REAL ANALYSIS
LECTURE NOTES

Abstract. The Notes indicate what we do in the lectures, but are not a complete replacement of the book and lectures. The text is from two books of *Real Analysis*:
[1] Xingwei Zhou & Wenchang Sun: Real Variable Analysis, the third edition, Science Press, 2014.
[2] E. Stein & R. Shakarchi: Real Analysis, Princeton University Press, 2005.

L^p space (continued)

1. The dual space of L^p when $1 \leq p < \infty$

Suppose that $\mathcal{B}$ is a complete normed linear space (Banach space) over $\mathbb{R}$ equipped with a norm $\|\cdot\|$. A linear functional is a linear mapping $T : \mathcal{B} \to \mathbb{R}$ which satisfies

$$T(af + bg) = aT(f) + bT(g), \quad \forall\ a, b \in \mathbb{R}, \text{ and } f, g \in \mathcal{B}.$$

A linear functional T is continuous if given $\varepsilon > 0$ there is $\delta > 0$ so that

$$|T(f) - T(g)| \leq \varepsilon \quad \text{whenever } \|f - g\| \leq \delta.$$

We say that a linear functional is bounded if there is $M > 0$ with

$$|T(f)| \leq M\|f\|, \quad \forall\ f \in \mathcal{B}.$$

The linearity of T shows that these two notions are equivalent.

Theorem 1.1. *A linear functional on a Banach space is continuous if and only if it is bounded.*

Proof. Observe that T is continuous if and only if it is continuous at the origin. This is a direct consequence of the linearity.

Indeed if T is continuous, we choose $\varepsilon = 1$ and $g = 0$ so that $|T(f)| \leq 1$ whenever $\|f\| \leq \delta$ for some $\delta > 0$. Hence, given any non-zero $h \in \mathcal{B}$, we see that

$$|T(\delta h/\|h\|) \leq 1 \Longrightarrow |T(h)| \leq \frac{1}{\delta}\|h\|.$$

Conversely, if T is bounded it is clearly continuous at the origin, hence continuous.

□

The set of all continuous linear functionals over $\mathcal{B}$ is a vector space via the following addition and scalar multiplication:

$$(T_1 + T_2)(\cdot) := T_1(\cdot) + T_2(\cdot) \quad \text{and} \quad (aT)(\cdot) := aT(\cdot).$$

This vector space may be equipped with a norm as follows:

$$\|T\| := \sup_{\|f\| \le 1} |T(f)| = \sup_{\|f\|=1} |T(f)| = \sup_{f \ne 0} \frac{|T(f)|}{\|f\|}.$$

The vector space of all continuous linear functionals on $\mathcal{B}$ equipped with $\|\cdot\|$ above is called the dual space of $\mathcal{B}$, and is denoted by $\mathcal{B}^*$.

Theorem 1.2. *The vector space $\mathcal{B}^*$ is a Banach space.*

Proof. Suppose $\{T_k\}$ is a Cauchy sequence in $\mathcal{B}^*$. Then for each $f \in \mathcal{B}$,

$$|T_k(f) - T_j(f)| = |(T_k - T_l)(f)| \le \|T_k - T_l\| \|f\|.$$

This means that the sequence $\{T_k(f)\}$ is Cauchy in $\mathcal{B}^*$, hence converges to a limit, denoted by $T(f)$.

Step 1. We verify the above-defined $T \in \mathcal{B}^*$.

Clearly the mapping $T : f \mapsto T(f)$ is linear. If M is so that $\|T_k\| \le M$ for all k, then

$$|T(f)| \le |(T - T_k)(f)| + |T_k(f)| \le |(T - T_k)(f)| + M\|f\|.$$

Therefore, taking $k \to \infty$ we find

$$|T(f)| \le M\|f\| \quad \forall\, f \in \mathcal{B}.$$

Thus T is bounded.

Step 2. We show $T_k \to T$ in $\mathcal{B}^*$.

Fix any $\varepsilon > 0$. One has

$$|(T - T_k)(f)| \le |(T - T_j)(f)| + |(T_k - T_j)(f)| \le |(T - T_j)(f)| + \frac{\varepsilon}{2}\|f\|,$$

when k, j are large. Choose j (dependent on f) so that

$$|(T - T_j)(f)| \le \frac{\varepsilon}{2}\|f\|.$$

In the end, we find that for k large, independent of f,

$$|(T - T_k)(f)| \le \varepsilon\|f\|.$$

This proves that $\|T - T_k\| \to 0$, as desired.

□

Suppose that $p \in [1, \infty]$ and q is the conjugate exponent of p, that is

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

The key observation to make is the following: Hölder inequality shows that every function $g \in L^q(E)$ gives rise to a bounded linear functional on $L^p(E)$ by

$$T_g(f) = \int_E f(x)g(x)dx,$$

and that $\|T_g\| \leq \|g\|_q$. Therefore, $L^q(E) \ni g \mapsto T_g \in [L^p(E)]^*$ implies that $L^q(E) \subset (L^p)^*(E)$ when $p \in [1, \infty]$.

The main result is that when $1 \leq p < \infty$, every linear functional on $L^p(E)$ is of the form (1.1). We remark that this result is in general not true when $p = \infty$; the dual of $L^\infty(E)$ contains $L^1(E)$, i.e. $(L^\infty(E))^* \supset L^1(E)$, but it is larger.

Theorem 1.3. *Suppose $p, q \geq 1$ and $1/p + 1/q = 1$. Then*

$$(L^p(E))^* = L^q(E), \quad 1 \leq p < \infty,$$

in the following sense: for every bounded linear functional T on $L^p(E)$, there is a unique $g \in L^q(E)$ so that $T = T_g$, which is given by

$$T_g(f) = \int_E f(x)g(x)dx, \tag{1.1}$$

Moreover $\|T\| = \|g\|_q$.

Lemma 1.1. *Suppose $1 \leq p, q \leq \infty$ are conjugate exponents.*

(i) *If $g \in L^q(E)$, then*

$$\|g\|_q = \sup\Big\{\Big|\int_E fg\Big| : \|f\|_p \leq 1\Big\}.$$

(ii) *Suppose g is integrable on all sets of finite measure, and*

$$\sup\Big\{\Big|\int_E fg\Big| : \|f\|_p \leq 1,\ f \text{ is simple}\Big\} = M < \infty.$$

Then $g \in L^q(E)$ and $\|g\|_q = M$.

Proof. We start with (i). Assume without loss of generality $g \neq 0$. By Hölder inequality

$$\sup\Big\{\Big|\int_E fg\Big| : \|f\|_p \leq 1\Big\} \leq \|g\|_q.$$

For the reverse inequality we consider the following cases.

(C1) $q = 1$ and $p = \infty$. Take $f(x) = \text{sign } g(x)$. Then $\|f\|_\infty = 1$ and

$$\Big|\int_E fg\Big| = \|g\|_1.$$

(C2) $1 < p, q < \infty$. Take $f(x) = |g(x)|^{q-1}\text{sign } g(x)/\|g\|_q^{q-1}$. Since $p(q-1) = q$,

$$\|f\|_p^p = \int_E |g(x)|^{p(q-1)}/\|g\|_q^{p(q-1)} = 1 \quad \text{and} \quad \Big|\int_E fg\Big| = \|g\|_q.$$

(C3) $q = \infty$ and $p = 1$. Fix any $\varepsilon > 0$.

Let E_ε be a set of finite positive measure such that

$$|g(x)| \geq \|g\|_\infty - \varepsilon \quad \text{on } E_\varepsilon.$$

Take $f(x) = \chi_{E_\varepsilon}(x)\text{sign } g(x)/m(E_\varepsilon)$. Then $\|f\|_1 = 1$ and

$$\Big|\int_E fg\Big| = ⨍_{E_\varepsilon} |g| \geq \|g\|_\infty - \varepsilon.$$

This finishes the proof of part (i).

Next we show (ii). The direction

$$M \leq \|g\|_q$$

is implied by Hölder inequality.

(C1) $p > 1$. Take a sequence of simple functions g_k so that $|g_k(x)| \leq |g(x)|$ and $g_k(x) \to g(x)$ for every x. Let $f_k(x) = |g_k(x)|^{q-1}\text{sign } g(x)/\|g_k\|_q^{q-1}$. Then $\|f_k\|_p = 1$. So

$$M \geq \Big|\int_E f_k g\Big| \geq \frac{1}{\|g_k\|_q^{q-1}} \int_E |g_k|^q = \|g_k\|_q.$$

By Fatou's Lemma, it follows that

$$\|g\|_q \leq \Big(\liminf_{k\to\infty} \int_E |g_k|^q\Big)^{\frac{1}{q}} \leq M.$$

(C2) $p = 1$. Take $f_k(x) = (\text{sign } g(x))\chi_{E_k}(x)/m(E_k)$, where E_k is an increasing sequence of sets of finite measure whose union is E. Then $\|f_k\|_1 = 1$ and

$$M \geq \lim_{k\to\infty} \Big|\int_E f_k g\Big| \geq \|g\|_\infty - \varepsilon.$$

Since ε is arbitrary, we obtain $M \geq \|g\|_\infty$.

This proves part (ii).

$\square$

We now state the the following celebrated theorem.

Theorem (Lebesgue-Radon-Nikodym Theorem). *Suppose μ is a σ-finite positive measure on the measure space $(X, \mathcal{M})$ and ν a σ-finite signed measure on $\mathcal{M}$. Then there exist unique signed measure ν_a and ν_s on $\mathcal{M}$ such that*

$$\nu = \nu_a + \nu_s, \quad \textit{with } \nu_s \perp \mu, \ \nu_a \ll \mu.$$

In addition, then measure ν_a takes the form $d\mu_a = f d\mu$, that is

$$\nu_a(E) = \int_E f(x) d\mu(x)$$

for some extended μ-integrable function f.

A proof can be found in *Real Analysis* by Stein-Shakarchi, see Theorem 4.3 in Chapter 6 of the book. The argument therein is due to von Neumann, which exploits a linear transform $T : L^2(X, \rho) \to L^2(X, \rho)$, where $\rho = \nu + \mu$ and

$$T(\psi) = \int_X \psi(x) d\nu(x),$$

with an elegent use of the Riesz representation theorem (see Theorem 2.4 below). The interested readers are referred to Stein-Shakarchi's book for the details.

Proof of Theorem 1.3.

Firstly, let us suppose $m(E) < \infty$.

Given a functional T on $L^p(E)$, we define a set function ν by

$$\nu(G) = T(\chi_G), \quad \text{where } G \text{ is measurable set of } E.$$

The definition makes sense because $\chi_G \in L^p(E)$ since $m(E) < \infty$. Observe that

$$|\nu(G)| \leq \|T\| \|\chi_G\|_p = \|T\| [m(G)]^{\frac{1}{p}}. \tag{1.2}$$

Now the linearity of T implies that ν is finitely-additive. If $\{G_k\}$ is a countable collection of disjoint measurable sets, and we set $G = \cup_{k \geq 1} G_k$, $G_N^* = \cup_{k \geq N+1} G_k$, then

$$\chi_G = \chi_{G_N^*} + \sum_{k=1}^N \chi_{G_k}.$$

Therefore

$$\nu(G) = \nu(E_N^*) + \sum_{k=1}^N \nu(G_k).$$

However $\nu(E_N^*) \to 0$ as $N \to \infty$, because of the continuity of T, see (1.2). This shows that ν is countably additive, and absolutely continuous w.r.t. the Lebesgue measure.

By the Lebesgue-Radon-Nikodym Theorem, there is $g \in L^1(E)$ so that

$$\nu(G) = \int_G g dx \quad \text{for every measurable set } G \text{ of } E.$$

Therefore

$$T(\chi_G) = \int_E \chi_G g dx$$

By the linear extension, the above formula holds for all simple functions. Employing part (ii) of Lemma 1.1, $g \in L^q(E)$. Since simple functions are dense in L^p, $1 \le p < \infty$, each $f \in L^p(E)$ can be written as $f = \sum_{k=1}^{\infty} a_k \chi_{G_k}$. By the continuity of T, we see that

$$T(f) = \lim_{N\to\infty} \int_E (S_N f) g dx, \quad \text{where } S_N f = \sum_{k=1}^{N} a_k \chi_{G_k}.$$

By the Hölder inequality, $|fg| \in L^1(E)$. Since $|(S_N f)g| \le |fg|$, the dominated convergence theorem yields

$$T(f) = \int_E f g dx, \quad \forall\ f \in L^p(E). \tag{1.3}$$

By Lemma 1.1, we see that $\|g\|_q = \|T\|$.

We next deal with the case $m(E) = \infty$.

Take an increasing sequence $\{E_k\}$ of sets of finite measure that exhaust E. By the previous argument, for each k there is $g_k \in L^1(E_k)$ and $g_k = 0$ outside E_k, so that

$$T(f) = \int_E f g_k dx, \quad \text{whenever } f \text{ is supported in } E_k \text{ and } f \in L^p(E_k).$$

By Lemma 1.1, $\|g_k\|_q \le \|T\|$.

Because of (1.3), $g_l = g_k$ a.e. on E_k, whenever $l \ge k$. Therefore $\lim_{k\to\infty} g_k(x) = g(x)$ exists for a.e. x, and by Fatou's Lemma,

$$\|g\|_q \le \Big(\liminf_{k\to\infty} \int_E |g_k|^q \Big)^{\frac{1}{q}} \le \|T\|.$$

As a result,

$$T(f) = \int_E f g dx, \ \ \forall f \in L^p(E_k), \ \operatorname{supp} f \subset E_k. \tag{1.4}$$

Each $f \in L^p(E)$ can be written as $f = \lim_{N\to\infty} f\chi_{E_N}$. A simple limiting argument shows that (1.4) holds for all $f \in L^p(E)$. By Lemma 1.1, $\|T\| = \|g\|_q$, and so the theorem is proved.

$\square$

2. The Hilbert space L^2

Motivated by $\mathbb{R}^n$ and $\mathbb{C}^n$, an inner product space is defined to be a vector space V over either $\mathbb{R}$ or $\mathbb{C}$ (not necessarily finite dimensional), together with an inner product $\langle\cdot,\cdot\rangle : V\times V \to \mathbb{R}$ ($\mathbb{C}$) which is required to be:

(1) Conjugate Symmetric: for all $u, v \in V$,

$$\langle u, v\rangle = \overline{\langle v, u\rangle}.$$

In the real case, this is just usual the usual symmetry condition $\langle u, v\rangle = \langle v, u\rangle$.

(2) Linear in the first argument: for all scalars α and all $u, u_1, u_2, v \in V$,

$$\langle \alpha u, v\rangle = \alpha\langle u, v\rangle, \quad \langle u_1 + u_2, v\rangle = \langle u_1, v\rangle + \langle u_2, v\rangle.$$

By conjugate symmetry, it follows that for all scalars α and all $u, v_1, v_2, v \in V$

$$\langle u, \alpha v\rangle = \overline{\alpha}\langle u, v\rangle, \quad \langle u, v_1 + v_2\rangle = \langle u, v_1\rangle + \langle u, v_2\rangle.$$

(3) Positive Definite:

$$\langle u, u\rangle 0, \langle u, u\rangle = 0 \text{ if and only if } u = 0 \text{ (the zero vector)}.$$

In an inner product space the norm is defined by

$$\|u\| = \langle u, u\rangle^{\frac{1}{2}}.$$

It is a norm because of the following Cauchy-Schwartz inequality [1]

$$|\langle u, v\rangle| \leq \|u\|\|v\|,$$

[1] Assume $v \neq 0$ (otherwise the result is trivial) and let $\lambda = \langle u, v\rangle/\langle v, v\rangle$. Then a calculation gives

$$0 \leq \langle u - \lambda v, u - \lambda v\rangle = \langle u, u\rangle - \frac{|\langle u, v\rangle|^2}{\langle v, v\rangle}.$$

The result follows.

with equality if and only if u and v are linearly dependent. It follows that $\|u\|$ defines a norm, that is

$$\begin{aligned}
&\text{Homogeneity: } \ \|\alpha u\| = |\alpha|\|u\|;\\
&\text{Positivity: } \ \|u\| \geq 0, \quad \|u\| = 0 \ \text{ iff } u = 0;\\
&\text{Triangle Inequality: } \ \|u+v\| \leq \|u\| + \|v\|.
\end{aligned}$$

The triangle inequality is a consequence of Cauchy-Schwartz inequality:

$$\|u+v\|^2 = \|u\|^2 + \|v\|^2 + 2\mathrm{Re}\langle u, v\rangle \leq \|u\|^2 + \|v\|^2 + 2\|u\|\|v\| \leq (\|u\| + \|v\|)^2.$$

We observe that $L^2(E)$ is an inner product space with the L^2 inner product given by

$$\begin{aligned}
\langle f, g\rangle &:= \int_E f\bar{g} \quad \text{(complex-valued functions);}\\
\langle f, g\rangle &:= \int_E fg \quad \text{(real-valued functions).}
\end{aligned}$$

Definition 2.1. *A Hilbert space $\mathcal{H}$ is a complete, separable, inner product space.*

Theorem 2.1. *Let E be a measurable set of $\mathbb{R}^n$. Then $L^2(E)$ is a Hilbert space.*

The notion of an orthonormal basis is fundamental in the theory of Hilbert spaces.

Definition 2.2. *Let $\mathcal{H}$ is a Hilbert space with a countable orthonormal subset $\mathcal{E} = \{e_1, e_2, e_3, \ldots\}$, either finite or infinite. We say $\mathcal{E}$ is an orthonormal basis for $\mathcal{H}$ if every $f \in \mathcal{H}$ is of the form*

$$f = \sum_{k\geq 1} a_k e_k,$$

that is

$$\left\| f - \sum_{k\geq 1}^{N} a_k e_k \right\| \to 0 \quad \text{as } N \to \infty.$$

for some $a_k \in \mathbb{R}$ (or $\mathbb{C}$).

Fix a large N. Write $f = S_N + (f - S_N)$, where

$$S_N = \sum_{k=1}^{N} a_k e_k.$$

Multiplying e_k at both sides, we obtain

$$\langle f, e_k\rangle = \langle S_N, e_k\rangle + \langle f - S_N, e_k\rangle = a_k + \langle f - S_N, e_k\rangle,$$

which yields

$$|\langle f, e_k\rangle - a_k| \leq \|f - S_N\|.$$

Sending $N \to \infty$, we find that

$$a_k = \langle f, e_k\rangle.$$

Theorem 2.2. *Every Hilbert space $\mathcal{H}$ has an orthonormal subset $\mathcal{E} = \{e_1, e_2, \ldots,\}$ such that finite linear combinations of the e_i are dense in $\mathcal{H}$.*

Suppose $\mathcal{E} = \{e_1, e_2, \ldots,\}$ is any orthonormal subset of $\mathcal{H}$ such that finite linear combinations of the e_i are dense in $\mathcal{H}$, then $\mathcal{E} = \{e_1, e_2, \ldots,\}$ is an orthonormal basis.

Given an orthonormal basis $\mathcal{E} = \{e_1, e_2, \ldots,\}$, for any $f \in \mathcal{H}$, we have

$$f = \sum_{k\geq 1} a_k e_k, \quad \text{where } a_k = \langle f, e_k\rangle.$$

Moreover

$$\|f\|^2 = \sum_{k\geq 1} |a_k|^2 \quad \text{(Parseval's identity).}$$

Suppose $\{e_1, e_2, \ldots,\}$ is an orthonormal set (but not necessarily a basis). Then

$$\sum_{k\geq 1} |\langle f, e_k\rangle|^2 \leq \|f\|^2 \quad \text{(Bessel's inequality)}$$

If equality holds then $\{e_1, e_2, \ldots\}$ is an orthonormal basis.

The Gram Schmidt process converts a basis of an inner product space into an orthonormal basis. Suppose $w_1, \ldots, w_m$ is a basis. Then we construct $e_1, \ldots, e_m$ as follows:

(1) $e_1 = w_1/\|w_1\|$. (Unit length vector in same direction as w_1.)
(2) $e_2 = (w_2 - \langle w_2, e_1\rangle e_1)/\|w_2 - \langle w_2, e_1\rangle e_1\|$. (Subtract from w_2 its projection on the space spanned by e_1, and then normalise to unit length. This gives a unit length vector orthogonal to e_1 and in the space spanned by w_1, w_2.)
(3) $e_k = (w_k - \sum_{j=1}^{k-1}\langle w_k, e_j\rangle e_j)/\|w_k - \sum_{j=1}^{k-1}\langle w_k, e_j\rangle e_j\|$.
(4) Repeat the above procedure.

Proof of Theorem 2.2.

Step 1. Because $\mathcal{H}$ is separable there is a countable dense subset $\{g_1, g_2, g_3, \ldots\}$. By successively removing any g_i which is a linear combination of previous g_j's, we can assume that each finite subset $\{g_1, \ldots, g_N\}$ is linearly independent. For this new sequence which we also denote by $\{g_1, g_2, g_3, \ldots\}$, finite linear combinations of the g_i will be dense.

Applying the Gram Schmidt process from we obtain an orthonormal set $\{e_1, e_2, e_3, \ldots\}$, such that for each N the sets $\{e_1, \ldots, e_N\}$ and $\{g_1, \ldots, g_N\}$ span the same N-dimensional subspace of $\mathcal{H}$. Since finite linear combinations of the g_i are dense it follows that finite linear combinations of the e_i are also dense.

From now until the end of Step 5 we assume that $\mathcal{E}$ is an orthonormal subset of $\mathcal{H}$ and that finite linear combinations of the e_i are dense in $\mathcal{H}$.

Step 2. If $\langle h, e_i \rangle = 0$ for all e_i then $h = 0$.

To prove this, use the result in Step 1 to obtain a sequence $(h_k)_{k \geq 1}$ of linear combinations of the e_i such that $h_k \to h$. Since $\langle h, e_i \rangle = 0$ for all e_i it follows that $\langle h, h_k \rangle = 0$ for all h_k. From this we have the following nice argument:

$$\|h\|^2 = |\langle h, h \rangle| = |\langle h, h - h_k \rangle| \leq \|h\| \|h - h_k\| \to 0 \quad \text{as } k \to \infty.$$

Hence $\|h\| = 0$ and so $h = 0$.

Step 3. Let $a_k = \langle f, e_k \rangle$. Then, for all $i = 1, \ldots, N$,

$$e_i \perp \Big(f - \sum_{k=1}^{N} a_k e_k \Big),$$

and

$$\|f\|^2 = \sum_{k=1}^{N} |a_k|^2 + \Big\| f - \sum_{k=1}^{N} a_k e_k \Big\|. \tag{2.1}$$

This is the Pythagoras theorem.

Step 4. We show that $f = \sum_{k=1}^{\infty} a_k e_k$.

It follows by (2.1) that

$$\sum_{k=1}^{\infty} |a_k|^2 \leq \|f\|^2 < \infty,$$

and so the series $\sum_{k=1}^{\infty} |a_k|^2$ converges. This in turn shows that $\{\sum_{k=1}^{N} a_k e_k\}_{N=1}^{\infty}$ is a Cauchy sequence in $\mathcal{H}$, because

$$\Big\| \sum_{k=1}^{M} a_k e_k - \sum_{j=1}^{N} a_j e_j \Big\| = \sum_{i=N+1}^{M} |a_i|^2 \to 0.$$

By the completeness, $\sum_{k=1}^{\infty} a_k e_k$ exists, denoted by g.

Fix i. Since

$$\langle f - \sum_{k=1}^{N} a_k e_k, e_i \rangle = 0, \quad \forall\ N \geq i.$$

Sending $N \to \infty$, we obtain $\langle f - g, e_i \rangle = 0$ for all i. Therefore $f = g$ by Step 2.

Step 5. Note that $\|f\|^2 = \sum_{k\geq 1} |a_k|^2$.

This now follows from (2.1), since $\big\| f - \sum_{k=1}^{N} a_k e_k \big\| \to 0$ as $N \to \infty$ by Step 4.

Step 6. If $\{e_1, e_2, \ldots\}$ is an orthonormal set, but not necessarily a basis, then Step 3 is still valid. It follows that $\sum_{k\geq 1} |a_k|^2 \leq \|f\|^2$. If equality holds, then $\|f - \sum_{k=1}^{N} a_k e_k\|^2 \to 0$ as $N \to \infty$ and so $f = \sum_{k\geq 1} a_k e_k$. In particular $\{e_1, e_2, \ldots\}$ is an orthonormal basis.

□

Definition 2.3. *Given* $1 \leq p < \infty$*, we define*

$$\ell^p_{\mathbb{R}} = \Big\{ x : \mathbb{N} \to \mathbb{R} : \sum_{k\geq 1} |x_k|^p < \infty,\ x_k = x(k) \Big\},$$

$$\ell^p_{\mathbb{C}} = \Big\{ x : \mathbb{N} \to \mathbb{C} : \sum_{k\geq 1} |x_k|^p < \infty,\ x_k = x(k) \Big\},$$

with the norm $\|\cdot\|_{\ell^p}$ *given by* $\|x\|_{\ell^p} = (\sum_{k\geq 1} |x_k|^p)^{\frac{1}{p}}$*. We also define*

$$\ell^\infty_{\mathbb{R}} = \Big\{ x : \mathbb{N} \to \mathbb{R} : \sup_{k\geq 1} |x_k| < \infty,\ x_k = x(k) \Big\},$$

$$\ell^\infty_{\mathbb{C}} = \Big\{ x : \mathbb{N} \to \mathbb{C} : \sup_{k\geq 1} |x_k| < \infty,\ x_k = x(k) \Big\}.$$

with the norm $\|\cdot\|_{\ell^\infty}$ *given by* $\|\alpha\|_{\ell^\infty} = \sup_{k\geq 1} |a_k|$*.*

There are Hölder and Minkowski inequalities for ℓ^p:

$$\sum_{i=1}^{\infty} |x_i y_i| \leq \|x\|_{\ell^p} \|y\|_{\ell^q}, \quad \frac{1}{p} + \frac{1}{q} = 1,\ 1 < p, q < \infty,$$

$$\|x + y\|_{\ell^p} \leq \|x\|_{\ell^p} + \|y\|_{\ell^p}, \quad 1 \leq p \leq \infty.$$

Hence ℓ^p $(1 \le p \le \infty)$ is a normed linear space. Moreover they are Banach spaces.

We can define an inner product for $\ell^2_{\mathbb{R}}$ or $\ell^2_{\mathbb{C}}$ by

$$\langle x, y\rangle = \sum_{i\ge 1} x_i y_i \;\text{ or }\; \langle x, y\rangle = \sum_{i\ge 1} x_i \bar{y}_i$$

It can be verified that ℓ^2 is a Hilbert space.

The following result is a consequence of Theorem 2.2.

Theorem 2.3. *Each Hilbert space over $\mathbb{R}$ (or $\mathbb{C}$) is isomorphic to $\mathbb{R}^n$ (or $\mathbb{C}^n$) for some n, or is isomorphic to $\ell^2_{\mathbb{R}}$ (or $\ell^2_{\mathbb{C}}$).*

2.1. **Riesz representation theorem.**

We present the following theorem which characterises the continuous linear functionals on each Hilbert space.

Theorem 2.4 (Riesz representation theorem)**.** *Let T be a continuous linear functional on a Hilbert space $\mathcal{H}$. Then, there is a unique $g \in \mathcal{H}$ such that*

$$T(f) = \langle f, g\rangle \;\; \textit{for all } f \in \mathcal{H}.$$

Moreover $\|T\| = \|g\|$.

Proof. We outline the proof by several steps.

Step 1. Let $\mathcal{S} = \{f : T(f) = 0\}$, i.e. the kernel of T. Then $\mathcal{S}$ is a subspace (since T is linear) and is closed (since T is continuous).

Step 2. If $\mathcal{S} = \mathcal{H}$ then T is the zero operator and we take $g = 0 \in \mathcal{H}$. Otherwise, let $\mathcal{S}^\perp = \{h : \langle h, f\rangle = 0 \;\; \forall\, f \in \mathcal{S}\}$, the orthogonal complement of $\mathcal{S}$. So $\mathcal{H} = \mathcal{S} \oplus \mathcal{S}^\perp$ [2]. Select

$$h \in \mathcal{S}^\perp \;\text{ with } \|h\| = 1.$$

Step 3. Check that $T(f)h - T(h)f \in \mathcal{S}$, for every $f \in \mathcal{H}$. Hence

$$0 = \langle T(f)h - T(h)f, h\rangle = T(f) - \langle f, \overline{T(h)}h\rangle.$$

[2] We have the following conclusion (not hard exercise): If $\mathcal{G}$ is a closed subspace of $\mathcal{H}$, then for any $f \in \mathcal{H}$, there is a unique $g_0 \in \mathcal{G}$ such that $\|f - g_0\| = \inf_{g\in\mathcal{G}} \|f - g\|$ and $(f - g_0) \perp g$ for all $g \in \mathcal{G}$. Since $\mathcal{S} = \ker(T)$ is a closed subspace of $\mathcal{H}$. Hence we can write $f = (f - g_0) + g_0$ where g_0 is the unique element in $\mathcal{S}$, so $f - g_0 \in \mathcal{S}^\perp$ by the above-mentioned conclusion.

By taking $g = \overline{T(h)}h$, we obtain

$$T(f) = \langle f, g \rangle, \quad \forall\ f \in \mathcal{H}.$$

By Cauchy-Schwartz inequality, $\|T(f)\| \leq \|f\|\|g\|$, so $\|T\| \leq \|g\|$. Using $T(g) = \|g\|^2$, we see that $\|T\| = \|g\|$.

Step 4. Uniqueness: Suppose $T(f) = \langle f, g_1 \rangle = \langle f, g_2 \rangle$ for all f. Then $\langle f, g_1 - g_2 \rangle = 0$. Taking $f = g_1 - g_2$ shows $g_1 = g_2$.

$\square$

Remark 2.1. *It follows that $\mathcal{S}^\perp$ is one dimensional. For this, suppose h_1 and h_2 are two unit vectors in $\mathcal{S}^\perp$. By the construction in the proof above and the uniqueness, we see that $\overline{T(h_1)}h_1 = \overline{T(h_2)}h_2$. So h_1 and h_2 are linearly dependent.*

Remark 2.2. *As a consequence of Theorem 2.4, we see that $(L^2)^* = L^2$, which was previously shown in Theorem 1.3.*

In general, consider a linear transformation $T : \mathcal{H}_1 \to \mathcal{H}_2$ where $\mathcal{H}_1$ and $\mathcal{H}_2$ are two Hilbert space. We say T is bounded if there is $M > 0$ so that

$$\|T(f)\|_{\mathcal{H}_2} \leq M\|f\|_{\mathcal{H}_1}.$$

The norm of T is given by

$$\begin{aligned} \|T\| &:= \sup\{\|T(f)\|_{\mathcal{H}_2}/\|f\|_{\mathcal{H}_1} :\ f \neq 0\} \\ &= \sup\{|\langle T(f), g\rangle| : \|f\| \leq 1,\ \|g\| \leq 1,\ f \in \mathcal{H}_1,\ g \in \mathcal{H}_2\}. \end{aligned}$$

The first identity is the definition, and the second needs a little argument.

We say a linear transformation T is continuous, if $\|T(f_k) - T(f)\|_{\mathcal{H}_2} \to 0$ as $\|f_k - f\|_{\mathcal{H}_1} \to 0$. It is not hard to show, by using the linearity, that T is continuous if and only if it is bounded.

An application of the Riesz representation theorem is to determine the existence of the "adjoint" of a linear transformation. Consider the continuous linear transformation $T : \mathcal{H} \to \mathcal{H}$, that is from $\mathcal{H}$ to itself. Given $g \in \mathcal{H}$, we have a continuous linear functional

$$\langle T(\cdot), g \rangle : \mathcal{H} \to \mathbb{R} \ \ (\text{or } \mathbb{C}).$$

By Riesz representation theorem, there is a unique $g^* \in \mathcal{H}$ so that

$$\langle T(\cdot), g \rangle = \langle \cdot, g^* \rangle.$$

Therefore we can define an operator $T^*:\mathcal{H}\to\mathcal{H}$ via $T^*(g)=g^*$. The Riesz representation theorem also gives us

$$\begin{aligned}\|T^*(g)\| &= \sup\{|\langle T(f),g\rangle|/\|f\| : f\in\mathcal{H},\ f\neq 0\}\\ &= \sup\{|\langle T(f),g\rangle|/\|f\| : f\in\mathcal{H},\ \|f\|\leq 1\}.\end{aligned}$$

It is not hard to see that T^* is continuous and linear. We list the following properties

$$(i)\ \langle T(f),g\rangle=\langle f,T^*(g)\rangle,\ \ \forall\ f,g\in\mathcal{H};\ \ (ii)\ \|T\|=\|T^*\|;\ \ (iii)\ (T^*)^*=T.$$

We will not go further. The content will be left to the course *Functional Analysis*.

2.2. **Fourier series in $L^2([-\pi,\pi])$.**

We work in $L^2([-\pi,\pi])$. Every interval $[a,b]$ is OK but the constants are messier.

We use the inner product on $L^2([-\pi,\pi])$ for convenience

$$\langle f,g\rangle=\fint_{-\pi}^{\pi} f\bar{g}.$$

In the real case one can check that

$$S:=S_{\mathbb{R}}=\{1,\sqrt{2}\cos kx,\sqrt{2}\sin kx:\ \ k\in\mathbb{N}\}$$

forms an orthonormal set in $L^2([-\pi,\pi])$.

In the complex case one can check that

$$S:=S_{\mathbb{C}}=\{e^{ik\pi}:\ \ k\in\mathbb{N}\}$$

forms an orthonormal set in $L^2([-\pi,\pi])$.

Theorem 2.5. *In the real or complex case, S defined above is an orthonormal basis.*

Proof. By Theorem 2.2, we show that $\mathrm{Span}(S)$, the set of finite liner combinations of functions from S, is dense in $L^2([-\pi,\pi])$, and so in the L^2 norm.

Step 1. We first consider the real case.

Observe that (i) $\mathrm{Span}(S_{\mathbb{R}})$ is an algebra, namely it is closed under addition, scalar multiplication, and products; (ii) $\mathrm{Span}(S_{\mathbb{R}})$ contains the constant functions; (iii) $\mathrm{Span}(S_{\mathbb{R}})$ separates points, namely for each pair $x\neq y$, one has $f(x)\neq f(y)$ for some $f\in\mathrm{Span}(S_{\mathbb{R}})$.

By the Stone-Weierstrass theorem (see Theorem 2.7 below), $\mathrm{Span}(S_{\mathbb{R}})$ is dense in $C([-\pi,\pi])$ in the sup norm.

Step 2. Next we consider the complex case.

Suppose $f = u + iv \in L^2([-\pi, \pi])$. Fix any $\varepsilon > 0$. From the real case, there are $\phi, \psi \in \mathrm{Span}(\mathbb{S}_{\mathbb{R}})$ such that $\|u - \phi\|_{L^2} < \varepsilon/2$ and $\|v - \psi\|_{L^2} < \varepsilon/2$. It follows that $\|f - g\| < \varepsilon$, where $f = \phi + i\psi$.

Note that $g \in \mathrm{Span}(S_{\mathbb{C}})$, which can be seen from

$$\cos kx = \frac{1}{2}(e^{ikx} + e^{-ikx}), \quad \cos kx = \frac{1}{2i}(e^{ikx} - e^{-ikx}), \quad \text{and} \quad 1 = e^{i0x}.$$

This completes the proof in the complex case.

□

We now summary some important facts for Fourier series in $L^2([-\pi, \pi])$ below. The result is a consequence of the combination of Theorem 2.2 and Theorem 2.5.

Theorem 2.6. *Suppose $f \in L^2([-\pi, \pi]; \mathbb{R})$. Let*

$$a_0 = \fint_{-\pi}^{\pi} f(x)dx, \;\; a_k = \sqrt{2}\fint_{-\pi}^{\pi} f(x)\cos kxdx, \;\; b_k = \sqrt{2}\fint_{-\pi}^{\pi} f(x)\sin kxdx, \;\; k > 0.$$

Then

(i) $f(x) = a_0 + \sqrt{2}\sum_{k\geq 1}(a_k \cos kx + b_k \sin kx)$ *with convergence in L^2 sense.*
(ii) $\fint_{-\pi}^{\pi} |f|^2 = a_0^2 + \sum_{k\geq 1}(a_k^2 + b_k^2)$.
(iii) $f = 0$ *a.e. if and only if $a_k = b_k = 0$ for all k.*
(iv) $f \mapsto (a_0, a_1, b_1, a_2, b_2, a_3, b_3, \ldots)$ *is an isomorphism between $L^2([-\pi, \pi]; \mathbb{R})$ and ℓ^2.*

Suppose $f \in L^2([-\pi, \pi]; \mathbb{C})$. Let $c_k = \fint_{-\pi}^{\pi} f(x)e^{-ikx}xdx$ for $k \in \mathbb{Z}$. Then

(i) $f(x) = \sum_{k-\infty}^{\infty} c_k e^{ikx}$ *with convergence in L^2 sense.*
(ii) $\fint_{-\pi}^{\pi} |f|^2 = \sum_{k-\infty}^{\infty} |c_k|^2$.
(iii) $f = 0$ *a.e. if and only if $c_k = 0$ for all k.*
(iv) $f \mapsto (c_0, c_1, c_{-1}, c_2, c_{-2}, \ldots)$ *is an isomorphism between $L^2([-\pi, \pi]; \mathbb{C})$ and ℓ^2.*

2.3. **Stone's generalisation of the Weierstrass theorem.**

Let K be a compact set, and $C(K) = C(K; \mathbb{R})$ (or $C(K) = C(K; \mathbb{C})$) be the set of all continuous real-valued (complex-valued) functions on K. We concern if some subset $\mathscr{A}$ with nice properties is dense in $C(K)$. That is the so-called approximation theorem.

The following Stone-Weierstrass theorem is used in the proof of Theorem 2.5, which is a generalisation of Theorem **??**.

Theorem 2.7 (Stone-Weierstrass Theorem). *Let $\mathscr{A}$ be an algebra of real continuous functions on a compact set K. Suppose*

(i) *$\mathscr{A}$ separates points on K;*
(ii) *$\mathscr{A}$ vanishes at no point of K.*

Then the uniform closure $\mathscr{B}$ of $\mathscr{A}$ is $C(K;\mathbb{R})$.

The presentation below is from the book *Principles of Mathematics Analysis* by Rudin.

Definition 2.4. *A family $\mathscr{A}$ of real (or complex) functions defined on a set E is said to be an algebra if*

$$(i)\ f+g \in \mathscr{A};\ (ii)\ fg \in \mathscr{A};\ (iii)\ cf \in \mathscr{A};\ \forall\ f,g \in \mathscr{A} \text{ and } c \in \mathbb{R} \text{ (or } \mathbb{C}).$$

That is, if $\mathscr{A}$ is closed under addition, product, and scalar multiplication.

Definition 2.5. *Let $\mathscr{B}$ be the set of all functions which are limits of uniformly convergent sequences of members of $\mathscr{A}$. Then $\mathscr{B}$ is called the uniform closure of $\mathscr{A}$.*

Definition 2.6. *Let $\mathscr{A}$ be a family of functions on a set E. Then $\mathscr{A}$ is said to separate points on E if to every pair of distinct points $x, y \in E$ there corresponds a function $f \in \mathscr{A}$ such that $f(x) \neq f(y)$.*

If to each $x \in E$ there corresponds a function $g \in \mathscr{A}$ such that $g(x) \neq 0$, we say that $\mathscr{A}$ vanishes at no point of E.

The algebra of all polynomials in one variable clearly separate points and vanishes at no point on $\mathbb{R}$. An example of an algebra which does not separate points is the set of all even polynomials, as $f(-x) = f(x)$ for each such function.

Lemma 2.1. *Suppose $\mathscr{A}$ is an algebra of functions on a set E, $\mathscr{A}$ separates points on E, and vanishes at no points of E. Suppose x, y are distinct points of E, and a, b are constants. Then $\mathscr{A}$ contains a function f such that*

$$f(x) = a \quad \text{and} \quad f(y) = b.$$

Proof. The assumptions show there are ϕ, ψ, η in $\mathscr{A}$ such that

$$\phi(x) \neq 0, \quad \psi(y) \neq 0, \quad \eta(x) \neq \eta(y).$$

Put

$$u(\xi) = \big(\eta(\xi) - \eta(x)\big)\psi(\xi) \quad \text{and} \quad v(\xi) = \big(\eta(\xi) - \eta(y)\big)\phi(\xi).$$

Then $u, v \in \mathscr{A}$ and

$$u(x) = v(y) = 0, \quad u(y) \neq 0 \ \text{ and } \ v(x) \neq 0.$$

Therefore

$$f(\xi) = \frac{av(\xi)}{v(x)} + \frac{bu(\xi)}{u(y)}$$

has the desired properties.

$\square$

Proof of Theorem 2.7. The proof is divided into several steps.

Step 1. If $f \in \mathscr{B}$, then $|f| \in \mathscr{B}$.

Let $M = \sup_{x \in K} |f(x)|$. Fix any $\varepsilon > 0$. By Weierstrass Theorem, there exist N and real numbers $a_1, \cdots a_N$ such that

$$\sup_{y \in [-M,M]} \Big| \sum_{k=1}^{N} a_k y^k - |y| \Big| < \varepsilon. \tag{2.2}$$

Since $\mathscr{B}$ is an algebra, the function

$$g_N(\xi) := \sum_{k=1}^{N} a_k f^k(\xi)$$

is an element of $\mathscr{B}$. We see from (2.2) that

$$\sup_{\xi \in K} \big| g_N(\xi) - |f(\xi)| \big| < \varepsilon.$$

Since $\mathscr{B}$ is uniformly closed, we have $|f| \in \mathscr{B}$.

Step 2. If $f, g \in \mathscr{B}$, then $\max(f,g), \min(f,g) \in \mathscr{B}$.

This is a consequence of Step 1 and the observation below

$$\max(f,g) = \frac{1}{2}(f+g) + \frac{1}{2}|f-g|,$$
$$\min(f,g) = \frac{1}{2}(f+g) - \frac{1}{2}|f-g|.$$

Step 3. Given $f \in C(K;\mathbb{R})$, a point $x \in K$, and $\varepsilon > 0$, there is a $g_x \in \mathscr{B}$ such that $g_x(x) = f(x)$ and

$$g_x(\xi) > f(\xi) - \varepsilon, \quad \forall\, \xi \in K. \tag{2.3}$$

Appying Lemma 2.1 to $\mathscr{B}$, we deduce, for each $y \in K$, there is $h_y \in \mathscr{B}$ such that

$$h_y(x) = f(x), \quad h_y(y) = f(y).$$

By the continuity of h_y, there exists an open set $\mathscr{U}_y \ni y$, such that

$$h_y(\xi) > f(\xi) - \varepsilon, \quad \forall\, \xi \in \mathscr{U}_y.$$

Since K is compact, there is a finite set $y_1, \dots, y_L$ such that

$$K \subset \bigcup_{j=1}^{L} \mathscr{U}_{y_j}.$$

By Step 2, the function defined below is in $\mathscr{B}$, and has the needed properties,

$$g_x(\xi) = \max(h_{y_1}(\xi), \dots, h_{y_L}(\xi)).$$

Step 4. Given $f \in C(K;\mathbb{R})$, and $\varepsilon > 0$, there is $h \in \mathscr{B}$ such that

$$\sup_{\xi \in K} |h(\xi) - f(\xi)| < \varepsilon \tag{2.4}$$

Once this is obtained, our theorem is proved.

For each $x \in K$, consider the function g_x constructed in Step 3. By the continuity, there are open sets $\mathscr{V}_x \ni x$, such that

$$g_x(\xi) < f(\xi) + \varepsilon, \quad \forall\, \xi \in \mathscr{V}_x. \tag{2.5}$$

By the compactness of K, we find $x_1, \dots, x_J$ such that

$$K \subset \bigcup_{j=1}^{J} \mathscr{V}_{x_j}.$$

Then define

$$h(\xi) = \min(g_{x_1}(\xi), \dots, g_{x_J}(\xi)).$$

Now (2.4) follows by combining (2.3) and (2.5).

$\square$

Theorem 2.7 does not hold for complex algebras in general. However the conclusion holds if $\mathscr{A}$ is self-adjoint, that is for every $f \in \mathscr{A}$ its complex conjugate $\bar{f} \in \mathscr{A}$, where $\bar{f}(x) = \overline{f(x)}$.

Theorem 2.8. *Suppose $\mathscr{A}$ is a self-adjoint algebra of complex continuous functions on a compact set K, $\mathscr{A}$ separates points of K, and $\mathscr{A}$ vanishes at no point of K. Then the uniform closure $\mathscr{B}$ of $\mathscr{A}$ is $C(K;\mathbb{C})$. In other words, $\mathscr{A}$ is dense in $C(K;\mathbb{C})$.*

Proof. This is a consequence of applying Theorem 2.7 to

$$\mathrm{Re}(f) = f + \bar{f} \quad \text{and} \quad \mathrm{Im}(f) = f - \bar{f}.$$

$\square$