FUNCTIONAL ANALYSIS 2009

V. LATVALA

Date: December 14, 2009.

Contents

1. Normed spaces

Throughout this text $\mathbb{F}$ is either $\mathbb{C}$ or $\mathbb{R}$.

1.1. **Definition and main examples.**

Definition 1.1.1. Let X be a vector space over $\mathbb{F}$. A *norm* on X is a function $\|\cdot\| : X \to \mathbb{R}$ such that $\forall x, y \in X$ $\forall \alpha \in \mathbb{F}$

(i) $\|x\| \geq 0$;
(ii) $\|x\| = 0 \iff x = 0_X$;
(iii) $\|\alpha x\| = |\alpha|\|x\|$;
(iv) $\|x + y\| \leq \|x\| + \|y\|$;

Note. If $\|\cdot\|$ is a norm on X, then $d : X \times X \to \mathbb{R}_+$,

$$d(x, y) := \|x - y\|,$$

defines a metric on X.

Example 1.1.2. Let $n \in \mathbb{N}$ and recall that $\mathbb{F}$ is $\mathbb{R}$ or $\mathbb{C}$. In both cases, $\|\cdot\| : \mathbb{F}^n$,

$$\|(x_1, \ldots, x_n)\| = \left(\sum_{j=1}^{n} |x_j|^2 \right)^{\frac{1}{2}} \qquad (*)$$

is a norm on $\mathbb{F}^n$ (the *standard norm* on $\mathbb{F}^n$).

The previous example is a special case of the following:

Example 1.1.3. Let X be a finite-dimensional vector space over $\mathbb{F}$ with basis $\{ e_1, \ldots, e_n \}$. Then any $x \in X$ can be written uniquely as

$$x = \sum_{j=1}^{n} \lambda_j e_j,$$

i.e. scalars λ_j are unique.

Claim: The function $\|\cdot\| : X \to \mathbb{R}$,

$$\|x\| = \left(\sum_{j=1}^{n} |\lambda_j|^2 \right)^{\frac{1}{2}} \qquad (**)$$

is a norm on X (Exercise).

Remark. If $X = \mathbb{R}^n$ (see Example 1.1.2) and $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, then $\lambda_j = x_j$ (with standard base) so $(*)$ and $(**)$ are equal. If $X = \mathbb{C}^n (= \mathbb{R}^{2n})$ and $x = (z_1, \ldots, z_n) \in \mathbb{C}^n$, then $z_j = x_j + iy_j$. In other words $x = (x_1, y_1, x_2, y_2, \ldots, x_n, y_n)$ and $(**)$ is (with standard base $e_1, \ldots, e_{2n}$)

$$\|x\| = \left(\sum_{j=1}^{n} x_j^2 + \sum_{j=1}^{n} y_j^2 \right)^{\frac{1}{2}} = \left(\sum_{j=1}^{n} \underbrace{x_j^2 + y_j^2}_{|z_j|^2} \right)^{\frac{1}{2}}$$

This equals $(*)$.

Note. Many normed function spaces are *not* finite-dimensional!

Example 1.1.4. Let (M, d) be a compact metric space and let

$$\mathbb{C}_F(M) := \{f : M \to F : f \text{ continuous}\}.$$

Then the function $\|\cdot\| \;:\; \mathbb{C}_F(M) \to \mathbb{R}$,

$$\|f\| := \sup\{|f(x)| : x \in M\}$$

is a norm (*standard norm* on $C_F(M)$) (Exercise).

Remarks: (a) If M is not compact, for example if $M =]0, 1[\subset \mathbb{R}$, then $f(x) = \frac{1}{x}$ is continous on M. However

$$\sup\{|f(x)| : x \in M\} = +\infty.$$

(b) Here $f + g$ and αf are defined pointwise, that is,

$$\left.\begin{array}{l}(f+g)(x) := f(x) + g(x)\\ (\alpha f)(x) := \alpha f(x)\end{array}\right\} \begin{array}{l}\forall x \in M, \forall f, g \in C_F(M)\\ \forall \alpha \in \mathbb{F}.\end{array}$$

(c) $(C_{\mathbb{F}}(M), \|\cdot\|)$ is *not* finite-dimensional.

Example 1.1.5. (a) Let $1 \le p < \infty$ and let

$$L^p(\mathbb{R}) := \{f : \mathbb{R} \to \bar{\mathbb{R}} : f \text{ measurable and } \int_{\mathbb{R}} |f|^p dx < \infty\}.$$

Then $\|\cdot\|_p : L^p(\mathbb{R}) \to \mathbb{R}$,

$$\|f\|_p := \left(\int_{\mathbb{R}} |f|^p dx\right)^{\frac{1}{p}},$$

is a norm ($L^p - norm\ on\ \mathbb{R}$). The *triangle-inequality*

$$\|f + g\|_p \le \|f\|_p + \|g\|_p$$

is called the *Minkowski inequality.*

If $1 < p < \infty$, then the *Hölder conjugate* of p is $1 < q < \infty$ so that

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

Hence

$$\frac{1}{q} = 1 - \frac{1}{p} = \frac{p-1}{p},$$

i.e.

$$q = \frac{p}{p-1} =: p'$$

(b) Let

$$L^\infty(\mathbb{R}) := \{\ f : \mathbb{R} \to \bar{\mathbb{R}}\ : f \text{ measurable and } \operatorname*{ess\,sup}_{x\in\mathbb{R}} |f(x)| < \infty\}$$

(Here $\operatorname{ess\,sup}_{x\in\mathbb{R}} |f(x)| < \infty$ means: $\exists M \in \mathbb{R}_+$ so that $|f(x)| \le M$ for a.e. $x \in \mathbb{R}$.)

Then $\|\cdot\|_\infty : L^\infty(\mathbb{R}) \to \mathbb{R}$,

$$\|f\|_\infty := \inf\{M > 0 \;:\; |f(x)| \le M \text{ for a.e. } x \in \mathbb{R}\},$$

is a norm on $L^\infty(\mathbb{R})$ (L^∞-*norm* on $\mathbb{R}$).

For $p = 1$, the Hölder conjugate is $q = \infty$. Conversely, for $p = \infty$, the Hölder conjugate is $q = 1$. Hence we write $1' = \infty, \infty' = 1$.

Here in (a) and (b), $f + g$ and αf are defined pairwise.

Lemma 1.1.6. *Let $1 \leq p \leq \infty$ and let q be the Hölder conjugate of p. Then for any $f \in L^p(\mathbb{R})$ and $g \in L^q(\mathbb{R})$*

$$\int_{\mathbb{R}} |fg| dx \leq \|f\|_p \|g\|_q.$$

Note. Hölder's inequality follows from Young's inequality:

$$|ab| \leq \frac{1}{p}|a|^p + \frac{1}{q}|b|^q \quad (a, b \in \mathbb{R},\ 1 < p < \infty,\ q = p')$$

with a trick. The Minkowski inequality follows from the Hölder inequality with a trick (see exercises).

Example 1.1.7. (a) Let $1 \leq p < \infty$ and let l^p be the set of all sequences $(a_n)_{n \in \mathbb{N}}$ in F so that

$$\sum_{n=1}^{\infty} |(a_n)|^p < \infty.$$

Then

$$\|(a_n)\|_p := \left(\sum_{n=1}^{\infty} |a_n|^p \right)^{\frac{1}{p}}$$

is a norm on l^p *(l^p-norm)*.

(b) Let l^∞ be the set of all sequences in F so that

$$\sup_{n \in \mathbb{N}} |a_n| < \infty \quad (\textit{bounded sequence}).$$

Then

$$\|(a_n)\|_\infty := \sup\{|a_n| : n \in \mathbb{N}\}$$

is a norm on l^∞ *(l^∞-norm)*. Here

$$(a_n) + (b_n) := (a_n + b_n) \quad \text{and} \quad \alpha(a_n) := (\alpha a_n).$$

Theorem 1.1.8. *Let $1 \leq p \leq \infty$ and let q be the Hölder conjugate of p. Then for any sequences $(a_n) \in l^p$, $(b_n) \in l^q$ we have*

$$\sum_{n=1}^{\infty} \|a_n b_n| \leq \|(a_n)\|_p \|(b_n)\|_q.$$

Proof. The case $p = 1$ or $q = 1$ is easy (Write the proof!). Assume that $1 < p < \infty$ and $1 < q < \infty$. We may also assume that $\|(a_n)\|_p > 0$ and $\|(b_n)\|_q > 0$. Indeed, if $\|(a_n)\|_p = (\sum_{n=1}^{\infty} \|a_n|^p)^{\frac{1}{p}} = 0$, then $|a_n| = 0$ for all $n \in \mathbb{N}$ and therefore the left-hand side $= 0$.

By Young's inequality with $a = \frac{|a_n|}{\|(a_n)\|_p}$, $b = \frac{|b_n|}{\|(b_n)\|_q}$,

$$\frac{|a_n|}{\|(a_n)\|_p} \frac{|b_n|}{\|(b_n)\|_q} \leq \frac{1}{p} \frac{|a_n|^p}{\|(a_n)\|_p^p} + \frac{1}{q} \frac{|b_n|^q}{\|(b_n)\|_q^q}.$$

By summing up and using the product + sum-rules for series:

$$\begin{aligned}
\frac{1}{\|(a_n)\|_p\|(b_n)\|_q}\sum_{n=1}^{\infty}|a_nb_n| &\leq \sum_{n=1}^{\infty}\frac{1}{p}\frac{|a_n|^p}{\|(a_n)\|_p^p}+\sum_{n=1}^{\infty}\frac{1}{q}\frac{|b_n|^q}{\|(b_n)\|_q^q}\\
&= \frac{1}{p}\frac{1}{\|(a_n)\|_p^p}\underbrace{\sum_{n=1}^{\infty}|a_n|^p}_{\|(a_n)\|_p^p}+\frac{1}{q}\frac{1}{\|(b_n)\|_q^q}\underbrace{\sum_{n=1}^{\infty}|b_n|^q}_{\|(b_n)\|_q^q}\\
&= 1.
\end{aligned}$$

The claim follows. □

1.2. **Convergence in normed spaces.** A normed space $(X, \|\cdot\|)$ is a vector space X Over F which is equipped with a norm $\|\cdot\|$. We assume throughout this subsection that $(X, \|\cdot\|)$ is a normed space and $x_n, x \in X$.

Definition 1.2.1. The sequence (x_n) converges to x in X, denote $\lim_{n\to\infty} x_n = x$, if $\forall\ \varepsilon > 0\ \exists n_\varepsilon \in \mathbb{N}$ such that

$$\|x_n - x\| < \varepsilon \quad \text{if } n \geq n_\varepsilon.$$

The sequence (x_n) is a *Cauchy sequence* if $\forall \varepsilon > 0\ \exists n_\varepsilon \in \mathbb{N}$ such that

$$\|x_m - x_n\| < \varepsilon \quad \text{if } m, n \geq n_\varepsilon.$$

Lemma 1.2.2. *Assume that* $\lim_{n\to\infty} x_n = x$. *Then*

(i) *The limit x is unique;*
(ii) $\lim_{n\to\infty} x_{n_i} = x$ *for any subsequence; that is, if $i \to n_i$ is a strictly increasing function $\mathbb{N} \to \mathbb{N}$;*
(iii) *(x_n) is a Cauchy sequence.*

Proof. The proofs are as in the case $X = \mathbb{R}$ (replace $|\cdot| \leftrightarrow \|\cdot\|$). (ii),(iii) Exercise. □

A set $M \in X$ is *compact* if every sequence (x_n) in M contains a subsequence (x_{n_i}) such that $\lim_{n\to\infty} x_{n_i} = x \in M$.

A set $M \in X$ is *complete* if every Cauchy sequence in M converges to $x \in M$.

Example. $X = \mathbb{R} \to X$ is complete but not compact. For example $x_i = i \in \mathbb{R}$ does not have a convergent subsequence.

Remark. We regard the following known: If M is complete, then a sequence (x_n) converges in M if and only if (x_n) is a Cauchy sequence.

Theorem 1.2.3. *Suppose that (x_n) and (y_n) are sequences in X such that*

$$\lim_{n\to\infty} x_n = x \in X \quad \text{and} \quad \lim_{n\to\infty} y_n = y \in X.$$

Then

(i) $\Big|\|x\| - \|y\|\Big| \leq \|x - y\|$;
(ii) $\lim_{n\to\infty} \|x_n\| = \|x\|$;
(iii) $\lim_{n\to\infty}(x_n + y_n) = x + y$;
(iv) $\lim_{n\to\infty} \alpha_n x_n = \alpha x$.

Proof. (i)-(ii) exercise, (iii) skip. Proofs are as in $(\mathbb{R}, |\cdot|)$.

(iv) Since (α_n) converges, it forms a bounded sequence. Hence $\exists M > 0$ such that $|\alpha_n| \le M$ for $\forall n \in \mathbb{N}$. By Definition 1.1.1 (iii), (iv),

$$\begin{aligned}
\|\alpha_n x_n - \alpha x\| &= \|\alpha_n x_n - \alpha_n x + \alpha_n x - \alpha x\|^{(*)} \\
&= \|\alpha_n(x_n - x) + (\alpha_n - \alpha)x\| \\
&\overset{(iv)}{\le} \|\alpha_n(x_n - x)\| + \|(\alpha_n - \alpha)x\| \\
&\overset{(iii)}{=} |\alpha_n|\ \|x_n - x\| + |\alpha_n - \alpha)x\| \\
&\le M\|x_n - x\| + |\alpha_n - \alpha|\|x\|.
\end{aligned}$$

Now, for given $\varepsilon > 0$, $\exists n_1 \in \mathbb{N}$ such that $\|x_n - x\| < \frac{\varepsilon}{2M}$ wherever $n \ge n_1$ & $\exists n_2 \in \mathbb{N}$ such that $|\alpha_n - \alpha| < \frac{\varepsilon}{2\|x\|}$ (assuming that $\|x\| \neq 0$). If $n \ge \max(n_1, n_2)$, then $\|\alpha_n x_n - \alpha x\| < \varepsilon$.

(*) We use the fact that $\forall \alpha \forall x$ holds $-\alpha x = (-\alpha)x = \alpha(-x)$. □

Definition 1.2.4. *Banach space* is a complete normed space $(X, \|\cdot\|)$, that is, each Cauchy sequence in X converges to an element of X.

Example. $(\mathbb{Q}, |\cdot|)$ is a normed space which is *not* Banach. For instance the sequence

$$x_n = \sum_{k=1}^{n} \frac{1}{k!} \in \mathbb{Q}$$

converges to $e \in \mathbb{R} \notin \mathbb{Q}$. By Lemma 1.2.2 (iii), (x_n) is a Cauchy sequence. By 1.2.2 (i), (x_n) can not converge to an element in $\mathbb{Q}$.

Theorem 1.2.5. *All the normed spaces in Examples 1.1.2, 1.1.4, 1.1.5 and 1.1.7 are Banach spaces.*

Proof. We skip the proof, see Analysis 4 / Rynne & Youngson. □

2. Linear operators

2.1. **Continuous linear transformations.**

Let V and W be vector spaces over the same scalar field F. A mapping $T : V \to W$ is called a *linear transformation* if $\forall \alpha, \beta \in F$ and $x, y \in V$,

$$T(\alpha x + \beta y) = \alpha T(x) + \beta T(y).(*)$$

Remark 2.1.1. Let V,W be vector spaces and $T : V \to W$ be linear; see Rynne & Youngson, p.3, (a)-(e). Let $x \in V$ and $\alpha \in F$; let 0_V be the zero-element in V and let 0_W be the zero-element in W.

Claim 1. $0x = 0_V$, $\alpha 0_V = 0_V$.
Proof. By (e), $0_X = (0+0)x = 0x + 0x$. We add $-0x$ on both sides $\Rightarrow 0_V = 0x$. similarly $\alpha 0_V = \alpha(0_V + 0_V) = \alpha 0_V + \alpha 0_V$.

Claim 2. $\alpha x = (-\alpha)x = \alpha(-x)$.
Proof. By (e)
$\alpha x + (-\alpha)x = (\alpha + (-\alpha))x = 0x = 0_V$,
$\alpha x + \alpha(-x) = \alpha(x + (-x)) = \alpha 0 = 0_V$

Claim 3. $T(0_V) = 0_W$ and $T(-x) = -T(x)$
Proof. By linearity (and Claim1):

$$T(00_V)) = T(00_V) + 00_V) = 0T(0_V) + 0T(0V),$$

that is, $T(0_V) = 0_W$. Moreover

$$T(0_V) = T(x + (-x)) = T(x) + T(-x)$$

that is, $T(-x) = -T(x)$.

Recall the necessary definitions:

Definition. Let X and Y be normed spaces. A function $F : X \to Y$ is *continuous at* $x \in X$ if $\forall \varepsilon > 0\ \exists \delta > 0$ such that

$$\|x - y\|_X < \delta \Rightarrow \|F(x) - F(y)\|_Y < \varepsilon.$$

F is *continuous on* X if F is continous at $x\ \forall x \in X$. F is *uniformly continous on* X if $\forall x \in X\ \forall \varepsilon > 0\ \exists \delta > 0$ not depending on x such that

$$\|x - y\|_X < \delta \Rightarrow \|F(x) - F(y)\|_Y < \varepsilon.$$

Lemma 2.1.2. *Let X and Y be normed spaces and let $T : X \to Y$ be a linear transformation. Then the following are equivalent:*

(a) *T is uniformly continous on X;*
(b) *T is continous on X;*
(c) *T is continous at 0_X;*
(d) *$\exists k \in \mathbb{R}_+$ such that $\|T(x)\| \le k$ whenever $x \in X and \|x\| \le 1$;*
(e) *$\exists k \in \mathbb{R}_+$ such that $\|T(x)\| \le k\|x\|\ \forall x \in X$.*

Proof. The implications $(a) \implies (b) \implies (c)$ are trivial.

$(c) \implies (d)$. Assume that T is continuous at 0_X. Then, for $\varepsilon = 1, \exists \delta > 0$ such that $\|T(x) - T(0_X)\| = \|T(x)\| < 1$ whenever $x \in X$ and $\|x - 0_X\| = \|x\| < \delta$. Let $w \in X$ with $\|w\| \leq 1$. As

$$\|\frac{\delta w}{2} = \frac{\delta}{2}\|w\| \leq \frac{\delta}{2} < \delta,$$

We have (T is linear)

$$1 > \|T(\frac{\delta w}{2})\| = \|\frac{\delta}{2}T(w)\| = \frac{\delta}{2}\|T(w)\|.$$

Hence $\|T(w)\| < \frac{2}{\delta}$ so that (d) holds with $k = \frac{2}{\delta}$

$(d) \implies (e)$. Let k be such that $\|T(x)\| \leq k$ whenever $x \in X$ and $\|x\| \leq 1$. Since $T(0_X) = 0_Y$, it is clear that $\|T(0_X)\| = \|0_Y\| = 0 \leq k\|0_X\|$. Let $x \in X, x \neq 0_X$. As $\|\frac{x}{\|x\|}\| = 1$, we have

$$k \leq \|T(\frac{x}{\|x\|})\| = \|\frac{1}{\|x\|}T(x)\| = \frac{1}{\|x\|}\|T(x)\|,$$

which implies $\|T(x)\| \leq k\|x\|$.

$(e) \implies (a)$. Assuming (e) we have by linearity $\forall x, y \in X$

$$(L) \quad \|T(x) - T(y)\| =^{2.1.1} \|T(x) + T(-y)\| = \|T(x - y)\| \leq k\|x - y\|.$$

Hence, for $\varepsilon > 0$ and $\delta := \frac{\varepsilon}{k}$ we have: If $x, y \in X$ and $\|x - y\| < \delta$, then

$$\|T(x) - T(y)\| \leq k\|x - y\| < k\delta = \varepsilon.$$

This shows that T is uniformly continous on X. □

Remark. In fact, (L) means that T is Lipschitz. This is more than just uniform continuity.

Example. Transformation $T : C_F[0, 1] \to F$ defined by

$$T(f) = f(0)$$

is linear, since $\forall \alpha, \beta \in F, \forall f, g \in C_F[0, 1]$

$$|T(f)| = |f(0)| \leq \sup_{x \in [0,1]} |f(x)| = \|f\|,$$

that is, 2.1.2 (c) holds with $k = 1$.

Lemma 2.1.3. *If $(c_n) \in l^\infty$ and $(x_n) \in l^p$, $1 \leq p < \infty$, then $(c_n x_n) \in l^p$ and*

$$\sum_{n=1}^{\infty} |c_n x_n|^p \leq \|(c_n)\|_\infty^p \sum_{n=1}^{\infty} |x_n|^p.$$

Proof. By assumptions, we have

$$\lambda := \sup\{\, |c_n| : \ n \in \mathbb{N} \,\} < \infty$$

and

$$\sum_{n=1}^{\infty} |x_n|^p = \|(x_n)\|_p^p < \infty.$$

Since for all $n \in \mathbb{N}$

$$|c_n x_n|^p \leq \lambda^p |x_n|^p$$

and $\sum_{n=1}^{\infty} < \infty$, the series $\sum_{n=1}^{\infty} |c_n x_n|^p$ converges and the claim follows. □

Example 2.1.4. If $(c_n) \in l^\infty$, then the transformation $T : l^1 \to F$,

$$T((x_n)) = \sum_{n=1}^{\infty} c_n x_n,$$

is linear and continous.

Proof. By Lemma 2.1.3, $(c_n x_n) \in l^1$ for all $(x_n) \in l^1$. Since (we regard as known)

$$\sum_{n=1}^{\infty} |c_n x_n| < \infty \qquad \Longrightarrow \sum_{n=1}^{\infty} c_n x_n < \infty,$$

T is well-defined. For all $\alpha\beta \in F$ and $(x_n), (y_n) \in l^1$,

$$\begin{aligned} T\big(\alpha(x_n) + \beta(y_n)\big) &= T\big((\alpha x_n + \beta y_n)\big) = \sum_{n=1}^{\infty} c_n(\alpha x_n + \beta y_n) \\ &= \alpha \sum_{n=1}^{\infty} c_n x_n + \beta \sum_{n=1}^{\infty} c_n x_n = \alpha T\big((x_n)\big) + \beta T\big((y_n)\big) \end{aligned}$$

since all the series converge. Hence T is linear. Moreover, for any $(x_n) \in l^1$,

$$|T\big((x_n)\big)| = |\sum_{n=1}^{\infty} c_n x_n| \leq \sum_{n=1}^{\infty} |c_n x_n| \leq^{2.1.3} \|(c_n)\|_\infty \|(x_n)\|_1.$$

Hence, Lemma 2.1.2 (e) holds with $k = \|(c_n)\|_\infty$. Thus T is continous. □

Example 2.1.5. If $(c_n) \in l^\infty$, then the transformation $T : l^2 \to l^2$,

$$T\big((x_n)\big) = (c_n x_n),$$

is linear and continous.

Proof. By Lemma 2.1.3, $(c_n x_n) \in l^2$ for any $(x_n) \in l^2$. Hence T is well-defined. For all $\alpha, \beta \in F$ and $(x_n), (y_n) \in l^2$

$$\begin{aligned} T\big(\alpha(x_n) + \beta(y_n)\big) &= T\big((\alpha x_n + \beta y_n)\big) = \big(c_n(\alpha x_n + \beta y_n)\big) \\ &= \alpha(c_n x_n) + \beta(c_n y_n) = \alpha T\big((x_n)\big) + \beta T\big((y_n)\big). \end{aligned}$$

Hence T is linear. Moreover, for any $(x_n) \in l^2$,

$$\|T\big((x_n)\big)\|_2^2 = \sum_{n=1}^{\infty} |c_n x_n|^2 \leq \|(c_n)\|_\infty^2 \sum_{n=1}^{\infty} |x_n|^2 = \|(c_n)\|_\infty^2 \|(x_n)\|_2^2.$$

Hence, Lemma 2.1.2 (e) holds with $k = \|(c_n)\|_\infty$. Thus T is continuous. □

Example 2.1.6. Let $P \subset C_{\mathbb{R}}[0,1]$ be the set of all real polynomials p restricted to $[0,1]$. It is evident that P is a vector space and clearly

$$\|p\| = \sup\{\, |p(t)| : t \in [0,1] \,\}$$

defines a norm in P. Let $T : P \to P$ be the linear operator

$$T(p) = p'. \ \ (\textit{derivative})$$

If $p_n \in P$ is defined by $p_n(t) = t^n$, then

$$\|p_n\| = \sup\big\{|t|^n \big| t \in [0,1]\big\} = 1 \ \ \forall n \in \mathbb{N}$$

while

$$\|T(p_n)\| = \sup\big\{\|nt^{n-1}|\big| t \in [0,1]\big\} = n \ \forall n \in \mathbb{N}$$

Hence Lemma 2.1.2 (e) does not hold for any $k \in \mathbb{R}_+$. It follows that T is not continous.

Definition 2.1.7. Let X and Y be normed spaces and let $T : X \to Y$ be a linear transformation. Then T is called *bounded* if $\exists k > 0$ such that

$$|T(x)\| \le k\|x\| \ \ \forall x \in X.$$

Remark. The function $T : \mathbb{R} \to \mathbb{R}, T(x) = x$, is a bounded transformation but not a bounded function. In fact, a linear transformation $T : X \to Y$ is a bounded function only if $T \equiv 0$.

Reason: If there is $x \in X$ such that $\|T(x)\| > 0$, then $\|T(\alpha x)\| = \|\alpha T(x)\| = |\alpha|\|T(x)\| \to \infty$ as $|\alpha| \to \infty$.

Notation. Let X and Y be normed spaces. Then $B(X,Y)$ denotes the set of all continous transformations $X \to Y$. Elements in $B(X,Y)$ are often called *bounded linear operators.*

Example 2.1.8. Let $a, b \in \mathbb{R}$, and let $k : [a,b] \times [a,b] \to \mathbb{R}$ be continuous. Denote

$$C[a,b] := \big\{f : [a,b] \to \mathbb{R} : f \text{ continuous}\big\}.$$

(a) If $f \in C[a,b]$, then $K : C[a,b] \to C[a,b]$ is defined by

$$Kf(s) := (K(f))(s) = \int_a^b k(s,t)f(t)dt, \quad s \in [a,b].$$

Claim. K is well-defined and linear.
Proof. For any $\alpha, \beta \in \mathbb{R}$ and $f, g \in C[a,b]$, we have

$$\begin{aligned} (K(\alpha f + \beta g))(s) &= \int_a^b k(s,t)\big(\alpha f(s) + \beta g(s)\big)dt \\ &= \alpha \int_a^b k(s,t)f(s)dt + \beta \int_a^b k(s,t)g(s)dt \\ &= \alpha(K(f))(s) + \beta(K(g))(s) \end{aligned}$$

This means that

$$K(\alpha f + \beta g) = \alpha K(f) + \beta K(g),$$

that is, K is linear.

We show next that $K(f) \in C[a,b] \ \forall f \in C[a,b]$. Let $\varepsilon > 0$. Since $[a,b] \times [a,b]$ is compact (closed and bounded in $\mathbb{R}^2$), k is uniformly continous (we regard this as known!). Hence $\exists \delta > 0$ such that $\forall (x,y), (x',y') \in [a,b] \times [a,b]$

$$|(x,y) - (x',y')| < \delta \Rightarrow |k(x,y) - k(x',y')| < \varepsilon.$$

In particular, if $|s-s'|<\varepsilon$, then $|(s,t)-(s',t)|=|s-s'|<\delta$, and $|k(s,t)-k(s',t)|<\varepsilon$. Hence, for $f\in C[a,b]$,

$$\begin{aligned}\Big|Kf(s)-Kf(s')\Big| &= \Big|\int_a^b k(s,t)f(t)dt-\int_a^b k(s',t)f(t)dt\Big| \\ &= \Big|\int_a^b \big(k(s,t)-k(s',t)\big)f(t)dt\Big| \\ &\le \int_a^b \Big|\underbrace{k(s,t)-k(s',t)}_{\le\,\varepsilon}\Big|\underbrace{(f(t))}_{\le\|f\|}dt\le\varepsilon\|f\|(b-a)\end{aligned}$$

whenever $|s-s'|<\delta$. Thus Kf is (uniformly) continous in [a,b].

(b) K is bounded, that is $K\in B(C[a,b],C[a,b])$. See exercise.

Linear transformations on finite-dimensional vector spaces are special in the following sense.

Theorem 2.1.9. *Let X be a finite-dimensional vector space, Y any normed space, and let $T:X\to Y$ be linear. Then $T\in B(X,Y)$.*

Proof. We define a new norm $\|\cdot\|_1$ on X by setting

$$\|x\|_1:=\|x\|+\|T(x)\|.$$

We leave it as an exercise to prove that $\|\cdot\|_1$ is a norm on X. Since X is finite-dimensional, the norms are equivalent (see Analysis 4/ Rynne & Youngson p.43). Hence $\exists$ a constant $K>0$ such that $\|x\|_1\le K\|x\|$ for all $x\in X$. Therefore

$$\|T(x)\|\le\|x\|_1\le K\|x\|\quad\forall x\in X,$$

i.e. T is bounded. □

Remark 2.1.10. Let V and W be vector spaces over the same field F. We denote by L(V,W) the set of all linear transformations $V\to W$ and define $+$ and $\cdot$ in L(V,W) by setting $\forall F,G\in L(V,W)$ and $\forall\lambda\in F$

$$(*)\quad\begin{cases}(F+G)(x):=\ F(x)+G(x), & x\in V\\ (\lambda F)(x):=\ \lambda F(x), & x\in V\end{cases}$$

For each $F,G\in L(V,W)$ and $\lambda\in F$ we have $F+G\in L(V,W)$ and $\lambda F\in L(V,W)$, since $x,y\in V$ and $\alpha,\beta\in F$

$$\begin{aligned}(F+G)(\alpha x+\beta y) &= F(\alpha x+\beta y)+G(\alpha x+\beta y)\\ &= \alpha F(x)+\beta F(y)+\alpha G(x)+\beta G(y)\\ &= \alpha(F(x)+G(x))+\beta(F(y)+G(y))\\ &= \alpha(F+G)(x)+\beta(F+G)(y)\end{aligned}$$

and

$$\begin{aligned}&(\lambda F)(\alpha x+\beta y)=\lambda F(\alpha x+\beta y)=\lambda(\alpha F(x)+\beta F(y))\\ =\ &\alpha\lambda F(x)+\beta\lambda F(y)=\alpha(\lambda F)(x)+\beta(\lambda F)(y).\end{aligned}$$

Hence $L(V, W)$ is a linear subspacce of $F(V, W)$ (= the vector space of all functions $V \to W$ with $+$ and $\cdot$ defined pointwise. We regard the existence of $F(V, W)$ known.

2.2. **The norm of a bounded linear operator.**

If X and Y are normed spaces, we know by Remark 2.1.10 that $B(X, Y)$ is a vector space. Next, we want to define a norm on $B(X, Y)$.

Definition 2.2.1. Let X and Y be normed spaces and let $T \in L(X, Y)$. Then we define

$$\|T\| := \sup\{\|T(x)\| : \|x\| \leq 1\}.$$

Remark 2.2.2. Let X and Y be normed spaces and $T \in L(X, Y)$. Recall from Lemma 2.1.2 that $T \in B(X, Y)$ iff $\|T\| < \infty$.

Proof. If $T \in B(X, Y)$, $\exists k \in \mathbb{R}_+$, such that $\|T\| \leq k\|x\| \, \forall x \in X$. Then

$$\|T\| \leq k. \quad (*)$$

Conversely, assume that $\|T\| < \infty$. Since $\|\frac{x}{\|x\|}\| = 1 \, \forall x \in X, x \neq 0_X$, we have

$$\frac{\|T(x)\|}{\|x\|} = \big\|\frac{1}{\|x\|}T(x)\big\| = \big\|T(\frac{x}{\|x\|})\big\| \leq \|T\|$$

for all $x \in X$, $x \neq 0_X$. Since $\|T(0_X)\| = \|0_Y\| = 0$, we have

$$(**) \qquad \|T(x)\| \leq \|T\|\|x\| \, \forall x \in X.$$

Hence T is bounded.

□

Remark 2.2.3. The proof of Remark 2.2.2 implies that

$$\|T\| = \inf\{k \in \mathbb{R}_+ : \|T(x)\| \leq k\|x\| \, \forall x \in X\}. \text{ (Exercise)}$$

Hence $\|T\|$ expresses the "minimal" bound for the boundedness of T.

Theorem 2.2.4. *Let X and Y be normed spaces. Then*

$$\|T\| := \sup\{\|T(x)\| : \|x\| \leq 1\}$$

defines a norm on $B(X, Y)$.

Proof. Recall that $B(X, Y)$ is a vector space by Lemma Let $S, T \in B(X, Y)$ and $\lambda \in \mathbb{F}$.
(i) Clearly $\|T\| \geq 0$. By Remark 2.2.2, $\|T\| \leq \infty$.
(ii)

$$\begin{aligned} \|T\| = 0 \qquad &\iff \big\|T(\frac{x}{\|x\|})\big\| = \frac{1}{\|x\|}\|T(x)\| = 0 \qquad \forall x \in X, x \neq 0_X \\ &\iff \|T(x)\| = 0 \qquad \forall x \in X, x \neq 0_X \\ &\iff T(x) = 0_Y \qquad \forall x \in X \\ &\iff T \text{ is the zero element in } L(X, Y). \end{aligned}$$

(iii) As $\|T(x)\| \leq \|T\|\|x\| \, \forall x \in X$ (Remark 2.2.2 $(**)$), we have (for $\lambda \in \mathbb{F}$)

$$\|(\lambda T)(x)\| = \|\lambda T(x)\| = |\lambda|\|T(x)\| \leq |\lambda|\|T\|\|x\|$$

for all $x \in X$ and hence

$$\|\lambda T\| = \sup_{\|x\|\leq 1} \|(\lambda T)(x)\| \leq |\lambda|\|T\|. \ (*)$$

If $\lambda = 0$, then $\|\lambda T\| = 0 = |\lambda|\|T\|$. If $\lambda \neq 0$, then

$$\|T\| = \|\lambda^{-1}(\lambda T)\| \overset{(*), T\leftrightarrow \lambda T}{\leq} |\lambda^{-1}|\|\lambda T\| \overset{(*)}{\leq} |\lambda^{-1}||\lambda|\|T\| = \|T\|$$

Hence

$$\|T\| = |\lambda^{-1}|\|\lambda T\| \iff |\lambda|\|T\| = \|\lambda T\|.$$

(iv) For each $x \in X$, we have

$$\|(S+T)(x)\| \overset{def}{=} \|S(x) + T(x)\| \overset{\Delta-ineq.}{\leq} \|S(x)\| + \|T(x)\|$$
$$\overset{Rem. 2.2.2(**)}{\leq} \|S\|\|x\| + \|T\|\|x\| = (\|S\| + \|T\|)\|x\|.$$

By taking sup over $\|x\| \leq 1$ yields

$$\|S + T\| \leq \|S\| + \|T\|.$$

□

There is no general procedure for finding the norm of a bounded linear operator! It is also possible that the supremum in the definition is not attained.

Example 2.2.5. Let $T : \mathcal{C}_{\mathbb{F}}[0,1] \to \mathbb{F}$ be the bounded linear operator defined by

$$T(f) = f(0).$$

Claim: $\|T\| = 1$.

Proof. We have

$$|T(f)| = |f(0)| \leq \sup\{|f(x)| : x \in [0,1]\} = \|f\|.$$

By Remark 2.2.3, $\|T\| \leq 1$.

On the other hand, if $g : [0,1] \to \mathbb{F}$ is defined by $g(x) = 1, x \in [0,1]$, then

$$\|g\| = \sup |g(x)| : x \in [0,1] = 1.$$

Since

$$|T(g)| = |g(0)| = 1,$$

we have

$$\|T\| = \sup\{|T(f)| : \|f\| \leq 1\},$$

The claim follows. □

Definition 2.2.6. Let X and Y be normed spaces and let $T \in L(X,Y)$. Then T is called an *isometry* if $\|T\| = \|x\|$ for all $x \in X$.

Example 2.2.7. (a) If X is a normed space and I is the identity transformation $I(x) = x, x \in X$, then I is an isometry $X \to X$.

(b) We define an operator $S : \ell^2 \to \ell^2$ by

$$S(x_1, x_2, x_3, \ldots) = (0, x_1, x_2, x_3, \ldots)$$

(S is called *unilateral shift*).

Claim: S is an isometry $\ell^2 \to \ell^2$.

Proof. It is easy to show that S is linear. If $(x_n) \in \ell^2$ and $(y_n) = S((x_n))$, then

$$\sum_{n=1}^{\infty} |x_n|^2 = \sum_{n=1}^{\infty} |y_n|^2 = 0^2 + \sum_{n=1}^{\infty} |x_n|^2.$$

Hence $\|S((x_n))\|_2 = \|(x_n)\|_2$, i.e S is an isometry. □

Remark 2.2.8. Let X and Y be normed spaces and let $T : X \to Y$ be an isometry. Then $\|T\| = 1$ if $X \neq \{0_X\}$. Indeed, $\|T(x)\| = \|x\| \ \forall x \in X$ and therefore

$$\|T\| = \sup\{\|T(x)\| : \|x\| \leq 1\} = \sup\{\|x\| : \|x\| \leq 1\} \leq 1,$$

if only $X \neq 0_X$. In this case $\exists x \in X$ such that $\|x\| > 0$ and hence for $y := \frac{x}{\|x\|}$ we have $\|y\| = 1$.

The converse does not hold, i.e. $\|T\| = 1$ does not imply that T is an isometry. In fact, for $T : \mathcal{C}_{\mathbb{F}}[0,1] \to \mathbb{F}, T(f) = f(0)$, we have $\|T\| = 1$ (2.2.5). However, for the function $h(x) = x, x \in [0,1], \|h\| = 1$, but $\|T(h)\| = |h(0)| = 0$.

Conclusion: T is an isometry is not the same as $\|T\| = 1$.

3. Inner product spaces

3.1. Inner products.

Definition 3.1.1. Let X be a real vector space, i.e. $\mathbb{F}=\mathbb{R}$. An *inner product* on X is a function $\langle\cdot\,,\cdot\rangle : X\times X\to\mathbb{R}$ such that $\forall x,y,z\in X$ and $\forall\alpha,\beta\in\mathbb{R}$

(a) $\langle x,x\rangle\geq 0$;
(b) $\langle x,x\rangle=0 \iff x=0_X$;
(c) $\langle \alpha x+\beta y,z\rangle=\alpha\langle x,z\rangle+\beta\langle y,z\rangle$;
(d) $\langle x,y\rangle=\langle y,x\rangle$.

Example 3.1.2. (a) The function $\langle\cdot\,,\cdot\rangle:\mathbb{R}^k\times\mathbb{R}^k\to\mathbb{R}$,

$$\langle x,y\rangle=\sum_{n=1}^k x_n y_n$$

is an inner product on $\mathbb{R}^k$ (known!). This is called the *standard inner product* on $\mathbb{R}^k$.

(b) The function $\langle\cdot\,,\cdot\rangle : L^2(\mathbb{R})\times L^2(\mathbb{R})\to\mathbb{R}$,

$$\langle x,y\rangle=\int_{\mathbb{R}} fg\,dx,$$

is an inner product on $L^2(\mathbb{R})$ (Analysis 4). Notice here that we regard $L^p(\mathbb{R})$-spaces as real vector spaces.

Definition 3.1.3. Let X be a complex vector space, i.e. $\mathbb{F}=\mathbb{C}$. An inner product on X is a function $\langle\cdot\,,\cdot\rangle : X\times X\to\mathbb{C}$ such that $\forall x,y,z\in X$ and $\forall\alpha,\beta\in\mathbb{C}$

(a) $\langle x,x\rangle\in\mathbb{R}$ & $\langle x,x\rangle\geq 0$;
(b) $\langle x,x\rangle=0 \iff x=0_X$;
(c) $\langle \alpha x+\beta y,z\rangle=\alpha\langle x,z\rangle+\beta\langle y,z\rangle$;
(d) $\langle x,y\rangle=\overline{\langle y,x\rangle}$.

Here $\bar z$ is the conjugate of $z=a+bi$, i.e. $\overline{z}=a-bi$.

Note. Recall that for all $z,w\in\mathbb{C}$ we have

$$\overline{z+w}=\overline{z}+\overline{w},\quad \overline{zw}=\overline{z}\cdot\overline{w},\quad \overline{\overline{z}}=z,\quad z+\overline{z}=2Re\,z,\quad z\overline{z}=|z|^2.$$

Example 3.1.4. (a) The function $\langle\cdot\,,\cdot\rangle:\mathbb{C}^k\times\mathbb{C}^k\to\mathbb{C}$ defined by

$$\langle x,y\rangle=\sum_{n=1}^k x_n\overline{y_n}$$

is an inner product on $\mathbb{C}^k$ (*standard inner product* on $\mathbb{C}^k$). Here $x=(x_1,\dots,x_k)$, $y=(y_1,\dots,y_k)\in\mathbb{C}^k$, i.e. $x_i,y_i\in\mathbb{C}$. We skip the proof.

(b) If $(a_n),(b_n)\in\ell^2(\mathbb{F}=\mathbb{C})$, then the function $\langle\cdot\,,\cdot\rangle:\ell^2\times\ell^2\to\mathbb{C}$ defined by

$$\langle a,b\rangle=\sum_{n=1}^k a_n\overline{b_n}$$

is an inner product on ℓ^2 (exercise).

Definition 3.1.5. A real or complex vector space X with an inner product $\langle\cdot\,,\cdot\rangle$ is called an *inner product space*.

Note. Concerning general abstract results, we always consider axioms for complex inner product. This covers the case that X happens to be a real vector space. In the real case the complex conjugate can be ignored.

Lemma 3.1.6. *Let X be an inner product space, $x, y, z \in X$ and $\alpha, \beta \in \mathbb{F}$. Then*

(a) $\langle 0_X, y\rangle = \langle x, 0_X\rangle = 0$;
(b) $\langle x, \alpha y + \beta z\rangle = \overline{\alpha}\langle x, y\rangle + \overline{\beta}\langle x, z\rangle$;
(c) $\langle \alpha x + \beta y, \alpha x + \beta y\rangle = |\alpha|^2\langle x, x\rangle + \alpha\overline{\beta}\langle x, y\rangle + \beta\overline{\alpha}\langle y, x\rangle + |\beta|^2\langle y, y\rangle$.

Proof. Exercise. □

Lemma 3.1.7. *Let X be an inner product space, $x, y \in X$. Then*

(a) $|\langle x, y\rangle| \leq \langle x, y\rangle\langle x, y\rangle$;
(b) *the function* $\|\cdot\| : X \to \mathbb{R}, \|x\| = \sqrt{\langle x, x\rangle}$ *defines a norm on* X.

Proof. (a) We are free to assume that $x \neq 0_X$ and $y \neq 0_X$. Choose $\alpha = -\frac{\overline{\langle y,x\rangle}}{\langle y,x\rangle}$ (see L. 3.1.6(a) & Def. 3.1.3(b)) and $\beta = 1$ in (c) of Lemma 3.1.6. We obtain

$$\begin{aligned}
0 &\leq \langle \alpha x + y, \alpha x + y\rangle \\
&= \frac{|\overline{\langle x, y\rangle}|^2}{|\langle x, x\rangle|^2}\langle x, x\rangle - \frac{\overline{\langle x, y\rangle}}{\langle x, x\rangle}\langle x, y\rangle - \frac{\langle x, y\rangle}{\langle x, x\rangle}\langle y, x\rangle + \langle y, y\rangle \\
&= \frac{|\overline{\langle x, y\rangle}|^2}{\langle x, x\rangle} - 2\frac{|\overline{\langle x, y\rangle}|^2}{\langle x, x\rangle} + \langle y, y\rangle = -\frac{|\overline{\langle x, y\rangle}|^2}{|\langle x, x\rangle|^2}\langle x, x\rangle + \langle y, y\rangle.
\end{aligned}$$

The claim follows by multiplying the inequality with $\langle x, x\rangle > 0$.

(b)

(i) $\|x\| = \sqrt{\langle x, x\rangle} \in \mathbb{R}_+$ (3.1.3(a));
(ii) $\|x\| = \sqrt{\langle x, x\rangle} = 0 \iff \langle x, x\rangle = 0 \iff x = 0_X$ (3.1.3(b));
(iii) For $\alpha \in \mathbb{F}, x \in X$

$$\|\alpha x\| = \sqrt{\langle \alpha x, \alpha x\rangle} \overset{3.1.6(c)}{=} \sqrt{|\alpha|^2\langle x, x\rangle} = |\alpha|\|x\|;$$

(iii) For $x, y \in X$

$$\begin{aligned}
\|x + y\|^2 &= \langle x + y, x + y\rangle = \langle x, x\rangle + \langle x, y\rangle \overbrace{\langle y, x\rangle}^{\overline{\langle x,y\rangle}} + \langle y, y\rangle \\
&= \langle x, y\rangle + 2Re\,\langle x, y\rangle + \langle y, y\rangle \\
&= \|x\|^2 + 2|\langle x, y\rangle| + \|y\|^2 \overset{(a)}{\leq} \|x\|^2 + 2|\langle x, y\rangle| + \|y\|^2 \\
&= (\|x\| + \|y\|)^2.
\end{aligned}$$

The claim follows. □

Remark. Lemma 3.1.7(a) is usually written in a form

$$|\langle x, y\rangle| \leq \|x\|\|y\|. \quad \textit{(Cauchy-Schwarz-inequality)}$$

Every inner product space is a normed space! How about the converse? The answer is <u>no</u>!

Lemma 3.1.8. *Let X be an inner product space with the norm $\|\cdot\|$ induced by the inner product (i.e. $\|x\| = \sqrt{\langle x, x\rangle}$). Then for all $u, v, x, y \in X$*

(a) $\langle u+v, x+y\rangle - \langle u-v, x-y\rangle = 2\langle u, y\rangle + 2\langle v, x\rangle$;
(b) $\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2)$ *(The parallelogram rule)*

Proof. Exercise. □

Example 3.1.9. In $\mathbb{R}^2$: *(Kuva suunnikkaasta.)*

The parallelogram rule can be used to prove that the given norm is not induced by any inner product.

Example 3.1.10. We show that the standard norm in $\mathcal{C}$ is not induced by any inner product. Choose $f(x) = 1,\ g(x) = x, x \in [0,1]$. Then

$$(f+g)(x) = 1 + x, \quad (f-g)(x) = 1 - x,$$

and

$$\|f+g\| = 2, \quad \|f-g\| = 1, \quad \|f\| = \|g\| = 1.$$

Hence

$$\|f+g\|^2 + \|f-g\|^2 = 5 \neq 4 = 2(\|f\|^2 + \|g\|^2).$$

This is not possible, if $\|\cdot\|$ were induced by some inner product.

Remark. Since an inner product space X is a normed space with the induced norm, X is also a metric space. Any metric space concepts on X will be understood in terms of the metric induced by the induced norm.

3.2. **Orthogonality.**

Let X be a real inner product space and $x, y \in X$ non-zero vectors. By the Cauchy-Schwarz inequality

$$-1 \leq \frac{\langle x, y\rangle}{\|x\|\|y\|} \leq 1.$$

Hence we can define an 'angle' θ between x and y by

$$\theta = \arccos \frac{\langle x, y\rangle}{\|x\|\|y\|}.$$

For complex inner products, the concept of angle is not relevant but we still talk about orthogonality.

Definition 3.2.1. Let X be an inner product space. Then $x, y \in X$ are *orthogonal* if $\langle x, y\rangle = 0$.

Definition 3.2.2. Let X be an inner product space. The set $\{e_1, ..., e_k\} \subseteq X$ is called *orthonormal* if

(a) $\|e_n\| = 1 \quad \forall n = 1, ..., k$;
(b) $\langle e_m, e_n\rangle = 0 \quad \forall m, n \in \{1, ..., k\}, \quad m \neq n$.

Lemma 3.2.3. *Let X be an inner product space. Then any orthonormal set $\{e_1, ..., e_k\} \subset X$ is linearly independent. In particular, if X is k-dimensional then the set $\{e_1, ..., e_k\}$ is a basis for X and any $x \in X$ can be expressed in the form*

$$x = \sum_{n=1}^{k} \langle x, e_n\rangle e_n.$$

Proof. Suppose that $\sum_{n=1}^k \alpha_n e_n = 0_X$, where $\alpha_n \in \mathbb{F}$. Then for any $m = 1, ..., k$

$$0 \overset{3.1.6}{=} \langle \sum_{n=1}^{k} \alpha_n e_n, e_m \rangle \overset{3.1.3}{=} \sum_{n=1}^{k} \alpha_n \langle e_n, e_m \rangle = \alpha_m \langle e_m, e_m \rangle = \alpha_m.$$

Hence $\{e_1, ..., e_k\}$ is linearly independent.
Suppose that $\dim X = k$. Since $\{e_1, ..., e_k\}$ is linearly independent and $\dim X = k$, $\{e_1, ..., e_k\}$ forms a basis for X (this is regarded as known from linear algebra!). Then for any $x \in X$ $\exists\ \lambda_n \in \mathbb{F}$ such that $x = \sum_{n=1}^k \alpha_n e_n$. It follows that

$$\langle x, e_m \rangle = \langle \sum_{n=1}^{k} \lambda_n e_n, e_m \rangle = \sum_{n=1}^{k} \lambda_n \langle e_n, e_m \rangle = \lambda_m$$

for any $m = 1, ..., k$. □

Lemma 3.2.4. *Let X be an inner product space and let $\{x_1, ..., x_k\} \subset X$ be linearly independent. Let*

$$S = Sp\{x_1, ..., x_k\} = \{\sum_{n=1}^{k} \lambda_n x_n : \lambda_n \in \mathbb{F}\}.$$

Then there is an orthonormal basis $\{e_1, ..., e_k\}$ for S.

Proof. Proof by Gram-Schmidt method (see linear algebra). □

Lemma 3.2.5. (Pythagoras) *Let X be an inner product space and let $x_1, ..., x_k \in X$ be pairwise orthogonal, i.e. $\langle x_i, x_j \rangle = 0$ for all $i, j \in \{1, ..., k\}, i \neq j$. Then*

$$\|x_1 + x_2 + ... + x_k\|^2 = \|x_1\|^2 + \|x_2\|^2 + ... + \|x_k\|^2.$$

Proof. Exercise. □

Definition 3.2.6. Let X be an inner product space and let $A \subset X$. *The orthogonal complement* of A is the set

$$A^\perp := \{x \in X : \langle x, a \rangle = 0\ \forall\ a \in A\}.$$

Example. If $X = \mathbb{R}^3$ and $A = \{(a_1, a_2, 0) : a_1, a_2 \in \mathbb{R}\}$, then

$$x = (x_1, x_2, x_3) \in A^\perp \iff \langle x, a \rangle = x_1 a_1 + x_2 a_2 = 0 \quad \forall\ a_1, a_2 \in \mathbb{R}.$$

Assume that $x \in A^\perp$. Choosing $a_1 = x_1$ and $a_2 = x_2$, we have $x_1^2 + x_2^2 = 0$ and hence $x_1 = x_2 = 0$. On the other hand, if $x_1 = x_2 = 0$ (and $x_3 \in \mathbb{R}$) then $x \in A^\perp$. We conclude that $A^\perp = \{(0, 0, x_3) : x_3 \in \mathbb{R}\}$.

Example 3.2.7. Let X be k-dimensional inner product space and let $\{e_1, ..., e_k\}$ be an orthonormal basis for X. If $A = Sp\{e_1, ..., e_p\}$ for all $1 \leq p < k$, then $A^\perp = Sp\{e_{p+1}, ..., e_k\}$. (Exercise)

Note. It appears below that $A^\perp$ is always a linear subspace. Therefore Example 3.2.7 essentially solves the problem of finding $A^\perp$ for $A \subset X$ whenever X is *finite-dimensional.*

Lemma 3.2.8. *Let X be an inner product space and suppose that $(x_n), (y_n)$ are sequences in X such that $\lim_{n\to\infty} x_n = x \in X$ and $\lim_{n\to\infty} y_n = y \in X$. Then*

$$\lim_{n\to\infty} \langle x_n, y_n \rangle = \langle x, y \rangle.$$

Proof. We have (by using Δ-inequality in $\mathbb{F}$ and Cauchy-Schwarz)

$$\begin{aligned} |\langle x_n, y_n\rangle - \langle x, y\rangle| &= |\langle x_n, y_n\rangle - \langle x_n, y\rangle + \langle x_n, y\rangle - \langle x, y\rangle| \\ &\overset{\Delta-ineq}{\leq} |\langle x_n, y_n\rangle - \langle x_n, y\rangle| + |\langle x_n, y\rangle - \langle x, y\rangle| \\ &\overset{3.1.6(b)}{=} |\langle x_n, y_n - y\rangle| + |\langle x_n - x, y\rangle| \\ &\overset{C-S}{\leq} \|x_n\|\|y_n - y\| + \|x_n - x\|\|y\|. \end{aligned}$$

Since (x_n) converges in X, (x_n) is bounded, i.e. $\exists\, M > 0$ such that $\|x_n\| \leq M \quad \forall\, n \in \mathbb{N}$. (Reason: $\exists\, n_1 \in \mathbb{N}$ such that

$$n \geq n_1 \Rightarrow \|x_n - x\| < 1 \Rightarrow \|x_n\| = \|x_n - x + x\| \leq \|x_n - x\| + \|x\| \leq 1 + \|x\|.$$

Hence we may choose $M := \max\{1 + \|x\|, \|x_1\|, ..., \|x_{n_1-1}\|\}$.) Therefore

$$0 \leq |\langle x_n, y_n\rangle - \langle x, y\rangle| \leq M\|y_n - y\| + \|x_n - x\|\|y\|.$$

By assumptions,$\lim_{n\to\infty} M\|y_n - y\| = 0$ and $\lim_{n\to\infty} \|y\|\|x_n - x\| = 0$. Therefore $\lim_{n\to\infty}(M\|y_n - y\| + \|y\|\|x_n - x\|) = 0$. By the sandwich principle

$$\lim_{n\to\infty} |\langle x_n, y_n\rangle - \langle x, y\rangle| = 0.$$

□

Lemma 3.2.9. *Let X be an inner product space and $A \subset X$, $A \neq \emptyset$.*

(a) $0_X \in A^\perp$;

(b) $A \cap A^\perp = \begin{cases} \{0_X\} & \text{if} \quad 0_X \in A \\ \emptyset & \text{if} \quad 0_X \notin A; \end{cases}$

(c) $\{0_X\}^\perp = X$ *and* $X^\perp = \{0_X\}$;

(d) $A^\perp$ *is a closed linear subspace of* X.

Proof. (a) Since $\langle 0_X, a\rangle = 0 \quad \forall\, a \in A$, we have $0_X \in A^\perp$.
(b) Suppose that $x \in A \cap A^\perp$. Then $\langle x, x\rangle = 0$ and $x = 0_X$. The claim follows since $0_X \in A^\perp$ by (a).
(c) If $A = \{0_X\}$, then $\forall\, x \in X$ we have $\langle x, 0_X\rangle = 0$. Hence $A^\perp = X$.
If $A = X$ and $x \in A^\perp$, then $\langle x, x\rangle = 0$ and hence $x = 0_X$. Therefore $A^\perp = \{0_X\}$ by (a).
(d)To show that $A^\perp$ is a linear subspace of X, let $x, y \in A^\perp$ and $\alpha, \beta \in \mathbb{F}$. Then $\forall\, a \in A$

$$\langle \alpha x + \beta y, a\rangle \overset{3.1.3}{=} \alpha\langle x, a\rangle + \beta\langle y, a\rangle = 0$$

so that $\alpha x + \beta y \in A^\perp$. To show that $A^\perp$ is closed, let (x_n) be a sequence in $A^\perp$ such that $\lim_{n\to\infty} x_n = x \in X$. By Lemma 3.2.8, for all $a \in A$

$$0 = \langle 0_X, a\rangle = \langle \lim_{n\to\infty}(x_n - x), a\rangle = \lim_{n\to\infty}\langle x_n - x, a\rangle = \lim_{n\to\infty}(\langle x_n, a\rangle - \langle x, a\rangle) = -\langle x, a\rangle.$$

Since $x_n \in A^\perp \Rightarrow \langle x, a\rangle = 0$. Hence $x \in A^\perp$ and $A^\perp$ is closed (see Rynne & Youngson, Theorem 1.25(c)). □

Minimization on Hilbert spaces.

Definition 3.2.10. Let X be an inner product space. If X is complete as a metric space induced by the induced norm, we call X a *Hilbert space*.

Lemma 3.2.11. *Let Y be a linear subspace of an inner product space X. Then*

$$x \in y^\perp \Leftrightarrow \|x - y\| \geq \|x\| \quad \forall\, x \in Y.$$

Proof. . For all $x \in X, y \in Y$ and $\alpha \in \mathbb{F}$ (by Lemma 3.1.6(c))

$$\|x - \alpha y\|^2 = \langle x - \alpha y, x - \alpha y\rangle = \|x\|^2 - \overline{\alpha}\langle x, y\rangle - \alpha\langle y, x\rangle + |\alpha|\|y\|^2 \quad (*).$$

($\Rightarrow$) Suppose that $x \in Y^\perp$ and $y \in Y$. Then $\langle x, y\rangle = 0 = \langle y, x\rangle$. So choosing $\alpha = 1$ in $(*)$ we have

$$\|x - y\|^2 = \|x\|^2 + \|y\|^2 \geq \|x\|^2.$$

($\Leftarrow$) Suppose that $x \in X$ and $\|x - y\|^2 \geq \|x\|^2 \quad \forall\, y \in Y$. Since Y is a linear subspace, $\alpha y \in Y \quad \forall\, \alpha \in \mathbb{F}, y \in Y$, and $(*)$ implies that

$$-\overline{\alpha}\langle x, y\rangle - \alpha\langle y, x\rangle + |\alpha|^2\|y\|^2 \geq 0. \quad (**)$$

For given $y \in Y$, we want to prove that $\langle x, y\rangle = 0$. Assume that $\langle x, y\rangle \neq 0$. Denote $\alpha := t\frac{|\langle x,y\rangle|}{\langle y,x\rangle}$ for $t > 0$. We replace α in $(**)$ and obtain

$$-t\frac{|\langle x, y\rangle|}{\overline{\langle y, x\rangle}}\langle x, y\rangle - t\frac{|\langle x, y\rangle|}{\langle y, x\rangle}\langle y, x\rangle + t^2\frac{|\langle x, y\rangle|^2}{|\langle y, x\rangle|^2}\|y\|^2 \geq 0$$

$$\Leftrightarrow \qquad |\langle x, y\rangle| \leq \frac{1}{2}t\|y\|^2 \quad \forall\, t > 0.$$

Hence $\langle x, y\rangle = 0$ and $x \in Y^\perp$. □

Example. Let $Y = \mathbb{R}^2 \times \{0\} \subset \mathbb{R}^3$ and $Y^\perp = \{0\}^2 \times \mathbb{R}$, see Example after Definition 3.2.6.

Definition 3.2.12. A subset A of a vector space X is *convex* if for all $x, y \in A$ and $\lambda \in [0, 1]$ we have $\quad \lambda x + (1 - \lambda)y \in A$.

Example. $A = \{x \in \mathbb{R}^2 : \|x\| \leq 1\}$ is convex but $B = \{x \in \mathbb{R}^2 : \|x\| = 1\}$ is not convex.

Theorem 3.2.13. *Let A be a non-empty closed convex subset of a Hilbert space $\mathcal{H}$ and let $p \in \mathcal{H}$. Then there exists a unique $q \in A$ such that*

$$\|p - q\| = \inf\{\|p - a\| : a \in A\}(= \min\{\|p - a\| : a \in A\}).$$

Proof. Exercise. □

Remark. In any metric space X and for any $A \subset X, A \neq \emptyset$, we may define the *distance between A and x* by

$$\mathrm{d}(x, A) = \inf\{\mathrm{d}(x, a) : a \in A\}.$$

If A is compact, inf is attained since we can prove that $x \mapsto \mathrm{d}(x, A)$ is continuous. The point is that the convexity quarantees *uniqueness*, which is important for applications e.g. convex optimization and variational calculus.

Example. Let $A = \{x \in \mathbb{R}^2 : \|x\| = 1\}$ and let $x = (0, 0)$. Then all points in A are distance-minimizing!

Theorem 3.2.14. *Let Y be a closed linear subspace of a Hilbert space $\mathcal{H}$. Then for any $x \in \mathcal{H}$ exists unique $y \in Y$ and $z \in Y^\perp$ such that $x = y + z$. Moreover, $\|x\|^2 = \|y\|^2 + \|z\|^2$.*

Proof. Exercise. □

Example. Let $\mathcal{H} = \mathbb{R}^2$ and $Y = \mathbb{R} \times \{0\}$. It is easy to prove that $Y^\perp = \{0\} \times \mathbb{R}$. In this case Theorem 3.2.14 is a version of the classical Pythagoras Theorem.

Suppose that Y is closed linear subspace of a Hilbert space $\mathcal{H}$ and $x \in \mathcal{H}$. The decomposition

$$x = y + z, \; y \in Y, z \in Y^\perp$$

is called the orthogonal decomposition of x with respect to Y. We denote $Y^{\perp\perp} = (Y^\perp)^\perp$.

Corollary 3.2.15. *If Y is a closed linear subspace of a Hilbert space $\mathcal{H}$, then $Y^{\perp\perp} = Y$.*

Proof. Exercise. □

Remark. We can also prove that $Y^{\perp\perp} = \overline{Y}$ (closure of Y) if Y is a linear subspace of $\mathcal{H}$ (see Rynne & Youngson p.71).

3.3. **Orthonormal bases in infinite dimensions.**

Definition 3.3.1. Let X be an inner product space. A sequence (e_n) in X is called an *orthonormal* sequence if

(i) $\|e_n\| = 1 \quad \forall\, n \in \mathbb{N}$;
(ii) $\langle e_n, e_m \rangle = 0 \quad \forall\, n, m \in \mathbb{N}, \; n \neq m$.

Example 3.3.2. (a) Let $\tilde{e_1} = (1, 0, 0, ...), \tilde{e_2} = (0, 1, 0, ...), \dots \tilde{e_n} = (\overbrace{0, ..., 0}^{n-1}, 1, 0, ...) \quad n \in \mathbb{N}$. Then $\tilde{e_n} \in l^p, 1 \leq p \leq \infty \quad (\|e_n\| = 1 \quad \forall\, p)$, and $(\tilde{e_n})$ forms an orhonormal sequence in l^2, since

(i) $\|e_n\|_2 = \langle e_n, e_n \rangle = 1 \cdot \overline{1} = 1$
(ii) $\langle e_n, e_m \rangle = 0 \quad$ if $n \neq m$.

(b) For any $[a, b] \subset \mathbb{R}$ we define the space $L^p([a, b])$ by setting $f \in L^p([a, b])$ iff $\tilde{f} \in L^p(\mathbb{R})$, where

$$\tilde{f} = \begin{cases} f & \text{in } [a, b] \\ 0 & \text{in } \mathbb{R} \setminus [a, b]. \end{cases}$$

Moreover, for any $f : [a, b] \to \mathbb{C}$, $f = (f_1, f_2)$, we write

$$f \in L^p_{\mathbb{C}}[a, b] \Leftrightarrow f_i \in L^p[a, b], \quad i = 1, 2.$$

The norm in $L^p_{\mathbb{C}}[a, b]$ is defined as

$$\|f\| = \|f\|_{L^p_{\mathbb{C}}[a,b]} = \left(\int_a^b |f_1(t)|^p \, dt + \int_a^b |f_2(t)|^p \, dt \right)^{\frac{1}{p}}.$$

We define the sequence $(e_n), \; e_n : [-\pi, \pi] \to \mathbb{C}$ by

$$e_n(x) = \frac{1}{\sqrt{2\pi}} e^{inx}, \quad n \in \mathbb{N}.$$

By Euler's formula $e_n(x) = \frac{1}{\sqrt{2\pi}} \Big(\cos(nx) + i \sin(nx) \Big)$. Hence the coodinate function

$$e_n^1(x) = \cos(nx), \quad e_n^2 = \sin(nx)$$

are bounded (and continuous). Therefore $e_n \in L^p_{\mathbb{C}}[-\pi,\pi] \quad \forall\, p$. We claim that (e_n) is an orthonormal sequence in $L^2_{\mathbb{C}}[-\pi,\pi]$ once $L^2_{\mathbb{C}}[-\pi,\pi]$ is equipped with the complex inner product

$$\langle f, g\rangle = \int_{-\pi}^{\pi} f\overline{g}dx.$$

(We omit an "easy" proof that $\langle\cdot,\cdot\rangle$ is an inner product.)

(i) $\|e_n\|_2 = \langle e_n, e_n\rangle = \int_{-\pi}^{\pi} \frac{1}{\sqrt{2\pi}}e^{inx}\cdot\frac{1}{\sqrt{2\pi}}\overline{e^{inx}}dx = \frac{1}{2\pi}\int_{-\pi}^{\pi} \underbrace{e^{inx}\cdot e^{-inx}}_{e^0}dx = \frac{1}{2\pi}\cdot 2\pi = 1$

(ii) Let $m, n \in \mathbb{Z},\ m \neq n$. Then

$$\begin{aligned}
\langle e_m, e_n\rangle &= \int_{-\pi}^{\pi} \frac{1}{\sqrt{2\pi}}e^{imx}\cdot\frac{1}{\sqrt{2\pi}}\overline{e^{inx}}dx \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(m-n)x}dx \\
&= \frac{1}{2\pi}\Big(\int_{-\pi}^{\pi}\cos(m-n)xdx, \int_{-\pi}^{\pi}\sin(m-n)xdx\Big) \\
&= \frac{1}{2\pi}(0,0) \\
&= (0,0)
\end{aligned}$$

Remark 3.3.3. (a) It is clear that X is infinite-dimensional if it contains an orthonormal sequence. Indeed, if (e_n) is an orthonormal sequence in X and $\dim X = k < \infty$, then $\{e_1, ..., e_k\}$ is a basis for X and (Lemma 3.2.3)

$$e_{k+1} = \sum_{i=1}^{k}\langle e_{k+1}, e_i\rangle e_i = 0_X.$$

This contradicts with $\|e_{k+1}\| = 1$.
(b) Also the converse is true: Any infinite-dimensional inner product space contains an orthonormal sequence. We omit the proof, see Rymme & Youngson, Chapter 3.4.

Question. Let (e_n) be an orthonormal sequence in an infinite-dimensional inner product space X. Then it is natural to ask whether the formula

$$x = \sum_{n=1}^{\infty}\langle x, e_n\rangle e_n \qquad (*)$$

holds? There are two major problems associated with $(*)$:

(a) Does the series converge?
(b) Does it converge to x?

Lemma 3.3.4. *Let $\{e_1, ..., e_k\}$ be an orthonormal subset of an inner product space X. Then, for any $\alpha_n \in \mathbb{F}$, n=1,...,k*

$$\|\sum_{n=1}^{k}\alpha_n e_n\|^2 = \sum_{n=1}^{k}|\alpha_n|^2.$$

Proof. By orthonormality

$$\begin{aligned}
\|\sum_{n=1}^{k}\alpha_n e_n\|^2 &= \sum_{n=1}^{k}\alpha_n e_n, \sum_{m=1}^{k}\alpha_m e_m\rangle \overset{3.1.3}{=} \sum_{n=1}^{k}\alpha_n\langle e_n, \sum_{m=1}^{k}\alpha_m e_m\rangle \\
&\overset{3.1.6}{=} \sum_{n=1}^{k}\alpha_n\sum_{m=1}^{k}\overline{\alpha_m}\langle e_n, e_m\rangle = \sum_{n=1}^{k}\sum_{m=1}^{k}\alpha_n\overline{\alpha_m}\langle e_n, e_m\rangle \\
&= \sum_{n=1}^{k}\alpha_n\overline{\alpha_n} = \sum_{n=1}^{k}|\alpha_n|^2.
\end{aligned}$$

□

Lemma 3.3.5. (Bessel's inequality) *Let X be an inner product space and let (e_n) be an orthonormal sequence in X. Then, for any $x \in X$ the series $\sum_{n=1}^{\infty}|\langle x, e_n\rangle|^2$ converges and*

$$\sum_{n=1}^{\infty}|\langle x, e_n\rangle|^2 \leq \|x\|^2.$$

Proof. Let $x \in X$. For each $k \in \mathbb{N}$, let $y_k := \sum_{n=1}^{k}\langle x, e_n\rangle e_n$. Then (by Lemma 3.3.4)

$$\begin{aligned}
\|x - y_k\|^2 &= \langle x - y_k, x - y_k\rangle \overset{3.1.6(c)}{=} \langle x, x\rangle - \langle x, y_k\rangle - \langle y_k, x\rangle\langle y_k, y_k\rangle \\
&= \|x\|^2 - \sum_{n=1}^{k}\overline{\langle x, e_n\rangle}\langle x, e_n\rangle - \sum_{n=1}^{k}\langle x, e_n\rangle\underbrace{\langle x, e_n\rangle}_{\overline{\langle x, e_n\rangle}} + \|y\|^2 \\
&\overset{3.3.4}{=} \|x\|^2 - 2\sum_{n=1}^{k}|\langle x, e_n\rangle|^2 + \sum_{n=1}^{k}|\langle x, e_n\rangle|^2 \\
&= \|x\|^2 - \sum_{n=1}^{k}|\langle x, e_n\rangle|^2
\end{aligned}$$

Therefore

$$\sum_{n=1}^{k}|\langle x, e_n\rangle|^2 = \|x\|^2 - \|x - y_k\|^2 \leq \|x\|^2.$$

Hence the sequence $(\sum_{n=1}^{k}|\langle x, e_n\rangle|^2)$ is upper bounded, $\|x\|^2$ as an upper bound. The partial sums form an increasing sequence and therefore

$$\sum_{n=1}^{\infty}|\langle x, e_n\rangle|^2 = \lim_{k\to\infty}\sum_{n=1}^{k}|\langle x, e_n\rangle|^2 = \sup_{k\in\mathbb{N}}\sum_{n=1}^{k}|\langle x, e_n\rangle|^2 \leq \|x\|^2.$$

□

Note. A series $\sum_{n=1}^{\infty}x_n$ in a normed space X converges if $\exists\ x \in X$ such that

$$x = \lim_{k\to\infty}\sum_{n=1}^{k}x_n \Leftrightarrow \lim_{k\to\infty}\|\sum_{n=1}^{k}x_n - x\| = 0.$$

In this case we write $x = \sum_{n=1}^{\infty}x_n$.

Theorem 3.3.6. *Let $\mathcal{H}$ be a Hilbert space and let (e_n) be an orthonormal sequence in $\mathcal{H}$. Then the series $\sum_{n=1}^{\infty} \alpha_n e_n$ converges iff $\sum_{n=1}^{\infty} |\alpha_n|^2 < \infty, \alpha_n \in \mathbb{F}$. If this holds, then*

$$\sum_{n=1}^{\infty} |\langle x, e_n\rangle|^2 = \|x\|^2.$$

Proof. ($\Rightarrow$) Exercise.
($\Leftarrow$) Suppose that $\sum_{n=1}^{\infty} |\alpha_n|^2 < \infty$. For each $k \in \mathbb{N}$, let $x_k := \sum_{n=1}^{k} \alpha_n e_n$. Since $\sum_{n=1}^{\infty} |\alpha_n|^2 < \infty$, the partial sums of this series form a Cauchy sequence. Therefore, for each $\varepsilon > 0$, $\exists\ n_\varepsilon$ so that

$$\text{if } k > j \geq n_\varepsilon, \text{ then} \qquad \|\sum_{n=1}^{k} |\alpha_n|^2 - \sum_{n=1}^{j} |\alpha_n|^2\| = \sum_{n=j+1}^{k} |\alpha_n|^2 < \varepsilon.$$

By Lemma 3.3.4, for $k > j$,

$$\|x_k - x_j\|^2 = \|\sum_{n=j+1}^{k} \alpha_n e_n\|^2 \overset{3.3.4}{=} \sum_{n=j+1}^{k} |\alpha_n|^2 < \varepsilon$$

whenever $j \geq n_\varepsilon$. Hence (x_k) is a Cauchy sequence in $\mathcal{H}$ and by completeness it converges in $\mathcal{H}$. Finally, by Lemma 1.2.3(ii) and Lemma 3.3.4

$$\|\sum_{n=1}^{\infty} \alpha_n e_n\|^2 = \|\lim_{k\to\infty} \sum_{n=1}^{k} \alpha_n e_n\|^2 \overset{1.2.3}{=} \lim_{k\to\infty} \|\sum_{n=1}^{k} \alpha_n e_n\|^2 \overset{3.3.4}{=} \lim_{k\to\infty} \|\sum_{n=1}^{k} \alpha_n\|^2 = \sum_{n=1}^{\infty} |\alpha_n|^2.$$

□

Remark. In other words, Theorem 3.3.6 says that $\sum_{n=1}^{\infty} \alpha_n e_n$ converges iff $(a_n) \in l^2$.

Corollary 3.3.7. *Let (e_n) be an orthonormal sequence in a Hilbert space $\mathcal{H}$. Then $\sum_{n=1}^{\infty} \langle x, e_n\rangle e_n$ converges in $\mathcal{H}$ for any $x \in \mathcal{H}$.*

Proof. By Bessel's inequality,

$$\sum_{n=1}^{\infty} |\langle x, e_n\rangle|^2 < \infty \qquad \forall\ x \in \mathcal{H}.$$

Hence, by Theorem 3.3.6 $\sum_{n=1}^{\infty} \langle x, e_n\rangle e_n$ converges. □

By Corollary 3.3.7, the answer to Question (a) is always positive in *Hilbert spaces.* The answer to Question (b) requires some additional assumptions on (e_n):

Example. Let (e_n) be an orthonormal sequence in a Hilbert space and let s be the sequence $s = (e_{2n})$. Then s is an orthonormal sequence in $\mathcal{H}$.

Claim. $e_1 \neq \sum_{n=1}^{\infty} \langle e_1, e_{2n}\rangle e_{2n}$

Proof. Suppose that $e_1 = \sum_{n=1}^{\infty} \alpha_n e_{2n}$ for $\alpha_n \in \mathbb{F}$. Then, by Lemma 3.2.8, for all $m \in \mathbb{N}$

$$0 = \langle e_1, e_{2m}\rangle \overset{3.2.8}{=} \lim_{k\to\infty} \langle \sum_{n=1}^{k} \alpha_n e_{2n}, e_{2m}\rangle = \lim_{k\to\infty} \sum_{n=1}^{k} \alpha_n \langle e_{2n}, e_{2m}\rangle \overset{k\geq m}{=} \lim_{k\to\infty} \alpha_m = \alpha_m.$$

Hence $e_1 = 0_{\mathcal{H}}$ which contradicts with $\|e_1\| = 1$. □

Definition 3.3.8. Let X be a normed space and let $E \subset X,\ E \neq \emptyset$. Then the *closed linear span* of E, denoted by $\overline{Sp}E$, is the intersection of all closed linear subspaces which contain E.

Definition 3.3.8 makes sense since any intersection

- of linear subspaces is a linear subspace
- of closed sets is closed

Thus $\overline{Sp}E$ is the smallest closed linear subspace that contains E.

Theorem 3.3.9. *Let $\mathcal{H}$ be a Hilbert space and let (e_n) be an orthonormal sequence. The following are equivalent:*

(a) $\{e_n : n \in \mathbb{N}\}^{\perp} = \{0_{\mathcal{H}}\}$
(b) $\overline{Sp}\{e_n : n \in \mathbb{N}\} = \mathcal{H}$
(c) $\|x\|^2 = \sum_{n=1}^{\infty} |\langle x, e_n\rangle|^2$ *for all* $x \in \mathcal{H}$
(d) $x = \sum_{n=1}^{\infty} \langle x, e_n\rangle e_n$ *for all* $x \in \mathcal{H}$

Proof. We proof that (a)⇒(d)⇒(b)⇒(a) and (a)⇒(d)⇒(c)⇒(a).

(a)⇒(d) Let $x \in \mathcal{H}$ and let $y = x - \sum_{n=1}^{\infty}\langle x, e_n\rangle e_n$ (see Corollary 3.3.7). For each $n \in \mathbb{N}$, by Lemma 3.2.8,

$$\begin{aligned}
\langle y, e_n\rangle &= \langle x, e_m\rangle - \langle \lim_{k\to\infty} \sum_{n=1}^{k} \langle x, e_n\rangle e_n, e_m\rangle \\
&\overset{3.2.8}{=} \langle x, e_m\rangle - \lim_{k\to\infty} \langle \sum_{n=1}^{k} \langle x, e_n\rangle e_n, e_m\rangle \\
&= \langle x, e_m\rangle - \lim_{k\to\infty} \sum_{n=1}^{k} \underbrace{\langle x, e_n\rangle\langle e_n, e_m\rangle}_{\langle x, e_m\rangle \ for\ k\ \geq\ m} \\
&= \langle x, e_m\rangle - \langle x, e_m\rangle = 0.
\end{aligned}$$

Hence $y \in \{e_m : m \in \mathbb{N}\}^{\perp} = \{0_{\mathcal{H}}\}$ so that $y = 0_{\mathcal{H}}$ and (d) holds.

(d)⇒(b) By assumption, for any $x \in \mathcal{H}$, we have $x = \lim_{k\to\infty} \sum_{n=1}^{k} \langle x, e_n\rangle e_n$. But

$$\sum_{n=1}^{k} \langle x, e_n\rangle e_n \in \ Sp\{e_1, ..., e_k\} \subset \overline{Sp}\{e_n : n \in \mathbb{N}\}$$

and therefore $x \in \overline{Sp}\{e_n : n \in \mathbb{N}\}$ since $\overline{Sp}\{e_n : n \in \mathbb{N}\}$ is closed.
Hence $\mathcal{H} \subset \overline{Sp}\{e_n : n \in \mathbb{N}\}$.

(d)⇒(c) Since $x = \lim_{k\to\infty} \sum_{n=1}^{k} \langle x, e_n\rangle e_n$ for any $x \in \mathcal{H}$, we have

$$\|x\|^2 \overset{1.2.3}{=} \lim_{k\to\infty} \| \sum_{n=1}^{k} \langle x, e_n\rangle e_n\|^2 \overset{3.3.4}{=} \lim_{k\to\infty} \sum_{n=1}^{k} |\langle x, e_n\rangle|^2 = \sum_{n=1}^{\infty} |\langle x, e_n\rangle|^2$$

by Lemma 1.2.3 and Lemma 3.3.4.

(b)⇒(a) Suppose that (b) holds and let $y \in \{e_n : n \in \mathbb{N}\}^{\perp}$. Then $\langle y, e_n\rangle = 0 \quad \forall\ n \in \mathbb{N}$, so that $e_n \in \{y\}^\perp$ for all $n \in \mathbb{N}$. By Lemma 3.2.9 (d) $\{y\}^\perp$ is a closed linear subspace. Hence

$$\mathcal{H} = \overline{Sp}\{e_n : n \in \mathbb{N}\} \subset \{y\}^\perp$$

and so $y \in \{y\}^\perp$. Therefore $\langle y, y\rangle = 0$ i.e. $y = 0_{\mathcal{H}}$.

(c)⇒(a) If $x \in \{e_n : n \in \mathbb{N}\}^\perp$, then $\langle x, e_n\rangle = 0$ for any $n \in \mathbb{N}$. Hence by (c),

$$\|x\|^2 = \sum_{n=1}^{\infty} |\langle x, e_n\rangle|^2 = 0,$$

so that $x = 0_{\mathcal{H}}$. We have proved that $\{e_n : n \in \mathbb{N}\}^\perp \subset \{0_{\mathcal{H}}\}$. The converse is clear. □

Definition 3.3.10. Let $\mathcal{H}$ be a Hilbert space and let (e_n) be an orthonormal sequence in $\mathcal{H}$. Then (e_n) is called *orthonormal basis* for $\mathcal{H}$ if the conditions (a)-(d) of Theorem 3.3.9 hold.

The scalars $\langle x, e_n\rangle$ in Theorem 3.3.9 (d) are often called the *Fourier coefficients* of x with respect to the basis (e_n).

Example. The orthonormal sequence $(\tilde{e_n})$ in l^2,

$$\tilde{e_n} = (0, ..., 0, \underbrace{1}_{n}, 0, ...)$$

is an orthonormal basis in l^2 (the standard orthonormal basis in l^2).

Proof. Let $x := (x_n) \in l^2$. By definitions,

$$\|x\|_2^2 = \sum_{n=1}^{\infty} |x_n|^2 = \sum_{n=1}^{\infty} |\langle x, \tilde{e_n}\rangle|^2,$$

i.e. Theorem 3.3.9(c) holds. □

Note. It is usually not so easy to decide whether the given orthonormal sequence is a basis or not, see Fourier series below.

Definition 3.3.11. A metric space X is called *separable* if it has a countable subset $E \subset X$ such that $\overline{E} = X$ (i.e. E is dense in X).

Example. It is well known that $\mathbb{Q}$ is dense in $\mathbb{R}$. Hence $\mathbb{R}$ is separable with respect to euclidean metric.

Theorem 3.3.12.
(a) *Finite dimensional normed spaces are separable.*
(b) *Infinite dimensional Hilbert space $\mathcal{H}$ is separable iff $\mathcal{H}$ has an orthonormal basis.*

Proof. (a) Let X be a finite-dimensional, real normal space and let $\{e_1, .., e_k\}$ be a basis for X. Then the set

$$E = \{\sum_{n=1}^{k} \alpha_n e_n : \alpha_n \in \mathbb{Q}\}$$

is countable since $\mathbb{Q}^k$ is countable. The claim $\overline{E} = X$ can be proved as in the proof of (b) below. In the complex case we define E similarly by using scalars

$$\alpha_n = p_n + iq_n, \qquad \text{where } p_n, q_n \in \mathbb{Q}.$$

Such numbers α_n are called *complex rationals.*

(b) Suppose that $\mathcal{H}$ has an orthonormal basis (e_n). For fixed $k \in \mathbb{N}$, let

$$E_k = \{\sum_{n=1}^{k} \alpha_n e_n : \alpha_n \text{rational (complex rational)}\}.$$

Then E_k is countable and also $E = \cup_{k=1}^{\infty} E_k$ is countable. We show that $\overline{E} = \mathcal{H}$.
Let $y \in \mathcal{H}$. By assumptions (and Theorem 3.3.9(d))

$$y = \sum_{n=1}^{\infty} \beta_n e_n, \quad \sum_{n=1}^{\infty} |\beta_n|^2 < \infty, \quad \beta_n = \langle y, e_n \rangle.$$

For any $\varepsilon > 0$, there is $N \in \mathbb{N}$ such that $\sum_{n=N+1}^{\infty} |\beta_n|^2 < \frac{\varepsilon^2}{2}$. For each $n = 1, ..., N$ choose rational (complex rational) coefficients such that $|\beta_n - \alpha_n|^2 < \frac{\varepsilon^2}{2N}$, and let $x = \sum_{n=N}^{\infty} \alpha_n e_n \in E$. Then

$$y - x = \sum_{n=1}^{\infty} \gamma_n e_n, \qquad \text{where } \gamma_n = \begin{cases} \beta_n - \alpha_n, & \text{if } 1 \le n \le N \\ \beta_n, & \text{if } n \ge N+1 \end{cases}$$

We obtain that (see Theorem 3.3.9; the proof of (d)$\Rightarrow$(c))

$$\|y - x\|^2 = \sum_{n=1}^{\infty} |\gamma_n|^2 = \sum_{n=1}^{N} |\beta_n - \alpha_n|^2 + \sum_{n=N+1}^{\infty} |\beta_n|^2 < N \cdot \frac{\varepsilon^2}{2N} + \frac{\varepsilon^2}{2} = \varepsilon^2,$$

i.e. $\|y - x\| < \varepsilon$. Hence $y \in \overline{E}$ and $\overline{E} = \mathcal{H}$. We skip the proof that every separable Hilbert space has an orthonormal basis, see Rynne & Youngson p.80. $\square$

Corollary 3.3.13. *The Hilbert space l^2 is separable.*

Example 3.3.14. *(Briefly on Fourier series; no details)* One can prove that

$$C = (c_n), \qquad \text{where } c_0(x) = \sqrt{\frac{1}{\pi}} \qquad \text{and } c_n(x) = \sqrt{\frac{2}{\pi}} \cos nx, \; n \in \mathbb{N},$$

is an orthonormal basis in $L^2[0, \pi]$.

The idea of the proof:

(1) Orthonormality is a calculus-exercise.
(2) By Theorem 3.3.9(d) it suffices to show that SpC (finite linear combinations of functions in C) is dense in $L^2[0, \pi]$.
(3) Suppose that $f \in L^2[0, \pi]$. Recall that f is real valued. It is well-known fact in L^p-theory that $\mathcal{C}[0, \pi]$ is dense in $L^2[0, \pi]$, i.e. for a given $\varepsilon > 0$ there is $g_1 \in \mathcal{C}[0, \pi]$ such that $\|f - g_1\|_2 < \frac{\varepsilon}{2}$.
(4) Using the Stone-Weierstrass theorem (see Rymme & Youngson, Theorem 1.39) polynomials are dense in $\mathcal{C}[0, \pi]$ with respect to sup-norm plus some trigonometry one can prove that

$$\exists \; g_2, \; g_2(x) = \sum_{n=0}^{m} \beta_n (\cos nx) \qquad \text{such that } \|g_1 - g_2\| < \frac{\varepsilon}{2}.$$

(5) It then follows that $\|f - g_2\| < \varepsilon$.

As a consequence we conclude that $L^2[0, \pi]$ is separable! Moreover, any function $f \in L^2[0, \pi]$ (for example any $f \in \mathcal{C}[0, \pi]$) can be written as a sum

$$f = \sum_{n=0}^{\infty} \langle f, c_n \rangle c_n.$$

Here the convergence of the series is understood in L^2-*sense*.

One can also proof that

$$S = (s_n), \qquad s_n(x) = \sqrt{\frac{2}{\pi}} \sin nx$$

is an orthonormal basis in $L^2[0, \pi]$ and

$$E = (e_n), \qquad e_n(x) = \frac{1}{\sqrt{2\pi}} e^{inx}$$

in $L^2_{\mathbb{C}}[-\pi, \pi]$.

4. Dual spaces

4.1. **The space $B(X,Y)$.** Recall that $B(X,Y)$ denotes the normed space of bounded linear operators $T : X \to Y$ whenever X and Y are normed spaces, see Theorem 2.2.4. The norm of T is defined by

$$\|T\| = \sup\{\|T(x)\| : \|x\| \le 1\}.$$

Theorem 4.1.1. *If X is a normed space and Y is a Banach space, then $B(X,Y)$ is a Banach space.*

Proof. We have to show that $B(X,Y)$ is complete. Let (T_n) be a Cauchy sequence in $B(X,Y)$. Then (T_n) is a bounded sequence, so. $\exists M > 0$ such that

$$\|T_n\| \le M \quad \forall n \in \mathbb{N}.$$

Let $x \in X$. As

$$\|T_n(x) - T_m(x)\| = \|(T_n - T_m)(x)\| \le \|T_n - T_m\|\|x\|$$

(see Remark 2.2.2 (∗∗)), it follows that $(T_n(x))$ is a Cauchy sequence in Y. (In fact, for $\varepsilon > 0, \exists n_\varepsilon \in \mathbb{N}$ such, that $\|T_n - T_m\| < \frac{\varepsilon}{\|x\|}$ if $m, n \ge n_\varepsilon$ and $\|x\| > 0$.) Since Y is complete, $(T_n(x))$ converges in Y, so we may define a mapping $T : X \to Y$ by

$$T(X) = \lim_{n\to\infty} T_n(x).$$

We show first that T is linear. For any $x, y \in X$ and $\alpha, \beta \in \mathbb{F}$ (scalar field of X) we have

$$\begin{aligned} T(\alpha x + \beta y) &= \lim_{n\to\infty} T_m(\alpha x + \beta y) \overset{T_m\, lin.}{=} \lim_{m\to\infty} \alpha T_m(x) + \beta T_m(y) \\ &= \alpha \lim_{n\to\infty} T_n(x) + \beta \lim_{m\to\infty} T_m(x) = \alpha T(x) + \beta T(y). \end{aligned}$$

Next we show that T is bounded. As

$$\|T(x)\| = \lim_{n\to\infty} \|T_n(x)\|$$

by Lemma 1.2.3, we obtain

$$\begin{aligned} \|T(x)\| \quad &\le \sup\{\|T_n(x)\| : n \in \mathbb{N}\} \\ &\overset{2.2.2}{\le} \sup\{\|T_n(x)\| : n \in \mathbb{N}\} \\ &\le M\|x\|. \end{aligned}$$

Hence $T \in B(X,Y)$.

Finally we show that $\lim_{n\to\infty} T_n = T$ in $\|\cdot\|$. Let $\varepsilon > 0$. Since (T_n) is a Cauchy sequence $\exists n_1 \in \mathbb{N}$ such that

$$\|T_n - T_m\| < \frac{\varepsilon}{2} \quad \text{if } m, n \ge n_1.$$

Hence, for any $x \in X$ with $\|x\| \le 1$,

$$\|T_n(x) - T_m(x)\| \le \|T_n - T_m\|\|x\| < \frac{\varepsilon}{2}$$

whenever $m, n \ge n_1$. As $T(x) = \lim_{n\to\infty} T_n(x)$, there is $n_2 \ge n_1$ depending on $x \in X$ such that

$$\|T(x) - T_m(x)\| < \frac{\varepsilon}{2} \quad \text{if } m \ge n_2.$$

Hence, if $\|x\| \le 1, n \ge n_1$ and $m \ge n_2$, we conclude that

$$\|T(x) - T_n(x)\| \le \|T(x) - T_m(x)\| + \|T_n(x) - T_m(x)\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Therefore

$$\|T - T_m\| = \sup\{\|T(x) - T_n(x)\| : \|x\| \le 1\} \le \varepsilon$$

if $n \ge n_\varepsilon$. This shows that $\lim_{n\to\infty} T_n = T$, i.e $B(X,Y)$ is a Banach space. □

Lemma 4.1.2. *Let X, Y and Z be normed spaces and let $T \in B(X,Y)$ and $S \in B(Y,Z)$. Then $S \circ T \in B(X,Z)$ and*

$$\|S \circ T\| \le \|S\|\|T\|.$$

Proof. Exercise. □

In finite-dimensional spaces X, Y and Z, the matrix of the composite $S \circ T$ is the product of the matrixes of S and T. Hence the function composition is a natural candidate for the product of bounded linear operators.

Definition 4.1.3. Let X, Y, Z be normed spaces and let $T \in B(X,Y), S \in B(Y,Z)$. Then $S \circ T$ is called *product of S and T*. We denote

$$ST := S \circ T.$$

In general, ST and TS are both defined only if $X = Y = Z$. Even in this case, in general holds

$$TS \ne ST.$$

Notation. If X is a normed space, we denote $B(X) := B(X,X)$.

Lemma 4.1.4. *Let X be a normed space. Then*

(a) *$B(X)$ is a ring with the identity I ($I(x) = x$);*
(b) *If (T_n) and (S_n) are sequences in $B(X)$ such that $\lim_{n\to\infty} T_n = T$ and $\lim_{n\to\infty} S_n = S$, then*

$$\lim_{n\to\infty} S_n T_n = ST.$$

Proof. (a) Since $B(X)$ is a vector space, $B(X)$ is an Abelian group with respect to $+$ (pointwise sum). We should show that $\forall R, S, T \in B(X)$

(1) $R(ST) = (RS)T$,
(2) $R(S+T) = RS + RT$ and $(R+S)T = RT + ST$,
(3) $IR = RI = R$.

Here (1) and (3) are trivial. For all $x \in X$, we have

$$\begin{aligned} (R(S+T))(x) &= (R \circ (S+T))(x) = R((S+T)(x)) = R(S(x) + T(x)) \\ &\overset{R\,lin.}{=} R(S(x)) + R(T(x)) = (R \circ S)(x) + (R \circ T)(x) \\ &= (RS + RT)(x). \end{aligned}$$

The other equality in (2) is similar.
(b) Exercise. □

Notation. Let X ba a normed space and let $T \in B(X)$.

(a) Then $T^2 = T \circ T,\ T^3 = T^2 \circ T,\ \ldots,\ T^n = T^{n-1} \circ T$.

(b) If $a_0, \dots, a_n \in \mathbb{F}$ and $p : \mathbb{F} \to \mathbb{F}$ is polynomial $p(x) = a_n x^n + \ldots + a_1 x + a_0$, then we define $p(T)$ by $p(T) = a_n T^n + \ldots + a_1 T + a_0$.

Definition 4.1.5. Let X be a normed space over $\mathbb{F}$. The space $B(X, \mathbb{F})$ is called the *dual space* of X. We denote $X' := B(X, \mathbb{F})$.

Corollary 4.1.6. *If X is a normed space, then X is a Banach space.*

Proof. Since $\mathbb{F} = \mathbb{R}$ of $\mathbb{F} = \mathbb{C}$, the claim follows from Theorem 4.1.1. □

Example 4.1.7. Let $\mathcal{H}$ be a Hilbert space over $\mathbb{F}$ and let $y \in \mathcal{H}$. Define $f : \mathcal{H} \to \mathbb{F}$ by

$$f(x) = \langle x, y \rangle.$$

Then $f \in \mathcal{H}'$ and $\|f\| = \|y\|$ (Exercise).

Theorem 4.1.8. (Riesz-Frechet Theorem). *If $\mathcal{H}$ is a Hilbert space and $f \in \mathcal{H}'$, then there is a unique $y \in \mathcal{H}$ such that*

$$f(x) = \langle x, y \rangle$$

for all $x \in \mathcal{H}$. Moreover, $\|f\| = \|y\|$.

For the proof we need a simple lemma.

Lemma 4.1.9. *If X and Y are normed spaces and $T \in B(X, Y)$, then*

$$Ker(T) = \{x \in X : T(x) = 0_Y\} = T^{-1}(\{0_Y\})$$

is a closed linear subspace of X.

Proof. $Ker(T)$ is a linear subspace, since for all $x, x' \in Ker(T)$ and for all $\alpha, \beta \in \mathbb{F}$

$$T(\alpha x + \beta x') \overset{T\,lin.}{=} \alpha \underbrace{T(x)}_{0_Y} + \beta \underbrace{T(x')}_{0_Y} = 0_Y.$$

Hence $\alpha x + \beta x' \in Ker(T)$. Since T is a bounded operator, $T : X \to Y$ is continuous (Lemma 2.1.2). Since $\{0_Y\}$ is closed, $Ker(T)$ is closed (we regard known that the pre-image of a closed set is closed if the mapping is continuous.) □

Proof of Theorem 4.1.8. (1) Existence: If $f = 0$, then $y = 0_{\mathcal{H}}$ will do. Assume that $f \neq 0$. Then $Ker(f)$ is a proper closed subspace of $\mathcal{H}$, which implies that $Ker(f)^\perp \neq \{0_{\mathcal{H}}\}$. In fact, if $Ker(f)^\perp = \{0_{\mathcal{H}}\}$, then

$$Ker(f)^{\perp\perp} = \{0_{\mathcal{H}}\}^\perp = \mathcal{H}$$

(L. 3.2.9 (c)). By corollary 3.2.15,

$$Ker(f) = Ker(f)^{\perp\perp} = \mathcal{H},$$

which is a contradiction, since $Ker(f)$ is a proper subset of $\mathcal{H}$. Hence $\exists z' \in Ker(f)^\perp \setminus \{0_{\mathcal{H}}\}$. Now $f(z') \neq 0$ (see Lemma 3.2.9 (b)) and for

$$z = \frac{z'}{f(z')}$$

it holds $z \neq 0_{\mathcal{H}}$,

$$f(z) = f\left(\frac{z'}{f(z')}\right) \overset{f\,lin.}{=} \frac{1}{f(z')} f(z') = 1.$$

Choose $y = \frac{z}{\|z\|^2}$. By linearity of f,

$$f(x - f(x)z) = f(x) - f(x)f(z) = 0,$$

and hence $x - f(x)z \in Ker(f)$ $\forall x \in \mathcal{H}$. Since $z \in Ker(f)^\perp$ $(z = \alpha z')$, we have

$$\langle x - f(x)z, z\rangle = 0 \iff \langle x, z\rangle - f(x)\langle z, z\rangle = 0.$$

It then follows that

$$f(x) = \frac{\langle x, z\rangle}{\|z\|^2} = \langle x, \frac{z}{\|z\|^2}\rangle = \langle x, y\rangle$$

for all $x \in \mathcal{H}$. The claim $\|f\| = \|y\|$ is an exercise.

(2) Uniqueness: If $y_1, y_2 \in \mathcal{H}$ are such that

$$f(x) = \langle x, y_1\rangle = \langle x, y_2\rangle \quad \forall x \in \mathcal{H}.$$

Then $\langle x, y_1 - y_2\rangle = 0$ $\forall x \in \mathcal{H}$. By choosing $x = y_1 - y_2$ we get $\|y_1 - y_2\|^2 = 0$. Hence $y_1 = y_2$. □

It is often a challenge to characterize the dual of a given space. However, the dual of ℓ^1 is relatively easy to identify:

Theorem 4.1.10. *Let $c = (c_n) \in \ell^\infty$.*

(a) *If $(x_n) \in \ell^1$, then $(c_n x_n) \in \ell^1$. If the linear transformation $f_c : \ell^1 \to \mathbb{F}$ is defined by*

$$f_c((x_n)) = \sum_{n=1}^{\infty} c_n x_n,$$

then $f_c \in (\ell^1)'$ with

$$\|f_c\| \leq \|c\|_\infty.$$

(b) *If $f \in (\ell^1)'$, there exists $c \in \ell^\infty$ such that $f = f_c$ and $\|c\|_\infty \leq \|f\| = \|f_c\|$.*

(c) *There is a bijective isometry between ℓ^∞ and $(\ell^1)'$.*

Proof. (a) The assertions are included in Example 2.1.4, see also Lemma 2.1.3.
(b) Let $(\tilde{e}_n)$ be the standard orthomormal sequence in ℓ^1. Let $c_n := f(\tilde{e}_n)$, $n \in \mathbb{N}$. Then

$$|c_n| = |f(\tilde{e}_n)| \overset{2.1.1}{\leq} \|f\|\|\tilde{e}_n\|_1 = \|f\|$$

for all $n \in \mathbb{N}$, so that $\|c\|_\infty \leq \|f\|$ (take sup over $n \in \mathbb{N}$). Let S be the linear subspace of ℓ^1 consisting of sequences with only finitely many non-zero terms. Then S is dense in ℓ^1 since for each $x := (x_n) \in \ell^1$ and for each $\varepsilon > 0$ we have $n_\varepsilon \in \mathbb{N}$ such that if $y = (x_1, \ldots, x_{n_\varepsilon}, 0, \ldots) \in S$, then

$$\|x - y\|_1 = \sum_{n=n_{\varepsilon+1}}^{\infty} |x_n| < \varepsilon.$$

For any $z := (z_1, \ldots, z_n, 0, \ldots) \in S$, we have

$$\begin{aligned} f(z) &= f(\sum_{j=1}^{n} z_j \tilde{e}_j) \overset{f\,lin.}{=} \sum_{j=1}^{n} z_j f(\tilde{e}_j) \\ &= \sum_{j=1}^{n} z_j c_j = f_c(z). \end{aligned}$$

Hence the continuous functions f and f_c are equal in a dense subset S of ℓ^1, which implies that $f = f_c$ in ℓ^1 (see Lemma 4.1.11 below).

(c) The mapping $T : \ell^\infty \to (\ell^1)'$, $T(c) = f_c$ for $c := (c) \in \ell^\infty$, is linear (exercise). By (b), T is surjective, and

$$\|c\|_\infty \leq \|f_c\| = \|T(c)\|.$$

By (a),

$$\|f_c\| = \|T(c)\| \leq \|c\|_\infty.$$

Hence $\|T(c)\| = \|c\|_\infty$ for all $c \in \ell^\infty$, i.e. T is an isometry. An isometry is always injective, see Exercise 6.

□

Lemma 4.1.11. *Let X be a metric space and E a dense subset of X. Let $f, g : X \to Y$ be continuous functions (Y is a metric space) such that $f = g$ in E. Then $f = g$.*

Proof. Exercise. □

4.2. **Inverses of operators.** In finite-dimensional vector spaces, the matrix equation

$$Ax = y$$

is solved by $x = A^{-1}y$ whenever A^{-1} exists and y is given. In this subsection, we study the existence of an inverse operator in the case of an infinite-dimensional space.

The basic question is: How to solve $x \in X$ if $T(x) = y$ and$T \in B(X, Y,\ y \in Y$ are given?

Definition 4.2.1. Let X be normed space. An operator $T \in B(X)$ is called *invertible* if $\exists S \in B(X)$ such that $ST = I = TS$. Such an S is called the *inverse* of T. We denote T^{-1} for the inverse of T.

Lemma 4.2.2. *Let X be a normed space and let $T_1, T_2 \in B(X)$ be invertible. Then*
 (a) *T_1^{-1} is invertible with $(T_1^{-1})^{-1} = T_1$;*
 (b) *T_1T_2 is invertible with $(T_1T_2)^{-1} = T_2^{-1}T_1^{-1}$.*

Proof. (a) Clear since

$$T_1^{-1}T_1 = T_1T_1^{-1} = I.$$

(b) Since the product is associative, we have

$$T_2^{-1}T_1^{-1}T_1T_2 = T_2^{-1}IT_2 = T_2^{-1}T_2 = I.$$

Similarly $T_1T_2T_2^{-1}T_1^{-1} = I$. □

Remark 4.2.3. Recall also that if X is a normed space, then for every $R, S, T \in B(X)$
 (a) $R(-S) = (-R)S = -RS$;
 (b) $(-R)(-S) = RS$;
 (c) $(R - S)T = RT - ST$ and $R(S - T) = RS - RT$.

These properties hold true in every ring, see Algebra.

Example 4.2.4. For any $h \in \mathcal{C}[0, 1]$, we define $T_h \in B(L^2[0, 1])$ by

$$(T_hg)(t) = h(t)g(t), \quad t \in [0, 1].$$

(a) If $f \in \mathcal{C}[0, 1]$ is defined by $f(t) = 1 + t$, then T_f is invertible.

Proof. We showed in Exercise 3/1 that T_h is bounded for any $h \in \mathcal{C}[0,1]$. Let $k(t) = \frac{1}{1+t}$. Then $k \in \mathcal{C}[0,1]$ and for any $g \in L^2[0,1]$

$$(T_k T_f g)(t) = T_k(fg)(t) = \underbrace{k(t)f(t)}_{1}g(t) = g(t).$$

Thus

$$(T_k T_f)(g) = g \quad \forall g \in L^2[0,1].$$

Hence $T_k T_f = I_{L^2[0,1]}$.

Similarly, we have $T_f T_k = I_{L^2[0,1]}$, i.e $T_f^{-1} = T_k$. □

(b) Let $f \in \mathcal{C}[0,1]$ be defined by $f(t) = t$. Then the idea in (a) would give the function $k(t) = \frac{1}{t}$. But k is not continuous (or bounded) in $[0,1]$! We can *not* directly conclude that T_f is not invertible as T_f could have an inverse not of the form T_k for $k \in \mathcal{C}[0,1]$.

Theorem 4.2.5. *Let X be a Banach space. If $T \in B(X)$ is an operator with $\|T\| < 1$, $I - T$ is invertible and the inverse is given by*

$$(I-T)^{-1} = \sum_{n=0}^{\infty} T^n.$$

Proof. Because X is Banach, $B(X)$ is Banach (Cor. 4.1.6). Since $\|T\| < 1$, the series $\sum_{n=0}^{\infty} \|T\|^n$ converges, and

$$\|T^n\| \leq \|T\|^n$$

for all $n \in \mathbb{N}$ (Lemma 4.1.2), the series $\sum_{n=0}^{\infty} \|T^n\|$ converges. By Exercise 7/6, the series $\sum_{n=0}^{\infty} T^n$ converges in $B(X)$. Let $S := \sum_{n=0}^{\infty} T^n$ and let $S_k := \sum_{n=0}^{k} T^n$. Hence $\lim_{k\to\infty} S_k = S$ in $B(X)$. We have

$$\begin{aligned} \|(I-T)S_k - I\| &= \|\sum_{n=0}^{k} T^n - \sum_{n=1}^{k+1} T^n - I\| \\ &= \|I - T^{k+1} - I\| = \|-T^{k+1}\| \\ &\overset{4.1.2}{\leq} \|T\|^{k+1}. \end{aligned}$$

Since $\|T\| < 1$, we deduce that

$$\lim_{k\to\infty} (I-T)S_k - I = 0_{B(X)} \iff \lim_{k\to\infty} (I-T)S_k = I. \quad (*)$$

By Lemma 4.1.4 (b)

$$(I-T)S = (I-T)\lim_{k\to\infty} S_k \overset{4.1.4}{=} \lim_{k\to\infty} (I-T)S_k \overset{(*)}{=} I.$$

Similarly, $S(I-T) = I$. Hence $S = (I-T)^{-1}$.

□

Note. The series $\sum_{n=0}^{\infty} T^n$ in Theorem 4.2.5 is called the *Neumann series.*

Example 4.2.6. Let $\lambda \in \mathbb{R}$ and let $k : [0,1] \times [0,1] \to \mathbb{R}$ be defined by

$$k(x,y) = \lambda \sin(x-y)$$

Claim. If $|\lambda| < 1$, then $\forall f \in \mathcal{C}[0,1]$ $\exists g \in \mathcal{C}[0,1]$ such that

$$\begin{aligned} g(x) &= f(x) + \int_0^1 k(x,y)g(y)\,dy \\ &= f(x) + \lambda \int_0^1 \sin(x-y)g(y)\,dy. \quad (*) \end{aligned}$$

Proof. In Example 2.1.8 and Exercise 2/4 we showed that the linear transformation $K : \mathcal{C}[0,1] \to \mathcal{C}[0,1]$,

$$(K(g))(s) = \int_0^1 k(s,t)g(t)\,dt,$$

is bounded and $\|K(g)\| \le |\lambda|\|g\|$. Hence $\|K\| \le |\lambda|$. Since the integral equation $(*)$ can be written as

$$(I-K)g = f$$

and $I-K$ is invertible by Theorem 4.2.5, the equation $(*)$ has the unique solution

$$g = (I-K)^{-1}f.$$

□

Corollary 4.2.7. *Let X be a Banach space. Then the set $\mathcal{A}$ of invertible elements in $B(x)$ is open.*

Proof. The set $\mathcal{A}$ is non-empty since $I \in \mathcal{A}$. Let $T \in \mathcal{A}$ and let $r := \|T^{-1}\|^{-1}$. Notice that $r > 0$ since $\|T^{-1}\|$ implies $T^{-1} \equiv 0$. This contradicts with $TT^{-1} = I$. It suffices to show that $S \in \mathcal{A}$ whenever $\|S-T\| < r$.

Let $S \in B(X)$, $\|T-S\| < r$. Then (Lemma 4.1.2)

$$\begin{aligned} \|(T-S)T^{-1}\| &= \|T-S\|\|T^{-1}\| \\ &< \|T^{-1}\|^{-1}\|T^{-1}\| = 1. \end{aligned}$$

Hence $I-(T-S)T^{-1}$ is invertible by Theorem 4.2.5. However,

$$\begin{aligned} I-(T-S)T^{-1} &= I - TT^{-1} + ST^{-1} \\ &= I - I + ST^{-1} = ST^{-1}. \end{aligned}$$

Therefore ST^{-1} is invertible and $S = (ST^{-1})T$ is invertible (Lemma 4.2.2 (b)). Hence $S \in \mathcal{A}$.

□

Lemma 4.2.8. *Let V, W be vector spaces and let $T \in L(V,W)$.*

(a) *T is injective iff $Ker(T) = \{0_V\}$;*
(b) *T is surjective iff $Im(T) = T(V) = W$;*
(c) *T is bijective iff $\exists S \in L(W,V)$, which is bijective and $S \circ T = I_V$, $T \circ S = I_W$.*

Proof. (a) See Algebra or Linear Algebra.
(b) Trivial.
(c) If T is bijective, $\exists T^{-1} : W \to V$ such, that $T^{-1} \circ T = I_V$ and $T \circ T^{-1} = I_W$. Let us recall that $T^{-1} \in L(W,V)$, i.e. T^{-1} is linear. Let $\alpha, \beta \in \mathbb{F}$ and $x, y \in W$. Then $T^{-1}(\alpha x + \beta y) \in V$ and

$$(*) \quad T(T^{-1}(\alpha x + \beta y)) = \alpha x + \beta y.$$

On the other hand, $T^{-1}(x),\ T^{-1}(y) \in V$ and

$$(**) \quad T(\alpha T^{-1}(x) + \beta T^{-1}(y)) \overset{T\,lin.}{=} \alpha T(T^{-1}(x)) + \beta T(T^{-1}(y)) = \alpha x + \beta y.$$

Since T is injective, we conclude from $(*)$ and $(*)$ that

$$T^{-1}(\alpha x + \beta y) = \alpha T^{-1}(x) + \beta T^{-1}(y).$$

The converse is well-known.

□

Note. Suppose that T is a bijective element in $B(X, Y)$. Then, by Lemma 4.2.8 there is $T^{-1} \in L(Y, X)$. However, we do not know that T^{-1} is a bounded operator. This additional knowledge is studied in the next subsection.

4.3. **Uniform boundedness principle and open mapping theorem.** To prove two corner-stones of functional analysis (open mapping theorem and uniform boundedness principle) we need a deep topological result called Baire's category theorem. The proof of this is omitted, see Väisälä: Topologia II.

Theorem 4.3.1. *Let X be a complete metric space. If $V_j \subset X$, $j \in \mathbb{N}$ is a countable collection of open subsets, then $\bigcap_{j=1}^{\infty} V_j$ is dense in X.*

Corollary 4.3.2. *Let X be a complete metric space and let $F_j \subset X$ be closed for all $j \in \mathbb{N}$ such that*

$$X = \bigcup_{j=1}^{\infty} F_j.$$

Then there is $j_0 \in \mathbb{N}$ such that F_{j0} contains an open ball.

Proof. Denote $V_j = X \setminus F_j, j \in \mathbb{N}$. Then V_j is open for all $j \in \mathbb{N}$. Assume, on the contrary, that none of the sets F_j contains an open ball, that is,

$$V_j \cap B(x, r) \neq \varnothing \quad \forall j \in \mathbb{N}, \forall x \in X, \forall r > 0.$$

This implies that V_j is dense in X for all $j \in \mathbb{N}$. By Theorem 4.3.1, $\bigcap_{j=1}^{\infty} V_j$ is dense in X. In particular, $\bigcap_{j=1}^{\infty} \neq \varnothing$, so there is $x \in X$ such that

$$x \in \bigcap_{j=1}^{\infty} V_j = \bigcap_{j=1}^{\infty} (X \setminus F_j) = X \setminus \bigcap_{j=1}^{\infty} F_j.$$

This contradicts with the assumption $X = \bigcup_{j=1}^{\infty} F_j$. □

Theorem 4.3.3. *Let X be a Banach space, Y a normed space and $(T_\alpha)_{\alpha \in J}$ an arbitrary collection of elements $T_\alpha \in B(X, Y)$. If*

$$M(x) := \sup_{\alpha \in J} \|T_\alpha(x)\| < \infty$$

for all $x \in X$, then

$$\sup_{\alpha \in J} \|T_\alpha\| = \sup_{\alpha \in J} \sup\{\|T_\alpha(x)\| : \|x\| \leq 1\} < \infty$$

Note. Observe that J is an arbitrary index set, J is not necessarily countable.

Before we prove Theorem 4.3.3, let us consider some applications of it.

Theorem 4.3.4. *Let X be a Banach space, Y a normed space and $(T_n)_{n\in\mathbb{N}}$ a sequence of elements in $B(X,Y)$ such that*

$$T(x) = \lim_{n\to\infty} T_n(x)$$

exists for every $x \in X$. Then $T \in B(X,Y)$.

Proof. The mapping T is linear (see the proof of Theorem 4.1.1). By assumption $(T_n(x))$ converges for all $x \in X$. Hence $(T_n(x))$ is a bounded sequence for all $x \in X$, so that

$$M(x) := \sup_{n\in\mathbb{N}} \|T_n(x)\| < \infty \ \forall x \in X$$

By Theorem 4.3.3, there is $M \in \mathbb{R}_+$ such that $\|T_n\| \leq M \forall n \in \mathbb{N}$. We obtain

$$\|T(x)\| = \|\lim_{n\to\infty} T_n(x)\| = \lim_{n\to\infty} \|T_n(x)\| \leq \sup_{n\in\mathbb{N}} \|T_n(x)\| \leq \sup_{n\in\mathbb{N}} \|T_n\|\|x\| \leq M\|x\|.$$

□

Note. In Theorem 4.1.1 Y is Banach, in Theorem 4.3.3 X is Banach. Otherwise Theorem 4.1.1 has stronger assumptions.

Example 4.3.5. Let $\mathcal{P} = \{\, x : x \textit{ is a real polynomial} \,\}$ and let

$$\|x\|_\infty = \sup\{|x(t)| : t \in [0,1]\}, \quad x \in \mathcal{P}.$$

For each $n \in \mathbb{N}$, we define $T_n : \mathcal{P} \to \mathbb{R}$ by

$$T_n(x) = n(x(1) - x(1 - \frac{1}{n})).$$

Then $T_n \in B(\mathcal{P}, \mathbb{R})$ since linearity is obvious and

$$|T_n(x)| \leq 2n\|x\|_\infty.$$

Hence $\|T_n\| \leq 2n$. Moreover,

$$\lim_{n\to\infty} T_n(x) = \lim_{n\to\infty} \frac{x(1) - x(1-\frac{1}{n})}{\frac{1}{n}} = x'(1)$$

so that $\lim_{n\to\infty} T_n(x) = T(x)$ for all $x \in \mathcal{P}$, where $T(x) = x'(1)$. However, T is not continous, since for $x_n(t) = t^n$ we have $\|x_n\|_\infty = 1$ but

$$|T(x_n)| = |x_n'(1)| = n.$$

Conclusions:

(1) Theorem 4.3.4 implies that P is not complete with respect to $\|x\|_\infty$.
(2) We infer that the completeness assumption for X is necessary in Theorem 4.3.4.

Proof of Theorem 4.3.3. Let

$$F(n,\alpha) := \{x \in X : \|T_\alpha(x)\| < n\}, \quad \alpha \in J,\ n \in \mathbb{N}.$$

The function $f_\alpha(x) = \|T_\alpha(x)\|$ is continuous as a composite function of continuous functions T_α and $\|\cdot\|$. Therefore $F(n,\alpha) = f_\alpha^{-1}([0,n])$ is closed X since the pre-image of an open (closed) set is a continuous function is open (closed). Hence the set

$$F_n := \bigcap_{\alpha\in J} F(n,\alpha)$$

is closed in X.

Assume that

$$\sup_{\alpha \in J} \|T_\alpha(x)\| < \infty$$

for all $x \in X$. Let $x \in X$ be arbitrary. Then $\exists n \in \mathbb{N}$ such that

$$\sup_{\alpha \in J} \|T_\alpha(x)\| \leq n. \quad (\Leftrightarrow f_\alpha(x) \leq n \ \forall \alpha)$$

Hence $x \in F(n, \alpha) \forall \alpha \in J$, that is, $x \in F_n$. It follows that

$$X = \bigcup_{n \in \mathbb{N}} F_n.$$

Since X is Banach, Corollary 4.3.2 implies that $\exists n_0 \in \mathbb{N}$ and an open ball $B(x_0, r_0) \subset X$ such that $B(x_0, r_0) \subset F_{n0}$. We are free to assume (by choosing a smaller r_0) that

$$\overline{B}(x_0, r_0) \subset F_{n0}. \quad (*)$$

It is enough to prove that $\|T_\alpha(x)\| \leq \frac{2n_0}{r_0} \forall \alpha \in J$ and $x \in X, \|x\| \leq 1$. Let $x \in X$ with $\|x\| \leq 1$. Then $x_0 + r_0 x \in \overline{B}(x_0, b_0)$ (since $\|x_0 + r_0 x - x_0\| = r_0 \|x\| \leq r_0$) and $(*)$ implies that

$$\|T_\alpha(x_0 + r_0 x\| \leq n_0.$$

Therefore

$$\begin{aligned} \|T_\alpha(x)\| &= \frac{1}{r_0} \|T_\alpha(r_0 x)\| = \frac{1}{r_0} \|T_\alpha(x_0 + r_0 x) - T_\alpha(x_0)\| \\ &\leq \frac{1}{r_0} \Big(\|T_\alpha(x_0 + r_0 x)\| + \|T_\alpha(x_0)\| \Big) \leq \frac{2n_0}{r_0} \end{aligned}$$

for all $\alpha \in J$. □

To understand the idea of the open mapping theorem we first recall some topological background.

Definition 4.3.6. Let X Y be normed spaces. A mapping $f : X \to Y$ is called <u>open</u> if $f(U)$ is open in Y whenever U is open in X.

Recall here that $U \subset X$ is <u>open</u> in a normed space $(X, \|\cdot\|)$ if for each $x \in U$ $\exists r > 0$ so that, $B_X(x, r) = \{y \in X : \|x - y\| < r\} \subset U$.

Lemma 4.3.7. *Let X and Y be normed spaces with norms $\|\cdot\|_X$ $\|\cdot\|_Y$ respectively. Then $f : X \to Y$ is an open mapping if and only if for each $x \in X$ and $r > 0$ there is $r' > 0$ such that $B_Y(f(x), r') \subset f(B_X(x, r))$.*

Proof. $(\Rightarrow)$. Assume that $f : X \to Y$ is open. Let $x \in X$ and $r > 0$. Then $B_X(x, r)$ is open in X and hence by assumption $f(B_X(x, r))$ is open in Y. Since $f(x) \in f(B_X(x, r))$, there is $r' > 0$ so that $B_Y(f(x), r') \subset f(B_X(x, r))$.
$(\Leftarrow)$. Let $U \subset X$ be open and assume that the (r, r')-condition holds. Let $y \in f(U)$ be arbitrary. Choose $x \in U$ so that $y = f(x)$. Since U is open, $\exists r > 0$ so that $B_X(x, r) \subset U$. By assumption, $\exists r' > 0$ such that

$$B_Y(y, r') = B_Y(f(x), r') \subset f(B_X(x, r)) \subset f(U).$$

Hence $f(U)$ is open in Y. □

In what follows, we say that $f : X \to Y$ (X, Y normed spaces) is *open at* $x \in X$ if $\forall r > 0\ \exists r' > 0$ so that

$$B_Y(f(x), r') \subset f(B_X(x, r)).$$

Example. (a) The function $f : \mathbb{R} \to \mathbb{R}, f(x) = (x)$, is not open. In fact, f is not open zero, since $f(]-\varepsilon, \varepsilon[) = [0, \varepsilon[$ does not contain any open neighborhood of $f(0) = 0$.

(b) The function $f : \mathbb{R} \to \mathbb{R}, f(x) = (1)$, is not open at any point $x \in \mathbb{R}$.

Remark 4.3.8. Lemma 4.3.7 is analogical to the well-known characterization of continuity which says that $f : X \to Y$ (X, Y normed spaces) is continous at each point $x \in X$ ($\forall \varepsilon > 0\ \exists r > 0$ so that $f(B_X(x, r)) \subset B(f(x), \varepsilon)$) if and only if for each $V \subset Y$ open the pre-image $f^{-1}(V)$ is open in X.

Lemma 4.3.9. *Let X and Y be normed spaces and $T \in L(X, Y)$. Then T is an open mapping if and only if T is open at 0_X.*

Proof. ($\Rightarrow$). This is included in Lemma 4.3.7.
($\Leftarrow$). Assume that T is open 0_X. By Lemma 4.3.7, it suffices to show that T is open at x for any $x \in X$. Let $x \in X$ and $r > 0$. By assumption, there is $r' > 0$ such that

$$B(T(0_X), r') = B(0_Y, r') \subset T(B(0_X, r)). \quad (*)$$

We claim that

$$T(B(x, r)) = T(x + B(0_X, r)) = T(x) + T(B(0_X, r)),$$

where (by definition of the direct sum)

$$x + B(0_X, r) = x + y : y \in B(0_X, r).$$

(1) $B(x, r) = x + B(0_X, r)$: If $y \in B(0_X, r)$, then $\|x - y\| < r$. Hence $y = x + (y - x)$, where $y - x \in B(0_X, r)$ So $y \in x + B(0_X, r)$. Conversely, if $y \in x + B(0_X, r)$, then $y = x + z$, where $\|z\| < r$. Then $\|y - x\| = \|z\| < r$, so that $y \in B(x, r)$.
(2) $T(x + B(0_X, r)) = T(x) + T(B(0_X, r))$: For any $x \in B(0_X, r)$ we have by linearity $T(x + y) = T(x) + T(y)$. Now, by using (1) and (2) together with ($*$) gives

$$T(B(x, r)) = T(x + B(0_X, r)) = T(x) + T(B(0_X, r)) \supset T(x) + B(0_Y, r') = B(T(x), r').$$

Hence the claim follows. □

As an exercise we obtain that an open mapping $T \in L(X, Y)$ (where X and Y normed spaces) is always surjective, that is, $T(x) = Y$. The open mapping theorem states that the converse is true if X and Y are Banach spaces and $T \in B(X, Y)$.

Theorem 4.3.10. *Let X and Y be Banach spaces and let $T \in B(X, Y)$ be surjective. Then T is an open mapping.*

We obtain Theorem 4.3.10 as a consequence of the following result whose proof we skip (see Rynne & Youngson, p. 115–117).

Theorem 4.3.11. *Let X and Y be Banach spaces and let $T \in B(X, Y)$ be surjective. Then there is $t > 0$ such that*

$$\{y \in Y : \|y\| \le t\} \subset T(\{x \in X : \|x\| \le 1\}) \quad (*)$$

To conclude Theorem 4.3.10, we infer from Theorem 4.3.11 that T is open at 0_X (see Lemma 4.3.9). Let $r > 0$ and let $y \in Y$ such that $\|y\| < \frac{r}{2}t$. Then

$$\|\frac{2}{r}y\| = \frac{2}{r}\|y\| < t$$

and $(*)$ implies that $\frac{2}{r}y = T(x)$ for some $x \in X, \|x\| \leq 1$. Now

$$y = \frac{r}{2}T(x) = T(\frac{r}{2}x),$$

where $\|\frac{r}{2}x\| \leq \frac{r}{2} < r$. We conclude that

$$B(0_Y, \frac{r}{2}t) \subset T(B(0_X, r)),$$

that is, T is open at 0_X.

Corollary 4.3.12. *Let X and Y be Banach spaces and let $T \in B(X, Y)$ be surjective. Then $T^{-1} \in B(Y, X)$.*

Proof. Exercise. □

Definition 4.3.13. Let X and Y be normed spaces and let $F : X \to Y$ be a mapping. Then the *graph* of F, denoted by $G(F)$, is defined as

$$G(F) = \{ (x, F(x)) : x \in X \}.$$

Theorem 4.3.14. *Let X and Y be normed spaces and let $F : X \to Y$ be continous. Then $G(F)$ is a closed subset of $X \times Y$, whose vector sum and scalar multiplication are defined by*

$$(x_1, y_1) + (x_2, y_2) := (x_1 + x_2, y_1 + y_2)$$

and

$$a(x_1, y_1) := (ax_1, ay_1)$$

for all $x_i, x_2 \in X$, $y_1, y_2 \in Y$, $a \in F$, and whose norm $\|\cdot\|$ is defined by

$$\|(x, y)\| := \|x\|_X + \|y\|_Y.$$

Here $\|x\|_X$ (resp. $\|y\|_Y$) is the norm of X (resp. Y).

Proof. We leave as an exercise to prove that $X \times Y, \|\cdot\|)$ is a normed space. To prove that $G(F)$ is closed in $X \times Y$, let $((x_n, y_n))$ be a sequence in $X \times Y$ such that $(x_n, y_n) \to (x, y) \in X \times Y$. This implies that $\lim_{n\to\infty} x_n = x$ in X and $\lim_{n\to\infty} y_n = y$ in Y. On the other hand, $y_n = F(x_n)$, so that

$$y = \lim_{n\to\infty} y_n = \lim_{n\to\infty} F(x_n) = F(x)$$

by continuity of F, see Remark 4.3.15 below. Therefore $(x, y) = (x, F(x)) \in G(F)$ and so $G(F)$ is closed. □

Remark 4.3.15. If X and Y are normed spaces and $T : X \to Y$ is linear, then $G(T)$ is a subspace of $X \times Y$. Indeed, for any $(x, y), (x', y') \in G(T)$ and for any $\alpha, \beta \in F$, we have

$$\begin{aligned}\alpha(x, y) + \beta(x', y') &= \alpha(x, T(X)) + \beta(x', T(x')) = (\alpha x + \beta x', \alpha T(x) + \beta T(x')) \\ &= (\alpha x + \beta x', T(\alpha x + \beta x')),\end{aligned}$$

which implies that $\alpha x + \beta x' \in G(T)$.

The closed graph theorem states that the converse for Theorem 4.3.14 holds if X and Y are Banach spaces and T is linear.

Theorem 4.3.16. *Let X and Y be Banach spaces and let $T : X \to Y$ be linear such that the graph $G(T)$ is closed. Then $T \in B(X, Y)$, that is, T is continuous.*

Proof. As $X \times Y$ is a Banach space (see exercise), $G(T)$ is a Banach space since it is a closed subspace of $X \times Y$. (In fact, a Cauchy sequence in $G(T)$ converges to an element of $X \times Y$ by completeness. But this limit is contained in $G(T)$ since $G(T)$ is closed.) Let $\phi : G(T) \to X$ be the mapping

$$\phi(x, T(x)) = x.$$

Then ϕ is linear since $\forall x, y \in X,\ \alpha, \beta \in F$

$$\begin{aligned} \phi(\alpha(x, T(x)) + \beta(y, T(y))) &= \phi(\alpha x + \beta y, \alpha T(x) + \beta T(y)) \\ &= \phi(\alpha x + \beta y, T(\alpha x + \beta y)) \\ &= \alpha x + \beta y = \alpha\phi(x, T(x)) + \beta\phi(y, T(y))). \end{aligned}$$

The mapping ϕ is clearly bijective. Since

$$\|\phi(x, T(x))\|_X = \|x\|_X \le \|x\|_X + \|T(x)\|_Y = \|(x, T(x))\|_{X\times Y}$$

we obtain that ϕ is bounded with $\|\phi\| \le 1$. By Corollary 4.3.12, $\phi^{-1} : X \to G(T)$ is a bounded linear operator. Since $\phi^{-1}(x) = (x, T(x))\ \forall x \in X$, we obtain

$$\|T(x)\|_Y \le \|x\|_X + \|T(x)\|_Y = \|(x, T(x)))\|_{X\times Y} = \|\phi^{-1}(x)\|_{X\times Y} \le \|\phi^{-1}\|\|x\|_X.$$

Hence T is a bounded operator. □

We continue the study of invertibility by using the open mapping theorem. This requires some lemmas.

Lemma 4.3.17. *If X is a normed linear space and $T \in B(X)$ is invertible, then for all $x \in X$*

$$\|T(x)\| \ge \|T^{-1}\|^{-1}\|x\|$$

Proof. Exercise. □

By Lemma 4.3.17, an invertible operator $T \in B(X)$ has the property that $\exists$ constants $\alpha > 0, \beta > 0$ such, that

$$\alpha\|x\| \le \|T(x)\| \le \beta\|x\|$$

for all $x \in X$.

Lemma 4.3.18. *If X is a Banach space and $T \in B(X)$ has the property that there is a constant $\alpha > 0$ such that*

$$\|T(x)\| \ge \alpha\|x\| \quad \forall x \in X,$$

then $Im(T) = T(X)$ is a closed set.

Proof. Let (y_n) be a sequence in $Im(T)$ such that, $\lim_{n\to\infty} y_n = y \in Y$. As $y_n \in Im(T)$, there exists $x_n \in X$ such that $T(x_n) = y_n$. As (y_n) converges, it is a Cauchy sequence by Lemma 1.2.2. Since

$$\|y_m - y_n\| = \|T(x_m) - T(x_n)\| = \|T(x_m - x_n)\| \ge \alpha\|x_m - x_n\|,$$

it is easy to see that (x_n) is a Cauchy sequence as well. By the completeness of X, there is $x \in X$ so that $\lim_{n\to\infty} x_n = x$. Therefore, by continuity of T, see Remark 4.3.15,

$$T(x) = \lim_{n\to\infty} T(x_n) = \lim_{n\to\infty} y_n = y.$$

Hence $y = T(x) \in Im(x_n)$ and so $Im(T)$ is closed. □

Remark 4.3.19. Let X and Y be normed spaces and let $f : X \to Y$ be continuous. Assume that $x_n, y_n \in X$ so that $\lim_{n\to\infty} x_n = x$. Then $\lim_{n\to\infty} f(x_n) = f(x)$.

Proof. Let $\varepsilon > 0$. By continuity of f, $\exists \delta > 0$ so that

$$|x_n + x| < \delta \Rightarrow |f(x_n) - f(x)| < \varepsilon.$$

Since $\lim_{n\to\infty} x_n = x, \exists n_\delta \in \mathbb{N}$ such that

$$n \geq n_\delta \to |x_n - x| < \delta.$$

Hence $n \geq n_\delta$ implies that $|f(x_n) - f(x)| < \varepsilon$. The claim $\lim_{n\to\infty} f(x_n) = f(x)$ follows. □

Theorem 4.3.20. *Let X be a Banach space and let $T \in B(X)$. The following are equivalent:*

(a) *T is invertible in $B(X)$;*
(b) *$Im(T)$ is dense in X and there is a constant $\alpha > 0$ so that $\|T(x)\| \geq \alpha\|x\|$ for all $x \in X$.*

Proof. $(a) \Rightarrow (b)$. This follows from 4.3.17 since $Im(T) = X$ if T is invertible.
$(b) \Rightarrow (a)$. By hypothesis $Im(T)$ is dense in X. We claim first that $Im(T) = X$. For any $x \in X$, we find a sequence $x_n \in Im(T)$ such that $\lim_{n\to\infty} x_n = x$ by picking $x_n \in B(x, \frac{1}{n}) \bigcap Im(T)$. By assumption and Lemma 4.3.18, $Im(T)$ is closed. Therefore $x \in Im(T)$ and so $Im(T) = X$. Hence T is surjective. To prove that T is injective, let $x \in Ker(T)$. Then $T(x) = 0_X$ so that

$$0 = \|T(x)\| \geq \alpha\|x\|$$

Hence $x = 0_X$ and $Ker(T) = \{0_X\}$. Lemma 4.2.8 implies that T is bijective. Corollary 4.3.12 yields that T is invertible in X. □

Theorem 4.3.20 can be used to show that an operator $T \in B(X)$ is not invertible. For this purpose we first reformulate Theorem 4.3.20.

Corollary 4.3.21. *Let X be a Banach space and let $T \in B(X)$. Then T is not invertible if and only if $Im(T)$ is not dense or*

$$\exists (x_n) \subset X,\ \|x_n\| = 1\ \forall n \in \mathbb{N} \textit{ such that } \lim_{n\to\infty} T(x_n) = 0. \quad (*)$$

Proof. The condition $\|T(x)\| \geq \alpha\|x\|$ does not hold for any $\alpha > 0$ if and only if

$$\exists (x_n') \subset X \setminus \{0_X\} \text{ with } \|T(x_n')\| < \frac{1}{n}\|x_n'\|. \quad (**)$$

If $(**)$ holds, then for $x_n = \frac{x_n'}{\|x_n'\|}$,

$$\|T(x_n)\| = \|T(\frac{x_n'}{\|x_n'\|})\| = \frac{1}{\|x_n'\|\|T(x_n')\|} < \frac{1}{\|x_n'\|}\frac{1}{n}\|x_n'\|.$$

It follows that $\lim_{n\to\infty} T(x_n) = 0_X$. Hence $(*)$ holds. The implication $(*) \Rightarrow (**)$ is similar. □

Example 4.3.22. In Example 4.2.4 we studied for any $h \in C[0,1]$ an operator $T_h \in B(L^2[0,1])$,

$$(T_h g)(t) = h(t)g(t), \quad t \in [0,1].$$

We show now that T_f is not invertible if $f \in C[0,1]$. For each $n \in \mathbb{N}$, let $g_n = \sqrt{n}\chi_{[0,\frac{1}{n}]}$. Then $g_n \in L^2[0,1]$ and

$$\|g_n\|_2^2 = \int_0^1 (\sqrt{n}\chi_{[0,\frac{1}{n}]})^2(t)dt = \int_0^{\frac{1}{n}} ndt = 1$$

for all $n \in \mathbb{N}$. However

$$\|T_f(g_n)\|^2 = \int_0^1 (f(t)g_n(t))^2 dt = \int_0^{\frac{1}{n}} nt^2 dt = \frac{n}{3}n^3$$

Hence

$$\lim_{n\to\infty} \|T_f(g_n)\| = 0$$

and Corollary 4.3.21 implies that T is not invertible.

5. Linear operators on Hilbert spaces

5.1. The adjoint of an operator.

We consider next a linear $T : \mathcal{H} \to \mathcal{K}$, where $\mathcal{H}$ and $\mathcal{K}$ are Hilbert spaces. For simplicity we denote inner products in each of the spaces $\mathcal{H}$ and $\mathcal{K}$ by $\langle \cdot, \cdot \rangle$. Throughout this section we assume that $\mathbb{F} = \mathbb{C}$.

Theorem 5.1.1. *Let $\mathcal{H}$ and $\mathcal{K}$ be complex Hilbert spaces and let $T \in B(\mathcal{H}, \mathcal{K})$. Then there is a unique operator $T^* \in B(\mathcal{K}, \mathcal{H})$ such that*

$$\langle T(x), y \rangle = \langle x, T^*(y) \rangle$$

for all $x \in \mathcal{H}$ and $y \in \mathcal{K}$. Moreover $\|T^*\| \leq \|T\|$.

Proof. Let $y \in \mathcal{K}$ and let $f : \mathcal{H} \to \mathbb{C}$ be defined by

$$f(x) = \langle T(x), y \rangle.$$

Then f is linear, since for all $\alpha, \beta \in \mathbb{C}$ and $x, x' \in \mathcal{H}$,

$$\begin{aligned} f(\alpha x + \beta x') &= \langle T(\alpha x + \beta x'), y \rangle \\ &= \langle \alpha T(x) + \beta T(x'), y \rangle \\ &= \alpha \langle T(x), y \rangle + \beta \langle T(x'), y \rangle \\ &= \alpha f(x) + \beta f(x'). \end{aligned}$$

By Cauchy-Schwarz and by the boundedness of T,

$$|f(x)| = |\langle T(x), y \rangle| \leq \|T(x)\| \|y\| = \|T\| \|x\| \|y\|$$

for all $x \in \mathcal{H}$. Hence f is bounded and Riesz-Frechet theorem (Theorem 4.1.8) implies that there exists unique $z \in \mathcal{H}$ such that

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

We define $T^* : \mathcal{K} \to \mathcal{H}$ by $T^*(y) = z$. Then

$$\langle T(x), y \rangle = \langle x, T^*(y) \rangle \qquad (*)$$

for all $x \in \mathcal{H}, y \in \mathcal{K}$. Now it is enough to show that T^* is linear, bounded and unique.
T is linear: Let $y_1, y_2 \in \mathcal{K}$, let $\alpha, \beta \in \mathbb{C}$ and let $x \in \mathcal{H}$. By $(*)$,

$$\begin{aligned} \langle x, T^*(\alpha y_1 + \beta y_2) \rangle &\overset{(*)}{=} \langle T(x), \alpha y_1 + \beta y_2 \rangle \\ &\overset{3.1.6}{=} \overline{\alpha} \langle T(x), y_1 \rangle + \overline{\beta} \langle T(x), y_2 \rangle \\ &\overset{(*)}{=} \overline{\alpha} \langle x, T^*(y_1) \rangle + \overline{\beta} \langle x, T^*(y_2) \rangle \\ &\overset{3.1.6}{=} \langle x, \alpha T^*(y_1) + \beta T^*(y_2) \rangle. \end{aligned}$$

This holds for all $x \in \mathcal{H}$ and therefore (Exercise 4/1)

$$T^*(\alpha y_1 + \beta y_2) = \alpha T^*(y_1) + \beta T^*(y_2).$$

Boundedness with $\|T^*\| \leq \|T\|$ and uniqueness exercise. □

Definition 5.1.2. If $\mathcal{H}$ and $\mathcal{K}$ are complex Hilbert spaces and $T \in B(\mathcal{H}, \mathcal{K})$, then the operator T^* of Theorem 5.1.1 is called the *adjoint of T*.

The uniqueness part of Theorem 5.1.1 is very useful when finding the adjoint of an operator. If we find a mapping S which satisfies

$$\langle T(x), y\rangle = \langle x, S(y)\rangle \quad \forall\, x \in \mathcal{H}, y \in \mathcal{K},$$

then $S = T^*$.

Example 5.1.3. Recall that the inner product in $\mathbb{C}^2$ is defined by

$$\langle x, y\rangle = x_1\overline{y_1} + x_2\overline{y_2}\ ; \qquad x_i, y_i \in \mathbb{C},\ i = 1, 2.$$

We denote by $M_{2x2}(\mathbb{C})$ the set of 2×2 matrices with complex entries a_{ij}.
Let $T : \mathbb{C}^2 \to \mathbb{C}^2$ be a linear mapping. Then T is continuous (Theorem 2.1.9) and (by linear algebra) there is $A = (a_{ij}) \in M_{2x2}(\mathbb{C})$ such that

$$T(x) = Ax = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

for all $x_1, x_2 \in \mathbb{C}$. To find the adjoint T^*, we write equation

$$\langle T(x), y\rangle = \langle x, T^*(y)\rangle$$

in a form ($T^*(y) = By$)

$$\begin{aligned}
\left\langle \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right\rangle &= \left\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right\rangle \\
\Leftrightarrow \quad \left\langle \begin{pmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right\rangle &= \left\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} b_{11}y_1 + b_{12}y_2 \\ b_{21}y_1 + b_{22}y_2 \end{pmatrix} \right\rangle \\
\Leftrightarrow \quad a_{11}x_1\overline{y_1} + a_{12}x_2\overline{y_1} + a_{21}x_1\overline{y_2} + a_{22}x_2\overline{y_2} &= x_1\overline{b_{11}}\,\overline{y_1} + x_1\overline{b_{12}}\,\overline{y_2} + x_2\overline{b_{21}}\,\overline{y_1} + x_2\overline{b_{22}}\,\overline{y_2}.
\end{aligned}$$

Since this holds for all $x_i, y_i \in \mathbb{C}$, we may choose $x_1 = y_1 = 1$ and $x_2 = y_2 = 0$, so that $a_{11} = \overline{b_{11}}$. Similarly $a_{12} = \overline{b_{21}}, a_{21} = \overline{b_{12}}, a_{22} = \overline{b_{22}}$. In general $b_{ij} = \overline{a_{ji}}$.

The result can be proved similarly for any $T \in B(\mathbb{C}^n, \mathbb{C}^m)$. Hence if

$$T(x) = Ax,$$

where $A \in M_{m\times n}(\mathbb{C}), A = (a_{ij})$, then

$$T^*(x) = Bx,$$

where $B = (b_{ij})$ and $b_{ij} = \overline{a_{ji}}$. We also denote $B = A^*$.

Warning. Here $A^* \neq adjA$. We call the matrix A^* *conjugate transpose (adjucate, Hermitean adjucate).*

Example 5.1.4. For any $k \in \mathcal{C}_{\mathbb{C}}[0, 1]$, let $T_k \in B(L^2_{\mathbb{C}}[0, 1])$ be defined by

$$(T_k g)(t) = k(t)g(t), \quad t \in [0, 1].$$

Note here that the proof of Exercise 3/1 applies also in complex case. Hence $\|T_k\| \leq \|k\|_\infty$.

$$\left(\|T_k g\|_2^2 = \int_0^1 |k(t)|^2 |g(t)|^2 dt \leq \|k\|_\infty^2 \int_0^1 |g(t)|^2 dt = \|k\|_\infty^2 \|g\|_2^2. \right)$$

Claim. If $f \in \mathcal{C}_{\mathbb{C}}[0, 1]$, then $(T_f)^* = T_{\overline{f}}$, where $f = f_1 + if_2$ and $\overline{f} = f_1 - if_2$.

Proof. Let $g, h \in L^2_{\mathbb{C}}[0,1]$ and let $k = (T_f)^* h$. By definition

$$\langle T_f g, h\rangle = \langle g, (T_f)^* h\rangle = \langle g, k\rangle$$

so that (See Example 3.3.2)

$$\int_0^1 f(t)g(t)\overline{h(t)}dt = \int_0^1 g(t)\overline{k(t)}dt.$$

This clearly holds if $\overline{k(t)} = f(t)\overline{h(t)}$, that is

$$k(t) = \overline{f(t)}h(t) = (T_{\overline{f}}h)(t).$$

By the uniqueness of adjoint, we deduce that $(T_f)^* = T_{\overline{f}}$. □

Example 5.1.5. Let $S \in B(l^2)$ be the unilateral shift

$$S(x_1, x_2, x_3, ...) = (0, x_1, x_2, x_3, ...).$$

Claim. $S^*(y_1, y_2, y_3, ...) = (y_2, y_3, y_4, ...)$.

Proof. Let $x = (x_n), y = (y_n) \in l^2$ and let $z = (z_n) = S^*(y)$. By definition

$$\langle S(x), y\rangle = \langle x, S^*(y)\rangle$$

so that

$$\langle (0, x_1, x_2, x_3, ...), (y_1, y_2, y_3, ...)\rangle = \langle (x_1, x_2, x_3, ...), (z_1, z_2, z_3, ...)\rangle.$$

Therefore

$$0 \cdot \overline{y_1} + x_1\overline{y_2} + x_2\overline{y_3} + \ldots = x_1\overline{z_1} + x_2\overline{z_2} + x_3\overline{z_3} + \ldots$$

holds true for all $x = (x_n) \in l^2$ if and only if $z_1 = y_2, z_2 = y_3, ...$ Hence by the uniqueness of the adjoint

$$S^*(y) = z = (y_2, y_3, y_4, ...).$$

□

In what follows, we also call S a *forward shift* and S^* a *backward shift.*

Example 5.1.6. Let $\mathcal{H}$ be a complex Hilbert space. If I is the identity operator on $\mathcal{H}$, then

$$I^* = I.$$

Proof. If $x, y \in \mathcal{H}$, then

$$\langle I(x), y\rangle = \langle x, I^*(y)\rangle \Leftrightarrow \langle x, y\rangle = \langle x, I^*(y)\rangle.$$

Therefore, by the uniqueness of the adjoint, $I^* = I$. □

Lemma 5.1.7. *Let $\mathcal{H}, \mathcal{K}$ and $\mathcal{L}$ be complex Hilbert spaces and let $R, S \in B(\mathcal{H}, \mathcal{K})$ and $T \in B(\mathcal{K}, \mathcal{L})$. Then*

(a) $(\mu R + \lambda S)^* = \overline{\mu}R^* + \overline{\lambda}S^*$ *for all* $\mu, \lambda \in \mathbb{C}$*;*
(b) $(TR)^* = R^*T^*$.

Proof. Exercise. □

Theorem 5.1.8. *Let $\mathcal{H}$ and $\mathcal{K}$ be complex Hilbert spaces and let $T \in B(\mathcal{H}, \mathcal{K})$. Then*

(a) $(T^*)^* = T$*;*
(b) $\|T^*\| = \|T\|$*;*

(c) *the function* $f : B(\mathcal{H},\mathcal{K}) \to B(\mathcal{K},\mathcal{H})$, $f(T) = T^*$, *is continuous;*
(d) $\|T^*T\| = \|T\|^2$.

Proof. (a) Exercise.
(b) By Theorem 5.1.1, we have $\|T^*\| \le \|T\|$. Applying this result to T^* and using (a) gives

$$\|T\| \overset{(a)}{=} \|(T^*)^*\| \le \|T^*\|.$$

Hence $\|T^*\| = \|T\|$.
(c) Let $\varepsilon > 0$ and choose $\delta = \varepsilon$. If $R, S \in B(\mathcal{H},\mathcal{K})$ and $\|R - S\| < \delta = \varepsilon$, then by Lemma 5.1.7 and (b)

$$\|f(R) - f(S)\| = \|R^* - S^*\| \overset{5.1.7}{=} \|(R - S)^*\| \overset{(b)}{=} \|R - S\| < \varepsilon.$$

Hence f is uniformly continuous in $B(\mathcal{H},\mathcal{K})$.
(d) Since $\|T\| = \|T^*\|$, we have

$$\|T^*T\| \le \|T^*\|\|T\| = \|T\|^2.$$

On the other hand, by the definition of T^*, (a) and Cauchy-Schwarz inequality,

$$\|T(x)\|^2 = \langle T(x), T(x)\rangle \overset{def.of T^*}{=} \langle T^*(T(x)), x\rangle \overset{C-S}{\le} \|T^*(T(x))\|\|x\| \le \|T^*T\|\|x\|^2.$$

By taking sup over $\|x\| \le 1$, we obtain

$$\|T\|^2 \le \|T^*T\|.$$

The claim follows. □

Note. By the proof of (c), we have in particular

$$\|f(R)\| = \|R\| \qquad \forall\ R \in B(\mathcal{H},\mathcal{K}),$$

since $0^* = 0$. However, f is not isometry since f is not (quite) linear, see Lemma 5.1.7 (a).

Next, we obtain an improved characterization for invertibility in the case of Hilbert spaces.

Lemma 5.1.9. *Let* $\mathcal{H}$ *and* $\mathcal{K}$ *be complex Hilbert spaces and let* $T \in B(\mathcal{H},\mathcal{K})$. *Then*
(a) $Ker(T) = Im(T^*)^\perp$*;*
(b) $Ker(T^*) = Im(T)^\perp$.

Proof. (a) 1° $Ker(T) \subset Im(T^*)^\perp$:
Let $x \in Ker(T)$ and $z \in Im(T^*)$. As $z \in Im(T^*)$, $\exists\ y \in \mathcal{K}$ such that $T^*(y) = z$. Then

$$\langle x, z\rangle = \langle x, T^*(y)\rangle = \langle T(x), y\rangle = \langle 0_{\mathcal{H}}, y\rangle = 0.$$

Hence $x \subset Im(T^*)^\perp$.
2° $Im(T^*)^\perp \subset Ker(T)$:
Let $x \in Im(T^*)^\perp$. As $T^*T(x) = T^*(T(x)) \in Im(T^*)$, we have

$$\|T(x)\|^2 = \langle T(x), T(x)\rangle = \langle \underbrace{T^*(T(x))}_{\in Im(T^*)}, x\rangle = 0.$$

Thus $\|T(x)\| = 0$ so that $T(x) = 0_{\mathcal{K}}$. Therefore $x \in Ker(T)$.
(b) By (a) and Theorem 5.1.8 (a) we have

$$Ker(T^*) \overset{(a)}{=} (Im(T^*)^*)^\perp \overset{5.1.8}{=} Im(T)^\perp.$$

□

Lemma 5.1.10. *If X is any linear subspace of a Hilbert space $\mathcal{H}$, then $X^{\perp\perp} = \overline{X}$.*

Proof. Since $X \subset \overline{X}$, it follows from Exercise 5/1 that $\overline{X}^{\perp} \subset X^{\perp}$ and $X^{\perp\perp} \subset \overline{X}^{\perp\perp}$. But X is closed and therefore by Corollary 3.2.15 $\overline{X}^{\perp\perp} = \overline{X}$. Hence we conclude that $X^{\perp\perp} \subset \overline{X}$.
By Exercise 5/1, $X \subset X^{\perp\perp}$. Since $X^{\perp\perp}$ is closed (Lemma 3.2.9), we have $\overline{X} \subset X^{\perp\perp}$. The last conclusion is regarded as known from topology. □

Theorem 5.1.11. *Let $\mathcal{H}$ and $\mathcal{K}$ be complex Hilbert spaces and let $T \in B(\mathcal{H}, \mathcal{K})$. Then $Ker(T^*) = \{0_\mathcal{K}\}$ if and only if $Im(T)$ is dense in $\mathcal{K}$.*

Proof. 1° Assume that $Ker(T^*) = \{0_\mathcal{K}\}$. By Lemma 5.1.9

$$(Im(T)^\perp)^\perp = Ker(T^*)^\perp = \{0_\mathcal{K}\}^\perp = \mathcal{K}.$$

By Lemma 5.1.10, $\overline{Im(T)} = \mathcal{K}$, so that $Im(T)$ is dense in $\mathcal{K}$.
2° Assume that $Im(T)$ is dense in $\mathcal{K}$. By Lemma 5.1.10

$$(Im(T)^\perp)^\perp = \overline{Im(T)} = \mathcal{K}.$$

Since $Im(T)$ is closed (Lemma 3.2.9), we obtain by Lemma 5.1.9 and Corollary 3.2.15 that

$$Ker(T^*) \overset{5.1.9}{=} Im(T)^\perp \overset{3.2.9,3.2.15}{=} ((Im(T)^\perp)^\perp)^\perp = \mathcal{K}^\perp = \{0_\mathcal{K}\}.$$

□

Corollary 5.1.12. *Let $\mathcal{H}$ be a complex Hilbert space and let $T \in B(\mathcal{H})$. The following are equivalent:*

(a) *T is invertible;*
(b) *$Ker(T^*) = \{0_\mathcal{H}\}$ and $\exists\ \alpha > 0$ such that $\|T(x)\| \geq \alpha\|x\| \quad \forall\ x \in \mathcal{H}$.*

Proof. Follows from Theorem 5.1.11 and Theorem 4.3.20. □

Despite having to do one more step it is often easier to find the adjoint of an operator T and then decide whether $Ker(T^*) = \{0_\mathcal{H}\}$ than show that $Im(T)$ is dense in $\mathcal{H}$.

Example 5.1.13. The forward shift $S \in B(l^2)$,

$$S(x_1, x_2, x_3, ...) = (0, x_1, x_2, x_3, ...) \quad \forall\ (x_n) \in l^2,$$

is not invertible.

Proof. We showed in Example 5.1.5 that

$$S^*(y_1, y_2, y_3, ...) = (y_2, y_3, y_4, ...) \quad \forall\ (y_n) \in l^2.$$

Hence $(1, 0, 0, 0, ...) \in Ker(S^*)$ and the claim follows from Corollary 5.1.12. □

5.2. **Normal, self-adjoint and unitary operators.** Adjoint can be used to define particular classes of operators which frequently arise in applications and for which much more than above is known.

Definition 5.2.1. If $\mathcal{H}$ is a complex Hilbert space and $T \in B(\mathcal{H})$, then T is *normal* if $TT^* = T^*T$.

Note. A complex $n \times n$-matrix A is called *normal* if $AA^* = A^*A$.

Example. Complex numbers can be regarded as $|x|$-matrices. What is the set of normal matrices? Now $a^* = \overline{a}$, so that the set of all normal operators $\mathbb{C} \to \mathbb{C}$ consists of mappings $z \to az$, where $a\overline{a} = \overline{a}a$. Hence any $a \in \mathbb{C}$ will do since

$$a\overline{a} = \overline{a}a = |a|^2.$$

Example 5.2.2. For any $k \in \mathcal{C}_{\mathbb{C}}[0,1]$, let $T_k \in B(L^2_{\mathbb{C}}[0,1])$ be defined by $T_k g = gk$. We claim that T_k is normal.

Proof. From Example 5.1.4 we know that $T_k^* = T_{\bar{k}}$ for any $k \in \mathcal{C}_{\mathbb{C}}[0,1]$. Hence, for all $g \in L^2_{\mathbb{C}}[0,1]$,

$$(T_k(T_k^*))(g) = T_k(T_k^* g) = T_k(T_{\bar{k}} g) = T_k(g\bar{k}) = g\bar{k}k,$$
$$(T_k^* T_k)(g) = T_k^*(T_k g) = T_{\bar{k}}(gk) = gk\bar{k},$$

So $T_k^* T_k = T_k T_k^*$. □

Example 5.2.3. The forward shift $S \in B(\ell^2)$ of Example 5.1.5 is not normal.

Proof. We know that

$$S^*(y_1, y_2, y_3, \ldots) = (y_2, y_3, y_4, \ldots) \quad \forall (y_n) \in \ell^2.$$

Hence for any $(x_n) \in \ell^2$,

$$\begin{aligned} S^*(S(x_1, x_2, x_3, \ldots)) &= S^*(0, x_1, x_2, \ldots)) = (x_1, x_2, x_3, \ldots), \\ S(S^*(x_1, x_2, x_3, \ldots)) &= S(x_2, x_3, x_4, \ldots)) = (0, x_2, x_3, \ldots). \end{aligned}$$

If $x_1 \neq 0$, then $S^*(S((x_n))) \neq S(S^*((x_n)))$. Hence $S^*S \neq SS^*$.

□

Example 5.2.4. If $\mathcal{H}$ is a complex Hilbert space, I is the identity on $\mathcal{H}$, $\lambda \in \mathbb{C}$, and $T \in B(\mathcal{H})$ is normal, then $T - \lambda I$ is normal.

Proof. By Lemma 5.1.7 and Example 5.1.6,

$$(T - \lambda I)^* \overset{5.1.7}{=} T^* - \overline{\lambda} I^* \overset{5.1.6}{=} T^* - \overline{\lambda} I.$$

We obtain

$$\begin{aligned} (T - \lambda I)(T - \lambda I)^* &= (T - \lambda I)(T^* - \overline{\lambda} I) \\ &= TT^* - T\overline{\lambda} I - \lambda I T^* + \lambda I \overline{\lambda} I \\ &= TT^* - \overline{\lambda} T - \lambda T^* + |\lambda|^2 I \end{aligned}$$

and similarly

$$\begin{aligned} (T - \lambda I)^*(T - \lambda I) &= (T^* - \overline{\lambda} I)(T - \lambda I) \\ &= T^*T - \lambda T^* - \overline{\lambda} T + |\lambda|^2 I. \end{aligned}$$

By assumption $TT^* = T^*T$ and the claim follows. □

Notice above e.g. that

$$(T\overline{\lambda} I)(x) = T(\overline{\lambda} I(x)) = T(\overline{\lambda} x) \overset{T lin.}{=} \overline{\lambda} T(x) = (\overline{\lambda} T)(x).$$
$$(\lambda I \overline{\lambda} I)(x) = \lambda I(\overline{\lambda} x) = \lambda \overline{\lambda} x = (|\lambda|^2 I)(x).$$

We study next the basic properties of normal operators.

Lemma 5.2.5. *Let $\mathcal{H}$ be a complex Hilbert space, let $T \in B(\mathcal{H})$ be normal. Then*

(a) $\|T(x)\| = \|T^*(x)\| \quad \forall x \in \mathcal{H}$*;*

(b) *If* $\|T(x)\| \geq \alpha\|x\|$ *for some* $\alpha > 0$ *and for all* $x \in \mathcal{H}$*, then* $Ker(T^*) = \{0_{\mathcal{H}}\}$.

Proof. (a) Let $x \in \mathcal{H}$. AS $T^*T = TT^*$, we obtainby the definition of the adjoint and Theorem 5.1.8 (a)

$$\begin{aligned} \|T(x)\|^2 - \|T^*(x)\|^2 &= \langle T(x), T(x)\rangle - \langle T^*(x), T^*(x)\rangle \\ &\overset{5.1.8(a)}{=} \langle x, T^*(T(x))\rangle - \langle x, T(T^*(x))\rangle \\ &= \langle x, T^*(T(x)) - T(T^*(x))\rangle = \langle x, 0_{\mathcal{H}}\rangle = 0. \end{aligned}$$

Therefore

$$\|T(x)\| = \|T^*(x)\| \quad \forall x \in \mathcal{H}.$$

(b) Let $y \in Ker(T^*)$, i.e. $T^*(y) = 0_{\mathcal{H}}$. Then by (a) and the assumption

$$0 = \|T^*(y)\| \overset{(a)}{=} \|T(y)\| \geq \alpha\|y\| \geq 0.$$

Therefore $\|y\| = 0$ and hence $y = 0_{\mathcal{H}}$. Hence $Ker(T^*) = \{0_{\mathcal{H}}\}$. □

Corollary 5.2.6. *Let* $\mathcal{H}$ *be a complex Hilbert space and let* $T \in B(\mathcal{H})$ *be a normal operator. The following are equivalent:*

(a) T *is invertible;*
(b) $\exists \alpha > 0$ *such that* $\|T(x)\| \geq \alpha\|x\| \quad \forall x \in \mathcal{H}$.

Proof. Corollary 5.1.12 and Lemma 5.2.5. □

Definition 5.2.7. If $\mathcal{H}$ is a complex Hilbert space and $T \in B(\mathcal{H})$, then T is *self-adjoint* if $T = T^*$.

Note. A complex $n \times n$-matrix A is *self-adjoint* if $A = A^*$.

Example. What is the set of self-adjoint operators $z \to az;\ z \in \mathbb{C},\ a \in \mathbb{Z}$? Now we require that $a^* = \overline{a} = a$, which holds iff $a \in \mathbb{R}$.

There are two natural ways to show that a given operator is self-adjoint.

Example 5.2.8. The matrix

$$A = \begin{bmatrix} 2 & i \\ -i & 3 \end{bmatrix}$$

is self adjoint. This is clear since

$$A^* = \overline{A^T} = \overline{\begin{bmatrix} 2 & -i \\ i & 3 \end{bmatrix}} = \begin{bmatrix} 2 & i \\ -i & 3 \end{bmatrix} = A.$$

The second approach is to show that

$$\langle T(x), y\rangle = \langle x, T(y)\rangle$$

$\forall x, y \in \mathcal{H}$. The uniqueness of adjoint then gives $T = T^*$.

Example 5.2.9. It is clear that $I \in B(\mathcal{H})$ satisfies

$$\langle I(x), y\rangle = \langle x, I(y)\rangle \quad \forall x, y \in \mathcal{H}.$$

Hence I is self-adjoint.

Example 5.2.10. For any $k \in \mathcal{C}[0,1]$, let $T_k \in B(L^2_{\mathbb{C}}[0,1])$ be defined by $T_k g = gk$. Hence we assume that k is real-valued. In this case T_k is self-adjoint.

Proof. Let $k \in \mathcal{C}[0,1]$. Now $(T_k)^* = T_{\overline{k}} = T_k$ since k is real (i.e. $k = k_1 + ik_2$, where $k_2 \equiv 0$). □

Lemma 5.2.11. *Let $\mathcal{H}$ be a complex Hilbert space and let $\mathcal{S}$ be the set of self-adjoint operators in $B(\mathcal{H})$. Then*

(a) $\alpha T_1 + \beta T_2 \in S \quad \forall T_1, T_2 \in S,\ \alpha, \beta \in \mathbb{R}$;
(b) *$\mathcal{S}$ is a closed subset of $B(\mathcal{H})$.*

Proof. (a) As T_1 and T_2 are self-adjoint, Lemma 5.1.7 gives

$$(\alpha T_1 + \beta T_2)^* \overset{5.1.7}{=} \overline{\alpha} T_1^* + \overline{\beta} T_2^* \overset{\alpha,\beta \in \mathbb{R}}{=} \alpha T_1 + \beta T_2.$$

(b) Exercise. □

An alternative way of stating Lemma 5.2.11 is to say that the set of salf-adjoint operators in $B(\mathcal{H})$ is a real Banach space.

Lemma 5.2.12. *Let $\mathcal{H}$ be a complex Hilbert space and let $T \in B(\mathcal{H})$. Then*

(a) *T^*T and TT^* are self-adjoint;*
(b) *$T = R + iS$, where R and S are self-adjoint.*

Proof. (a) By Lemma 5.1.7 and Theorem 5.1.8 (a)

$$(T^*T)^* \overset{5.1.7}{=} T^*(T^*)^* \overset{5.1.8}{=} T^*T.$$

Hence T^*T is self-adjoint. Similarly TT^* is self-adjoint.
(b) Let $R = \frac{1}{2}(T + T^*)$ and $S = \frac{1}{2i}(T - T^*)$. Then

$$R + iS = \frac{1}{2}T + \frac{1}{2}T^* + i\frac{1}{2i}(T - T^*) = T.$$

On the other hand, by Lemma 5.1.7

$$R^* = \frac{1}{2}T^* + \frac{1}{2}(T^*)^* = \frac{1}{2}(T^* + T) = R$$

and

$$S^* = (\frac{1}{2i}T - \frac{1}{2i}T^*)^* = \overline{\frac{1}{2i}}T^* - \overline{\frac{1}{2i}}T = -\frac{1}{2i}T^* + -\frac{1}{2i}T = S,$$

since

$$\frac{1}{2i} = \frac{2i}{4i^2} = -\frac{i}{2} \Rightarrow \overline{\frac{1}{2i}} = \frac{i}{2} = -\frac{1}{2i}.$$

Hence R and S are self-adjoint. □

Note. By analogy with complex numbers, the operators R and S in Lemma 5.2.12 are sometimes called the *real* and *imaginary* parts of T.

Definition 5.2.13. If $\mathcal{H}$ is a complex Hilbert space and $T \in B(\mathcal{H})$, then T is *unitary* if $TT^* = T^*T = I$.

Note. (a) By definition, for unitary operators $T^* = T^{-1}$.
(b) A complex $n \times n$-matrix A is called *unitary* if $AA^* = A^*A = I$.

Example. What are the unitary operators of $\mathbb{C} \to \mathbb{C}$? Now we require that the mapping $z \to az$ is such that $aa^* = 1$. This holds iff $|a| = 1$. Hence a is the point of the unit circle.

Example 5.2.14. For any $k \in \mathcal{C}_{\mathbb{C}}[0,1]$, let $T_k \in B(L^2_{\mathbb{C}}[0,1])$ be defined by

$$T_k g = gk.$$

Claim. If $f \in \mathcal{C}_{\mathbb{C}}[0,1]$ satisfies $|f(t)| = 1 \ \forall t \in [0,1]$, then T_f is unitary.

Proof. We know from Example 5.1.4 that$(T_f)^* = T_{\overline{f}}$, where $\overline{f} = f_1 - if_2$ and $f = f_1 + if_2$. Let $g \in L^2_{\mathbb{C}}[0,1]$. Then

$$(T_f^* T_f)(g) = T_f^*(T_f g) = T_{\overline{f}}(gf) = gf\overline{f}.$$

Since $|f(t)| = 1 \ \forall t \in [0,1]$, we obtain

$$(f\overline{f})(t) = f(t)\overline{f}(t) = f_1^2(t) + f_2^2(t) = |f(t)|^2 = 1.$$

Hence $\forall t \in [0,1]$

$$(T_f^* T_f)(g)(t) = g(t),$$

so that $(T_f^* T_f)(g) = g$. The proof of $(T_f T_f^*)(g) = g$ is similar. □

For example, a natural choice in Example 5.2.14 for f would be $f : [0,1] \to \mathbb{C}$,

$$f(t) = e^{2i\pi t}.$$

We give next a more geometric characterization for unitary operators. This requires a lemma.

Lemma 5.2.15. *If X is a complex inner product space and $S, T \in B(X)$ are such that*

$$\langle S(x), x\rangle = \langle T(x), x\rangle$$

for all $x \in X$, then $S = T$.

Proof. By Lemma 3.1.8 for any $u, v, x, y \in X$

$$\langle u + v, x + y\rangle - \langle u - v, x - y\rangle = 2\langle u, y\rangle + 2\langle v, x\rangle. \quad (*)$$

Replacing here v by iv and y by iy gives

$$\begin{aligned} \langle u + iv, x + iy\rangle - \langle u - iv, x - iy\rangle &= 2\langle u, iy\rangle + \langle iv, x\rangle \\ &= -2i\langle u, y\rangle + 2i\langle v, x\rangle. \end{aligned}$$

Multiplying this with i and adding $(*)$ yields

$$\langle u + v, x + y\rangle - \langle u - v, x - y\rangle + i\langle u + v, x + y\rangle - i\langle u - v, x - y\rangle = 4\langle u, y\rangle \quad (**)$$

We replace $u = T(x),\ v = T(y)$ in $(**)$ and obtain by linearity and the assumption that

$$\begin{aligned} &4\langle T(x), y\rangle \\ &= \langle T(x + y), x + y\rangle - \langle T(x - y), x - y\rangle + i\langle T(x + iy), x + iy\rangle - \langle T(x - iy), x - iy\rangle \\ &= \langle S(x + y), x + y\rangle - \langle S(x - y), x - y\rangle + i\langle S(x + iy), x + iy\rangle - \langle S(x - iy), x - iy\rangle \\ &\overset{(**)}{=} 4\langle S(x), y\rangle \quad \forall x, y \in X. \end{aligned}$$

Hence $\langle T(x), y\rangle = \langle S(x), y\rangle \ \forall x, y \in X$ and Exercise 4/1 implies that $T(x) = S(x) \ \forall x \in X$. □

Theorem 5.2.16. *Let $\mathcal{H}$ be a complex Hilbert space and let $T, U \in B(\mathcal{H})$. Then*

(a) *$T^*T = I$ iff T is an isometry;*
(b) *U is unitary iff U is a bijective isometry $\mathcal{H} \to \mathcal{H}$.*

Proof. (a) Suppose first that $T^*T = I$. Then

$$\begin{aligned}\|T(x)\|^2 &= \langle T(x), T(x)\rangle = \langle x, T^*(T(x))\rangle = \langle x, I(x)\rangle = \langle x, x\rangle \\ &= \|x\|^2 \quad \forall x \in \mathcal{H}.\end{aligned}$$

Hence T is an isometry. Conversely, suppose that T is an isometry. Then

$$\begin{aligned}\langle (T^*T)(x), x\rangle &= \langle T^*(T)(x)), x\rangle \overset{(T^*)^*=T}{=} \langle T(x), T(x)\rangle \\ &= \|T(x)\|^2 = \|x\|^2 = \langle x, x\rangle = \langle I(x), x\rangle \quad \forall x \in \mathcal{H}.\end{aligned}$$

Now Lemma 5.2.15 implies that $T^*T = I$.
(b) Suppose first that U is unitary. Then U is an isometry by (a). Hence clearly U is injective. Moreover, if $y \in \mathcal{H}$, then $y = U(U^*(y))$, which gives $y \in Im(U)$. Hence $Im(U) = \mathcal{H}$ so that U is surjective.

Conversely, suppose that $U : \mathcal{H} \to \mathcal{H}$ is a bijective isometry. Then $U^*U = I$ by (a). Moreover, if $y \in \mathcal{H}$, then there is $x \in \mathcal{H}$ such that $y = U(x)$. Hence

$$(UU^*)(y) = U(U^*(Y)) = U(U^*(U(x))) \overset{U^*U=I}{=} U(x) = y.$$

Thus $UU^* = I$ so that U is unitary. □

Corollary 5.2.17. *Let $\mathcal{H}$ be a complex Hilbert space and let $\mathcal{U}$ be the set of unitary operators in $B(\mathcal{H})$. Then $U^* \in \mathcal{U}$ for all $U \in \mathcal{U}$ and*

$$\|U\| = \|U^*\| = 1.$$

Proof. Let $U \in \mathcal{U}$. Then $UU^* = U^*U = I$. In other words (by Theorem 5.1.8)

$$(U^*)^*U^* = U^*(U^*)^* = I,$$

so that $U^* \in \mathcal{U}$. By Theorem 5.2.16, $\|U\| = \|U^*\| = 1$ since U and U^* are isometres. □

Remark 5.2.18. Let $\mathcal{H}$ and $\mathcal{U}$ be as in Corollary 5.2.17. Then $u_1u_2 \in \mathcal{U}$ and $u_1^{-1} \in \mathcal{U}$ for all $u_1, u_2 \in \mathcal{U}$ (exercise). Hence $\mathcal{U}$ forms a group with respect to the operator product.

5.3. **The spectrum of an operator.** Given a complex $n \times n$-matrix A, a number $\lambda \in \mathbb{C}$ is called an *eigenvalue* of A if there exists a non-zero vector $x \in \mathbb{C}^n$ such that

$$Ax = \lambda x.$$

Here x is an *eigenvector*. It can be proved (see Linear Algebra) that λ is an eigenvalue if and only if $A - \lambda I$ is not invertible.

Definition 5.3.1. Let $\mathcal{H}$ be a complex Hilbert space, let $I \in B(\mathcal{H})$ be the identity and let $T \in B(\mathcal{H})$. The *spectrum* of T is defined as a set

$$\sigma(T) = \{\lambda \in \mathbb{C} : T - \lambda I \text{ is not invertible}\}.$$

A number $\mu \in \mathbb{C}$ is called an *eigenvalue* of T if there exists $x \in \mathcal{H},\ x \neq 0_{\mathcal{H}}$, such that

$$T(x) = \mu x.$$

Example 5.3.2. Let $\mathcal{H}$ be a complex Hilbert space and let I be the identity on $\mathcal{H}$. Then, for any $\mu \in \mathbb{C}$,

$$\sigma(\mu I) = \{\mu\}.$$

In fact, for any $\tau \in \mathbb{C}$, τI is invertible if and only if $\tau \neq 0$, since

$$\tau I \tau^{-1} I = \tau^{-1} I \tau I = I \quad \text{if } \tau \neq 0.$$

Clearly $0 \cdot I$ is not invertible. Hence

$$\begin{aligned} \sigma(\mu I) &= \{\lambda \in \mathbb{C} : \mu I - \lambda I \text{ is not invertible}\} \\ &= \{\lambda \in \mathbb{C} : (\mu - \lambda) I \text{ is not invertible}\} \\ &= \{\mu\}. \end{aligned}$$

Lemma 5.3.3. *Let $\mathcal{H}$ be a complex Hilbert space and let $T \in B(\mathcal{H})$. If λ is an eigenvalue of T, then $\lambda \in \sigma(T)$.*

Proof. Let $x \in \mathcal{H} \setminus \{0_{\mathcal{H}}\}$ be such that $T(x) = \lambda x$. Then

$$T(x) - \lambda x = 0_{\mathcal{H}} \quad \text{i.e.} \quad (T - \lambda I)(x) = 0_{\mathcal{H}}.$$

Hence $x \in Ker(T - \lambda I)$ and Lemma 4.2.8 (a) implies that $T - \lambda I$ is not invertible. □

It appears that on infinite-dimensional spaces the spectrum does not necessarily coincide with the set of eigenvalues.

Example 5.3.4. The forward shift $S \in B(l^2)$ has no eigenvalues. To see this, assume that $\lambda \in \mathbb{C}$ is an eigenvalue of S and $x = (x_n)$ is the corresponding non-zero eigenvector. Then

$$S(x) = (0, x_1, x_2, x_3, ...) = (\lambda x_1, \lambda x_2, \lambda x_3, ...) = \lambda x.$$

If $\lambda = 0$, then $x = (x_n) = 0_{l^2}$, which is a contradiction.

If $\lambda \neq 0$, then $\lambda x_1 = 0$ implies that $x_1 = 0$. Hence $\lambda x_2 = 0$ and again $x_2 = 0$. Continuing this way we conclude $x = 0_{l^2}$, a contradiction.

How to find the spectrum if there are no eigenvalues? The following two results can sometimes help.

Theorem 5.3.5. *Let $\mathcal{H}$ be a complex Hilbert space and let $T \in B(\mathcal{H})$. Then*
(a) *$\lambda \notin \sigma(T)$ if $|\lambda| > \|T\|$;*
(b) *$\sigma(T)$ is a closed set.*

Proof. (a) If $|\lambda| > \|T\|$, then

$$\overbrace{|\lambda^{-1}||\lambda|}^{1} > |\lambda^{-1}|\|T\| = \|\lambda^{-1}T\|.$$

Hence $\|\lambda^{-1}T\| < 1$ and so $I - \lambda^{-1}T$ is invertible by Theorem 4.2.5. Hence

$$\lambda I - T = \lambda(I - \lambda^{-1}T)$$

is invertible and so $T - \lambda I$ is invertible. Therefore $\lambda \notin \sigma(T)$.
(b) Define $F : \mathbb{C} \to B(\mathcal{H})$ by $F(\lambda) = T - \lambda I$. As

$$\|F(\mu) - F(\lambda)\| = \|T - \mu I - (T - \lambda I)\| = |\mu - \lambda|\|I\| = |\mu - \lambda|,$$

F is continuous. By Corollary 4.2.7, the set of invertible elements in $B(\mathcal{H})$ is open. Hence the set $\mathcal{C}$ concisting of non-invertible elements in $B(\mathcal{H})$ is closed. Since

$$\sigma(T) = F^{-1}(\mathcal{C}) \quad \text{(pre-image)}$$

we infer by continuity of F that $\sigma(T)$ is closed. □

Theorem 5.3.5 states that the spectrum of an operator T is a closed bounded (and hence compact) subset of $\mathbb{C}$ which is contained in an open disc with the center origin and the radius $\|T\|$.

Lemma 5.3.6. *If $\mathcal{H}$ is a complex Hilbert space and $T \in B(\mathcal{H})$, then*

$$\sigma(T^*) = \{\overline{\lambda} : \lambda \in \sigma(T)\}.$$

Proof. 1°) If $\lambda \in \sigma(T)$, then $T - \lambda I$ is invertible and so

$$(T - \lambda I)^* = T^* - \overline{\lambda} I$$

is invertible by Exercise 9/7. Hence $\overline{\lambda} \in \sigma(T^*)$.
2°) Conversely, if $\overline{\lambda} \notin \sigma(T^*)$, then $T^* - \overline{\lambda} I$ is invertible and so

$$(T^* - \overline{\lambda} I)^* = (T^*)^* - \lambda I = T - \lambda I$$

is invertible since $(T^*)^* = T$. Hence $\lambda \notin \sigma(T)$.
The claim follows by combining 1° and 2°. □

Example 5.3.7. If $S : l^2 \to l^2$ is the forward shift, then

(a) λ is an eigenvalue of S^* for any $\lambda \in \mathbb{C}, |\lambda| < 1$;
(b) $\sigma(S) = \{\lambda \in \mathbb{C} : |\lambda| < 1\}$.

Proof. (a) Let $\lambda \in \mathbb{C}$ with $|\lambda| < 1$. We have to find a non-zero vector $(x_n) \in l^2$ such that

$$S^*((x_n)) = \lambda(x_n).$$

By Example 5.1.5,

$$S^*(x_1, x_2, x_3, ...) = (x_2, x_3, x_4, ...),$$

so we need to find a non-zero $(x_n) \in l^2$ such that

$$(x_2, x_3, x_4, ...) = (\lambda x_1, \lambda x_2, \lambda x_3, ...),$$

that is $x_{n+1} = \lambda x_n$ for all $n \in \mathbb{N}$. This holds if $x_n = \lambda^{n-1}$. Here we agree that $0^0 = 1$. Then $(x_n) = (\lambda^{n-1})$ is non-zero even for $\lambda = 0$. Moreover, as $|\lambda| < 1$,

$$\sum_{n=1}^{\infty} |x_n|^2 = \sum_{n=0}^{\infty} |\lambda^n|^2 = \sum_{n=0}^{\infty} |\lambda|^{2n} < \infty,$$

and so $(x_n) \in l^2$. Thus λ is an eigenvalue of S^* with an eigenvector (λ^{n-1}), where $0^0 = 1$.

(b) We have $\{\lambda \in \mathbb{C} : |\lambda| < 1\} \subset \sigma(S^*)$ by (a) and Lemma 5.3.3. Thus $\{\overline{\lambda} \in \mathbb{C} : |\lambda| < 1\}$ is contained in $\sigma(S)$ by Lemma 5.3.6. Clearly

$$\{\overline{\lambda} \in \mathbb{C} : |\lambda| < 1\} = \{\lambda \in \mathbb{C} : |\lambda| < 1\}$$

and so

$$\{\lambda \in \mathbb{C} : |\lambda| < 1\} \subset \sigma(S).$$

As $\sigma(S)$ is closed, by Theorem 5.3.5, we infer that $\{\lambda \in \mathbb{C} : |\lambda| \leq 1\} \subset \sigma(S)$. On the other hand, if $|\lambda| > 1$, then $\lambda \notin \sigma(S)$ by Theorem 5.3.5 since $\|S\| = 1$. Hence

$$\sigma(S) = \{\lambda \in \mathbb{C} : |\lambda| \leq 1\}.$$

□

If we know the spectrum of T, it is easy to find the spectrum of powers of T and (if T is invertible) the inverse of T.

Theorem 5.3.8. *Let $\mathcal{H}$ be a complex Hilbert space and let $T \in B(\mathcal{H})$.*

(a) *If $p : \mathbb{C} \to \mathbb{C}$ is a polynomial, then*

$$\sigma(p(T)) = \{p(\mu) : \mu \in \sigma(T)\};$$

(b) *If T is invertible, then*

$$\sigma(T^{-1}) = \{\mu^{-1} : \mu \in \sigma(T)\}.$$

Here

$$p(T) = a_n T^n + a_{n-1}T^{n-1} + \ldots + a_1 T + a_0 I$$

whenever

$$p(z) = a_n z^n + a_{n-1}z^{n-1} + \ldots + a_1 z + a_0.$$

Proof. (a) Let $\lambda \in \mathbb{C}$ and let $q(z) = \lambda - p(z),\ z \in \mathbb{C}$. Then q is a polynomial, so by the fundamental theorem of algebra, it has a factorization

$$q(z) = c(z - \mu_1)\cdots(z - \mu_n),$$

where $c, \mu_i \in \mathbb{C}$ with $c \neq 0$ and μ_i are roots of q. Here we may assume that $p \neq \lambda$, since if $p \equiv \lambda$, then (Example 5.3.2)

$$\sigma(p(T)) = \sigma(\lambda I) = \{\lambda\} = \{p(\mu) : \mu \in \sigma(T)\}.$$

Hence

$$\begin{aligned}
\lambda \notin \sigma(p(T)) &\Leftrightarrow q(T) = \lambda I - p(T) \text{ is invertible}\\
&\Leftrightarrow c(T - \mu_1 I)\cdots(T - \mu_n I) \text{ is invertible}\\
&\overset{(*)}{\Leftrightarrow} T - \mu_j I \text{ is invertible for all } j = 1, ..., n\\
&\Leftrightarrow \mu_j \notin \sigma(T) \quad \forall\, j = 1, ..., n\\
&\Leftrightarrow q(\mu) \neq 0 \quad \forall\, \mu \in \sigma(T)\\
&\Leftrightarrow \lambda \neq p(\mu) \quad \forall\, \mu \in \sigma(T).
\end{aligned}$$

Hence $\sigma(p(T)) = \{p(\mu) : \mu \in \sigma(T)\}$. Here the equivalence $(*)$ is left as an exercise.

(b) As $T^{-1} = T^{-1} - 0 \cdot I$ is invertible, $0 \notin \sigma(T^{-1})$. Hence any element of $\sigma(T^{-1})$ is of the form μ^{-1} for some $\mu \in \mathbb{C} \setminus \{0\}$. For any $\mu \neq 0$,

$$\mu^{-1}I - T^{-1} = -\mu^{-1}T^{-1}(\mu I - T),$$

and $-\mu^{-1}T^{-1}$ is invertible. Hence

$$\begin{aligned}
\mu^{-1} \in \sigma(T^{-1}) &\Leftrightarrow \mu^{-1}I - T^{-1} \text{ is not invertible}\\
&\Leftrightarrow -\mu^{-1}T^{-1}(\mu I - T) \text{ is not invertible}\\
&\overset{(*)}{\Leftrightarrow} \mu I - T \text{ is not invertible}\\
&\Leftrightarrow \mu \in \sigma(T).
\end{aligned}$$

The proof of $(*)$:
1° If $\mu I - T$ is invertible, then $-\mu^{-1}T^{-1}(\mu I - T)$ is invertible by Lemma 4.2.2.
2° If $-\mu^{-1}T^{-1}(\mu I - T)$ is invertible, then

$$(-\mu^{-1}T^{-1})^{-1}(-\mu^{-1}T^{-1})(\mu I - T) = \mu I - T$$

is invertible by Lemma 4.2.2.

Thus $\sigma(T^{-1}) = \{\mu^{-1} : \mu \in \sigma(T)\}$. □

Notation. Let $\mathcal{H}$ be a complex Hilbert space and let $T \in B(\mathcal{H})$. If $p : \mathbb{C} \to \mathbb{C}$ is polynomial, we denote

$$p(\sigma(T)) = \{p(\mu) : \mu \in \sigma(T)\}.$$

Corollary 5.3.9. *If $\mathcal{H}$ is a complex Hilbert space and $U \in B(\mathcal{H})$ is unitary, then*

$$\sigma(U) \subset \{\lambda \in \mathbb{C} : |\lambda| = 1\}.$$

Proof. As U is unitary, $\|U\| = 1$ and Theorem 5.3.5 implies that

$$\sigma(U) \subset \{\lambda \in \mathbb{C} : |\lambda| \leq 1\}.$$

Similarly

$$\sigma(U^*) \subset \{\lambda \in \mathbb{C} : |\lambda| \leq 1\}$$

since U is unitary. However, $U^* = U^{-1}$ so that Theorem 5.3.8 (b) implies that ($0 \notin \sigma(U^*)$ since U^* is invertible)

$$\sigma(U) = \{\lambda^{-1} : \lambda \in \sigma(U^*)\} \subset \{\lambda \in \mathbb{C} : |\lambda| \geq 1\}.$$

The claim follows. □

Definition 5.3.10. Let $\mathcal{H}$ be a complex Hilbert space and let $T \in B(\mathcal{H})$. Then

(a) the *spectrum radius* of T, denoted by $r_\sigma(T)$, is defined as

$$r_\sigma(T) = \sup\{|\lambda| : \lambda \in \sigma(T)\};$$

(b) the *numerical range* of T, denoted by $V(T)$, is defined as

$$V(T) = \{\langle T(x), x\rangle : \|x\| = 1\}.$$

Note. In (a), $\sup = \max$ since $\sigma(T)$ is closed and bounded (i.e. compact).

Lemma 5.3.11. *If $\mathcal{H}$ is a complex Hilbert space and $T \in B(\mathcal{H})$ is normal, then*

$$\sigma(T) \subset \overline{V(T)}.$$

Proof. Let $\lambda \in \sigma(T)$. As $T - \lambda I$ is normal by Example 5.2.4 and $T - \lambda I$ is non-invertible, Corollary 5.2.6 implies that there exists $(x_n) \in \mathcal{H}$ such that $\|x\| = 1$ $\forall\, n \in \mathbb{N}$ and

$$\lim_{n\to\infty} \|(T - \lambda I)(x_n)\| = 0.$$

$\Big($ Corollary 5.2.6: For any $n \in \mathbb{N}$ $\exists\, x_n' \neq 0$ such that

$$\|\overbrace{S}^{T-\lambda I}(x_n')\| < \frac{1}{n}\|(x_n')\|.$$

Take $x_n' = \frac{x_n'}{\|x_n'\|}$. Hence $\|S(x_n')\| < \frac{1}{n}$.$\Big)$
By the Cauchy-Schwarz-inequality,

$$|\langle (T - \lambda I)(x_n), x_n\rangle| \overset{\|x_n\|=1}{\leq} \|(T - \lambda I)(x_n)\|$$

so that

$$0 = \lim_{n\to\infty} \langle \overbrace{(T - \lambda I)(x_n)}^{T(x_n)-\lambda(x_n)}, x_n\rangle = \lim_{n\to\infty} (\langle T(x_n), x_n\rangle - \lambda\langle x_n, x_n\rangle).$$

However, $\langle x_n, x_n\rangle = \|x_n\| = 1$ and so

$$\lim_{n\to\infty} \underbrace{\langle T(x_n), x_n\rangle}_{\in V(T)} = \lambda.$$

Therefore $\lambda \in \overline{V(T)}$. □

Theorem 5.3.12. *Let $\mathcal{H}$ be a complex Hilbert space and let $S \in B(\mathcal{H})$ be self-adjoint. Then*

(a) $V(S) \subset \mathbb{R}$;
(b) $\sigma(S) \subset \mathbb{R}$;
(c) *At least one of* $\|S\|$ *and* $-\|S\|$ *is contained in* $\sigma(S)$;
(d) $r_\sigma(S) = \sup\{|\tau| : \tau \in V(S)\} = \|S\|$.

Proof. (a) As S is self-adjoint,

$$\langle S(x), x\rangle \overset{S^*=S}{=} \langle x, S(x)\rangle = \overline{\langle S(x), x\rangle}$$

for all $x \in \mathcal{H}$. Hence $\langle S(x), x\rangle \in \mathbb{R} \;\; \forall\, x \in \mathcal{H}$ and hence $V(S) \subset \mathbb{R}$.
(b) Lemma 5.3.11; notice that $|\langle S(x), x\rangle| \overset{C-S}{\leq} \|S(x)\| \leq \|S\|$ if $\|x\| = 1$.
(c) Since $0 - 0 \cdot I$ is non-invertible, the claim holds for $S \equiv 0$. So by working with $\|S\|^{-1}S$, we may assume that $\|S\| = 1$. By the definition of $\|S\|$, there exists $(x_n) \in \mathcal{H}$ such that $\|x_n\| = 1$ and $\lim_{n\to\infty} \|S(x_n)\| = 1$. In fact, since $\|S\| = 1$, the definition of norm implies the existence of a sequence $(x'_n) \subset \mathcal{H} \setminus \{0_\mathcal{H}\}$ such that $\|x'_n\| \leq 1$ and $\lim_{n\to\infty} \|S(x'_n)\| = 1$. Since

$$\|S(x'_n)\| \leq \|S\|\|x'_n\| = \|x'_n\|,$$

we have $\lim_{n\to\infty} \|x'_n\| = 1$ as well. Choose $x_n = \frac{x'_n}{\|x'_n\|}$. Then $\|x_n\| = 1$ and

$$\|S(x_n)\| = \frac{\|S(x'_n)\|}{\|x'_n\|} \to 1$$

as $n \to \infty$.

Since S^2 is self-adjoint ($(S^2)^* = S^*S^* = S^2$), we have

$$\langle S^2(x), x\rangle = \langle x, S^2(x)\rangle \quad \forall\, x \in \mathcal{H}.$$

Therefore, by Lemma 3.1.6,

$$\begin{aligned}
\|(I - S^2)(x_n)\|^2 &= \langle (I - S^2)(x_n), (I - S^2)(x_n)\rangle = \langle x_n - S^2(x_n), x_n - S^2(x_n)\rangle \\
&\overset{3.1.6}{=} \|x_n\|^2 + \|S^2(x_n)\|^2 - \underbrace{\langle x_n, S^2(x_n)\rangle}_{\in\mathbb{R}} - \langle S^2(x_n), x_n\rangle \\
&\overset{\|S^2\|\leq\|S\|\|S\|=1}{\leq} 2 - 2\langle S^2(x_n), x_n\rangle \overset{S^*=S}{=} 2 - 2\langle S(x_n), S(x_n) \\
&= 2 - 2\|S(x_n)\|^2.
\end{aligned}$$

It follows that

$$\lim_{n\to\infty} \|(I - S^2)(x_n)\| = 0$$

and Corollary 5.2.6 implies that $I - S^2$ is non-invertible. Hence $1 \in \sigma(S^2)$ and Theorem 5.3.8 implies that $1 \in (\sigma(S))^2$. This is possible if either 1 or -1 is in $\sigma(S)$.
(d) Exercise. □

Example 5.3.13. (a) If A is a self-adjoint matrix with eigenvalues $\{\lambda_1, ..., \lambda_n\}$, then by (d) of Theorem 5.3.12

$$\|A\| = \max\{|\lambda_1|, ..., |\lambda_n|\}.$$

(b) If B is any square matrix, then B^*B is self-adjoint by Lemma 5.2.12 and

$$\|B\|^2 = \|B^*B\|$$

by Theorem 5.1.8. Hence $\|B\|$ can be calculated by using eigenvalues of B^*B.

6. Compact operators

6.1. Some general properties.

Definition 6.1.1. Let X and Y be normed spaces. A linear transformation $T \in L(X,Y)$ is *compact* if for any bounded sequence (x_n) in X the sequence $(T(x_n))$ in Y contains a convergent subsequence.

The set of compact transformations in $L(X,Y)$ is denoted by $K(X,Y)$.

Theorem 6.1.2. *Let X and Y be normed spaces and let $T \in K(X,Y)$. Then $T \in B(X,Y)$.*

Proof. Exercise. □

Theorem 6.1.3. *Let X,Y,Z be normed spaces. Then*
(a) *If $S,T \in K(X,Y)$ and $\alpha,\beta \in \mathbb{C}$, then $\alpha S + \beta T$ is compact.*
(b) *If $S \in B(X,Y)$, $T \in B(Y,Z)$, and at least one of the operators S,T is compact, then $TS \in B(X,Z)$ is compact.*

Proof. (a) Let (x_n) be a bounded sequence in X. Since S is compact, there is a subsequence (x_{n_j}) such that $(S(x_{n_j}))$ converges. Since the subsequence (x_{n_j}) is bounded and T is compact, there is a subsequence $(x_{n_{j_k}})$ of (x_{n_j}) such that $T(x_{n_{j_k}})$ converges. Hence, for the sequence $(x_{n_{j_k}})$, there exists $y, y' \in Y$ so that

$$\lim_{k\to\infty} S(x_{n_{j_k}}) = y \quad \text{and} \quad \lim_{k\to\infty} T(x_{n_{j_k}}) = y';$$

see Lemma 1.2.2 (iii). Therefore

$$\lim_{k\to\infty} (\alpha S + \beta T)(x_{njk}) = \lim_{k\to\infty} \alpha S(x_{njk}) + \beta T(x_{njk}) = \alpha y + \beta y' \in Y,$$

and so $\alpha s + \beta T$ is compact.
(b) Let (x_n) be a bounded sequence in X. If S is compact, there is a subsequence (x_{n_j}) so that $\lim_{j\to\infty} S(x_{n_j}) = y \in Y$. Since T is bounded, and hence continuous, $\lim_{j\to\infty} T(S(x_{n_j})) = T(y)$ by Remark 4.3.19. Thus TS is compact.

Suppose that S is bounded and T is compact. Then the sequence $(S(x_n))$ is bounded. Since T is compact, there is a subsequence (x_{n_j}) so that $(T(S(x_{n_j})))$ converges. Again TS is compact. □

Notation. When dealing with compact operators one often considers subsequences or subsequences of subsequences. For notational simplicity, it is common to write (x_n) for subsequences (and for subsequences of subsequences etc.) of the sequence (x_n).

Definition 6.1.4. Let V,W be vector spaces and let $T \in L(V,W)$. The *rank* of T is the number

$$r(T) = dim(Im(T)).$$

Moreover, T is called a *finite rank operator* (or T *has finite rank*) if $dim(Im(T)) < \infty$, that is, $Im(T)$ has a finite basis.

Theorem 6.1.5. *Let X and Y be normed spaces and let $T \in B(X,Y)$. If T has finite rank, then T is compact.*

The proof if based on the following *Bolzano-Weierstrass theorem*, which we recall without proof.

Lemma 6.1.6. *Any infinite and bounded set A in $\mathbb{C}^k$ has an accumulation point.*

The proof of Theorem 6.1.5. Since T has finite rank, the space $Im(T)$ is finite-dimensional. If (x_n) is a bounded sequence in X, then by boundedness of T, $(T(x_n))$ is a bounded sequence in $Im(T)$. Let $y_n = T(x_n)$. Then $y_n = \sum_{i=1}^{k} \lambda_{in} e_i$, where $\lambda_{in} \in \mathbb{C}$ and $\{ e_1, \dots, e_k \}$ is a base of $Im(T)$. Moreover, if

$$y = \sum_{i=1}^{k} \mu_i e_i \in Im(T),$$

then $y_n \to y$ in $Im(T)$ if and only if

$$\lambda_n := (\lambda_{1n}, \dots, \lambda_{kn}) \to (\mu_1, \dots, \mu_k)$$

in $\mathbb{C}^k$, see Example 1.1.3 and notice that all norms and equivalent in $Im(T)$, since $Im(T)$ is finite-dimensional (Analysis 4/Rynne & Youngson, p.43). Since (y_n) is a bounded sequence, (λ_n) is a bounded sequence in $\mathbb{C}^k$. If $\{\lambda_n \ : \ n \in \mathbb{N}\}$ is a finite set, (λ_n) contains a subsequence which is constant; hence converging. If $\{\lambda_n : n \in \mathbb{N}\}$ is infinite, Lemma 6.1.6 implies that (λ_n) contains a converging subsequence. In any case for some subsequence (λ_{n_j}), $(\lambda 1n_j, \dots, \lambda kn_j) \to (\mu_1, \dots, \mu_k) \in \mathbb{C}^k$, and then

$$y_{n_j} \to y = \sum_{i=1}^{k} \mu_i e_i \in Im(T). \quad \square$$

Remark 6.1.7. Let X, Y be normed spaces and let $T \in B(X, Y)$. If $dim(X) < \infty$, then T has finite rank (see Linear algebra). Hence T is compact.

In general, compact operators have analogical properties as bounded operators in finite-dimensional case! Many operators related to applications are compact.

Theorem 6.1.8. *Let X be normed spaces, Y a Banach space, and let $T_k)$ be a sequence in $K(X,Y)$ so that $T_k \to T$ in $B(X,Y)$. Then T is compact, that is, $K(X,Y)$ is a closed subset of $B(X,Y)$.*

Proof. Let (x_n) be a bounded sequence in X. Since T_1 is compact, there is a subsequence $(x_{n_j(1)})$ so that $(T_1(x_{n_j(1)}))$ converges. Again, since T_2 is compact, there is a subsequence $(x_{n_j(2)})$ of $(x_{n_j(1)})$ so that $(T_2(x_{n_j(2)}))$ converges. Clearly, $(T_1(x_{n_j(2)}))$ converges as well as a subsequence of a converging sequence. Continuing in this fashion, we find subsequences $(x_{n_j(k)})$, $k \in \mathbb{N}$ so that

$$\{ n_j(1) \} \supset \{ n_j(2) \} \supset \cdots \supset \{ n_j(k) \} \supset \cdots$$

and $(T_i(x_{n_j(k)}))$ converges for all $i = 1, \dots, k$ for each $k \in \mathbb{N}$.

Let $n_k := n_k(k)$ be the diagonal of indices, $k \in \mathbb{N}$. Now $(T_i(x_{n_k}))$ converges for all $i \in \mathbb{N}$. By completeness of Y, it is enough to show that $(T(x_{n_k}))$ is a Cauchy sequence. Let $\varepsilon > 0$. Since the subsequence (x_{n_k}) is bounded, $\exists M > 0$ so that $\|x_{nk}\| \leq M$ $\forall k \in \mathbb{N}$. Also, since $\|T_k - T\| \to 0$ as $k \to \infty$, $\exists k_1 \in \mathbb{N}$ so that

$$\|T_k - T\| < \frac{\varepsilon}{3M} \quad \text{whenever } k \leq k_1.$$

Next, since $(T_{k_1}(x_{n_k}))$ converges (and therefore is a Cauchy sequence), $\exists k_2 \in \mathbb{N}$ so that

$$\|T_{k_1}(x_{n_r}) - T_{k_1}(x_{n_s})\| < \frac{\varepsilon}{3} \quad \text{whenever } r, s \leq k_2.$$

Now, since

$$\|T_{k_1}(x_{n_i}) - T(x_{n_i})\| \leq \|T_{k_1} - T\| \|x_{n_i}\| < \frac{\varepsilon}{3}$$

for all $i \in \mathbb{N}$, we have for all $r, s \leq k_2$

$$\begin{aligned} &\|T(x_{nr}) - T(x_{ns})\| \\ &\leq \|T_{k1}(x_{nr}) - T(x_{nr})\| + \|T_{k1}(x_{nr}) - T_{k1}(x_{ns})\| + \|T_{k1}(x_{nr}) - T_{k1}(x_{ns})\| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

This proves the claim. □

Note. The process for selecting the subsequence in Theorem 6.1.8 is called *Cantor's diagonalization*. The same idea is used in *Ascoli-Arzela theorem*.

Corollary 6.1.9. *If X is a normed space, Y a Banach space and (T_k) is a sequence of finite rank operators in $B(X,Y)$ so that $T_k \to T$ in $B(X,Y)$, then T is compact.*

Example 6.1.10. We show that $T \in B(l^2)$,

$$T((a_n)) = (\frac{1}{n}a_n),$$

is compact.

Proof. We know by Example 2.1.5 that $T \in B(l^2)$. For each $k \in \mathbb{N}$, let $T_k : l^2 \to l^2$ be defined by

$$T_k((a_n)) = ((a_1, \frac{1}{2}a_2, \cdots, \frac{1}{k}a_k, 0, \cdots)).$$

Then T_k are bounded and linear, and have finite rank since $dim(Im(T_k)) = k$. For any $a := (a_n) \in l^2$,

$$\|(T_k - T)(a)\|^2 = \sum_{n=k+1}^{\infty} \frac{|a_n|^2}{n^2} \leq (k+1)^{-2} \sum_{n=k+1}^{\infty} |a_n|^2 \leq (k+1)^{-2}\|a\|^2.$$

It follows that (by taking sup over a, $\|a\| \leq 1$)

$$\|T_k - T\| \leq (k+1)^{-1}.$$

Hence $T_k \to T$ in $B(l^2)$ and T is compact by Corollary 6.1.9. □

Remark 6.1.11. It is possible to prove: If X is a normed space, $\mathcal{H}$ is a Hilbert space, and $T \in K(X, \mathcal{H})$, then there is a sequence (T_k) of finite rank operators so that $T_k \to T$ in $B(X, \mathcal{H})$. See Rynne & Youngson, p. 167.