Analysis

Will Troiani

August 2020

Contents

1 Elementary facts

Definition 1.0.1. The **real numbers**, denoted $\mathbb{R}$, are equivalence classes of Cauchy sequences.

Remark 1.0.2. It can be shown that $\mathbb{R}$ satisfies the axioms for the real numbers, thus equivalence classes of Cauchy sequences is an appropriate model for the real numbers. Another example is given by Dedekind cuts.

Throughout we work with the space $\mathbb{C}$, but the proofs can be amalgamated easily for the real case.

Proposition 1.0.3. *Let $(z_1, z_2, ...), (w_1, w_2, ...)$ be convergent sequences of complex numbers, then their sum $(z_1 + w_1, z_2 + w_2, ...)$, any scalar product $(kz_1, kz_2, ...)$, and their product $(z_1w_1, z_2w_2, ...)$ are also convergent.*

Proof. Say $z_n \overset{n\to\infty}{\longrightarrow} c_1$ and $w_n \overset{n\to\infty}{\longrightarrow} c_2$. Then for $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that for $n > N$ we have $|z_n - c_1| < \epsilon/2, |w_n - c_2| < \epsilon/2$. Thus

$$|z_n + w_n - (c_1 + c_2)| \leq |z_n - c_1| + |w_n - c_2| \leq \epsilon/2 + \epsilon/2 = \epsilon \tag{1}$$

Next, there exists $N > 0$ such that for all $n > N$ we have $|z_n - c_1| < \epsilon/|k|$, the second claim follows.

Lastly, there exists $N \in \mathbb{N}$ such that for all $n > N$ we have

$$|a_n - c_1| < \frac{\epsilon}{2(|c_1| + 1)}, \quad \text{and} \quad |b_n - c_2| < \frac{\epsilon}{2(|c_2| + 1)} \tag{2}$$

Also, N can be taken so that for $n > N$ we have

$$|z_n| - |c_1| \leq ||z_n| - |c_1|| \leq |z_n - c_1| < 1 \tag{3}$$

so that

$$|z_n|/(|c_1| + 1) < 1 \tag{4}$$

Thus,

$$\begin{aligned}
|z_n w_n - c_1 c_2| &= |z_n w_n + z_n c_2 - z_n c_2 - c_1 c_2| \\
&\leq |z_n||w_n - c_2| + |c_2||z_n - c_1| \\
&\leq \epsilon/2 + \epsilon/2 \\
&= \epsilon
\end{aligned}$$

□

In other words, convergent sequences forms a $\mathbb{C}$-algebra.

Corollary 1.0.4. *If $\sum_{i=1}^{\infty} z_i$ and $\sum_{i=1}^{\infty} w_j$ converge, then so do (for any $k \in \mathbb{C}$)*

$$\sum_{i=1}^{\infty}(z_i + w_i), \quad \sum_{i=1}^{\infty} k z_i, \quad \textit{and} \quad \sum_{i=1}^{\infty} z_i w_i \tag{5}$$

2 Taylor's Theorem

2.1 Taylor's Theorem for real analysis

This Theorem can be found by pressing on the *Intermediate Value Theorem*, which itself is a result of pressing on the completeness property of the real numbers:

Axiom 2.1.1 (Completeness Axiom)**.** Every non-empty, bounded subset of $\mathbb{R}$ admits a supremum and an infimum.

Definition 2.1.2. A subset $E \subseteq \mathbb{R}$ is **connected** if it cannot be written as the union of two disjoint, open subsets.

Lemma 2.1.3. *A subset $E \subseteq \mathbb{R}$ is connected if and only if it satisfies the following property: if $x < y \in E$ and $z \in \mathbb{R}$ is such that $x < z < y$ then $z \in E$.*

Proof. Say E is not connected, say $E = A \cup B$ with A, B disjoint, open subsets of $\mathbb{R}$ and assume without loss of generality that every element of B is an upper bound of A. Let $x \in A, y \in B$ and consider

$$z := \sup(A \cap [x, y])$$

Then z is not in E as if $z \in A$ then A being open implies there exists $\epsilon > 0$ such that $B_\epsilon(z) \subseteq A$, which implies $z < z + \epsilon/2 \in A \cap [x, y]$ contradicting that z is an upper bound. If $z \in B$ then there exists $\epsilon > 0$ such that $z > z - \epsilon/2 \in B$ with $z - \epsilon/2 \notin A$, contradicting that z is the least upper bound of $A \cap [x, y]$.

Conversely, let $x < z < y$ be such that $x \in E, y \in E, z \notin E$, then $E = \big((-\infty, z) \cap E\big) \cup \big(E \cap (z, \infty)\big)$. □

Lemma 2.1.4. *Let $f : A \longrightarrow \mathbb{R}$ be a continuous function where $A \subseteq \mathbb{R}$ is any set (the following holds if A is any metric space). Then A connected implies $f(A)$ is connected.*

Proof. Say $f(A)$ is not connected, so $f(A) = B \cup C$ with $B, C \supseteq \mathbb{R}$ open subsets. Then $A = f^{-1}(B) \cup f^{-1}(C)$. □

Theorem 2.1.5 (Intermediate Value Theorem)**.** *Let $f : [a, b] \longrightarrow \mathbb{R}$ be a continuous function satisfying $f(a) \leq f(b)$. Then for all $u \in [f(a), f(b)]$ there exists $c \in [a, b]$ such that $f(c) = u$.*

Proof. By Lemma 2.1.4 we have that $f([a, b])$ is connected, so the result follows from Lemma 2.1.3. □

Corollary 2.1.6 (Mean Value Theorem)**.** *Let $f : [a,b] \longrightarrow \mathbb{R}$ be a function which is differentiable for all $t \in (a,b)$. Then there exists $c \in (a,b)$ such that*

$$f'(x) = \frac{f(b) - f(a)}{b - a}$$

Proof. [?]. $\square$

Theorem 2.1.7 (Taylor's Theorem)**.** *Let $f : [a,b] \longrightarrow \mathbb{R}$ be n times differentiable in the interval (a,b), and assume $f^{(n-1)}$ is continuous on some closed interval $[\alpha, \beta] \subseteq [a,b]$. Define:*

$$P_n(t,\alpha) := \sum_{m=0}^{n-1} \frac{f^{(m)}(\alpha)}{m!}(t-\alpha)^m$$

Then there exists $\gamma \in (\alpha, \beta)$ such that

$$f(\beta) = P_n(\beta, \alpha) + \frac{f^{(n)}(\gamma)}{n!}(\beta - \alpha)^n$$

Proof. Let M be such that

$$f(\beta) = P_n(\beta, \alpha) + M(\beta - \alpha)^n$$

and define

$$g(t) = f(t) - P_n(t, \alpha) - M(t - \alpha)^n$$

so that

$$g^{(n)}(t) = f^{(n)}(t) - n!M$$

Notice that $g^{(n)}(\alpha) = 0$. The last step is to prove by induction on n that there exists $x_n \in (\alpha, \beta)$ such that $g^{(n)}(x_n) = 0$. If $n = 1$ then $g^{(1)}(\alpha) = g'(\alpha) = f'(\alpha) - f'(\alpha) = 0$. Also, $g(\beta) = 0$ by the defining property of M so there exists $x_1 \in (\alpha, \beta)$ such that $g'(x_1) = 0$. Now assume $n > 1$ and the result holds for $n - 1$. Then since there exists x_{n-1} such that $g^{(n-1)}(x_{n-1}) = 0$ we have by the mean value theorem that there exists $x_n \in (\alpha, \beta)$ such that $g^{(n)}(x_n) = 0$. $\square$

Definition 2.1.8. Given a function $f : [a,b] \longrightarrow \mathbb{R}$ and real number $\alpha \in (a,b)$ the function $P_n(t,\alpha)$ is the n^{th} **order Taylor polynomial about** α. (Notice the notation is a bit naughty as we suppress the function f).

Remark 2.1.9. If a function $f : [a,b] \longrightarrow \mathbb{R}$ can be written as a power series about $\alpha \in (a,b)$:

$$f(x) = \sum_{k=0}^{\infty} c_n (x - \alpha)^k$$

then the sequence of Taylor polynomials about α corresponding to f, $(P_1(x,\alpha), P_2(x,\alpha), \ldots)$ converges to $f(x)$.

3 Hilbert spaces

Definition 3.0.1. A **Hilbert space** is an inner product space which is complete with respect to the norm generated by its inner product.

3.1 The space ℓ^2

Definition 3.1.1. Let ℓ^2 denote the Hilbert product space of **square summable sequences**, that is, sequences $(z_1, z_2, ...)$ of complex numbers such that

$$\sum_{i=1}^{\infty} |z_i| < \infty \tag{6}$$

Define an inner product on this space by

$$\big\langle (z_1, z_2, ...), (w_1, w_2, ...) \big\rangle = \sum_{i=1}^{\infty} z_i \bar{w}_i \tag{7}$$

We check this is in fact a Hilbert space:

Proof. We have already seen that ℓ^2 is a vector space (Corollary 1.0.4). That we have a valid inner product is clear, so it remains to show that this space is complete with respect to this inner product.

Say we had a Cauchy sequence $\big((z_n^1), (z_n^2), ...\big)$ consisting of elements in ℓ^2. Consider for each i the sequence $(z_i^1, z_i^2, ...)$, which we claim is a Cauchy sequence in $\mathbb{C}$. This is easy to establish as for all $\epsilon > 0$ there exists $N > 0$ such that for $m, k > N$ we have

$$\Big(\sum_{n=1}^{\infty} |z_n^m - z_n^k|\Big)^{1/2} < \sqrt{\epsilon} \tag{8}$$

and clearly

$$|z_n^m - z_i^k| \leq \sum_{n=1}^{\infty} |z_n^m - z_n^k| \tag{9}$$

Since $\mathbb{C}$ is complete, we thus have $(z_i^1, z_i^2, ...)$ converges to some complex number, c_i say.

We now claim that $\big((z_n^1), (z_n^2), ...\big) \longrightarrow (c_1, c_2, ...)$.

Let $\epsilon > 0$ and $N > 0$ such that for all $m, k > N$ we have $||(z_n^m) - (z_n^k)|| < \epsilon/\sqrt{2}$. Fix $M > 0$ and consider the finite sum

$$\sum_{n=1}^{M} |z_n^m - c_n|^2 \tag{10}$$

We pick l such that $|z_n^m - z_n^l| < \epsilon/\sqrt{2M}$ for all $n = 1, ..., M$. Then we have

$$\begin{aligned}
\sum_{n=1}^{M} |z_n^m - c_n|^2 &= \sum_{n=1}^{M} |z_n^m + z_n^l - z_n^l - c_n|^2 \\
&\leq \sum_{n=1}^{M} |z_n^m - z_n^l|^2 + \sum_{n=1}^{M} |z_n^l - c_n|^2 \\
&< \epsilon^2/2M + \epsilon^2/2 \\
&= \epsilon^2
\end{aligned}$$

Since this holds true for arbitrary M, we have

$$||z_n^m - c_n||_2 < \epsilon \tag{11}$$

proving the claim. $\square$

4 Operator Theory

4.1 Adjoint operator

We will be chiefly concerned with the Hilbert space ℓ^2 but we work in a more general setting for now. A *Hilbert space* will always mean over $\mathbb{C}$. Associated to every operator between Hilbert spaces is an operator between their *dual spaces*:

In general, if $\mathcal{I}$ is any inner product space over $\mathbb{C}$ and we have two vectors $x, y \in I$ then we can consider the projection of y onto x which is given by

$$\mathrm{Proj}_y(x) := \frac{\langle x, y\rangle}{||y||}\frac{y}{||y||} \tag{12}$$

Thus, if $U \subseteq \mathcal{I}$ is a one dimensional subspace spanned by a unit vector $u \in U$ then the projection of any $x \in \mathcal{I}$ onto u is given by the simple formula $\langle x, u\rangle u$. The following Lemma shows what we can say when the subspace is of arbitrary dimension but with U closed:

Lemma 4.1.1. *Let $\mathbb{H}$ be a Hilbert space and $U \subseteq \mathbb{H}$ a closed subspace. Then*

$$\mathbb{H} = U \oplus U^{\perp}$$

Proof. We will define a projection

$$\begin{aligned} P_U : \mathbb{H} &\longrightarrow U \\ x &\longmapsto \inf\{||x - y|| \mid y \in U\} \end{aligned}$$

We let d denote $\inf\{||x-y|| \mid y \in U\}$. By definition of inf there exists a sequence $(x_n)_{n=0}^{\infty}$ of elements in U such that $\lim_{n\to\infty} ||x - x_n|| = d$. Since U is closed it is complete and the norm is continuous so it suffices to show that the sequence $(x_n)_{n=0}^{\infty}$ is Cauchy. This can be done for example using the parallelogram identity: for all $n, m \geq 0$:

$$||x_n - x_m||^2 + ||(x - x_n) + (x - x_m)||^2 = 2||x - x_n||^2 + 2||x - x_m||^2 \tag{13}$$

As given $\epsilon > 0$ there exists $N \geq 0$ such that $||x - x_n||^2 < d^2 + \epsilon^2/4$, for $n \geq N$. Thus

$$\begin{aligned} ||x_n - x_m||^2 &= 2||x - x_n||^2 + 2||x - x_m||^2 - 4||x = ||1/2(x_n + x_m)||^2 \\ &\leq 4d^2 + \epsilon^2 - 4||x - 1/2(x_n + x_m)||^2 \end{aligned}$$

which since $1/2(x_n + x_m) \in C$ we have $d \leq ||x - 1/2(x_n + x_m)||$, proving $(x_n)_{n=0}^{\infty}$ is Cauchy. This also shows linearity.

It remains to show $x - P_U(x) \in U^{\perp}$. To do this, we will consider the family of vectors $c(t) = (1-t)P_U(x) + ty$, $(t \in \mathbb{R})$ and analyse the derivative of $||x - y_t||^2$ at $t = 0$.

Consider the composition

$$\begin{aligned} \gamma : \mathbb{R} &\longrightarrow \mathbb{R} \qquad (14) \\ t &\longmapsto ||x - c(t)||^2 \qquad (15) \end{aligned}$$

We can write γ in a more explicit form:

$$\begin{aligned} \gamma(t) &= ||x - P_U(x) + t(y - P_U(x))||^2 \\ &= \big\langle x - P_U(x) + t(y - P_U(x)), x - P_U(x) + t(y - P_U(x))\big\rangle \\ &= ||x - P_U(x)||^2 - 2t\,\mathrm{Re}\langle x - P_U(x), y - P_U(x)\rangle + t^2||y - P_U(x)|| \end{aligned}$$

which is clearly differentiable and has derivative $-2\,\mathrm{Re}\langle x - P_U(x), y - P_U(x)\rangle$ at $t = 0$. Since $P_U(x)$ (which equals $c(0)$) is a minimum of $\gamma(t)$ we have that $\mathrm{Re}\langle x - P_U(x), y - P_U(x)\rangle = 0$. This holds true for arbitrary $y \in U$ and lastly we have

$$\{y - P_U(x) \mid y \in U\} = U$$

thus for all $y \in U$:

$$\mathrm{Re}\langle x - P_U(x), y\rangle = 0 \tag{16}$$

This shows that $x - P_U(x) \in U^\perp$. $\square$

Given a Hilbert space $\mathbb{H}$ there is a map

$$\begin{aligned} \Phi : \mathbb{H} &\longrightarrow \mathbb{H}^* && (17)\\ b &\longmapsto \langle \_, b\rangle && (18) \end{aligned}$$

Notice that in order to produce a *linear* functional, it was important we put b in the second argument, we must define Φ so that $\Phi(b) \neq \langle b, \_\rangle$. By anti-linearity of the second argument of the inner product we have that Φ is anti-linear, and moreover is injective as

$$\begin{aligned} \Phi(b) = \Phi(b') &\implies \langle \_, b\rangle = \langle \_, b'\rangle \\ &\implies \forall b'' \in \mathbb{H}, \langle b'', b - b'\rangle = 0 \\ &\implies \text{ in particular, } \langle b - b', b - b'\rangle = 0 \\ &\implies b - b' = 0 \end{aligned}$$

In the special case where $\mathbb{H}$ is finite dimensional, we automatically have that this map is surjective as it is injective, and any anti-linear, injective map between two finite dimensional spaces of equal dimension is automatically surjective. More generally, if $\mathbb{H}$ has arbitrary dimension, then for any $y \in \mathbb{H}$ the map $\langle \_, y\rangle$ is bounded (see Remark 4.1.4) so the image of Φ is contained in the set of continuous linear functionals, the following establishes the reverse inequality:

Theorem 4.1.2 (Riesz Representation Theorem)**.** *Let $\mathbb{H}$ be a Hilbert space. For every continuous linear functional $\varphi \in \mathbb{H}^*$ there exists a unique element $h_\varphi \in \mathbb{H}$ such that*

$$\varphi = \langle \_, h_\varphi\rangle \tag{19}$$

Moreover, we have

$$||\varphi||_{\mathbb{H}^*} = ||h_\varphi||_{\mathbb{H}} \tag{20}$$

We will use the following Lemma:

Lemma 4.1.3. *Let $\mathbb{H}$ be a Hilbert space and $\varphi \in \mathbb{H}^*$ be non-zero and continuous. Then $(\ker\varphi)^\perp$ is one dimensional.*

Proof. Since φ is continuous the set $\ker\varphi$ is closed and so by Lemma 4.1.1 we have $\mathbb{H} = \ker\varphi \oplus (\ker\varphi)^\perp$, which since $\varphi \neq 0$ implies there exists $v \neq 0 \in (\ker\varphi)^\perp$, so $\dim(\ker\varphi)^\perp > 0$. Now, say $v_1, v_2 \in (\ker\varphi)^\perp$ so that $\varphi(v_1) \neq 0$ and $\varphi(v_2) \neq 0$. These are complex numbers and so there exists $\lambda \in \mathbb{C}$ such that

$$0 = \lambda\varphi(v_1) - \varphi(v_2) = \varphi(\lambda v_1 - v_2)$$

which means $\lambda v_1 - v_2 \in \ker\varphi \cap (\ker\varphi)^\perp = \{0\}$. $\square$

Proof of Theorem 4.1.2. Clearly if $\ker\varphi = \mathbb{H}$ we can take $h_\varphi = 0$ so assume this is not the case. Since φ is continuous its kernel $\ker\varphi$ is a closed subset of $\mathbb{H}$. Thus, by Lemma 4.1.1 the Hilbert space $\mathbb{H}$ decomposes: $\mathbb{H} = \ker\varphi \oplus (\ker\varphi)^*$. Since $\ker\varphi$ is a proper subset it then follows that there exists a non-zero element $v \neq 0 \in (\ker\varphi)^*$, by normalising we may assume that v is a unit vector. We will show that $\overline{\varphi(v)}v$ is the appropriate unique choice for h_φ.

By Lemma 4.1.3 the subspace $(\ker\varphi)^\perp$ is one dimensional, hence we can use formula (12) for the projection of arbitrary x onto $(\ker\varphi)^\perp$. Observe the following calculation:

$$\begin{aligned}\varphi(x) &= \varphi(x - \langle x, v\rangle v + \langle x, v\rangle v)\\ &= \varphi(x - \langle x, v\rangle v) + \varphi(\langle x, v\rangle v)\\ &= 0 + \langle x, v\rangle \varphi(v)\\ &= \langle x, \overline{\varphi(v)} v\rangle\end{aligned}$$

For uniqueness, say h'_φ was another such element. Then

$$\begin{aligned}&\forall x \in \mathbb{H}, \langle x, h_\varphi\rangle = \langle x, h'_\varphi\rangle\\ &\implies \forall x \in \mathbb{H}, \langle x, h_\varphi - h'_\varphi\rangle = 0\\ &\implies ||h_\varphi - h'_\varphi|| = 0\\ &\implies h_\varphi = h'_\varphi\end{aligned}$$

For the second claim, we use the Cauchy-Schwartz inequality:

$$|\varphi(x)| = |\langle x, \overline{\varphi(v)} v\rangle| \leq ||x||||\overline{\varphi(v)}||v|| = ||x|||\varphi(v)|$$

and so if x has unit norm $|\varphi(x)| \leq |\varphi(v)|$, in other words, $||\varphi||_{\mathbb{H}^*} \leq |\varphi(v)|$ however v has unit norm itself, so $||\varphi||_{\mathbb{H}^*} = |\varphi(v)|$. The proof is now complete once it is noted that $||h_\varphi||_{\mathbb{H}} = |\varphi(v)|$. □

Remark 4.1.4. Let $y \in \mathbb{H}$ be an element of a Hilbert space $\mathbb{H}$ and consider the function $\langle \_, y\rangle$. This is bounded, as by Cauchy-Schwartz:

$$|\langle x, y\rangle| \leq ||x||||y||$$

thus $|\langle \_, y\rangle|/||x|| \leq ||y||$ and in fact this is equality as $|\langle y/||y||, y\rangle| = ||y||$.

Given an operator $u : \mathbb{H}_1 \longrightarrow \mathbb{H}_2$ there is for each $y \in \mathbb{H}_2$ an associated linear functional $x \longmapsto \langle u(x), y\rangle$ which we denote by $\langle u(\_), y\rangle$. By Theorem 4.1.2 there is thus an element $y^* \in \mathbb{H}_1$ such that $\langle u(\_), y\rangle = \langle \_, y^*\rangle$. The assignment $y \mapsto y^*$ is in fact linear, we show additivity:

$$\langle u(\_), y_1 + y_2\rangle = \langle \_, (y_1 + y_2)^*\rangle$$

and

$$\begin{aligned}\langle u(\_), y_1 + y_2\rangle &= \langle u(\_), y_1\rangle + \langle u(\_), y_2\rangle\\ &= \langle \_, y_1^*\rangle + \langle \_, y_2^*\rangle\\ &= \langle \_, y_1^* + y_2^*\rangle\end{aligned}$$

which implies $(y_1 + y_2)^* = y_1^* + y_2^*$. We define:

Definition 4.1.5. The **adjoint operator** associated to an operator $u : \mathbb{H}_1 \longrightarrow \mathbb{H}_2$ is the linear map:

$$\begin{aligned}u^* : \mathbb{H}_2 &\longrightarrow \mathbb{H}_1\\ y &\longmapsto y^*\end{aligned}$$

Its existence is established by the Riesz Representation Theorem (4.1.2) and it is uniquely determined by the property:

$$\forall x \in \mathbb{H}_1, y \in \mathbb{H}_2, \langle u(x), y\rangle = \langle x, u^*(y)\rangle \tag{21}$$

Remark 4.1.6. Let $(\mathbb{H}_1, \langle\cdot,\cdot\rangle_B), (\mathbb{H}_2, \langle\cdot,\cdot\rangle)_C$ be Hilbert spaces and let $u : \mathbb{H}_1 \longrightarrow \mathbb{H}_2$ be an operator. The **adjoint** to u, denoted u^* is the operator:

$$\_ \circ u : \mathbb{H}_2^* \longrightarrow \mathbb{H}_1^* \tag{22}$$

$$\varphi \longmapsto \varphi \circ u \tag{23}$$

The following diagram commutes:

$$\begin{tikzcd}
\mathbb{H}_2^* \arrow[rrrr, "_\circ u"] \arrow[ddd, "\cong"'] & & & & \mathbb{H}_1^* \arrow[ddd, "\cong"] \\
& \langle _, v\rangle \arrow[r, mapsto] & \langle u(_), v\rangle \arrow[r, leftrightarrow] & \langle _, v^*\rangle & \\
& v \arrow[u, mapsto] \arrow[r, mapsto] & u^*(v) \arrow[ur, mapsto] & & \\
\mathbb{H}_2 \arrow[rrrr, "u^*"'] & & & & \mathbb{H}_1
\end{tikzcd} \tag{24}$$

which explains the overloading of terminology.

Notation 4.1.7. Given a complex matrix A, the matrix given by conjugating each element $a \in A$ and then transposing the result, ie, the **conjugate transpose** is denoted $A^\dagger$. Due to Proposition 4.1.8 below, the conjugate transpose of a matrix is often referred to as the **adjoint**.

Proposition 4.1.8. *Let $\mathbb{H}_1, \mathbb{H}_2$ be finite dimensional, and let $v_1, ..., v_n \in \mathbb{H}_1, w_1, ..., w_m \in \mathbb{H}_2$ be orthonormal bases for $\mathbb{H}_1, \mathbb{H}_2$ respectively. If $\varphi : \mathbb{H}_1 \longrightarrow \mathbb{H}_2$ is a linear transformation and A its matrix representation with respect to these bases, then the matrix representation of the adjoint φ^* is $A^\dagger$, the conjugate transpose of A.*

Proof. For each $j = 1, ..., m$ write $w_j^* = \alpha_1 v_1 + \ldots + \alpha_n v_n$ and each $i = 1, ..., n$ write $\varphi(v_i) = \beta_1 w_1 + \ldots + \beta_m w_m$. We calculate:

$$\langle \varphi(v_i), w_j\rangle = \beta_m \langle w_1, w_j\rangle + \ldots + \beta_m \langle w_m, w_j\rangle = \beta_j \tag{25}$$

and

$$\langle v_i, w_j^*\rangle = \bar{\alpha}_1 \langle v_i, v_1\rangle + \ldots + \bar{\alpha}_n \langle v_i, v_n\rangle = \bar{\alpha}_i \tag{26}$$

Since by definition $\langle \varphi(v_i), w_j\rangle = \langle v_i, w_j^*\rangle$ the proof is complete. □

4.2 Hermitian operators

Throughout, V is a complex vector space.

Definition 4.2.1. A square, complex matrix A is **Hermitian** if it is self-adjoint, that is $A^\dagger = A$.

A matrix is **normal** if $AA^\dagger = A^\dagger A$

An operator $\varphi : V \longrightarrow V$ is **Hermitian** (**normal**) if a (and hence all) matrix representation(s) of V is Hermitian (normal).

Clearly, all Hermitian matrices are normal.

Theorem 4.2.2 (Spectral decomposition)**.** *Let V be a finite dimensional, complex vector space and A a matrix representation of an operator on V. The matrix A is normal if and only if it is diagonalisable with respect to some orthonormal basis for V.*

Proof. We prove that normal matrices are diagonalisable.

We proceed by induction on the size of the matrix. If the matrix is 1×1 then there is nothing to prove. Now for the inductive step. Let λ be an eigenvalue of A, and P the matrix which projects onto the λ-eigenspace. We let Q denote $I - P$, the projector onto the complement subspace. We notice that

$$A = (P + Q)A(P + Q) = PAP + QAP + PAQ + QAQ \tag{27}$$

We have that $QAP = 0$ because A maps the λ-eigenspace onto itself, and we claim moreover that $PAQ = 0$. To see this, let v be an eigenvector with eigenvalue λ, then

$$AA^\dagger v = A^\dagger Av = A^\dagger \lambda v = \lambda A^\dagger v \tag{28}$$

which means $A^\dagger$ maps the λ-eigenspace onto itself. This implies $QA^\dagger P = 0$, taking the transpose of which we end at $PAQ = 0$ as claimed.

Thus $A = PAP + QAQ$. The matrix PAP is diagonalisable with respect to some orthonormal basis for P. Since $P \cap Q = 0$ it remains to show that QAQ is diagonalisable with respect to some orthonormal basis for Q. The space Q has strictly smaller size than A and so this follows by induction once we have shown that QAQ is normal. This is a simple calculation:

$$\begin{aligned} QAQQA^\dagger Q &= QAQA^\dagger Q \\ &= QA(P+Q)A^\dagger Q \\ &= QAA^\dagger Q \\ &= QA^\dagger AQ \\ &= QA^\dagger(P+Q)AQ \\ &= QA^\dagger QAQ \\ &= QA^\dagger QQAQ \end{aligned}$$

□

Definition 4.2.3. A matrix U is **unitary** if $U^\dagger U = I$.

Lemma 4.2.4. *A unitary matrix U satisfies $UU^\dagger = I$.*

Proof. We calculate:

$$U^\dagger U = (\langle v_i, v_j\rangle)_{ij} = (\overline{\langle v_j, v_i\rangle})_{ij} = (\langle v_j, v_i\rangle)_{ij} = (\langle v_j, v_i\rangle)_{ji} = UU^\dagger \tag{29}$$

where the third and fourth equality follow from the fact that $U^\dagger U = I$. □

Notice that the spectral decomposition (4.2.2) states that the matrix A is such that $A = U^\dagger DU$ for a diagonal matrix D and a unitary matrix U.

Corollary 4.2.5. *A normal matrix A is Hermitian if and only if its eigenvalues are real.*

Proof. First notice that if a matrix is Hermitian then for any eigenvector v with eigenvalue λ:

$$\lambda|v|^2 = \langle \lambda v, v\rangle = \langle Av, v\rangle = \langle v, Av\rangle = \bar{\lambda}|v|^2 \tag{30}$$

Now we prove the other direction. Let D be diagonal and U a unitary matrix such that $A = U^{-1}DU$. Then

$$A^\dagger = U^\dagger D^\dagger U^{-1^\dagger} = U^{-1}DU = A \tag{31}$$

□

Definition 4.2.6. An operator $\varphi : V \longrightarrow V$ is **positive** if:

$$\forall v \in V, \langle v, \varphi v\rangle \geq 0 \tag{32}$$

which means, $\langle v, \varphi v\rangle$ is real and non-negative. If the inequality is strict, then φ is **positive definite**.

Example 4.2.7. Let A be any operator. Then for any $v \in V$:

$$\langle v, A^\dagger A v\rangle = \langle Av, Av\rangle = ||Av||^2 \geq 0 \tag{33}$$

Thus $A^\dagger A$ is positive.

Proposition 4.2.8. *A positive operator on a finite dimensional vector space is necessarily Hermitian.*

Proof. Let A be a matrix representation of the positive operator. Notice the following calculation:

$$\begin{aligned}
0 \leq \langle v, (A - A^\dagger)v\rangle &= \langle (A^\dagger - A)v, v\rangle \\
&= \overline{\langle v, (A^\dagger - A)v\rangle} \\
&= \langle v, (A^\dagger - A)v\rangle \\
&= -\langle v, (A - A^\dagger)v\rangle \geq 0
\end{aligned}$$

and so for all $v \in V$ we have $\langle v, (A - A^\dagger)v\rangle = 0$.

Moreover, we notice that $A - A^\dagger$ is normal and hence diagonalisable, by the Spectral decomposition. It follows from these two observations that $A - A^\dagger = 0$. $\square$

Definition 4.2.9. Let A, B be matrices, then the **commutator** is $[A, B] := AB - BA$. The **anticommutator** is $\{A, B\} = AB + BA$.

Theorem 4.2.10 (Simultaneous Diagonalisation Theorem)**.** *Let A, B be Hermitian operators. Then $[A, B] = 0$ if and only if A and B are simultaneously diagonalisable.*

Proof. If A and B are simultaneously diagonalisable, then let U be a unitary matrix and D_1, D_2 diagonal matrices such that

$$A = U^{-1}D_1U, \qquad B = U^{-1}D_2U \tag{34}$$

We then have:

$$AB = U^{-1}D_1UU^{-1}D_2U = U^{-1}D_1D_2U = U^{-1}D_2D_1U = U^{-1}D_2UU^{-1}D_1U = BA \tag{35}$$

Conversely, say $[A, B] = 0$. We have that A is Hermitian and so admits a spectral decomposition. Let $a_1, ..., a_n$ be the eigenvalues corresponding to this decomposition and let V_{a_i} denote the a_i-eigenspace. We first notice that B maps V_{a_i} into itself: for any $v \in V_{a_i}$

$$ABv = BAv = a_iBv \tag{36}$$

Now, since B is Hermitian, it follows that $B_{V_{a_i}} : V_{a_i} \longrightarrow V_{a_i}$ is and so there exists a spectral decomposition of $B_{V_{a_i}}$ for each vector space V_{a_i}. Denote by $b_1^{a_i}, ..., b_{k_{a_i}}^{a_i}$ an orthonormal basis for V_{a_i}. We then have that

$$\{b_1^{a_i}, ..., b_{k_{a_i}}^{a_i}\}_{i=1}^n \tag{37}$$

is a basis of eigenvectors of both A and B for the whole space V. $\square$

There is another decomposition which is often helpful:

Observation 4.2.11. Let $T : V \longrightarrow V$ be a linear operator on an n-dimensional vector space V. We could ask if T can be factored $T = UT'$ where U is unitary? Say this was possible, then

$$T^\dagger T = T'^\dagger U^\dagger U T' \tag{38}$$

so if T' were Hermitian we would have $T^\dagger T = T'^2$ which would imply $T' = \sqrt{T^\dagger T}$, in fact $T^\dagger T$ is Hermitian (indeed it is positive) and thus so is $\sqrt{T^\dagger T}$ and so our assumption that T' be Hermitian is not too much to ask for, and if U were to exist it must be that $T' = \sqrt{T^\dagger T}$. Thus we are prompted to make the following calculation: let $v_1, ..., v_n$ be a basis for V such that (we write P_{v_i} for the projection onto v_i)

$$\sqrt{T^\dagger T} = \sum_{i=1}^{n} \lambda_i P_{v_i} \tag{39}$$

then

$$\sqrt{T^\dagger T} v_i \lambda_i \tag{40}$$

and indeed we want U such that $\lambda_i U v_i = T v_i$. One might suggest defining $U v_i = T v_i / \lambda_i$ at this point, however there is no reason for this to be unitary. Instead we define

$$U = \sum_{j=1}^{n} T v_j P_{v_j} / \sqrt{\lambda_j} \tag{41}$$

which indeed is unitary. In fact we read off from this that $\{T v_1/\sqrt{\lambda_1}, ..., T v_n/\sqrt{\lambda_n}\}$ is an orthonormal basis for V. Notice however that this assumes $\lambda_i \neq 0$ for all i. This can be fixed by doing this process first for all $\lambda_i \neq 0$, and to construct an orthonormal set $\{T v_1/\sqrt{\lambda_1}, ..., T v_j/\sqrt{\lambda_j}\}$ and then extending this to an orthonormal basis for V via the Gram-Schmidt process.

We have proven the first half of:

Theorem 4.2.12 (Polar decomposition). *Let $T : V \longrightarrow V$ be a linear operator on an n-dimensional vector space V. Then there exists a unitary operator U and positive operators J, K such that*

$$T = UJ = KU \tag{42}$$

with $J = \sqrt{A^\dagger A}, K = \sqrt{AA^\dagger}$.

To obtain K we simply notice

$$A = JU = UJU^\dagger U \tag{43}$$

so we set $K = UJU^\dagger$, which is a positive operator. Then $AA^\dagger = KUU^\dagger K = K^2$.

If we have such a decomposition $T = UJ$, then J is diagonalisable, being positive, thus $T = USDS^\dagger$ for unitary S and diagonal D. Setting $V = S^\dagger$ we obtain:

Corollary 4.2.13 (Singular value decomposition). *Let $T : V \longrightarrow V$ be a linear operator on an n-dimensional vector space, then there exists unitary operators U, V and a diagonal operator D such that*

$$T = UDV \tag{44}$$