach space. Then being a closed subspace, Y is also complete. To Show X/Y is a Banach space, let $\{x_n + Y\}$ be an absolutely summable sequence in X/Y with respect to the quotient norm $\|| \ \ \||$. This means that $\sum_n \||x_n + Y\|| < \infty$. Now, for each n we can find $y_n \in Y$ such that $\|x_n + y_n\| \leq \||x_n + Y\|| + 1/2^n$. Therefore, the sequence $\{x_n + y_n\}$ is absolutely summable in X. Since X is assumed to be Banach, let $x = \sum_n (x_n + y_n)$. We claim $x + Y$ is the sum $\sum_n (x_n + Y)$. For

$$|||(\sum_1^m (x_n + Y)) - (x + Y)||| = |||(\sum_1^m (x_n + y_n)) - (x + Y)||| \leq \|(\sum_1^m (x_n + y_n) - x\| \to 0.$$

Conversely, assume that Y and X/Y are Banach and let $\{x_n\}$ be a Cauchy sequence in X. It follows easily that $\{x_n + Y\}$ is Cauchy in X/Y and hence is convergent to say $x + Y$. Choose a sequence (!) $y_n \in Y$ such that $x_n + y_n \to x$. But then

$$\|y_n - y_m\| = \|(y_n + x_n - x) - (x_n - x_m) - (x_m + y_m - x)\|$$

it follows that $\{y_n\}$ is Cauchy in Y and hence converges to say $y \in Y$. This means $x_n \to x - y \in X$. ♠

Theorem 2.3.4 *A finite product $X_1 \times \cdots \times X_k$ of NLSs is a Banach space iff each of X_i is.*

Proof: Exercise.

Theorem 2.3.5 *Let X, Y be NLSs. Then $B(X, Y)$ is a Banach space iff Y is.*

Proof: Let Y be a Banach space. Given a Cauchy sequence $\{f_n\}$ in $B(X, Y)$. Then for each $x \in X$ and for all n, m we have

$$\|f_n(x) - f_m(x)\| \leq \|f_n - f_m\| \|x\|$$

for all n, m. Given $\epsilon > 0$ choose n_0 such that for $n, m \geq n_0$, $\|f_n - f_m\| \leq \epsilon$. Then

$$\|f_n(x) - f_m(x)\| \leq \epsilon \|x\|, n, m \geq n_0. \tag{2.3}$$

In particular, we have shown that $\{f_n(x)\}$ is a Cauchy sequence in Y and hence we may define $f(x) = \lim_n f_n(x)$. It is easily checked that this $f : X \to Y$ is linear. Fixing $n = n_0$ and letting $m \to \infty$ in (2.3), we get

$$\|(f_{n_0} - f)(x)\| \leq \epsilon \|x\|. \tag{2.4}$$

This implies $f_{n_0} - f \in B(X, Y)$ and therefore, $f \in B(X, Y)$. Moreover, (2.4) implies that $f_n \to f$ in $B(X, Y)$.

Conversely, assume that $B(X, Y)$ is a Banach space. Given a Cauchy sequence $\{y_n\}$ in Y, we want to show that it is convergent. Choose any non zero $f \in X'$. (See Corollary 2.2.2.) Define $F_n(x) = f(x) y_n, x \in X$. Then $F_n \in B(X, Y)$ and

$$\|F_n(x) - F_m(x)\| = \|y_n - y_m\| \|f(x)\| \leq \|y_n - y_m\| \|f\| \|x\|$$

which shows $\{F_n\}$ is Cauchy. Therefore, $F_n \to F \in B(X, Y)$. In particular, $y_n = F_n(a) \to F(a) \in Y$. ♠

Exercise 2.3.6

1. Let Y be a subspace of a NLS X such that the codimension of Y in X is at least 2. Then show that $X \setminus Y$ is path connected.

2. Let Y be a subspace of X and $a \in X \setminus \bar{Y}$. Then there exists $f \in X'$ such that $\|f\| = 1,\ f|_Y \equiv 0$ and $f(a) = d(a, Y)$.

3. Let T be a set with at least two points, and $t_0 \in T$. Let $X = B(T, \mathbb{C})$ be the space of all complex valued bounded functions on T. Take $H = \{x \in X \ : \ x(t_0) \in \mathbb{R}\}$. Then H is a hypersurface of $X_{\mathbb{R}}$ and $H \cap \imath H = \{x \in X \ : \ x(t_0) = 0\}$.

4. Let K, F be two disjoint convex subsets of X with K compact and F closed. Then there exists $f \in X'$ and real numbers α, β such that
$$\Re f(x) < \alpha < \beta < \Re f(y), \quad x \in K,\ y \in F.$$

5. A NLS X is said to be *smooth* if for every $0 \neq x \in X$ there is $f \in X'$ such that $f(x) = \|x\|$ and $\|f\| = 1$. Show that if X' is strictly convex (resp. smooth) then X is smooth (resp. strictly convex).

6. Let $\mathbb{D}$ denote the open unit disc in $\mathbb{C}$. Let $X = \{f \in C(\bar{\mathbb{D}}) \ : \ f \text{ is analytic in } \mathbb{D}\}$. Define $\|f\| = \text{Sup } \{|f(z)| \ : \ |z| = 1\}$. Then X is a Banach space.

7. Given $f \in \ell_p$ for some $p < \infty$, show that $\|f\|_t \to \|f\|_\infty$ as $t \to \infty$.

8. For any $x \in \ell_p, y \in \ell_q, z \in \ell_r$, where $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 1$, show that
$$\sum_i |x_i y_i z_i| \leq \|x\|_p \|y\|_q \|z\|_r.$$

9. Let X be a Banach space and $F \in B(X, X)$ be such that $\|F\| < 1$. Then show that $I_X \pm F$ are invertible.

10. Let X and Y be Banach spaces, $F \in B(X, Y), G \in (Y, X)$ are such that $G \circ F = I_X$. Given any $T \in B(X, Y)$ such that $\|T - F\|\|G\| < 1$, show that there is $S \in B(Y, X)$ such that $ST = I_X$.

11. Let X be the infinite dimensional vector space of all polynomials $p(T) = a_0 + a_1 T + \ldots + a_n T^n$. Define
$$\|p\| = \text{Sup } \{|p(t)|; 0 \leq t \leq 1\}; \quad \|p\|_1 = \sum_0^n |a_i|.$$
Then $\|\ \|, \|\ \|_1$ are norms on X such that $\|p\| \leq \|p\|_1$. However, there are no constants $\alpha > 0$ such that $\|p\|_1 \leq \alpha \|p\|$ for all $p \in X$.

12. Let X be a real NLS and $f : X \to \mathbb{R}$ be a continuous function such that $f(x + y) = f(x) + f(y)$ for all $x, y \in X$. Show that f is linear.

Chapter 3

Hilbert Spaces

IPSs and Hilbert Spaces: definition and examples, polarization identify and parallelogram law, Schwarz inequality, strict and uniform convexity,), more examples, orthonormality and Pythagoras theorem, Gram-Schmidt, Examples, Bessel inequality and Riesz-Fisher theorem, Fourier expansion and Parseval identity, some examples, separability.

3.1 Definition and Examples

Definition 3.1.1 An inner product on a vector space X over $\mathbb{K}$ is a map $\langle -, -\rangle : X \times X \to \mathbb{K}$ satisfying the following axioms:
(i) $\langle x, x\rangle \geq 0,\ x \in X$ and equality holds iff $x = 0$.
(ii) $\langle x_1 + x_2, y\rangle = \langle x_1, y\rangle + \langle x_2, y\rangle$ and $\langle kx, y\rangle = k\langle x, y\rangle$.
(iii) $\langle y, x\rangle = \overline{\langle x, y\rangle}$.

Remark 3.1.2 (i) is referred to as positive definiteness; (ii) is called linearity in the first slot; (iii) Note that the 'bar' on the right hand side of rule (iii) denotes the complex conjugate when $\mathbb{K} = \mathbb{C}$ and the identity when $\mathbb{K} = \mathbb{R}$. Thus, it is called conjugate symmetry, or symmetry accordingly. Note that combing (ii) and (iii), you get conjugate linearity (resp. linearity) in the second slot. In some books (especially in some physics books, the slots are interchanged but this does not influence the the rest of the theory in any way.)

A vector space X together with an inner product is called an inner product space (IPS) or a pre-Hilbert space. As you may have anticipated, an inner product induces a norm on X as follows:

$$\|x\|^2 = \langle x, x\rangle.$$

Thus, an IPS is a NLS and therefore, a metric space as well. If this metric is complete then the IPS is called a Hilbert space. The norm given by an inner product satisfies the so called **Parallelogram Law:**

$$\|x + y\|^2 + \|x - y\|^2 = 2[\|x\|^2 + \|y\|^2$$

which can be verified easily.

A natural question that arises here is: can one recover the inner product from a norm? The key step toward an answer is the so called

Polarization Identity:

$$4\langle x, y\rangle = \|x+y\|^2 - \|x-y\|^2 \tag{3.1}$$

or

$$4\langle x, y\rangle = \|x+y\|^2 - \|x-y\|^2 + \imath\|x+\imath y\|^2 - \imath\|x-\imath y\|^2 \tag{3.2}$$

respectively if X is real or complex NLS. Here the norm on the right hand side is the one corresponding to the inner product on the left. Therefore, beginning with a norm, which should satisfy the parallelogram law above, if we want to get an inner product such that the corresponding norm is the given one, we must define the inner product as above. It turns out that this definition indeed gives us an inner product. The verification is left to the reader as an exercise and should skip reading the proof and go ahead. After giving a good trial, she may come back to this.

[Let us first consider the real case. Checking symmetry is easy. Putting $x = y$, we also see easily that $\langle x, x\rangle = \|x|^2$. In order to prove the linearity in the first slot, fix $y \in X$ and consider the function $f(x) = \langle x, y\rangle$ as given by (3.1). Then f is continuous and from exercise 2.3.6.12, it is enough to show that f is an additive homomorphism. That is $f(x+z) = f(x)+f(z)$. For this, we apply parallelogram law to $\|x+z+y\|^2$ in two different ways, viz., writing $x+z+y = (x+y)+z = x+(y+z)$, we obtain

$$\|x+z+y\|^2 = 2(\|x+y\|^2 + \|z\|^2) - \|x+y-z\|^2 = 2(\|x\|^2 + \|z+y\|^2) - \|x-y-z\|^2.$$

Likewise we also get

$$\|x+z-y\|^2 = 2(\|x-y\|^2 + \|z\|^2) - \|x-y-z\|^2 = 2(\|x\|^2 + \|z-y\|^2) - \|x+y-z\|^2.$$

Therefore

$$2(\|x+z+y\|^2 - \|x+z-y\|^2) = 2(\|x+y\|^2 - \|x-y\|^2 + 2(\|z+y\|^2 - \|z-y\|^2)$$

as required.

In the case of complex NLS, we consider the real and imaginary parts of RHS of (3.2) separately. What we have proved so far tells us that the real part defines a real inner product, which we may now denote by $\langle x, y\rangle_{\mathbb{R}}$. Consider the imaginary part:

$$\langle x, y\rangle_{\mathbb{C}} := \langle x, \imath y\rangle_{\mathbb{R}}.$$

Then $\langle x, x\rangle_{\mathbb{C}} = 0$ because $|1+\imath|^2 = |1-\imath|^2$, and

$$4\langle y, x\rangle_{\mathbb{C}} = 4\langle y, \imath x\rangle_{\mathbb{R}} = 4\langle \imath x, y\rangle_{\mathbb{R}} = \|\imath x + y\|^2 - \|\imath x - y\|^2 = \|x-\imath y\|^2 - \|x+\imath\|^2 = -4\langle x, y\rangle_{\mathbb{C}}$$

which proves that $\langle y, x\rangle_{\mathbb{C}} = -\langle x, y\rangle_{\mathbb{C}}$. However, the linearity in the first slot for $\langle x, y\rangle_{\mathbb{C}}$ is proved exactly similarly as in the case earlier case. It now follows that

$$\langle x, y\rangle = \langle x, y\rangle_{\mathbb{R}} + \imath\langle x, y\rangle_{\mathbb{C}}$$

is an inner product.]

3.2 Orthonormal Sets

Definition 3.2.1 Let X be an IPS. Two elements $x, y \in X$ are said to be orthogonal to each other if $\langle x, y\rangle = 0$. We write this by $x \perp y$. Similarly, if every element of a subset A is orthogonal to every element of another subset B we write $A \perp B$. A itself is called an orthogonal set if $a \neq b \in A$ implies $a \perp b$, i.e., any two distinct elements of A are orthogonal to each other. If further, each element of A is of norm 1 then A is called an orthonormal set.

The following properties are immediate:
(A) **Pythagoras:** If $\{x_1, \ldots, x_k\}$ is an orthogonal set then

$$\|x_1 + \cdots + x_n\|^2 = \|x_1\|^2 + \cdots + \|x_n\|^2.$$

(B) Every orthogonal subset of $X \setminus \{0\}$ is linearly independent.
(C) If $x \neq y$ belong to an orthonormal set, then $\|x - y\|^2 = 2$.

Theorem 3.2.2 (Gram-Schmidt) *Given a linearly independent set* $\{x_n : n = 1, 2, \ldots\}$, *define* $y_1 = x_1; u_1 = y_1/\|y_1\|$; *and for* $n \geq 2$, *define*

$$y_n = x_n - \sum_{i=1}^{n-1}\langle x_n, u_i\rangle u_i, \quad u_n = y_n/\|y_n\|.$$

Then $\{u_n : n = 1, 2, \ldots\}$ *is orthonormal and*

$$Span\ \{u_1, \ldots, u_n\} = Span\ \{x_1, x_2, \ldots, x_n\},\ for\ all\ n.$$

Remark 3.2.3 Indeed, if x_is are all continuous functions from some topological space into X, then so are the resulting u_is.

Example 3.2.4 We know that a polynomial $p(t)$ in one variable over $\mathbb{K}$ has only finitely many roots. This means that we can consider them as genuine elements of of $L^2[a, b]$ and then it follows that $\{1, t, t^2 \ldots,\}$ is a linearly independent set spanning the space of all polynomials. However, it is clear that this is not an orthogonal set. So, we can apply GS process to it and get an orthonormal set $\{l_0, l_1, \ldots, l_n \ldots\}$ of polynomials. For the special case when $[a, b] = [-1, 1]$ these l_i are called Legendre polynomials. A simple computation (exercise) will show that $l_0 = 1/\sqrt{2}, l_1 = \sqrt{3}t/\sqrt{2}, l_2 = \sqrt{10}(3t^2 - 1)/4$, and so on.

More generally, given a continuous positive function $\omega : [a, b] \to \mathbb{R}$, we can consider the space $L^2[a, b]_\omega$ of all measurable functions $x : [a, b] \to \mathbb{K}$ such that $\int_a^b |x|^2\omega\, d\mu < \infty$. If we define

$$\langle x, y\rangle_\omega = \int_a^b x\bar{y}\omega\, d\mu$$

then it is a routine to check that $L^2[a,b]_\omega$ becomes an inner product space. If our choice of ω is such that for all $n \in \mathbb{Z}^+$, $\int_a^b |t|^{2n}\omega\, d\mu < \infty$, then we can talk about orthonormal set of polynomials with respect to the weight function ω. Classically interesting cases occur when $[a,b] = [-1,1]$. The case of Legendre polynomials corresponds to the weight function 1. For $\omega(t) = 1/\sqrt{1-t^2}; \sqrt{1-t^2}$, they are respectively called Tchebychev polynomials of I kind and II kind. For $\omega(t) = e^{-t}, e^{-t^2}$, they are called Laguerre polynomials and Hermite polynomials. There are fat books written on them and they play some major role in various branches of mathematics.

Theorem 3.2.5 (Bessel's inequality) *: Let E be a finite or countable orthonormal set in an IPS X. Then for every $x \in X$ we have,*

$$\sum_{e\in E} |\langle x,\, e\rangle|^2 \leq \|x\|^2.$$

Moreover equality in the above expression holds iff $x \in$ Span E.

Proof: First consider the case when E is finite. Put $z = \sum_{e\in E}\langle x,\, e\rangle e$. Then

$$\langle x,\, z\rangle = \langle z,\, x\rangle = \langle z,\, z\rangle = \sum_{e\in E} |\langle x,\, e\rangle|^2.$$

Therefore,

$$\begin{aligned} 0 \leq \|x-z\|^2 &= \langle x-z,\, x-z\rangle \\ &= \langle x,\, x\rangle + \langle z,\, z\rangle - \langle x,\, z\rangle - \langle z,\, x\rangle \\ &= \|x\|^2 - \textstyle\sum_{e\in E} |\langle x,\, e\rangle|^2, \end{aligned}$$

which takes care of the statement when E is finite. Upon taking the limit we obtain the result when E is countably infinite also. The last assertion is obvious from the steps in the proof: equality holds iff $\|x-z\|^2 = 0$ iff $x = z$ iff $x \in$ Span E. ♠

Corollary 3.2.6 Let E be any orthonormal set in an IPS X. Then for any $x \in X$ the set $E_x = \{e \in E \;:\; \langle x,\, e\rangle \neq 0\}$ is a countable set. If $\{e_1, e_2, \ldots,\}$ is an enumeration of E_x then $\langle x,\, e_n\rangle \to 0$ as $n \to \infty$.

Proof: For every $m \geq 1$, let $E_m = \{e \in E \;:\; |\langle x,\, e\rangle|^2 \geq \|x\|^2/m.\}$ Then it follows easily from Bessel inequality that $\#(E_m) \leq m$. Now $E_x = \cup_m E_m$ is a countable union of finite sets and hence is countable. The last part follows since for any convergent series of real numbers, the n^{th} term must tend to zero. ♠

The next theorem tells the role of a countable infinite orthonormal set in a Hilbert space.

Theorem 3.2.7 Riesz-Fisher: *Let $\{e_n : n = 1, 2\ldots,\}$ be an orthonornal set in a Hilbert space H and let $\lambda_n \in K$ be a sequence of scalars. Then the following are equivalent:*
(A) *There exists $x \in H$ such that $\langle x,\, e_n\rangle = \lambda_n$, for all n.*
(B) $\sum_n |\lambda_n|^2 < \infty$.
(C) $\sum_1^\infty \lambda_n e_n$ *is convergent in H.*

Proof: (A) $\Longrightarrow$ (B): Use Bessel's inequality.
(B) $\Longrightarrow$ (C): The sequence of partial sums $s_n = \sum_1^n \lambda_m e_m$ is Cauchy, due to the fact that $\{e_n\}$ is orthonormal.
(C) $\Longrightarrow$ (A) Take $x = \sum_n \lambda_n e_n$. ♠

Definition 3.2.8 An orthonormal set in an IPS is called an orthonormal basis if it is not contained in any larger orthonormal set. A simple application of Zorn's lemma tells you that every orthonormal set is contained in an orthonormal basis.

Theorem 3.2.9 *Let $\{e_\alpha\}$ be an orthonormal set in a Hilbert space. Then the following are equivalent:*
(A) *$\{e_\alpha\}$ is an orthonormal basis for H.*
(B) **(Fourier)** *Given $x \in H$, the set $\{\alpha : \langle x, e_\alpha\rangle \neq 0\}$ is countable and $x = \sum_n \langle x, e_n\rangle e_n$.*
(C) **Parseval identity***: For $x \in H$,*

$$\|x\|^2 = \sum_1^\infty |\langle x, e_n\rangle|^2.$$

(D) *For $x \in H$, $\langle x, e_\alpha\rangle = 0$ for all α implies $x = 0$.*

Proof: (A) $\Longrightarrow$ (B) We have seen that E_x is countable for each x. By Bessel's inequality,

$$\sum_1^\infty |\langle x, e_n\rangle|^2 \leq \|x\|^2 < \infty.$$

By Riesz-Fisher, $\sum_i \langle x, e_i\rangle e_i$ makes sense as an element, say, $y \in H$. If $y \neq x$ then take $e = \frac{y-x}{\|y-x\|}$ and show that $\{e_\alpha\} \cup \{e\}$ is an orthonormal set, contradicting the maximality of $\{e_\alpha\}$.

Other implications are left as an exercise to the reader. ♠

3.3 Separability

Theorem 3.3.1 *Let X be a Hilbert space. Then the following are equivalent:*
(A) *H has a countable orthonormal basis.*
(B) *H is isometric to ℓ_2^n for some $0 \leq n \leq \infty$.*
(C) *H is separable.*

Proof: (A) $\Longrightarrow$ (B): We may assume that $H \neq (0)$ and let $\{e_n : n = 1, 2 \ldots,\}$ be an orthonormal basis for H. Define $F : H \to \ell_2^n$ by the formula

$$F(x)_n = \langle x, e_n\rangle,$$

i.e., the n^{th} entry of the sequence $F(x)$ is equal to $\langle x, e_n\rangle$. Clearly, F is linear. Appeal to Bessel's inequality to see that F is continuous. Use Parseval's formula to see that F is an isometry. Appeal to Riesz-Fisher to see that F is onto.
(B) $\Longrightarrow$ (C): Obvious.

(C) $\Longrightarrow$ (A): Start with a countable dense subset $\{a_1, \ldots, a_n, \ldots\}$. We construct a sequence $\{k_n\}$ finite or infinite as follows: Put $k_1 = 1$. Having chosen k_n let k_{n+1} be the first number such that $\{a_{k_1}, a_{k_2}, \ldots, a_{k_n}, a_{k_{n+1}}\}$ is an independent set, if it exists; otherwise, stop. Rename a_{k_n}as b_n. Check that $\{b_1, b_2, \ldots, \}$ is an independent set such that

$$\text{Span } \{b_1, b_2 \ldots, \} = \text{Span } \{a_1, a_2, \ldots\}.$$

Apply Gram-Schmidt to $\{b_1, b_2, \ldots\}$ to obtain an orthonormal set $\{e_1, e_2, \ldots\}$. To show that this is a basis, it is enough to show that $\langle x, e_n\rangle = 0$ for all n implies $x = 0$. But this condition implies $\langle x, b_n\rangle = 0$ for all n which in turn implies $\langle x,\, a_n\rangle = 0$ for all n. Since $\{a_1, a_2, \ldots, \}$ is dense there exists a subsequence a_{m_n} converging to x. By continuity of the inner product, we get,

$$\langle x,\, x\rangle = \langle x,\, \lim_n a_{m_n}\rangle = \lim_n \langle x,\, a_{m_n}\rangle = 0.$$

Therefore $x = 0$. ♠

Example 3.3.2 Weierstrass's approximation theorem says that given any continuous function $f : [a, b] \to \mathbb{K}$ and an $\epsilon > 0$, there exists a polynomial function $p(t)$ such that $\|f - p\|_\infty < \epsilon$. This in turn implies that the space of all polynomials is dense in $L^2[a, b]$. In particular, this means that the various set of special polynomials such as Legendre polynomials etc. that we have discussed in example 3.2.4 are all orthonormal basis for the respective Hilbert spaces.

Example 3.3.3 Another important class of orthonormal sets occurs when we are studying Fourier analysis of periodic functions on $\mathbb{R}$ (which are the same as functions on $\mathbb{S}^1$). Consider the space $L^2[-\pi, \pi]$, and the sequence of functions

$$u_n(t) = e^{\imath nt}/\sqrt{2\pi}, \quad n \in \mathbb{Z}$$

Check that this is an orthonormal set. A classical theorem in Fourier analysis asserts that this is indeed an orthonormal basis. This is equivalent to saying that given a measurable function x on $[-\pi, \pi]$, there are unique $\hat{x} \in \mathbb{K}$ (called Fourier coefficients of x) such that

$$x = \sum_{-\infty}^{\infty} \hat{x}_n u_n$$

as an element of $L^2[-\pi, \pi]$.

Exercise 3.3.4

1. Let $A \subset \mathbb{R}^+$ be such that for every countable (or finite) subset $B \subset A$, we have $\sum_{r \in B} r < \infty$. Then A is countable. (Compare Corollary 3.2.6.)

2. Let $X = C[a, b]$. For $f, g \in X$, let

$$\langle f,\, g\rangle_0 = \int_a^b \int_a^b f(s)\overline{g(t)}\, ds\, dt,$$

and

$$\langle f,\, g\rangle = \int_a^b \int_a^b \frac{\sin(s-t)}{s-t} f(s)\overline{g(t)}\, ds\, dt.$$

Then $\langle\ ,\ \rangle_0$ is not an inner product but $\langle\ ,\ \rangle$ is. Hint:

$$\langle f,\, g\rangle = \frac{1}{2}\int_{-1}^{1}\left(\int_a^b e^{\imath us} f(s)\, ds\right)\left(\int_a^b e^{-\imath ut} g(t)\, dt\right).$$

3. In any inner product space, we can introduce the notion of angle: For $x = 0$ or $y = 0$, let $\Theta_{x,y} = 0$. For $x \neq 0 \neq y$, let $\Theta_{x,y} = arc\cos\frac{\langle x,\, y\rangle}{\|x\|\|y\|}$. Show that $\Theta_{x,y}$ satisfying the cosine formula:

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2\|x\|\|y\|\cos\Theta_{x,y}.$$

4. Show that a NLS X is an inner product space iff every two-dimensional subspace of X is so.

5. Let X, Y be two inner product spaces. We say they are isomorphic if there is a linear map $f : X \to Y$ which preserves the inner product. Show that X is isomorphic to Y iff as NLSs they are isometric.

6. Let D be a domain in $\mathbb{C}$ and let $L^2(D)$ be the space of equivalence classes of square integrable complex valued functions on D with the inner product given by

$$\langle f,\, g\rangle = \int_D f(z)\overline{g(z)}\, dx\, dy.$$

Let $A^2(D)$ denote the subspace of analytic functions. Show that $L^2(D)$ is a Hilbert space and $A^2(D)$ is a closed subspace. In fact, if $f_n \to f$ in L^2 norm, then f_n converges to f uniformly on every closed disc in D. (Hint: If $B_r(z_0) \subset D$ then by mean value theorem for analytic functions we have

$$f(z_0) = \frac{1}{\pi r^2}\int_{|z-z_0|\le r} f(z)\, dx\, dy.$$

Now use Hölder's inequality to obtain $|f(z_0)| \le \|f\| r\sqrt{\pi}$.)