16. Banach Spaces II

Theorem 16.1 (Open Mapping Theorem)**.** *Let X, Y be Banach spaces, $T \in L(X,Y)$. If T is surjective then T is an open mapping, i.e. $T(V)$ is open in Y for all open subsets $V \subset X$.*

Proof. For all $\alpha > 0$ let $B_\alpha^X = \{x \in X : \|x\|_X < \alpha\} \subset X$, $B_\alpha^Y = \{y \in Y : \|y\|_Y < \alpha\} \subset Y$ and $E_\alpha = T(B_\alpha^X) \subset Y$. The proof will be carried out by proving the following three assertions.

(1) There exists $\delta > 0$ such that $B_{\delta\alpha}^Y \subset \overline{E_\alpha}$ for all $\alpha > 0$.
(2) For the same $\delta > 0$, $B_{\delta\alpha}^Y \subset E_\alpha$, i.e. we may remove the closure in assertion 1.
(3) The last assertion implies T is an open mapping.

1. Since $Y = \bigcup\limits_{n-1}^{\infty} E_n$, the Baire category Theorem 15.16 implies there exists n such that $\overline{E}_n^0 \neq \emptyset$, i.e. there exists $y \in \overline{E}_n$ and $\epsilon > 0$ such that $\overline{B^Y(y,\epsilon)} \subset \overline{E}_n$. Suppose $\|y'\| < \epsilon$ then y and $y + y'$ are in $B^Y(y,\epsilon) \subset \overline{E}_n$ hence there exists $x', x \in B_n^X$ such that $\|Tx' - (y+y')\|$ and $\|Tx - y\|$ may be made as small as we please, which we abbreviate as follows

$$\|Tx' - (y+y')\| \approx 0 \text{ and } \|Tx - y\| \approx 0.$$

Hence by the triangle inequality,

$$\begin{aligned}\|T(x'-x) - y'\| &= \|Tx' - (y+y') - (Tx - y)\| \\ &\leq \|Tx' - (y+y')\| + \|Tx - y\| \approx 0\end{aligned}$$

with $x' - x \in B_{2n}^X$. This shows that $y' \in \overline{E_{2n}}$ which implies $B^Y(0,\epsilon) \subset \overline{E_{2n}}$. Since the map $\phi_\alpha : Y \to Y$ given by $\phi_\alpha(y) = \frac{\alpha}{2n} y$ is a homeomorphism, $\phi_\alpha(E_{2n}) = E_\alpha$ and $\phi_\alpha(B^Y(0,\epsilon)) = B^Y(0, \frac{\alpha\epsilon}{2n})$, it follows that $B_{\delta\alpha}^Y \subset \overline{E_\alpha}$ where $\delta \equiv \frac{\epsilon}{2n} > 0$.

2. Let δ be as in assertion 1., $y \in B_\delta^Y$ and $\alpha_1 \in (\|y\|/\delta, 1)$. Choose $\{\alpha_n\}_{n=2}^\infty \subset (0,\infty)$ such that $\sum_{n=1}^\infty \alpha_n < 1$. Since $y \in B_{\alpha_1\delta}^Y \subset \overline{E_{\alpha_1}} = \overline{T\left(B_{\alpha_1}^X\right)}$ by assertion **1**. there exists $x_1 \in B_{\alpha_1}^X$ such that $\|y - Tx_1\| < \alpha_2\delta$. (Notice that $\|y - Tx_1\|$ can be made as small as we please.) Similarly, since $y - Tx_1 \in B_{\alpha_2\delta}^Y \subset \bar{E}_{\alpha_2} = \overline{T\left(B_{\alpha_2}^X\right)}$ there exists $x_2 \in B_{\alpha_2}^X$ such that $\|y - Tx_1 - Tx_2\| < \alpha_3\delta$. Continuing this way inductively, there exists $x_n \in B_{\alpha_n}^X$ such that

$$\|y - \sum_{k=1}^{n} Tx_k\| < \alpha_{n+1}\delta \text{ for all } n \in \mathbb{N}. \tag{16.1}$$

Since $\sum\limits_{n=1}^\infty \|x_n\| < \sum\limits_{n=1}^\infty \alpha_n < 1$, $x \equiv \sum\limits_{n=1}^\infty x_n$ exists and $\|x\| < 1$, i.e. $x \in B_1^X$. Passing to the limit in Eq. (16.1) shows, $\|y - Tx\| = 0$ and hence $y \in T(B_1^X) = E_1$. Therefore we have shown $B_\delta^X \subset E_1$. The same scaling argument as above then shows $B_{\alpha\delta}^X \subset E_\alpha$ for all $\alpha > 0$.

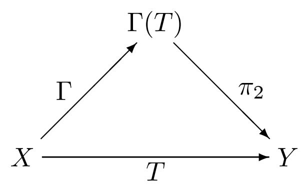
3. If $x \in V \subset_o X$ and $y = Tx \in TV$ we must show that TV contains a ball $B^Y(y,\epsilon) = Tx + B_\epsilon^Y$ for some $\epsilon > 0$. Now $B^Y(y,\epsilon) = Tx + B_\epsilon^Y \subset TV$ iff $B_\epsilon^Y \subset TV - Tx = T(V - x)$. Since $V - x$ is a neighborhood of $0 \in X$, there exists $\alpha > 0$ such that $B_\alpha^X \subset (V - x)$ and hence by assertion 2., $B_{\alpha\delta}^Y \subset TB_\alpha^X \subset T(V-x)$ and therefore $B^Y(y,\epsilon) \subset TV$ with $\epsilon := \alpha\delta$. ■

Corollary 16.2. *If X, Y are Banach spaces and $T \in L(X,Y)$ is invertible (i.e. a bijective linear transformation) then the inverse map, T^{-1}, is* ***bounded****, i.e. $T^{-1} \in L(Y,X)$. (Note that T^{-1} is automatically linear.)*

Theorem 16.3 (Closed Graph Theorem)**.** *Let X and Y be Banach space $T : X \to Y$ linear is continuous iff T is closed i.e. $\Gamma(T) \subset X \times Y$ is closed.*

Proof. If T continuous and $(x_n, Tx_n) \to (x,y) \in X \times Y$ as $n \to \infty$ then $Tx_n \to Tx = y$ which implies $(x,y) = (x,Tx) \in \Gamma(T)$.

Conversely: If T is **closed** then the following diagram commutes



where $\Gamma(x) := (x, Tx)$.

The map $\pi_2 : X \times Y \to X$ is continuous and $\pi_1|_{\Gamma(T)} : \Gamma(T) \to X$ is continuous bijection which implies $\pi_1|_{\Gamma(T)}^{-1}$ is bounded by the open mapping Theorem 16.1. Hence $T = \pi_2 \circ \pi_1|_{\Gamma(T)}^{-1}$ is bounded, being the composition of bounded operators. ■

As an application we have the following proposition.

Proposition 16.4. *Let H be a Hilbert space. Suppose that $T : H \to H$ is a linear (not necessarily bounded) map such that there exists $T^* : H \to H$ such that*

$$\langle Tx, Y\rangle = \langle x, T^* Y\rangle \ \forall\, x, y \in H.$$

Then T is bounded.

Proof. It suffices to show T is closed. To prove this suppose that $x_n \in H$ such that $(x_n, Tx_n) \to (x,y) \in H \times H$. Then for any $z \in H$,

$$\langle Tx_n, z\rangle = \langle x_n, T^* z\rangle \longrightarrow \langle x, T^* z\rangle = \langle Tx, z\rangle \text{ as } n \to \infty.$$

On the other hand $\lim_{n\to\infty} \langle Tx_n, z\rangle = \langle y, z\rangle$ as well and therefore $\langle Tx, z\rangle = \langle y, z\rangle$ for all $z \in H$. This shows that $Tx = y$ and proves that T is closed. ■

Here is another example.

Example 16.5. Suppose that $\mathcal{M} \subset L^2([0,1], m)$ is a closed subspace such that each element of $\mathcal{M}$ has a representative in $C([0,1])$. We will abuse notation and simply write $\mathcal{M} \subset C([0,1])$. Then

(1) There exists $A \in (0,\infty)$ such that $\|f\|_\infty \le A\|f\|_{L^2}$ for all $f \in \mathcal{M}$.
(2) For all $x \in [0,1]$ there exists $g_x \in \mathcal{M}$ such that

$$f(x) = \langle f, g_x\rangle \text{ for all } f \in \mathcal{M}.$$

Moreover we have $\|g_x\| \le A$.
(3) The subspace $\mathcal{M}$ is finite dimensional and $\dim(\mathcal{M}) \le A^2$.

Proof. 1) I will give a two proofs of part 1. Each proof requires that we first show that $(\mathcal{M}, \|\cdot\|_\infty)$ is a complete space. To prove this it suffices to show $\mathcal{M}$ is a closed subspace of $C([0,1])$. So let $\{f_n\} \subset \mathcal{M}$ and $f \in C([0,1])$ such that $\|f_n - f\|_\infty \to 0$ as $n \to \infty$. Then $\|f_n - f_m\|_{L^2} \le \|f_n - f_m\|_\infty \to 0$ as $m, n \to \infty$, and since $\mathcal{M}$ is closed in $L^2([0,1])$, $L^2 - \lim_{n\to\infty} f_n = g \in \mathcal{M}$. By passing to a subsequence if necessary we know that $g(x) = \lim_{n\to\infty} f_n(x) = f(x)$ for m - a.e. x. So $f = g \in \mathcal{M}$.

i)Let $i : (\mathcal{M}, \|\cdot\|_\infty) \to (\mathcal{M}, \|\cdot\|_2)$ be the identity map. Then i is bounded and bijective. By the open mapping theorem, $j = i^{-1}$ is bounded as well. Hence there exists $A < \infty$ such that $\|f\|_\infty = \|j(f)\| \le A\|f\|_2$ for all $f \in \mathcal{M}$.

ii) Let $j : (\mathcal{M}, \|\cdot\|_2) \to (\mathcal{M}, \|\cdot\|_\infty)$ be the identity map. We will shows that j is a closed operator and hence bounded by the closed graph theorem. Suppose that $f_n \in \mathcal{M}$ such that $f_n \to f$ in L^2 and $f_n = j(f_n) \to g$ in $C([0,1])$. Then as in the first paragraph, we conclude that $g = f = j(f)$ a.e. showing j is closed. Now finish as in last line of proof i).

2) For $x \in [0,1]$, let $e_x : \mathcal{M} \to \mathbb{C}$ be the evaluation map $e_x(f) = f(x)$. Then

$$|e_x(f)| \le |f(x)| \le \|f\|_\infty \le A\|f\|_{L^2}$$

which shows that $e_x \in \mathcal{M}^*$. Hence there exists a unique element $g_x \in \mathcal{M}$ such that

$$f(x) = e_x(f) = \langle f, g_x \rangle \text{ for all } f \in \mathcal{M}.$$

Moreover $\|g_x\|_{L^2} = \|e_x\|_{\mathcal{M}^*} \le A$.

3) Let $\{f_j\}_{j=1}^n$ be an L^2 – orthonormal subset of $\mathcal{M}$. Then

$$A^2 \ge \|e_x\|_{\mathcal{M}^*}^2 = \|g_x\|_{L^2}^2 \ge \sum_{j=1}^n |\langle f_j, g_x \rangle|^2 = \sum_{j=1}^n |f_j(x)|^2$$

and integrating this equation over $x \in [0,1]$ implies that

$$A^2 \ge \sum_{j=1}^n \int_0^1 |f_j(x)|^2 dx = \sum_{j=1}^n 1 = n$$

which shows that $n \le A^2$. Hence $\dim(\mathcal{M}) \le A^2$. ■

Remark 16.6. Keeping the notation in Example 16.5, $G(x,y) = g_x(y)$ for all $x, y \in [0,1]$. Then

$$f(x) = e_x(f) = \int_0^1 f(y)\overline{G(x,y)}dy \text{ for all } f \in \mathcal{M}.$$

The function G is called the reproducing kernel for $\mathcal{M}$.

The above example generalizes as follows.

Proposition 16.7. *Suppose that $(X, \mathcal{M}, \mu)$ is a finite measure space, $p \in [1,\infty)$ and W is a closed subspace of $L^p(\mu)$ such that $W \subset L^p(\mu) \cap L^\infty(\mu)$. Then* $\dim(W) < \infty$.

Proof. With out loss of generality we may assume that $\mu(X) = 1$. As in Example 16.5, we shows that W is a closed subspace of $L^\infty(\mu)$ and hence by the open mapping theorem, there exists a constant $A < \infty$ such that $\|f\|_\infty \le A\|f\|_p$ for all $f \in W$. Now if $1 \le p \le 2$, then

$$\|f\|_\infty \le A\|f\|_p \le A\|f\|_2$$

and if $p \in (2,\infty)$, then $\|f\|_p^p \le \|f\|_2^2 \|f\|_\infty^{p-2}$ or equivalently,

$$\|f\|_p \le \|f\|_2^{2/p} \|f\|_\infty^{1-2/p} \le \|f\|_2^{2/p} \left(A\|f\|_p\right)^{1-2/p}$$

from which we learn that $\|f\|_p \leq A^{1-2/p}\|f\|_2$ and therefore that $\|f\|_\infty \leq AA^{1-2/p}\|f\|_2$ so that in any case there exists a constant $B < \infty$ such that $\|f\|_\infty \leq B\|f\|_2$.

Let $\{f_n\}_{n=1}^N$ be an orthonormal subset of W and $f = \sum_{n=1}^N c_n f_n$ with $c_n \in \mathbb{C}$, then

$$\left\|\sum_{n=1}^N c_n f_n\right\|_\infty^2 \leq B^2 \sum_{n=1}^N |c_n|^2 \leq B^2 |c|^2$$

where $|c|^2 := \sum_{n=1}^N |c_n|^2$. For each $c \in \mathbb{C}^N$, there is an exception set E_c such that for $x \notin E_c$,

$$\left|\sum_{n=1}^N c_n f_n(x)\right|^2 \leq B^2 |c|^2.$$

Let $\mathbb{D} := (\mathbb{Q} + i\mathbb{Q})^N$ and $E = \cap_{c\in\mathbb{D}} E_c$. Then $\mu(E) = 0$ and for $x \notin E$, $\left|\sum_{n=1}^N c_n f_n(x)\right| \leq B^2 |c|^2$ for all $c \in \mathbb{D}$. By continuity it then follows for $x \notin E$ that

$$\left|\sum_{n=1}^N c_n f_n(x)\right|^2 \leq B^2 |c|^2 \text{ for all } c \in \mathbb{C}^N.$$

Taking $c_n = f_n(x)$ in this inequality implies that

$$\left|\sum_{n=1}^N |f_n(x)|^2\right|^2 \leq B^2 \sum_{n=1}^N |f_n(x)|^2 \text{ for all } x \notin E$$

and therefore that

$$\sum_{n=1}^N |f_n(x)|^2 \leq B^2 \text{ for all } x \notin E.$$

Integrating this equation over x then implies that $N \leq B^2$, i.e. $\dim(W) \leq B^2$. ■

Theorem 16.8 (Uniform Boundedness Principle)**.** *Let X and Y be a normed vector spaces, $\mathcal{A} \subset L(X,Y)$ be a collection of bounded linear operators from X to Y,*

$$F = F_\mathcal{A} = \{x \in X : \sup_{A\in\mathcal{A}} \|Ax\| < \infty\} \text{ and}$$

$$R = R_\mathcal{A} = F^c = \{x \in X : \sup_{A\in\mathcal{A}} \|Ax\| = \infty\}. \tag{16.2}$$

1. *If $\sup_{A\in\mathcal{A}} \|A\| < \infty$ then $F = X$.*
2. *If F is not meager, then $\sup_{A\in\mathcal{A}} \|A\| < \infty$.*
3. *If X is a Banach space, F is not meager iff $\sup_{A\in\mathcal{A}} \|A\| < \infty$. In particular if $\sup_{A\in\mathcal{A}} \|Ax\| < \infty$ for all $x \in X$ then $\sup_{A\in\mathcal{A}} \|A\| < \infty$.*
4. *If X is a Banach space, then $\sup_{A\in\mathcal{A}} \|A\| = \infty$ iff R is residual. In particular if $\sup_{A\in\mathcal{A}} \|A\| = \infty$ then $\sup_{A\in\mathcal{A}} \|Ax\| = \infty$ for x in a dense subset of X.*

Proof. 1. If $M := \sup_{A\in\mathcal{A}} \|A\| < \infty$, then $\sup_{A\in\mathcal{A}} \|Ax\| \leq M\|x\| < \infty$ for all $x \in X$ showing $F = X$.

2. For each $n \in \mathbb{N}$, let $E_n \subset X$ be the closed sets given by

$$E_n = \{x : \sup_{A \in \mathcal{A}} \|Ax\| \le n\} = \bigcap_{A \in \mathcal{A}} \{x : \|Ax\| \le n\}.$$

Then $F = \cup_{n=1}^{\infty} E_n$ which is assumed to be non-meager and hence there exists an $n \in \mathbb{N}$ such that E_n has non-empty interior. Let $B_x(\delta)$ be a ball such that $\overline{B_x(\delta)} \subset E_n$. Then for $y \in X$ with $\|y\| = \delta$ we know $x - y \in \overline{B_x(\delta)} \subset E_n$, so that $Ay = Ax - A(x-y)$ and hence for any $A \in \mathcal{A}$,

$$\|Ay\| \le \|Ax\| + \|A(x-y)\| \le n + n = 2n.$$

Hence it follows that $\|A\| \le 2n/\delta$ for all $A \in \mathcal{A}$, i.e. $\sup_{A \in \mathcal{A}} \|A\| \le 2n/\delta < \infty$.

3. If X is a Banach space, $F = X$ is not meager by the Baire Category Theorem 15.16. So item 3. follows from items 1. and 2 and the fact that $F = X$ iff $\sup_{A \in \mathcal{A}} \|Ax\| < \infty$ for all $x \in X$.

4. Item 3. is equivalent to F is meager iff $\sup_{A \in \mathcal{A}} \|A\| = \infty$. Since $R = F^c$, R is residual iff F is meager, so R is residual iff $\sup_{A \in \mathcal{A}} \|A\| = \infty$. ■

Example 16.9. Suppose that $\{c_n\}_{n=1}^{\infty} \subset \mathbb{C}$ is a sequence of numbers such that

$$\lim_{N \to \infty} \sum_{n=1}^{N} a_n c_n \text{ exists in } \mathbb{C} \text{ for all } a \in \ell^1.$$

Then $c \in \ell^\infty$.

Proof. Let $f_N \in \left(\ell^1\right)^*$ be given by $f_N(a) = \sum_{n=1}^{N} a_n c_n$ and set $M_N := \max\{|c_n| : n = 1, \dots, N\}$. Then

$$|f_N(a)| \le M_N \|a\|_{\ell^1}$$

and by taking $a = e_k$ with k such $M_N = |c_k|$, we learn that $\|f_N\| = M_N$. Now by assumption, $\lim_{N\to\infty} f_N(a)$ exists for all $a \in \ell^1$ and in particular,

$$\sup_N |f_N(a)| < \infty \text{ for all } a \in \ell^1.$$

So by the Theorem 16.8,

$$\infty > \sup_N \|f_N\| = \sup_N M_N = \sup\{|c_n| : n = 1, 2, 3, \dots\}.$$

■

16.1. **Applications to Fourier Series.** Let $T = S^1$ be the unit circle in S^1 and m denote the normalized arc length measure on T. So if $f : T \to [0, \infty)$ is measurable, then

$$\int_T f(w)dw := \int_T f dm := \frac{1}{2\pi} \int_{-\pi}^{\pi} f(e^{i\theta}) d\theta.$$

Also let $\phi_n(z) = z^n$ for all $n \in \mathbb{Z}$. Recall that $\{\phi_n\}_{n\in\mathbb{Z}}$ is an orthonormal basis for $L^2(T)$. For $n \in \mathbb{N}$ let

$$\begin{aligned} s_n(f,z) &:= \sum_{k=-n}^{n} \langle f, \phi_n\rangle \phi_k(z) = \sum_{k=-n}^{n} \langle f, \phi_n\rangle z^k = \sum_{k=-n}^{n} \left(\int_T f(w)\bar{w}^k dw\right) z^k \\ &= \int_T f(w) \left(\sum_{k=-n}^{n} \bar{w}^k z^k\right) dw = \int_T f(w) d_n(z\bar{w}) dw \end{aligned}$$

where $d_n(\alpha) := \sum_{k=-n}^{n} \alpha^k$. Now $\alpha d_n(\alpha) - d_n(\alpha) = \alpha^{n+1} - \alpha^{-n}$, so that

$$d_n(\alpha) := \sum_{k=-n}^{n} \alpha^k = \frac{\alpha^{n+1} - \alpha^{-n}}{\alpha - 1}$$

with the convention that

$$\frac{\alpha^{n+1} - \alpha^{-n}}{\alpha - 1}|_{\alpha=1} = \lim_{\alpha\to 1} \frac{\alpha^{n+1} - \alpha^{-n}}{\alpha - 1} = 2n+1 = \sum_{k=-n}^{n} 1^k.$$

Writing $\alpha = e^{i\theta}$, we find

$$\begin{aligned} D_n(\theta) := d_n(e^{i\theta}) &= \frac{e^{i\theta(n+1)} - e^{-i\theta n}}{e^{i\theta} - 1} = \frac{e^{i\theta(n+1/2)} - e^{-i\theta(n+1/2)}}{e^{i\theta/2} - e^{-i\theta/2}} \\ &= \frac{\sin(n+\frac{1}{2})\theta}{\sin\frac{1}{2}\theta}. \end{aligned}$$

Recall by Hilbert space theory, $L^2(T) - \lim_{n\to\infty} s_n(f,\cdot) = f$ for all $f \in L^2(T)$. We will now show that the convergence is not pointwise for all $f \in C(T) \subset L^2(T)$.

Proposition 16.10. *For each $z \in T$, there exists a residual set $R_z \subset C(T)$ such that $\sup_n |s_n(f,z)| = \infty$ for all $f \in R_z$. Recall that $C(T)$ is a complete metric space, hence R_z is a dense subset of $C(T)$.*

Proof. By symmetry considerations, it suffices to take $z = 1 \in T$. Let $\Lambda_n f := s_n(f, 1)$. Then

$$|\Lambda_n f| = \left|\int_T f(w) d_n(\bar{w}) dw\right| \le \int_T |f(w) d_n(\bar{w})|\, dw \le \|f\|_\infty \int_T |d_n(\bar{w})|\, dw$$

showing

$$\|\Lambda_n\| \le \int_T |d_n(\bar{w})|\, dw.$$

Since $C(T)$ is dense in $L^1(T)$, there exists $f_k \in C(T,\mathbb{R})$ such that $f_k(w) \to \mathrm{sgn} d_k(\bar{w})$ in L^1. By replacing f_k by $(f_k \wedge 1) \vee (-1)$ we may assume that $\|f_k\|_\infty \le 1$. It now follows that

$$\|\Lambda_n\| \ge \frac{|\Lambda_n f_k|}{\|f_k\|_\infty} \ge \left|\int_T f_k(w) d_n(\bar{w}) dw\right|$$

and passing to the limit as $k \to \infty$ implies that $\|\Lambda_n\| \ge \int_T |d_n(\bar{w})|\, dw$. Hence we have shown that

$$\|\Lambda_n\| = \int_T |d_n(\bar{w})|\, dw = \frac{1}{2\pi}\int_{-\pi}^{\pi} \left|d_n(e^{-i\theta})\right| d\theta = \frac{1}{2\pi}\int_{-\pi}^{\pi} \left|\frac{\sin(n+\frac{1}{2})\theta}{\sin\frac{1}{2}\theta}\right| d\theta. \tag{16.3}$$

Since

$$\sin x = \int_0^x \cos y dy \le \int_0^x |\cos y|\, dy \le x$$

for all $x \ge 0$. Since $\sin x$ is odd, $|\sin x| \le |x|$ for all $x \in \mathbb{R}$. Using this in Eq. (16.3) implies that

$$\begin{aligned}\|\Lambda_n\| &\ge \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\frac{\sin(n+\frac{1}{2})\theta}{\frac{1}{2}\theta}\right| d\theta = \frac{2}{\pi}\int_0^{\pi}\left|\sin(n+\frac{1}{2})\theta\right|\frac{d\theta}{\theta}.\\ &= \frac{2}{\pi}\int_0^{\pi}\left|\sin(n+\frac{1}{2})\theta\right|\frac{d\theta}{\theta}\end{aligned}$$

Since

$$\int_0^{\pi}\left|\sin(n+\frac{1}{2})\theta\right|\frac{d\theta}{\theta} = \int_0^{(n+\frac{1}{2})\pi}|\sin y|\frac{dy}{y} \to \infty \text{ as } n \to \infty,$$

we learn that $\sup_n \|\Lambda_n\| = \infty$. So by Theorem 16.8,

$$R_1 = \{f \in C(T) : \sup_n |\Lambda_n f| = \infty\}$$

is a residual set. ■

See Rudin Chapter 5 for more details.

Lemma 16.11. *For $f \in L^1(T)$, let*

$$\hat{f}(n) := \langle f, \phi_n\rangle = \int_T f(w)\bar{w}^n dw.$$

Then $\hat{f} \in c_0$ and the map $f \in L^1(T) \to c_0$ is a one to one bounded linear transformation into but not onto c_0.

Proof. By Bessel's inequality, $\sum_{n\in\mathbb{Z}}\left|\hat{f}(n)\right|^2 < \infty$ for all $f \in L^2(T)$ and in particular $\lim_{|n|\to\infty}\left|\hat{f}(n)\right| = 0$. Given $f \in L^1(T)$ and $g \in L^2(T)$ we have

$$\left|\hat{f}(n) - \hat{g}(n)\right| = \left|\int_T [f(w) - g(w)]\,\bar{w}^n dw\right| \le \|f - g\|_1$$

and hence

$$\limsup_{n\to\infty}\left|\hat{f}(n)\right| = \limsup_{n\to\infty}\left|\hat{f}(n) - \hat{g}(n)\right| \le \|f - g\|_1$$

for all $g \in L^2(T)$. Since $L^2(T)$ is dense in $L^1(T)$, it follows that $\limsup_{n\to\infty}\left|\hat{f}(n)\right| = 0$ for all $f \in L^1$, i.e. $\hat{f} \in c_0$.

Since $\left|\hat{f}(n)\right| \le \|f\|_1$, we have $\left\|\hat{f}\right\|_{c_0} \le \|f\|_1$ showing that $\Lambda f := \hat{f}$ is a bounded linear transformation from $L^1(T)$ to c_0.

To see that Λ is injective, suppose $\hat{f} = \Lambda f \equiv 0$, then $\int_T f(w)p(w,\bar{w})dw = 0$ for all polynomials p in w and $\bar{w}$. By the Stone - Wierestrass and the dominated convergence theorem, this implies that

$$\int_T f(w)g(w)dw = 0$$

for all $g \in C(T)$. Lemma 9.7 now implies $f = 0$ a.e.

If Λ were surjective, the open mapping theorem would imply that $\Lambda^{-1} : c_0 \to L^1(T)$ is bounded. In particular this implies there exists $C < \infty$ such that

$$\|f\|_{L^1} \le C \left\|\hat{f}\right\|_{c_0} \text{ for all } f \in L^1(T). \tag{16.4}$$

Taking $f = d_n$, we find $\left\|\hat{d}_n\right\|_{c_0} = 1$ while $\lim_{n\to\infty} \|d_n\|_{L^1} = \infty$ contradicting Eq. (16.4). Therefore $\mathrm{ran}(\Lambda) \neq c_0$. ■

16.2. **Hahn Banach Theorem.** Our next goal is to show that continuous dual X^* of a Banach space X is always large. This will be the content of the Hahn – Banach Theorem 16.15 below.

Proposition 16.12. *Let X be a complex vector space over $\mathbb{C}$. If $f \in X^*$ and $u = \mathrm{Re} f \in X^*_{\mathbb{R}}$ then*

$$f(x) = u(x) - iu(ix). \tag{16.5}$$

Conversely if $u \in X^*_{\mathbb{R}}$ and f is defined by Eq. (16.5), then $f \in X^*$ and $\|u\|_{X^*_{\mathbb{R}}} = \|f\|_{X^*}$. More generally if p is a semi-norm on X, then

$$|f| \le p \text{ iff } u \le p.$$

Proof. Let $v(x) = \mathrm{Im}\ f(x)$, then

$$v(ix) = \mathrm{Im}\ f(ix) = \mathrm{Im}(if(x)) = \mathrm{Re} f(x) = u(x).$$

Therefore

$$f(x) = u(x) + iv(x) = u(x) + iu(-ix) = u(x) - iu(ix).$$

Conversely for $u \in X^*_{\mathbb{R}}$ let $f(x) = u(x) - iu(ix)$. Then

$$f((a+ib)x) = u(ax+ibx) - iu(iax - bx) = au(x) + bu(ix) - i(au(ix) - bu(x))$$

while

$$(a+ib)f(x) = au(x) + bu(ix) + i(bu(x) - au(ix)).$$

So f is complex linear.

Because $|u(x)| = |\mathrm{Re} f(x)| \le |f(x)|$, it follows that $\|u\| \le \|f\|$. For $x \in X$ choose $\lambda \in S^1 \subset \mathbb{C}$ such that $|f(x)| = \lambda f(x)$ so

$$|f(x)| = f(\lambda x) = u(\lambda x) \le \|u\|\ \|\lambda x\| = \|u\|\|x\|.$$

Since $x \in X$ is arbitrary, this shows that $\|f\| \le \|u\|$ so $\|f\| = \|u\|$.[34]

34

Proof. To understand better why $\|f\| = \|u\|$, notice that

$$\|f\|^2 = \sup_{\|x\|=1} |f(x)|^2 = \sup_{\|x\|=1} (|u(x)|^2 + |u(ix)|^2).$$

Supppose that $M = \sup_{\|x\|=1} |u(x)|$ and this supremum is attained at $x_0 \in X$ with $\|x_0\| = 1$. Replacing x_0 by $-x_0$ if necessary, we may assume that $u(x_0) = M$. Since u has a maximum at x_0,

$$\begin{aligned} 0 &= \frac{d}{dt}\Big|_0 u\left(\frac{x_0 + itx_0}{\|x_0 + itx_0\|}\right) \\ &= \frac{d}{dt}\Big|_0 \left\{\frac{1}{|1+it|}\ (u(x_0) + tu(ix_0))\right\} = u(ix_0) \end{aligned}$$

since $\frac{d}{dt}|_0 |1+it| = \frac{d}{dt}|_0 \sqrt{1+t^2} = 0$.This explains why $\|f\| = \|u\|$. ■

For the last assertion, it is clear that $|f| \leq p$ implies that $u \leq |u| \leq |f| \leq p$. Conversely if $u \leq p$ and $x \in X$, choose $\lambda \in S^1 \subset \mathbb{C}$ such that $|f(x)| = \lambda f(x)$. Then

$$|f(x)| = \lambda f(x) = f(\lambda x) = u(\lambda x) \leq p(\lambda x) = p(x)$$

holds for all $x \in X$. ■

Definition 16.13 (Minkowski functional)**.** $p : X \to \mathbb{R}$ is a Minkowski functional if

(1) $p(x + y) \leq p(x) + p(y)$ for all $x, y \in X$ and
(2) $p(cx) = cp(x)$ for all $c \geq 0$ and $x \in X$.

Example 16.14. Suppose that $X = \mathbb{R}$ and

$$p(x) = \inf \{\lambda \geq 0 : x \in \lambda[-1, 2] = [-\lambda, 2\lambda]\}.$$

Notice that if $x \geq 0$, then $p(x) = x/2$ and if $x \leq 0$ then $p(x) = -x$, i.e.

$$p(x) = \begin{cases} x/2 & \text{if} \quad x \geq 0 \\ |x| & \text{if} \quad x \leq 0. \end{cases}$$

From this formula it is clear that $p(cx) = cp(x)$ for all $c \geq 0$ but not for $c < 0$. Moreover, p satisfies the triangle inequality, indeed if $p(x) = \lambda$ and $p(y) = \mu$, then $x \in \lambda[-1, 2]$ and $y \in \mu[-1, 2]$ so that

$$x + y \in \lambda[-1, 2] + \mu[-1, 2] \subset (\lambda + \mu)[-1, 2]$$

which shows that $p(x + y) \leq \lambda + \mu = p(x) + p(y)$. To check the last set inclusion let $a, b \in [-1, 2]$, then

$$\lambda a + \mu b = (\lambda + \mu) \left(\frac{\lambda}{\lambda + \mu} a + \frac{\mu}{\lambda + \mu} b \right) \in (\lambda + \mu)[-1, 2]$$

since $[-1, 2]$ is a convex set and $\frac{\lambda}{\lambda+\mu} + \frac{\mu}{\lambda+\mu} = 1$.

TODO: Add in the relationship to convex sets and separation theorems, see Reed and Simon Vol. 1. for example.

Theorem 16.15 (Hahn-Banach)**.** *Let X be a real vector space, $M \subset X$ be a subspace $f : M \to \mathbb{R}$ be a linear functional such that $f \leq p$ on M. Then there exists a linear functional $F : X \to \mathbb{R}$ such that $F|_M = f$ and $F \leq p$.*

Proof. Step (1) We show for all $x \in X \setminus M$ there exists and extension F to $M \oplus \mathbb{R}x$ with the desired properties. If F exists and $\alpha = F(x)$, then for all $y \in M$ and $\lambda \in \mathbb{R}$ we must have $f(y)+\lambda\alpha = F(y+\lambda x) \leq p(y+\lambda x)$ i.e. $\lambda\alpha \leq p(y+\lambda x)-f(y)$. Equivalently put we must find $\alpha \in \mathbb{R}$ such that

$$\alpha \leq \frac{p(y + \lambda x) - f(y)}{\lambda} \text{ for all } y \in M \text{ and } \lambda > 0$$
$$\alpha \geq \frac{p(z - \mu x) - f(z)}{\mu} \text{ for all } z \in M \text{ and } \mu > 0.$$

So if $\alpha \in \mathbb{R}$ is going to exist, we have to prove, for all $y, z \in M$ and $\lambda, \mu > 0$ that

$$\frac{f(z) - p(z - \mu x)}{\mu} \leq \frac{p(y + \lambda x) - f(y)}{\lambda}$$

or equivalently

$$f(\lambda z + \mu y) \leq \mu p(y + \lambda x) + \lambda p(z - \mu x) \tag{16.6}$$
$$= p(\mu y + \mu\lambda x) + p(\lambda z - \lambda\mu x).$$

But

$$\begin{aligned} f(\lambda z+\mu y) &= f(\lambda z+\mu\lambda x)+f(\lambda z-\lambda\mu x)\\ &\le p(\lambda z+\mu\lambda x)+p(\lambda z-\lambda\mu x)\end{aligned}$$

which shows that Eq. (16.6) is true and by working backwards, there exist an $\alpha\in\mathbb{R}$ such that $f(y)+\lambda\alpha\le p(y+\lambda x)$. Therefore $F(y+\lambda x):=f(y)+\lambda\alpha$ is the desired extension.

Step (2) Let us now write $F:X\to\mathbb{R}$ to mean F is defined on a linear subspace $D(F)\subset X$ and $F:D(F)\to\mathbb{R}$ is linear. For $F,G:X\to\mathbb{R}$ we will say $F<G$ if $D(F)\subset D(G)$ and $F=G|_{D(F)}$, that is G is an extension of F. Let

$$\mathcal{F}=\{F:X\to\mathbb{R}:M\subset D(F),\ F\le p\text{ on }D(F)\}.$$

Then $(\mathcal{F},<)$ is a partially ordered set. If $\Phi\subset\mathcal{F}$ is a chain (i.e. a linearly ordered subset of $\mathcal{F}$) then Φ has an upper bound $G\in\mathcal{F}$ defined by $D(G)=\bigcup\limits_{F\in\Phi}D(F)$ and $G(x)=F(x)$ for $x\in D(F)$. Then it is easily checked that $D(G)$ is a linear subspace, $G\in\mathcal{F}$, and $F<G$ for all $F\in\Phi$. We may now apply Zorn's Lemma (see Theorem B.7) to conclude there exists a maximal element $F\in\mathcal{F}$. Necessarily, $D(F)=X$ for otherwise we could extend F by step (1), violating the maximality of F. Thus F is the desired extension of f. ■

Corollary 16.16. *Suppose that X is a complex vector space, $p:X\to[0,\infty)$ is a semi-norm, $M\subset X$ is a linear subspace, and $f:M\to\mathbb{C}$ is linear functional such that $|f(x)|\le p(x)$ for all $x\in M$. Then there exists $F\in X'$ (X' is the **algebraic** dual of X) such that $F|_M=f$ and $|F|\le p$.*

Proof. Let $u=\mathrm{Re}f$ then $u\le p$ on M and hence by Theorem 16.15, there exists $U\in X'_{\mathbb{R}}$ such that $U|_M=u$ and $U\le p$ on M. Define $F(x)=U(x)-iU(ix)$ then as in Proposition 16.12, $F=f$ on M and $|F|\le p$. ■

Theorem 16.17. *Let X be a normed space $M\subset X$ be a closed subspace and $x\in X\setminus M$. Then there exists $f\in X^*$ such that $\|f\|=1, f(x)=\delta=d(x,M)$ and $f=0$ on M.*

Proof. Define $f:M\oplus\mathbb{C}x\to\mathbb{C}$ by $f(m+\lambda x)\equiv\lambda\delta$ for all $m\in M$ and $\lambda\in\mathbb{C}$. Notice that

$$\|m+\lambda x\|=|\lambda|\|x+m/\lambda\|\ge|\lambda|\delta$$

and hence

$$|f(m+\lambda x)|=|\lambda|\delta\le\|m+\lambda x\|$$

which shows $\|f\|\le 1$. In fact, since $|f(m+x)|=\delta=\inf\limits_{m\in M}\|x+m\|$, $\|f\|=1$. By Hahn-Banach theorem there exists $F\in X^*$ such that $F|_{M\oplus\mathbb{C}x}=f$ and $|F(x)|\le\|x\|$ for all $x\in X$, i.e. $\|F\|\le 1$. Since $1=\|f\|\le\|F\|\le 1$ we see $\|F\|=\|f\|$. ■

Corollary 16.18. *The linear map $x\in X\to\hat{x}\in X^{**}$ where $\hat{x}(f)=f(x)$ for all $x\in X$ is an isometry. (This isometry need not be surjective.)*

Proof. Since $|\hat{x}(f)|=|f(x)|\le\|f\|_{X^*}\|x\|_X$ for all $f\in X^*$, it follows that $\|\hat{x}\|_{X^{**}}\le\|x\|_X$. Now applying Theorem 16.17 with $M=\{0\}$, there exists $f\in X^*$ such that $\|f\|=1$ and $|\hat{x}(f)|=f(x)=\|x\|$, which shows that $\|\hat{x}\|_{X^{**}}\ge\|x\|_X$. This shows that $x\in X\to\hat{x}\in X^{**}$ is an isometry. Since isometries are necessarily injective, we are done. ■

Definition 16.19. A Banach space X is reflexive if the map $x \in X \to \hat{x} \in X^{**}$ is surjective.

Example 16.20. Every Hilbert space H is reflexive. This is a consequence of the Riesz Theorem, Proposition 10.15.

Example 16.21. Suppose that μ is a $\sigma-$ finite measure on a measurable space $(X, \mathcal{M})$, then $L^p(X, \mathcal{M}, \mu)$ is reflexive for all $p \in (1, \infty)$, see Theorem 13.14.

Example 16.22 (Following Riesz and Nagy, p. 214)**.** The Banach space $X := C([0,1])$ is not reflexive. To prove this recall that X^* may be identified with complex measures μ on $[0,1]$ which may be identified with right continuous functions of bounded variation (F) on $[0,1]$, namely

$$F \to \mu_F \to (f \in X \to \int_{[0,1]} f d\mu_F = \int_0^1 f dF).$$

Define $\lambda \in X^{**}$ by

$$\lambda(\mu) = \sum_{x \in [0,1]} \mu(\{x\}) = \sum_{x \in [0,1]} (F(x) - F(x-)),$$

so $\lambda(\mu)$ is the sum of the "atoms" of μ. Suppose there existed an $f \in X$ such that $\lambda(\mu) = \int_{[0,1]} f d\mu$ for all $\mu \in X^*$. Choosing $\mu = \delta_x$ for some $x \in (0,1)$ would then imply that

$$f(x) = \int_{[0,1]} f\delta_x = \lambda(\delta_x) = 1$$

showing f would have to be the constant function,1, which clearly can not work.

Example 16.23. The Banach space $X := L^1(\mathbb{R}, m)$ is not reflexive. As we have seen in Theorem 13.14, $X^* \cong L^\infty(\mathbb{R}, m)$. The argument in Example 13.15 shows $(L^\infty(\mathbb{R}, m))^* \ncong L^1(\mathbb{R}, m)$.

16.3. **Weak and Strong Topologies.**

Definition 16.24. Let X and Y be be a normed vector spaces and $L(X, Y)$ the normed space of bounded linear transformations from X to Y.

(1) The **weak topology** on X is the topology generated by X^*, i.e. sets of the form
$$N = \cap_{i=1}^n \{x \in X : |f_i(x) - f_i(x_0)| < \epsilon\}$$
where $f_i \in X^*$ and $\epsilon > 0$ form a neighborhood base for the weak topology on X at x_0.

(2) The **weak-$*$ topology** on X^* is the topology generated by X, i.e.
$$N \equiv \cap_{i=1}^n \{g \in X^* : |f(x_i) - g(x_i)| < \epsilon\}$$
where $x_i \in X$ and $\epsilon > 0$ forms a neighborhood base for the weak$-*$ topology on X^* at $f \in X^*$.

(3) The **strong operator topology** on $L(X, Y)$ is the smallest topology such that $T \in L(X, Y) \longrightarrow Tx \in Y$ is continuous for all $x \in X$.

(4) The **weak operator topology** on $L(X, Y)$ is the smallest topology such that $T \in L(X, Y) \longrightarrow f(Tx) \in \mathbb{C}$ is continuous for all $x \in X$ and $f \in Y^*$.

Theorem 16.25 (Alaoglu's Theorem)**.** *If X is a normed space the unit ball in X^* is weak - $*$ compact.*

Proof. For all $x \in X$ let $D_x = \{z \in \mathbb{C} : |z| \leq \|x\|\}$. Then $D_x \subset \mathbb{C}$ is a compact set and so by Tychonoff's Theorem $\Omega \equiv \prod_{x \in X} D_x$ is compact in the product topology. If $f \in \overline{B} := \{f \in X^* : \|f\| \leq 1\}$, $|f(x)| \leq \|f\| \, \|x\| \leq \|x\|$ which implies that $f(x) \in D_x$ for all $x \in X$, i.e. $\overline{B} \subset \Omega$. The topology on $\bar{B}$ inherited from the weak–$*$ topology on X^* is the same as that relative topology coming from the product topology on Ω. So to finish the proof it suffices to show $\bar{B}$ is a closed subset of the compact space Ω. To prove this let $\pi_x(f) = f(x)$ be the projection maps. Then

$$\begin{aligned}
\overline{B} &= \{f \in \Omega : f \text{ is linear}\} \\
&= \{f \in \Omega : f(x+cy) - f(x) - cf(y) = 0 \text{ for all } x, y \in X \text{ and } c \in \mathbb{C}\} \\
&= \bigcap_{x,y \in X} \bigcap_{c \in \mathbb{C}} \{f \in \Omega : f(x+cy) - f(x) - cf(y) = 0\} \\
&= \bigcap_{x,y \in X} \bigcap_{c \in \mathbb{C}} \left(\pi_{x+cy} - \pi_x - c\pi_y\right)^{-1}(\{0\})
\end{aligned}$$

which is closed because $(\pi_{x+cy} - \pi_x - c\pi_y) : \Omega \to \mathbb{C}$ is continuous. ■

Theorem 16.26 (Alaoglu's Theorem for separable spaces)**.** *Suppose that X is a separable Banach space, $C^* := \{f \in X^* : \|f\| \leq 1\}$ is the closed unit ball in X^* and $\{x_n\}_{n=1}^\infty$ is an countable dense subset of $C := \{x \in X : \|x\| \leq 1\}$. Then*

$$\rho(f, g) := \sum_{n=1}^{\infty} \frac{1}{2^n} |f(x_n) - g(x_n)| \tag{16.7}$$

defines a metric on C^* which is compatible with the weak topology on C^*, $\tau_{C^*} := (\tau_{w^*})_{C^*} = \{V \cap C : V \in \tau_{w^*}\}$. Moreover (C^*, ρ) is a compact metric space.

Proof. The routine check that ρ is a metric is left to the reader. Let τ_ρ be the topology on C^* induced by ρ. For any $g \in X$ and $n \in \mathbb{N}$, the map $f \in X^* \to (f(x_n) - g(x_n)) \in \mathbb{C}$ is τ_{w^*} continuous and since the sum in Eq. (16.7) is uniformly convergent for $f \in C^*$, it follows that $f \to \rho(f, g)$ is τ_{C^*} – continuous. This implies the open balls relative to ρ are contained in τ_{C^*} and therefore $\tau_\rho \subset \tau_{C^*}$.

We now wish to prove $\tau_{C^*} \subset \tau_\rho$. Since τ_{C^*} is the topology generated by $\{\hat{x}|_{C^*} : x \in C\}$, it suffices to show $\hat{x}$ is τ_ρ – continuous for all $x \in C$. But given $x \in C$ there exists a subsequence $y_k := x_{n_k}$ of $\{x_n\}_{n=1}^\infty$ such that such that $x = \lim_{k \to \infty} y_k$. Since

$$\sup_{f \in C^*} |\hat{x}(f) - \hat{y}_k(f)| = \sup_{f \in C^*} |f(x - y_k)| \leq \|x - y_k\| \to 0 \text{ as } k \to \infty,$$

$\hat{y}_k \to \hat{x}$ uniformly on C^* and using $\hat{y}_k$ is τ_ρ – continuous for all k (as is easily checked) we learn $\hat{x}$ is also τ_ρ continuous. Hence $\tau_{C^*} = \tau(\hat{x}|_{C^*} : x \in X) \subset \tau_\rho$.

The compactness assertion follows from Theorem 16.25. The compactness assertion may also be verified directly using: 1) sequential compactness is equivalent to compactness for metric spaces and 2) a Cantor's diagonalization argument as in the proof of Theorem 10.35. ■

16.4. **Supplement: Quotient spaces, adjoints, and more reflexivity.**

Definition 16.27. Let X and Y be Banach spaces and $A : X \to Y$ be a linear operator. The **transpose** of A is the linear operator $A^\dagger : Y^* \to X^*$ defined by

$\left(A^{\dagger}f\right)(x) = f(Ax)$ for $f \in Y^*$ and $x \in X$. The **null space** of A is the subspace $\mathrm{nul}(A) := \{x \in X : Ax = 0\} \subset X$. For $M \subset X$ and $N \subset X^*$ let

$$M^0 := \{f \in X^* : f|_M = 0\} \text{ and}$$
$$N^{\perp} := \{x \in X : f(x) = 0 \text{ for all } f \in N\}.$$

Proposition 16.28 (Basic Properties)**.**

(1) *$\|A\| = \left\|A^{\dagger}\right\|$ and $A^{\dagger\dagger}\hat{x} = \widehat{Ax}$ for all $x \in X$.*
(2) *M^0 and $N^{\perp}$ are always closed subspace of X^* and X respectively.*
(3) *$\left(M^0\right)^{\perp} = \bar{M}$.*
(4) *$\bar{N} \subset \left(N^{\perp}\right)^0$ with equality when X is reflexive.*
(5) *$\mathrm{nul}(A) = \mathrm{ran}(A^{\dagger})^{\perp}$ and $\mathrm{nul}(A^{\dagger}) = \mathrm{ran}(A)^0$. Moreover, $\overline{\mathrm{ran}(A)} = \mathrm{nul}(A^{\dagger})^{\perp}$ and if X is reflexive, then $\overline{\mathrm{ran}(A^{\dagger})} = \mathrm{nul}(A)^0$.*
(6) *X is reflexive iff X^* is reflexive. More generally $X^{***} = \widehat{X^*} \oplus \hat{X}^0$.*

Proof.

(1)

$$\|A\| = \sup_{\|x\|=1} \|Ax\| = \sup_{\|x\|=1} \sup_{\|f\|=1} |f(Ax)|$$
$$= \sup_{\|f\|=1} \sup_{\|x\|=1} \left|A^{\dagger}f(x)\right| = \sup_{\|f\|=1} \left\|A^{\dagger}f\right\| = \left\|A^{\dagger}\right\|.$$

(2) This is an easy consequence of the assumed continuity off all linear functionals involved.
(3) If $x \in M$, then $f(x) = 0$ for all $f \in M^0$ so that $x \in \left(M^0\right)^{\perp}$. Therefore $\bar{M} \subset \left(M^0\right)^{\perp}$. If $x \notin \bar{M}$, then there exists $f \in X^*$ such that $f|_M = 0$ while $f(x) \neq 0$, i.e. $f \in M^0$ yet $f(x) \neq 0$. This shows $x \notin \left(M^0\right)^{\perp}$ and we have shown $\left(M^0\right)^{\perp} \subset \bar{M}$.
(4) It is again simple to show $N \subset \left(N^{\perp}\right)^0$ and therefore $\bar{N} \subset \left(N^{\perp}\right)^0$. Moreover, as above if $f \notin \bar{N}$ there exists $\psi \in X^{**}$ such that $\psi|_{\bar{N}} = 0$ while $\psi(f) \neq 0$. If X is reflexive, $\psi = \hat{x}$ for some $x \in X$ and since $g(x) = \psi(g) = 0$ for all $g \in \bar{N}$, we have $x \in N^{\perp}$. On the other hand, $f(x) = \psi(f) \neq 0$ so $f \notin \left(N^{\perp}\right)^0$. Thus again $\left(N^{\perp}\right)^0 \subset \bar{N}$.
(5)

$$\mathrm{nul}(A) = \{x \in X : Ax = 0\} = \{x \in X : f(Ax) = 0 \ \forall \ f \in X^*\}$$
$$= \left\{x \in X : A^{\dagger}f(x) = 0 \ \forall \ f \in X^*\right\}$$
$$= \left\{x \in X : g(x) = 0 \ \forall \ g \in \mathrm{ran}(A^{\dagger})\right\} = \mathrm{ran}(A^{\dagger})^{\perp}.$$

Similarly,

$$\mathrm{nul}(A^{\dagger}) = \left\{f \in Y^* : A^{\dagger}f = 0\right\} = \left\{f \in Y^* : (A^{\dagger}f)(x) = 0 \ \forall \ x \in X\right\}$$
$$= \{f \in Y^* : f(Ax) = 0 \ \forall \ x \in X\}$$
$$= \left\{f \in Y^* : f|_{\mathrm{ran}(A)} = 0\right\} = \mathrm{ran}(A)^0.$$

(6) Let $\psi \in X^{***}$ and define $f_{\psi} \in X^*$ by $f_{\psi}(x) = \psi(\hat{x})$ for all $x \in X$ and set $\psi' := \psi - \hat{f}_{\psi}$. For $x \in X$ (so $\hat{x} \in X^{**}$) we have

$$\psi'(\hat{x}) = \psi(\hat{x}) - \hat{f}_{\psi}(\hat{x}) = f_{\psi}(x) - \hat{x}(f_{\psi}) = f_{\psi}(x) - f_{\psi}(x) = 0.$$

This shows $\psi' \in \hat{X}^0$ and we have shown $X^{***} = \widehat{X^*} + \hat{X}^0$. If $\psi \in \widehat{X^*} \cap \hat{X}^0$, then $\psi = \hat{f}$ for some $f \in X^*$ and $0 = \hat{f}(\hat{x}) = \hat{x}(f) = f(x)$ for all $x \in X$, i.e. $f = 0$ so $\psi = 0$. Therefore $X^{***} = \widehat{X^*} \oplus \hat{X}^0$ as claimed. If X is reflexive, then $\hat{X} = X^{**}$ and so $\hat{X}^0 = \{0\}$ showing $X^{***} = \widehat{X^*}$, i.e. X^* is reflexive. Conversely if X^* is reflexive we conclude that $\hat{X}^0 = \{0\}$ and therefore $X^{**} = \{0\}^\perp = \left(\hat{X}^0\right)^\perp = \hat{X}$, so that X is reflexive.

Alternative proof. Notice that $f_\psi = J^\dagger\psi$, where $J : X \to X^{**}$ is given by $Jx = \hat{x}$, and the composition

$$f \in X^* \xrightarrow{\hat{}} \hat{f} \in X^{***} \xrightarrow{J^\dagger} J^\dagger \hat{f} \in X^*$$

is the identity map since $\left(J^\dagger \hat{f}\right)(x) = \hat{f}(Jx) = \hat{f}(\hat{x}) = \hat{x}(f) = f(x)$ for all $x \in X$. Thus it follows that $X^* \xrightarrow{\hat{}} X^{***}$ is invertible iff $J^\dagger$ is its inverse which can happen iff $\text{nul}(J^\dagger) = \{0\}$. But as above $\text{nul}(J^\dagger) = \text{ran}(J)^0$ which will be zero iff $\overline{\text{ran}(J)} = X^{**}$ and since J is an isometry this is equivalent to saying $\text{ran}(J) = X^{**}$. So we have again shown X^* is reflexive iff X is reflexive.

■

Theorem 16.29. *Let X be a Banach space, $M \subset X$ be a proper closed subspace, X/M the quotient space, $\pi : X \to X/M$ the projection map $\pi(x) = x + M$ for $x \in X$ and define the quotient norm on X/M by*

$$\|\pi(x)\|_{X/M} = \|x + M\|_{X/M} = \inf_{m \in M} \|x + m\|_X .$$

Then

1. $\|\cdot\|_{X/M}$ *is a norm on X/M.*
2. *The projection map $\pi : X \to X/M$ has norm 1, $\|\pi\| = 1$.*
3. $(X/M, \|\cdot\|_{X/M})$ *is a Banach space.*
4. *If Y is another normed space and $T : X \to Y$ is a bounded linear transformation such that $M \subset \text{nul}(T)$, then there exists a unique linear transformation $S : X/M \to Y$ such that $T = S \circ \pi$ and moreover $\|T\| = \|S\|$.*

Proof. 1) Clearly $\|x + M\| \geq 0$ and if $\|x + M\| = 0$, then there exists $m_n \in M$ such that $\|x + m_n\| \to 0$ as $n \to \infty$, i.e. $x = \lim_{n\to\infty} m_n \in \bar{M} = M$. Since $x \in M$, $x + M = 0 \in X/M$. If $c \in \mathbb{C}\setminus\{0\}$, $x \in X$, then

$$\|cx + M\| = \inf_{m \in M} \|cx + m\| = |c| \inf_{m \in M} \|x + m/c\| = |c| \, \|x + M\|$$

because m/c runs through M as m runs through M. Let $x_1, x_2 \in X$ and $m_1, m_2 \in M$ then

$$\|x_1 + x_2 + M\| \leq \|x_1 + x_2 + m_1 + m_2\| \leq \|x_1 + m_1\| + \|x_2 + m_2\|.$$

Taking infinums over $m_1, m_2 \in M$ then implies

$$\|x_1 + x_2 + M\| \leq \|x_1 + M\| + \|x_2 + M\|.$$

and we have completed the proof the $(X/M, \|\cdot\|)$ is a normed space.

2) Since $\|\pi(x)\| = \inf_{m \in M} \|x + m\| \leq \|x\|$ for all $x \in X$, $\|\pi\| \leq 1$. To see $\|\pi\| = 1$, let $x \in X \setminus M$ so that $\pi(x) \neq 0$. Given $\alpha \in (0, 1)$, there exists $m \in M$ such that

$$\|x + m\| \leq \alpha^{-1} \|\pi(x)\| .$$

Therefore,

$$\frac{\|\pi(x+m)\|}{\|x+m\|} = \frac{\|\pi(x)\|}{\|x+m\|} \geq \frac{\alpha \|x+m\|}{\|x+m\|} = \alpha$$

which shows $\|\pi\| \geq \alpha$. Since $\alpha \in (0,1)$ is arbitrary we conclude that $\|\pi(x)\| = 1$.

3) Let $\pi(x_n) \in X/M$ be a sequence such that $\sum \|\pi(x_n)\| < \infty$. As above there exists $m_n \in M$ such that $\|\pi(x_n)\| \geq \frac{1}{2}\|x_n + m_n\|$ and hence $\sum \|x_n + m_n\| \leq 2\sum \|\pi(x_n)\| < \infty$. Since X is complete, $x := \sum_{n=1}^{\infty}(x_n+m_n)$ exists in X and therefore by the continuity of π,

$$\pi(x) = \sum_{n=1}^{\infty} \pi(x_n + m_n) = \sum_{n=1}^{\infty} \pi(x_n)$$

showing X/M is complete.

4) The existence of S is guaranteed by the "factor theorem" from linear algebra. Moreover $\|S\| = \|T\|$ because

$$\|T\| = \|S \circ \pi\| \leq \|S\| \|\pi\| = \|S\|$$

and

$$\begin{aligned}\|S\| &= \sup_{x \notin M} \frac{\|S(\pi(x))\|}{\|\pi(x)\|} = \sup_{x \notin M} \frac{\|Tx\|}{\|\pi(x)\|} \\ &\geq \sup_{x \notin M} \frac{\|Tx\|}{\|x\|} = \sup_{x \neq 0} \frac{\|Tx\|}{\|x\|} = \|T\|.\end{aligned}$$

■

Theorem 16.30. *Let X be a Banach space. Then*

(1) *Identifying X with $\hat{X} \subset X^{**}$, the weak – $*$ topology on X^{**} induces the weak topology on X. More explicitly, the map $x \in X \to \hat{x} \in \hat{X}$ is a homeomorphism when X is equipped with its weak topology and $\hat{X}$ with the relative topology coming from the weak-$*$ topology on X^{**}.*
(2) *$\hat{X} \subset X^{**}$ is dense in the weak-$*$ topology on X^{**}.*
(3) *Letting C and C^{**} be the closed unit balls in X and X^{**} respectively, then $\hat{C} := \{\hat{x} \in C^{**} : x \in C\}$ is dense in C^{**} in the weak – $*$ topology on X^{**}.*
(4) *X is reflexive iff C is weakly compact.*

Proof.

(1) The weak $-$ $*$ topology on X^{**} is generated by

$$\left\{\hat{f} : f \in X^*\right\} = \{\psi \in X^{**} \to \psi(f) : f \in X^*\}.$$

So the induced topology on X is generated by

$$\{x \in X \to \hat{x} \in X^{**} \to \hat{x}(f) = f(x) : f \in X^*\} = X^*$$

and so the induced topology on X is precisely the weak topology.

(2) A basic weak - $*$ neighborhood of a point $\lambda \in X^{**}$ is of the form

$$\mathcal{N} := \cap_{k=1}^{n} \{\psi \in X^{**} : |\psi(f_k) - \lambda(f_k)| < \epsilon\} \tag{16.8}$$

for some $\{f_k\}_{k=1}^n \subset X^*$ and $\epsilon > 0$. be given. We must now find $x \in X$ such that $\hat{x} \in \mathcal{N}$, or equivalently so that

$$|\hat{x}(f_k) - \lambda(f_k)| = |f_k(x) - \lambda(f_k)| < \epsilon \text{ for } k = 1, 2, \ldots, n. \tag{16.9}$$

In fact we will show there exists $x \in X$ such that $\lambda(f_k) = f_k(x)$ for $k = 1, 2, \ldots, n$. To prove this stronger assertion we may, by discarding some of the f_k's if necessary, assume that $\{f_k\}_{k=1}^n$ is a linearly independent set. Since the $\{f_k\}_{k=1}^n$ are linearly independent, the map $x \in X \to (f_1(x), \ldots, f_n(x)) \in \mathbb{C}^n$ is surjective (why) and hence there exists $x \in X$ such that

$$(f_1(x), \ldots, f_n(x)) = Tx = (\lambda(f_1), \ldots, \lambda(f_n)) \tag{16.10}$$

as desired.

(3) Let $\lambda \in C^{**} \subset X^{**}$ and $\mathcal{N}$ be the weak - $*$ open neighborhood of λ as in Eq. (16.8). Working as before, given $\epsilon > 0$, we need to find $x \in C$ such that Eq. (16.9). It will be left to the reader to verify that it suffices again to assume $\{f_k\}_{k=1}^n$ is a linearly independent set. (Hint: Suppose that $\{f_1, \ldots, f_m\}$ were a maximal linearly dependent subset of $\{f_k\}_{k=1}^n$, then each f_k with $k > m$ may be written as a linear combination $\{f_1, \ldots, f_m\}$.) As in the proof of item 2., there exists $x \in X$ such that Eq. (16.10) holds. The problem is that x may not be in C. To remedy this, let $N := \cap_{k=1}^n \mathrm{nul}(f_k) = \mathrm{nul}(T)$, $\pi : X \to X/N \cong \mathbb{C}^n$ be the projection map and $\bar{f}_k \in (X/N)^*$ be chosen so that $f_k = \bar{f}_k \circ \pi$ for $k = 1, 2, \ldots, n$. Then we have produced $x \in X$ such that

$$(\lambda(f_1), \ldots, \lambda(f_n)) = (f_1(x), \ldots, f_n(x)) = (\bar{f}_1(\pi(x)), \ldots, \bar{f}_n(\pi(x))).$$

Since $\{\bar{f}_1, \ldots, \bar{f}_n\}$ is a basis for $(X/N)^*$ we find

$$\begin{aligned}\|\pi(x)\| &= \sup_{\alpha \in \mathbb{C}^n \setminus \{0\}} \frac{\left|\sum_{i=1}^n \alpha_i \bar{f}_i(\pi(x))\right|}{\left\|\sum_{i=1}^n \alpha_i \bar{f}_i\right\|} = \sup_{\alpha \in \mathbb{C}^n \setminus \{0\}} \frac{\left|\sum_{i=1}^n \alpha_i \lambda(f_i)\right|}{\left\|\sum_{i=1}^n \alpha_i f_i\right\|} \\ &= \sup_{\alpha \in \mathbb{C}^n \setminus \{0\}} \frac{\left|\lambda(\sum_{i=1}^n \alpha_i f_i)\right|}{\left\|\sum_{i=1}^n \alpha_i f_i\right\|} \le \|\lambda\| \sup_{\alpha \in \mathbb{C}^n \setminus \{0\}} \frac{\left\|\sum_{i=1}^n \alpha_i f_i\right\|}{\left\|\sum_{i=1}^n \alpha_i f_i\right\|} = 1.\end{aligned}$$

Hence we have shown $\|\pi(x)\| \le 1$ and therefore for any $\alpha > 1$ there exists $y = x + n \in X$ such that $\|y\| < \alpha$ and $(\lambda(f_1), \ldots, \lambda(f_n)) = (f_1(y), \ldots, f_n(y))$. Hence

$$|\lambda(f_i) - f_i(y/\alpha)| \le \left|f_i(y) - \alpha^{-1} f_i(y)\right| \le (1 - \alpha^{-1}) |f_i(y)|$$

which can be arbitrarily small (i.e. less than ϵ) by choosing α sufficiently close to 1.

(4) Let $\hat{C} := \{\hat{x} : x \in C\} \subset C^{**} \subset X^{**}$. If X is reflexive, $\hat{C} = C^{**}$ is weak - $*$ compact and hence by item 1., C is weakly compact in X. Conversely if C is weakly compact, then $\hat{C} \subset C^{**}$ is weak $-*$ compact being the continuous image of a continuous map. Since the weak $-*$ topology on X^{**} is Hausdorff, it follows that $\hat{C}$ is weak $-*$ closed and so by item 3, $C^{**} = \overline{\hat{C}}^{\text{weak}-*} = \hat{C}$. So if $\lambda \in X^{**}$, $\lambda / \|\lambda\| \in C^{**} = \hat{C}$, i.e. there exists $x \in C$ such that $\hat{x} = \lambda / \|\lambda\|$. This shows $\lambda = (\|\lambda\| x)^{\hat{}}$ and therefore $\hat{X} = X^{**}$.

■

16.5. **Exercises.**

16.5.1. *More Examples of Banach Spaces.*

Exercise 16.1. Let $(X, \mathcal{M})$ be a measurable space and $M(X)$ denote the space of complex measures on $(X, \mathcal{M})$ and for $\mu \in M(X)$ let $\|\mu\| \equiv |\mu|(X)$. Show $(M(X), \|\cdot\|)$ is a Banach space. (Move to Section 14.)

Exercise 16.2. Folland 5.9.

Exercise 16.3. Folland 5.10.

Exercise 16.4. Folland 5.11.

16.5.2. *Hahn-Banach Theorem Problems.*

Exercise 16.5. Folland 5.17.

Exercise 16.6. Folland 5.18.

Exercise 16.7. Folland 5.19.

Exercise 16.8. Folland 5.20.

Exercise 16.9. Folland 5.21.

Exercise 16.10. Let X be a Banach space such that X^* is separable. Show X is separable as well. (Folland 5.25.) **Hint:** use the greedy algorithm, i.e. suppose $D \subset X^* \setminus \{0\}$ is a countable dense subset of X^*, for $\ell \in D$ choose $x_\ell \in X$ such that $\|x_\ell\| = 1$ and $|\ell(x_\ell)| \geq \frac{1}{2}\|\ell\|$.

Exercise 16.11. Folland 5.26.

16.5.3. *Baire Category Result Problems.*

Exercise 16.12. Folland 5.29.

Exercise 16.13. Folland 5.30.

Exercise 16.14. Folland 5.31.

Exercise 16.15. Folland 5.32.

Exercise 16.16. Folland 5.33.

Exercise 16.17. Folland 5.34.

Exercise 16.18. Folland 5.35.

Exercise 16.19. Folland 5.36.

Exercise 16.20. Folland 5.37.

Exercise 16.21. Folland 5.38.

Exercise 16.22. Folland 5.39.

Exercise 16.23. Folland 5.40.

Exercise 16.24. Folland 5.41.

16.5.4. *Weak Topology and Convergence Problems.*

Exercise 16.25. Folland 5.47.

Definition 16.31. A sequence $\{x_n\}_{n=1}^{\infty} \subset X$ is **weakly Cauchy** if for all $V \in \tau_w$ such that $0 \in V$, $x_n - x_m \in V$ for all m, n sufficiently large. Similarly a sequence $\{f_n\}_{n=1}^{\infty} \subset X^*$ is **weak–$*$ Cauchy** if for all $V \in \tau_{w^*}$ such that $0 \in V$, $f_n - f_m \in V$ for all m, n sufficiently large.

Remark 16.32. These conditions are equivalent to $\{f(x_n)\}_{n=1}^{\infty}$ being Cauchy for all $f \in X^*$ and $\{f_n(x)\}_{n=1}^{\infty}$ being Cauchy for all $x \in X$ respectively.

Exercise 16.26. Folland 5.48.

Exercise 16.27. Folland 5.49.

Exercise 16.28. land 5.50.

Exercise 16.29. Show every weakly compact subset X is norm bounded and every weak–$*$ compact subset of X^* is norm bounded.

Exercise 16.30. Folland 5.51.

Exercise 16.31. Folland 5.53.