Hilbert Spaces

Joseph Muscat

2014-5-27

(A revised and expanded version of these notes are now published by Springer.)

1 Inner Product Spaces

1.1 Introduction

Definition An **inner-product** on a vector space X is a map

$$\langle\ ,\ \rangle : X \times X \to \mathbb{C}$$

such that

$$\langle x, y+z\rangle = \langle x,y\rangle + \langle x,z\rangle,$$
$$\langle x, \lambda y\rangle = \lambda\langle x,y\rangle,$$
$$\langle y,x\rangle = \overline{\langle x,y\rangle},$$
$$\langle x,x\rangle \geqslant 0; \quad \langle x,x\rangle = 0 \Leftrightarrow x = 0.$$

In words, it is said to be a positive-definite sesquilinear form.

The simplest examples are $\mathbb{R}^N$ and $\mathbb{C}^N$ with $\langle \boldsymbol{x}, \boldsymbol{y}\rangle = \sum_{n=1}^N \bar{x}_n y_n$; the square matrices of size $N \times N$ also have an inner-product given by $\langle A, B\rangle = \sum_{i,j=1}^N \bar{A}_{ij} B_{ij}$.

Proposition 1.1

1. **$\langle x,x\rangle$ is real (and positive) and is denoted by $\|x\|^2$**
2. **$\langle x,y\rangle = 0\ \forall x \ \Rightarrow\ y = 0$ (put $x = y$)**
3. **$\langle \lambda x, y\rangle = \bar{\lambda}\langle x,y\rangle$ (anti-linear); $\langle x+y, z\rangle = \langle x,z\rangle + \langle y,z\rangle$;**

4. $\|\lambda x\| = |\lambda|\|x\|$**.**

Definition Two vectors x, y are **orthogonal** when $\langle x, y\rangle = 0$, written as $x \perp y$. The *angle* between two vectors is given by $\cos\theta = \langle x, y\rangle / \|x\|\,\|y\|$.

Proposition 1.2

1. $\|x+y\|^2 = \|x\|^2 + 2\mathbf{Re}\langle x, y\rangle + \|y\|^2$**.**
2. **(Pythagoras) If** $z = x + y$ **and** $\langle x, y\rangle = 0$ **then** $\|z\|^2 = \|x\|^2 + \|y\|^2$**.**
3. **For any orthogonal vectors** x, y**,** $\|x\| \leqslant \|x+y\|$**.**
4. **For any non-zero vectors** x, y**, there is a unique vector** z **and a unique scalar** λ **such that** $x = z + \lambda y$ **and** $z \perp y$**.**

Proof. The first three statements follow immediately from the axioms and the first proposition. For the last statement, let $z = x - \lambda y$ as required; then $z \perp y$ holds $\Leftrightarrow$ $\lambda = \langle y, x\rangle / \langle y, y\rangle$ (check).

□

Proposition 1.3 **Cauchy-Schwarz**

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

Proof. Decompose x into orthogonal parts $x = (x - \lambda y) + \lambda y$ where $\lambda = \frac{\langle y,x\rangle}{\langle y,y\rangle}$ (assuming $y \neq 0$); now use Pythagoras' theorem, and deduce that $|\lambda y| \leqslant \|x\|$.

□

Note that Pythagoras' theorem and Cauchy-Schwarz's inequality are still valid even if the 'inner-product' is not positive definite but just semi-definite, as long as $\|y\| \neq 0$.

Corollary

$$\|x+y\| \leqslant \|x\| + \|y\|$$

The proof is simply an application of the Cauchy-Schwarz inequality to the expansion of $\|x+y\|^2$.

□

Hence $\|x\|$ is a norm, and all the facts about normed vector spaces apply to inner-product spaces. In particular they are metric spaces with distance $d(x,y) = \|x-y\|$, convergence of sequences makes sense as $x_n \to x \Leftrightarrow \|x_n - x\| \to 0$, continuity and dual spaces also make sense. Inner-product spaces are special normed spaces which not only have a concept of length but also of angle.

Definition A **Hilbert** space is an inner-product space which is complete as a metric space.

Proposition 1.4 **The closure of a linear subspace remains a linear subspace.**

Proof. If $x_n, y_n \in A$, then $\lim_{n\to\infty} x_n + \lim_{n\to\infty} y_n = \lim_{n\to\infty}(x_n+y_n) \in \bar{A}$. $\square$

Recall that finite dimensional subspaces are always closed. (exercise)

From the theory of Banach spaces we know that addition and scalar multiplication are continuous operations.

Proposition 1.5 $+, \lambda\cdot, \langle , \rangle$ **are continuous in each variable.**

Proof. If $\|x-x'\| < \delta$ and $\|y-y'\| < \delta$ then

$$\|(x+y)-(x'+y')\| \leqslant \|x-x'\| + \|y-y'\| < 2\delta.$$

Similarly,

$$\|\lambda(x-x')\| = |\lambda|\|x-x'\| < |\lambda|\delta.$$

We can consider $\phi_y := \langle y, \ \rangle$ to be a functional mapping $x \to \langle y, x\rangle$. This functional is continuous because $|\phi_y(x)| = |\langle y,x\rangle| \leqslant \|y\|\|x\|$. In fact $\|\phi_y\| = \|y\|$.

$\square$

It follows that taking limits commutes with the inner-product:

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

Do all norms on vector spaces come from inner-products, and if not, which normed vector spaces are in fact inner-product spaces? The answer is given by

Proposition 1.6 **(Parallelogram law)**
A norm comes from an inner-product if, and only if, it satisfies

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

Proof. Expanding and adding $\|x+y\|^2$ and $\|x-y\|^2$ gives the parallelogram law. Conversely, subtracting the two gives $4\operatorname{Re}\langle x, y\rangle$. Hence, by noticing that $\operatorname{Im}\langle x, y\rangle = -\operatorname{Re} i\langle x, y\rangle = \operatorname{Re}\langle ix, y\rangle$, we get the *polarization identity*

$$\langle x, y\rangle = \frac{1}{4}\left(\|y+x\|^2 - \|y-x\|^2 + i\|y+ix\|^2 - i\|y-ix\|^2\right).$$

The converse is true but harder to prove.

□

This law can be generalized further. If $\omega^N = 1$ $(N > 2)$, then $\langle x, y\rangle = \frac{1}{N}\sum_{n=1}^{N}\omega^n\|y+\omega^n x\|^2$. Even more generally, $\langle x, y\rangle = \frac{1}{2\pi i}\int_{S^1}\|y+zx\|^2\,dz$. That is, the inner product $\langle x, y\rangle$ is a 'complex' average of lengths on a ball of radius $\|x\|$, centred at y.

Proposition 1.7 $L^2(\mathbb{R})$ **is a Hilbert space with inner-product** $\langle f, g\rangle = \int \bar{f}g$**.**

Proof. We assume $L^2(\mathbb{R})$ is a complete normed space. Verifying the parallelogram law,

$$\int |f+g|^2 + \int |f-g|^2 = \int (|f|^2 + \bar{f}g + f\bar{g} + |g|^2 + |f|^2 - \bar{f}g - f\bar{g} + |g|^2.$$

Hence find the associated inner-product from the previous proposition.

□

1.1.1 Exercises

1. Find the angle between sin and cos in the space $L^2[-\pi, \pi]$.
2. Show that ℓ^2 is a Hilbert space with inner-product $\langle x, y\rangle = \sum_{n=1}^{\infty}\bar{x}_n y_n$. Show further that l^1 and l^∞ are not inner-product spaces by finding two sequences which do not satisfy the parallelogram law.
3. Similarly show that L^∞ and L^1 are not inner-product spaces.

4. Show that the Cauchy-Schwarz inequality becomes an equality if, and only if, $x = \alpha y \ \exists \alpha$.

5. Let $d = \inf_\lambda \|x + \lambda y\|$; show that
$$|\langle x, y\rangle|^2 \leqslant (\|x\|^2 - d^2)\|y\|^2.$$

6. Show that any finite-dimensional subspace of a Hilbert space is closed. Deduce that the set of polynomials of degree at most m forms a closed linear subspace of $L^2[a, b]$ with dimension $m + 1$; find a basis for this space.

7. Show that the set of continuous functions $C[a, b]$ is a non-closed linear subspace of $L^2[a, b]$.

8. Starting from a norm $\|.\|$ that satisfies the parallelogram law, we can create an inner-product using Prop. 6, from which we can define a norm $\|.\|_1 := \sqrt{\langle ., .\rangle}$. Show that the two norms are identical.

9. Show that two inner-products, on the same vector space, are conformal i.e. give the same angles between vectors, if, and only if, $\langle x, y\rangle_1 = \lambda\langle x, y\rangle_2$ for some positive real λ.

10. We have shown that an inner-product is determined by a norm that satisfies the parallelogram law. More generally, show that any bilinear form $A(x, y)$ that satisfies $A(y, x) = \overline{A(x, y)}$, is determined by its real quadratic form $q(x) = A(x, x)$, as $A(x, y) = \frac{1}{4}(q(x + y) + q(x - y) + iq(x + iy) - iq(x - iy))$.

11. Show that the real and imaginary parts of an inner-product give symmetric/anti-symmetric real-valued inner-products. Thus a complex inner-product has more structure than a real (symmetric) inner-product e.g. $\mathbb{C}^2$ has more structure than $\mathbb{R}^4$.

12. Let $H_1 \times H_2$ be the product of two inner-product spaces. Show that
$$\langle \, (x_1, x_2) \, , (y_1, y_2) \, \rangle := \langle x_1, y_1\rangle_{H_1} + \langle x_2, y_2\rangle_{H_2}$$
defines an inner-product on $H_1 \times H_2$.

13. Show that the formula $\langle f, g\rangle := \int \overline{f(x)} g(x) w(x) \, dx$, where $w(x)$ is a positive real function, defines an inner-product. The resulting space of functions is called a *weighted* L^2 space.

14. * Consider the vector space of holomorphic functions $f : \mathbb{R} \times \mathbb{R}^+ \to \mathbb{C}$ which satisfy $\sup_{y>0} \int |f|^2 \, dx < \infty$. Show that the formula $\langle f, g \rangle := \lim_{\epsilon \to 0^+} \int \overline{f(x+i\epsilon)} g(x+i\epsilon) \, dx$ gives an inner-product.

1.2 Orthogonality

Definition The **orthogonal space** of a set A is the set

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

1.2.1 Exercises

1. Show that $0^\perp = X$, $X^\perp = 0$.
2. Show that if $\{ x \}^\perp = X$ then $x = 0$.
3. Show that if $\{ x \}^\perp = 0$ then X is one-dimensional.

Proposition 1.8 **$A^\perp$ is a closed linear subspace.**

Proof. That $A^\perp$ is a linear subspace follows from the linearity of the inner-product. Let $x \in \overline{A^\perp}$. That is, there is a sequence of vectors $x_n \in A^\perp$ such that $x_n \to x$. Now, for any $y \in A$, $\langle x, y \rangle = \langle \lim_n x_n, y \rangle = \lim_n \langle x_n, y \rangle = 0$. Hence $x \in A^\perp$.

□

Proposition 1.9

1. $A \cap A^\perp \subseteq 0$**;**

2. $A \subseteq B \ \Rightarrow \ B^\perp$ **is a closed subspace of** $A^\perp$**;**

3. $A \subseteq A^{\perp\perp}$**.**

Proof. (i) is left as an easy exercise. For (ii), let $x \in B^\perp$ i.e. $\langle b, x \rangle = 0 \ \forall b \in B$. In particular this is true for $b \in A$ so that $x \in A^\perp$. For (iii), let $x \in A$, then $\langle y, x \rangle = 0, \forall y \in A^\perp$. Hence $\langle x, y \rangle = 0 \ \forall y \in A^\perp$. Hence $x \in A^{\perp\perp}$.
□

Note that $A^{\perp\perp}$ is always a closed linear subspace even if A isn't. Question: if M is a closed linear subspace is it necessarily true that $M = M^{\perp\perp}$?

1.3 Least Distance in Hilbert spaces

Definition A **convex** set is a set A that includes all line segments between its points:

$$\forall x, y \in A, \forall t \in [0,1], \quad tx + (1-t)y \in A.$$

Exercise: Show that linear subspaces and balls are convex.

Theorem 1.10 **If M is a closed convex set and x is any point of a Hilbert space H, then there is a unique point $x_0 \in M$ which is closest to x.**

$$\forall x \in H\ \exists! x_0 \in M : \quad d(x, x_0) \leqslant d(x, y)\ \forall y \in M.$$

Proof. Let d be the smallest distance from M to x i.e. $d = \inf_{y \in M} d(x, y)$ (how do we know that d exists?) Then there is a sequence of vectors $y_n \in M$ such that $\|x - y_n\| = d(x, y_n) \to d$. Now, by the parallelogram law,

$$\begin{aligned}
\|y_n - y_m\|^2 &= 2\|y_n - x\|^2 + 2\|y_m - x\|^2 - \|(y_n + y_m) - 2x\|^2 \\
&\leqslant 2\|y_n - x\|^2 + 2\|y_m - x\|^2 - 4d^2 \\
&\to 0
\end{aligned}$$

as $n \to \infty$. Hence (y_n) is a Cauchy sequence. But M is closed (hence complete) and so $y_n \to x_0 \in M$. It follows, by continuity of the norm, that $\|x - x_0\| = d$.

Suppose $y \in M$ is another closest point to x, i.e. $d(x, y) = d$. Then

$$\begin{aligned}
\|x_0 - y\|^2 &= 2\|x_0 - x\|^2 + 2\|y - x\|^2 - \|(x_0 + y) - 2x\|^2 \\
&\leqslant 2\|x_0 - x\|^2 + 2\|y - x\|^2 - 4d^2 \\
&= 0
\end{aligned}$$

□

Note that this theorem, which does not refer to inner-products, is *not* true in Banach spaces in general. For example l^∞ has a closed subspace c_0, but there is no closest sequence in c_0 to the sequence $(1, 1, 1, \ldots)$. In fact, the smallest distance is $d = 1$ and this is achieved by any bounded sequence (a_n) with $0 \leqslant a_n \leqslant 2$.

Exercise: prove these assertions.

Let us concentrate on the case when M is a closed subspace of H.

Theorem 1.11 **For a closed linear subspace M of a Hilbert space H, $x_0 \in M$ is the closest point to $x \in H$ if, and only if, $x - x_0 \in M^\perp$.**

Proof. Let a be any point of M. Then $x - x_0 = \lambda a + z$ with $a \perp z$, where $\lambda = \langle a, x - x_0\rangle / \|a\|^2$ and $z = x - (x_0 + \lambda a)$. By Pythagoras' theorem, we get $\|z\| \leqslant \|x - x_0\|$ making $x_0 + \lambda a$ even closer to x than the closest point x_0. This is only possible if $\lambda = 0$ i.e. $\langle a, x - x_0\rangle = 0$. Since a is arbitrary we get $x - x_0 \perp M$.

Conversely, if $x - x_0 \perp a$ for any $a \in M$, then by Pythagoras' theorem, $\|x - a\|^2 = \|x - x_0\|^2 + \|x_0 - a\|^2$, so that $\|x - x_0\| \leqslant \|x - a\|$, making x_0 the closest point to x.

□

Corollary **If M is a closed linear subspace of a Hilbert space H, then $H = M \oplus M^\perp$.**

Proof. Write $x = x_0 + (x - x_0)$. These two components are in M and $M^\perp$ as proved previously. Moreover $M \cap M^\perp = 0$.

□

Note that this is false for Banach spaces e.g. the space $l^\infty \neq c_0 \oplus M$ for any linear subspace M.

Corollary **If M is a closed linear subspace of a Hilbert space H, then $M^{\perp\perp} = M$. More generally, for any set A, $A^{\perp\perp} = \overline{[\![A]\!]}$.**

Proof. Let $x \in M^{\perp\perp}$. Then $x = a + b$ where $a \in M$ and $b \in M^\perp$. Then $0 = \langle b, x\rangle = \langle b, a\rangle + \langle b, b\rangle = \|b\|^2$, making $b = 0$ and $x \in M$. For the second part, note that $\overline{[\![A]\!]}$ is the smallest closed linear subspace containing A. Therefore, since $A \subseteq \overline{[\![A]\!]}$, we get $\overline{[\![A]\!]}^\perp \subseteq A^\perp$ and hence $A^{\perp\perp} \subseteq \overline{[\![A]\!]}^{\perp\perp} = \overline{[\![A]\!]}$.
□

Corollary **$[\![A]\!]$ is dense in H if, and only if, $A^\perp = 0$.**

Proof. Suppose $[\![\bar{A}]\!] = H$. Then $A^\perp = A^{\perp\perp\perp} = H^\perp = 0$, since $A^\perp$ is a closed linear subspace. Conversely, $[\![\bar{A}]\!] = A^{\perp\perp} = 0^\perp = H$.

□

1.3.1 Exercises

1. Find the closest points x_0 to a closed linear subspace M from (i) a point $x \in M$, (ii) a point $x \in M^{\perp}$.

2. In the decomposition $x = a + b$ with $a \in M$ and $b \in M^{\perp}$, show that a and b are unique.

3. Deduce that if $H = M \oplus N$ where M is a closed linear subspace and $M \perp N$ then $N = M^{\perp}$.

4. Show that, if M is a closed linear subspace of N, then $M \oplus (M^{\perp} \cap N) = N$.

1.3.2 Projections

Let us take a closer look at the map from x to $x_0 \in M$.

Definition A **projection** is a linear map $P : H \to H$ such that $P^2 = P$. An **orthogonal** projection is one such that $\ker P \perp \mathrm{im} P$.

Theorem 1.12 **If M is a closed linear subspace of a Hilbert space H then the map $P : H \to H$ defined by $Px = x_0$ is a continuous orthogonal projection with im$P = M$ and $\ker P = M^{\perp}$.**

Proof. By the definition of P, for any $a \in H$, $(a - Pa) \in M^{\perp}$ and this property defines $Pa \in M$.

P is linear since $(x + y) - (Px + Py) = (x - Px) + (y - Py) \in M^{\perp}$ and $Px + Py \in M$, hence $P(x + y) = Px + Py$. Similarly $P(\lambda x) = \lambda Px$.

P is a projection since $Px \in M$ implies that $P^2 x = Px$.

P is onto M since for any $x \in M$, $Px = x$. Moreover $x \in \ker P \Leftrightarrow Px = 0 \Leftrightarrow x = x - Px \in M^{\perp}$.

P is continuous since $\|x\|^2 = \|x - Px\|^2 + \|Px\|^2$ by Pythagoras' theorem so that $\|Px\| \leqslant \|x\|$.

□

1.3.3 Exercises

1. Show that the map $(x, y, z) \mapsto (x, 0, 0)$ is an orthogonal projection in $\mathbb{R}^3$. Find a projection which is not orthogonal.

2. Let P, Q be two commuting projections. Show that PQ is also a projection.

3. Show that if P is a projection then $H = \ker P \oplus \operatorname{im} P$.

4. Show that any orthogonal projection is continuous, and $\operatorname{im} P \perp \ker P$. Hence note that there is a 1-1 correspondence between closed linear subspaces and orthogonal projections.

5. Show that if P is an orthogonal projection onto M then $I - P$ is an orthogonal projection onto $M^\perp$.

6. Suppose that $H = M \oplus N$. Show that, if P and Q denote the orthogonal projections onto M and N respectively, then $I = P + Q$ and $PQ = 0$. Extend this by induction to the case when $H = M_1 \oplus \ldots \oplus M_n$ to get $I = P_1 + \ldots P_n$ with $P_i P_j = P_i \delta_{ij}$.

7. Consider two closed linear subspaces M_1 and M_2 (with associated projections P_1 and P_2). Show that the iteration $x_{n+1} = P_2 P_1 x_n$ starting from $x_0 = y$ converges to the closest point $y_0 \in M_1 \cap M_2$.

8. Let $M =< y >$ (a closed linear subspace). Find the orthogonal projection P which maps any point x to its closest point in M. (Ans: $Px = \langle y, x \rangle y$)

9. Let $M = \langle y_1, \ldots, y_N \rangle$. Find P again. (Ans: $Px = \sum_i \alpha_i y_i$ where $\alpha_i = \langle y_i, x \rangle / \|y_i\|^2$.

1.3.4 Examples: Least Squares Approximation

1. Find the closest point in the plane $2x + y - 3z = 0$ to the point $\boldsymbol{x} = (5, 2, 0)$. Letting $\boldsymbol{n} = (2, 1, -3)$, the equation of the plane is given by $\boldsymbol{n} \cdot \boldsymbol{x} = 0$. Therefore we are looking for a point $\boldsymbol{x}_0$ in the plane such that $\boldsymbol{x} - \boldsymbol{x}_0$ is perpendicular to the plane i.e. is parallel to $\boldsymbol{n}$. Hence, $\boldsymbol{x}_0 = \boldsymbol{x} + \lambda \boldsymbol{n}$ such that it satisfies $\boldsymbol{n} \cdot \boldsymbol{x}_0 = 0$. Substituting gives λ.

2. Find the best-fitting (closest in the space L^2) cubic polynomial to the function sin in the region $[0, 2\pi]$.

   The space of cubic polynomials, $a + bx + cx^2 + dx^3$, is a four-dimensional closed linear subspace of the Hilbert space $L^2[0, 2\pi]$, with basis $1, x, x^2, x^3$. It must therefore have an element $p(x) = a + bx + cx^2 + dx^3$ which is closest to $\sin x$. In fact, p is characterized by the condition $p - \sin \perp q, \forall q \in M$. In particular we get four equations $\langle p(x) - \sin(x), x^i \rangle = 0$ for $i = 0, \ldots, 3$, or, equivalently, $\langle p(x), x^i \rangle = \langle \sin(x), x^i \rangle$. The right-hand sides can be worked out, giving four linear equations in the four unknowns a, b, c, d, which can be extracted by solving.

3. Finding the best-fitting (least-squares) line to a number N of points $(a_n, b_n) \in \mathbb{R}^2$.

   One can consider such a sequence of N pairs to be a single point $\boldsymbol{b} = (b_1, \ldots, b_N)$ in $\mathbb{R}^N$ which is a Hilbert space. We wish to find a line $y = mx + c$ for which $\boldsymbol{y} = (y_n)$ where $y_n = ma_n + c$) are collectively as close to b_n as possible. These lines form a closed subspace M spanned by the two vectors $\boldsymbol{e}_1 = (a_1, \ldots, a_N)$ and $\boldsymbol{e}_2 = (1, \ldots, 1)$ (check). Hence the condition $\langle \boldsymbol{y} - \boldsymbol{b}, M\rangle = 0$ becomes two equations $\langle \boldsymbol{y} - \boldsymbol{b}, \boldsymbol{e}_i\rangle = 0$, which translated becomes,

   $$m\sum_n a_n + c\sum_n 1 = \sum_n b_n,$$

   $$m\sum_n a_n^2 + c\sum_n a_n = \sum_n b_n a_n.$$

   Solving for m and c gives the usual regression line as used in statistics.

1.3.5 Exercises

1. Find that quadratic polynomial that is closest to $\sin x$ in $L^2[0, 2\pi]$.

2. Find that combination of sin and cos which is closest to $1 - x^3$ in $L^2[0, 1]$.

3. Implement the above examples as Mathematica programs.

4. Find, and implement, the best-fitting *circle* to a number N of points in $\mathbb{R}^2$.

2 Orthonormal Bases

Definition A **Hilbert basis**, also called a **total** set of vectors, is a set of vectors $\{e_n\}$ such that

$$\langle e_n \rangle \text{ is dense in } X,$$
$$\langle e_n, e_m \rangle = \delta_{nm}.$$

Note that the first condition is equivalent to $\langle e_n, x\rangle = 0 \,\forall n \;\Rightarrow\; x = 0$. Note also that the set of basis vectors need not be countable; when it is, the Hilbert space is called separable. From now on we will assume that the Hilbert basis is countable.

2.0.6 Example

The space of sequences ℓ^2 has a Hilbert basis consisting of $\boldsymbol{e}_n = (0, \ldots, 0, 1, 0, \ldots)$.

To prove the first condition, let $\boldsymbol{x} = (x_1, x_2, \ldots)$ be a sequence in ℓ^2 which is orthogonal to all the $\boldsymbol{e}_i$. Then for any i, $x_i = \langle \boldsymbol{e}_i, \boldsymbol{x}\rangle = 0$; hence $\boldsymbol{x} = 0$.

2.0.7 Gram-Schmidt orthogonalization

Of the two properties, it is the first one that is crucial; if the span of a countable number of vectors $\{a_n\}$ is dense in X but not orthonormal, then they can be made so using the usual Gram-Schmidt process:

$$\begin{array}{ll} b_0 := a_0, & e_0 := b_0/\|b_0\| \\ b_n := a_n - \sum_{i=0}^{n-1} \langle e_i, a_n\rangle, & e_n := b_n/\|b_n\|. \end{array}$$

2.1 Fourier Expansion

Proposition 2.1 **Let M be a closed linear subspace M, with a countable Hilbert basis e_n;**

$$(a_n) \in \ell^2 \;\Leftrightarrow\; \sum_{n=1}^{\infty} a_n e_n \textbf{ converges in M}.$$

Proof. Let $x_N := \sum_{n=1}^N a_n e_n$. Then, assuming $N > M$ wolog,

$$\begin{aligned} \|x_N - x_M\|^2 &= \langle x_N - x_M, x_N - x_M\rangle \\ &= \langle \textstyle\sum_{n=M+1}^N a_n e_n, \sum_{m=M+1}^N a_m em\rangle \\ &= \textstyle\sum_{n,m=M+1}^N \overline{a}_n a_m \langle e_n, e_m\rangle \\ &= \textstyle\sum_{n=M+1}^N |a_n|^2. \end{aligned}$$

Suppose that $(a_n) \in \ell^2$; then this last sum converges to 0, implying that (x_n) is a Cauchy sequence in M, which must therefore converge to a point $x \in M$. Conversely, suppose that the series x_N converges; then $\|x_N - x_M\| \to 0$, implying that $\sum_{n=1}^{N} |a_n|^2$ is a Cauchy sequence in $\mathbb{C}$, and hence must converge as $N \to \infty$.

$\square$

Theorem 2.2 **(Bessel's inequality)**
If a closed linear subspace M of a Hilbert space, has a countable Hilbert basis e_n, then

$$(i) \quad \sum_n |\langle e_n, x\rangle|^2 \leqslant \|x\|^2,$$

$$(ii) \quad Px = x_0 = \sum_n \langle e_n, x\rangle e_n.$$

Proof. Let $x_N := \sum_{n=1}^{N} \langle e_n, x\rangle e_n$. Then $x_N \in M$ and, letting $a_n = \langle e_n, x\rangle$, we get

$$\begin{aligned} 0 \leqslant \|x - x_N\|^2 &= \langle x - x_N, x - x_N\rangle \\ &= \|x\|^2 - \langle x_N, x\rangle - \langle x, x_N\rangle + \langle x_N, x_N\rangle \\ &= \|x\|^2 - 2\textstyle\sum_{n=1}^{N} \overline{a}_n a_n + \sum_{n,m=1}^{N} \overline{a}_n a_m \langle e_n, e_m\rangle \\ &= \|x\|^2 - \textstyle\sum_{n=1}^{N} |a_n|^2, \end{aligned}$$

hence

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

is an increasing series, bounded above by $\|x\|^2$. Hence the sum on the left-hand side must converge, proving the Bessel inequality.

By the previous proposition, the series of vectors $\sum_{n=1}^{N} \langle e_n, x\rangle e_n$ converges, say to $y \in M$.

Moreover,

$$\forall m \ \langle e_m, x - y\rangle = \langle e_m, x\rangle - \textstyle\sum_{n=1}^{\infty} \langle e_n, x\rangle\langle e_m, e_n\rangle = 0,$$

so that $x - y \in \{ e_m \}^\perp = M^\perp$. Hence y must be the closest point in M to x.

$\square$

Corollary **(Parseval's identity)**
If e_n is a countable Hilbert basis for X, then

$$\forall x \in X \qquad x = \sum_{n=1}^{\infty} \langle e_n, x\rangle e_n,$$

and

$$\langle x, y\rangle = \sum_n \overline{\alpha_n}\beta_n.$$

where $x = \sum_n \alpha_n e_n$ **and** $y = \sum_n \beta_n e_n$**. In particular**

$$\|x\| = (\sum_n |\alpha_n|^2)^{1/2}.$$

Proof. The first part is immediate from the theorem since $Px = x$ when $M = X$. The second part is a simple expansion of the two series in the inner-product making essential use of the linearity and continuity of $\langle,\rangle$.

$\square$

Proposition 2.3 **If X has a countable Hilbert basis, then $X \simeq \ell^2$.**

Proof. Let e_n be a countable Hilbert basis. Consider the map

$$\begin{aligned} J : X &\to \ell^2 \\ x &\mapsto (\alpha_n) = (\langle e_n, x\rangle) \end{aligned}$$

Then Bessel's inequality shows that the resulting sequence is truly in ℓ^2. Linearity is easily checked. The Parseval identity is precisely the fact that the inner-products are preserved i.e. $\langle x, y\rangle_X = \langle Jx, Jy\rangle_{\ell^2}$. Note that isometries must be 1-1 (their kernel must be trivial). J is onto ℓ^2 since it maps $\sum_n \alpha_n e_n$ to (α_n) and the former is a well-defined vector for $(\alpha_n) \in \ell^2$.

$\square$

2.1.1 Solving Linear Equations

One of the earliest methods for solving a linear equation $Ax = b$, for x, b in a function space, and for A a continuous linear operator, was to write the unknown element x and the known element b in terms of a Hilbert basis e_n to get $x = \sum_n x_n e_n$ and $b = \sum_n b_n e_n$ where $x_n = \langle e_n, x\rangle$ are unknown coefficients and $b_n = \langle e_n, b\rangle$ are known coefficients. Substituting into $Ax = b$ we get

$$\sum_n x_n A e_n = \sum_n b_n e_n.$$

Moreover we can expand the vectors $Ae_n = \sum_m A_{nm} e_m$ and, comparing coefficients, we get

$$\sum_{n,m} A_{nm} x_n = b_n.$$

This can be thought of as a matrix equation in ℓ^2 with the matrix $[A_{nm}]$ having a countable number of rows and columns. Effectively, we have transferred the problem from one on H to one on ℓ^2, via the map J.

If the Hilbert basis elements e_n are chosen to be *eigenvectors* of A, then the equation simplifies because of $Ae_n = \lambda_n e_n$; this gives $\lambda_n x_n = b_n$. When $\lambda_n \neq 0$, we must choose $x_n = b_n/\lambda_n$; when $\lambda_n = 0$ (i.e. the homogeneous equation $Ax = 0$ has solutions), then we get $0 = b_n = \langle e_n, b \rangle$; if this is false, then there are no solutions, otherwise we are free to choose x_n arbitrarily. Thus there will be a solution if, and only if, $b \perp \ker A$. In this case the eigenvectors of the zero eigenvalue $\{ e_m \}$ will span the homogeneous solutions, and the complete solution will be

$$x = \sum_m a_m e_m + \sum_n (b_n/\lambda_n) e_n,$$

where the a_m are arbitrary constants. Note that the latter part is the "particular solution" and, for the case of L^2, can be rewritten as

$$\sum_n (b_n/\lambda_n) e_n = \sum_n \langle e_n, b \rangle e_n/\lambda_n = \int \left(\sum_n \frac{1}{\lambda_n} \overline{e_n(s)} e_n(x) \right) f(s)\, ds$$

equivalent to the Green's function formulation of the particular solution.

2.1.2 Applications

Hilbert bases are widely used to approximate functions. The first N "large" coefficients can be used to "store" the function in a useful compressed way. Regenerating the function, or manipulating functions is easily done using the Parseval identity. This has been used in JPEG and MPEG compression, as well as in compressing images for Microsoft Encarta, and by the FBI to store millions of fingerprints for rapid retrieval.

Such bases are also used to filter out noise or pick out particular features in a function. First expand the function in an appropriate basis, then remove those coefficients which are smaller than a given threshold. Regenerate the function from the remaining coefficients.

2.2 Hilbert Bases for L^2.

There are various Hilbert bases suitable for the space of L^2 functions on different domains. Each basis has particular properties that can be utilized in specific contexts. One should treat these the same way that we treat bases in finite dimensional linear algebra. They are indispensable for actual calculations, but one has to be careful which basis to choose that makes the problem amenable. For example, for a problem that has spherical symmetry, it would make sense to use a Hilbert basis adapted to spherical symmetry.

2.2.1 $L^2[a,b]$: Fourier series

Theorem 2.4 **The functions $\frac{1}{\sqrt{2\pi}}e^{inx}$ for $n \in \mathbb{Z}$ form a countable Hilbert basis for $L^2[-\pi,\pi]$.**

It follows that $\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}}\cos(x), \frac{1}{\sqrt{\pi}}\sin(x), \ldots$ is a countable Hilbert basis.

Proof (NOT FOR EXAM) Suppose that $f \in \{\, e^{inx} \,\}^\perp$, i.e. $\int_{-\pi}^{\pi} e^{-iny} f(y)\, dy = 0, \forall n \in \mathbb{Z}$. Multiplying by e^{inx} gives $\int e^{in(x-y)} f(y)\, dy = 0$ and, changing variables, becomes $\int e^{-iny} f(x+y)\, dy = 0$. Now $(\cos(y/2))^{2n} = (\frac{1}{2}e^{iny/2} + \frac{1}{2}e^{-iny/2})^{2n}$ is a linear combination of exponentials of various frequencies, so that we have

$$\int_{-\pi}^{\pi} (\cos(y/2))^{2n} f(x+y)\, dy = 0.$$

But $\int_{-\pi}^{\pi} (\cos(y/2))^{2n}\, dy = 2/\pi(n+1)$ (use a reduction formula), so that we get

$$f(x) = \frac{\pi(n+1)}{2} \int_{-\pi}^{\pi} (\cos(y/2))^{2n} (f(x) - f(x+y))\, dy.$$

Therefore,

$$\begin{aligned} |f(x)| &\leqslant \tfrac{\pi(n+1)}{2} \int_{-\pi}^{\pi} (\cos(y/2))^{2n} |f(x) - f(x+y)|\, dy \\ &\leqslant \tfrac{\pi(n+1)}{2} \|(\cos(y/2))^{2n}\|_{L^2[-\pi,\pi]} \|f(x) - f(x+y)\|_{L^2[-\pi,\pi]} \\ &\leqslant c\|f(x) - f(x+y)\|_{L^2[-\pi,\pi]} \end{aligned}$$

It is a theorem from Lebesgue integration (but probably not-covered in the introductory course in the third year), that L^2 functions are L^2-continuous i.e. $\|f(x) - f(x+y)\|_{L^2[a,b]} \to 0$ as $y \to 0$. Hence the function f must be zero.

Orthonormality of the functions follows from

$$\langle e^{inx}, e^{imx} \rangle = \int_{-\pi}^{\pi} e^{ix(m-n)}\, dx = 2\pi\delta_{nm}.$$

□

Note that there is nothing special about the interval $[-\pi,\pi]$. Any other interval $[a,b]$ will do, except that the basis functions have to be modified accordingly. For example, $e^{2\pi inx}$ is a Hilbert basis for $L^2[0,1]$.

Corollary **The polynomials are dense in $L^2[a, b]$.**

Proof. The exponential functions can be approximated in $L^2[a, b]$ as closely as needed by Taylor polynomials. This means that the set of polynomials is dense in the set of exponential functions written above which in turn are dense in L^2.

□

The Parseval identity for this Hilbert basis becomes

$$f(x) = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \langle e^{inx}, f(x) \rangle e^{inx} = \sum_{n=-\infty}^{\infty} \alpha_n e^{inx}$$

where $\alpha_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-inx} f(x)\, dx$, thus giving the usual Fourier series expansion. This can also be written in terms of the basis consisting of cosines and sines. Notice that this equation holds if, and only if, the coefficients α_n are in ℓ^2. The classical Parseval identity is none other than

$$\int_{-\pi}^{\pi} |f(x)|^2\, dx = \|f\|^2_{L^2[-\pi,\pi]} = \|(\alpha_n)\|^2_{\ell^2} = \sum_n |\alpha_n|^2.$$

Moreover the Fourier operator $\mathcal{F}$ which sends functions in L^2 to sequences of Fourier coefficients (in ℓ^2) is a continuous isometry between Hilbert spaces.

This Fourier basis has the special property that they are eigenvectors of the differential operator:

$$De^{inx} = ine^{inx}.$$

It can therefore be used to turn a constant-coefficient differential equation into a polynomial equation as follows: expanding the unknown function $y(x) = \sum_n c_n e^{inx}$ and the known function $f(x) = \sum_n b_n e^{inx}$ turns the differential equation

$$(a_m D^m + \ldots + a_0)y = f$$

into the series

$$\sum_n (a_m (in)^m + \ldots + a_0) c_n e^{inx} = \sum b_n e^{inx}.$$

Comparing coefficients of the nth term gives

$$(a_m (in)^m + \ldots + a_0) c_n = b_n,$$

which can be solved to give the coefficients c_n in terms of the known coefficients b_n. Of course this method assumes that both the solution y and the given function f are in $L^2[-\pi, \pi]$, which need not be the case. Any other solution will not be found this way.

2.2.2 $L^2[-1,1]$: Legendre polynomials

We have just found that the set of polynomials is dense in $L^2[a,b]$ but they turn out to be non-orthogonal, as can be easily verified by calculating $\langle 1, x^2\rangle$. However we can make them orthonormal using the Gram-Schmidt process. On the interval $[-1,1]$, the resulting polynomials are called the **Legendre** polynomials.

The first few polynomials are $\frac{1}{\sqrt{2}}$, $\sqrt{\frac{3}{2}}x$, $\frac{3}{2}\sqrt{\frac{5}{2}}(x^2-\frac{1}{3})$, etc. The general formula is (!)

$$P_n(x) = \frac{\sqrt{n+\frac{1}{2}}}{2^n n!}\sum_{k=0}^{\mathrm{int}(n/2)} \frac{(-1)^k(2n-2k)!}{(n-2k)!}\binom{n}{k} x^{n-2k} = \frac{\sqrt{n+\frac{1}{2}}}{2^n n!}(\frac{d}{dx})^n(x^2-1)^n.$$

The Legendre functions are eigenvectors of $T = D(x^2-1)D = (x^2-1)D^2 + 2xD$:

$$Tp_n = n(n+1)p_n.$$

2.2.3 $L^2[0,\infty)$: Laguerre functions

This Hilbert space does *not* contain any polynomials x^n, but we can modify them to $x^n e^{-x/2}$ which do belong. A Gram-Schmidt orthonormalization gives the **Laguerre** functions. The first few terms are $e^{-x/2}$, $(1-x)e^{-x/2}$, $(1-2x+\frac{1}{2}x^2)e^{-x/2}$, etc. The general formula is

$$l_n(x) = \frac{1}{n!}\sum_{k=0}^{n}(-1)^k\frac{1}{k!}\binom{n}{k} x^k e^{-x/2} = \frac{1}{n!}e^{x/2}D^n(x^n e^{-x}).$$

The Laguerre function are eigenvectors of $S = xD^2 + D - x/4$

$$Sl_n = -(n+\frac{1}{2})l_n.$$

2.2.4 $L^2(\mathbb{R})$: Hermite functions

This time we perform orthonormalization starting with the functions $x^n e^{-x^2/2}$. We get the **Hermite** functions, $\frac{1}{\sqrt[4]{\pi}}e^{-x^2/2}$, $\frac{2}{\sqrt{2\sqrt{\pi}}}xe^{-x^2/2}$, etc.

$$h_n(x) = \frac{\sqrt{2^n}}{\sqrt{n!\sqrt{\pi}}}\sum_{k=0}^{\mathrm{int}(n/2)}\frac{(-1)^k(2k)!}{4^k k!}\binom{n}{2k} x^{n-2k}e^{-x^2/2} = \frac{(-1)^n}{\sqrt{2^n n!\sqrt{\pi}}}e^{x^2/2}D^n e^{-x^2}.$$

The Hermite functions are eigenvectors of $R = D^2 - x^2$:

$$Rh_n = -(2n+1)h_n.$$

2.2.5 $L^2(A)$

There are many other orthonormal Hilbert bases adapted to specific sets A or weights. The Jacobi functions, the Chebychev (on the circle), the modified Bessel functions on $L^2(0,\infty)$, the spherical harmonics on the sphere ($L^2(S^2)$) etc. It is a theorem of Rodriguez that the functions $f_n(x) = w(x)^{-\frac{1}{2}}D^n(w(x)p(x)^n)$ for any polynomial p and weight function $w \in L^2(A)$ are orthogonal: the Legendre, Laguerre, Hermite, Jacobi, Chebychev functions are all of this type.

2.2.6 Exercises

1. Suppose that e_n and $\tilde{e}_m$ form a Hilbert basis for X and Y respectively. Show that $(e_n, \tilde{e}_m)$ form a Hilbert basis for $X \times Y$.

2. Write down the Parseval identity for the basis $\frac{1}{\sqrt{2\pi}}$, $\frac{1}{\sqrt{\pi}}\cos(nx)$, $\frac{1}{\sqrt{\pi}}\sin(nx)$.

3. Show that the Fourier operator $\mathcal{F}$ sends functions in $L^1[-\pi,\pi]$ to ℓ^∞ continuously.

4. This exercise is to prove that the Legendre polynomials are orthonormal. Define $u_n(x) = (x^2-1)^n$. Note that $D^k u_n(\pm 1) = 0$. Show that $\langle D^n u_n, D^m u_m\rangle = -\langle D^{n-1}u_n, D^{m+1}u_m\rangle$ and hence show that this inner-product must be zero unless $n = m$.

5. (extra) Prove that the Legendre, Laguerre and Hermite functions are indeed eigenvectors of the corresponding differential operators.

6. Find all $L^2[-1,1]$ solutions of the equation $(x^2-1)y'' + 2xy' + y = \sin(\pi x)$.

7. Solve $y^{(4)} - 2x^2y'' - 4xy' + (x^4-2)y = g$ in series form, for functions in $L^2(\mathbb{R})$. (Hint: expand R^2)

8. Let u be a solution of $\partial_t u = \partial_x^2 u$ on $x \in [0,1]$ with $u(0) = u_0$; show that $u(t,x) = \sum_n \alpha_n(t)e_n(x)$ (justifies the separation of variables method).

9. Show that the Bessel inequality is still valid even if the orthonormal set of vectors e_n is not countable; deduce that $\langle e_n, x\rangle = 0$ except for a countable number of e_n.

10. Suppose that e_n are a set of vectors such that $\|\langle e_n, x\rangle\|_{\ell^2} = \|x\|$ for all $x \in X$. Show that the vectors must be dense in X and orthogonal.

2.2.7 Frequency-Time Hilbert Bases

A recent development in Hilbert bases are those bases for functions $f(t)$ that give information in both frequency and time. In contrast the coefficients that result from the Fourier operator, for example, only give information about the frequency content of the function. A large nth coefficient means that there is a substantial amount of the term e^{inx} i.e. of frequency n, *somewhere* in the function $f(x)$. The aim of frequency-time bases is to have coefficients a_{mn} depending on two parameters m and n, one of which is a frequency index, the other a "time" index.

Windowed Fourier Bases
The simplest way to achieve this is to define the basis functions by

$$h_{m,n}(x) = e^{2\pi inx}h(x-m),$$

where h is a carefully chosen function, with $\|h\|_{L^2} = 1$, such that h_{mn} are orthonormal. The most common choices are the *window*-function $h = \chi_{[0,1]}$ and the gaussian $h(x) = e^{-x^2/2}$. The m gives position (time) information, while the n gives frequency information.

Note that summing the coefficients in n gives the windowed function:

$$\begin{aligned}\sum_n |\langle h_{mn}, f\rangle|^2 &= \sum_n |\langle e^{2\pi inx}h(x-m), f(x)\rangle|^2\\ &= \sum_n |\langle e^{2\pi inx}, h(x-m)f(x)\rangle|^2\\ &= \|h(x-m)f(x)\|_{L^2}\\ &= |\langle h(x-m), |f(x)|\rangle|^2\end{aligned}$$

Similarly summing the coefficients in m gives a windowed fourier transform of the function.

Wavelet Bases
The basis in this case consists of the following functions in $L^2[0,1]$

$$\psi_{mn}(x) := 2^{m/2}\psi(2^m x - n),$$

together with $\phi(x) = \chi_{[0,1]}$. The classical Haar basis is generated by the case $\psi(x) = \chi_{[0,\frac{1}{2}]} - \chi_{[\frac{1}{2},1]}$. More recently, wavelets generated by a continuous function have been used. Such functions are necessarily nowhere differentiable.

2.2.8 Other Bases

There are many other bases used specifically for compression etc. A common one is the Walsh basis which consists of step functions that are the normalized column vectors of the matrix M_N where

$$M_1 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \qquad M_{n+1} = \begin{pmatrix} M_n & M_n \\ M_n & -M_n \end{pmatrix}.$$

2.2.9 Exercises

1. Prove that the windowed Fourier basis with the window function $h(x) = \chi_{[0,1]}(x)$ are orthonormal.

2. Show that the Haar basis is orthonormal.

3. Look up information on applications of wavelet and other Hilbert bases.

3 Dual Spaces

3.1 $X^* \equiv X$

In any inner-product space, every vector x gives rise to a continuous functional

$$x^* = \langle x, \ \rangle : \begin{array}{rcl} X & \to & \mathbb{C} \\ y & \mapsto & \langle x, y \rangle \end{array}$$

This is indeed linear, while continuity follows from the Cauchy-Schwarz inequality $|x^*(y)| = |\langle x, y\rangle| \leqslant \|x\|\|y\|$.

Are there any other functionals besides these? In normed spaces this is generally the case e.g. for the Banach space of continuous bounded functions $C[a,b]$, the dual space is $L^1[a,b]$ which contains many other functions besides the continuous ones. In Hilbert spaces however this is not the case:

Theorem 3.1 **(Riesz' theorem)**
Every continuous functional of X is of the form $\langle x, \ \rangle$ for some unique vector x i.e.

$$\forall \phi \in X^* \ \exists! x \in X \quad \phi = \langle x, \ \rangle.$$

Proof. First notice that for any z and y, $\phi(y)z - \phi(z)y \in \ker \phi$. Assuming $\phi \neq 0$, pick $z \perp \ker \phi$, non-zero, to get

$$0 = \langle z, \phi(y)z - \phi(z)y \rangle = \phi(y)\|z\|^2 - \phi(z)\langle z, y\rangle.$$

Hence

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

where $x = \overline{\phi(z)}z/\|z\|^2$. To show that it is unique, suppose $\tilde{x}$ is another such x then,

$$\langle x - \tilde{x}, y\rangle = \langle x, y\rangle - \langle \tilde{x}, y\rangle = \phi(y) - \phi(y) = 0, \ \forall y$$

so that $x = \tilde{x}$.

$\square$

Let us study this map, called the Riesz map, from X to X^* more closely:

Proposition 3.2 **The map**

$$\begin{aligned} J : X &\to X^* \\ x &\mapsto x^* \end{aligned}$$

is a bijective conjugate-linear isometry between X and X^*.

Proof. Let x and y be two vectors in X. Then $J(x+y) = (x+y)^* = x^*+y^* = Jx+Jy$ since $(x+y)^*(z) = \langle x+y, z\rangle = \langle x, z\rangle+\langle y, z\rangle = x^*(z)+y^*(z)$ for any $z \in X$. Similarly $J(\lambda x) = \bar{\lambda}Jx$ since $(\lambda x)^*(z) = \langle \lambda x, z\rangle = \bar{\lambda}\langle x, z\rangle = \bar{\lambda}x^*(z)$.

To show that J is isometric, for any $y \in X$,

$$\|x^*\|_{X^*} = \sup_y \frac{|x^*(y)|}{\|y\|} = \sup_y \frac{|\langle x, y\rangle|}{\|y\|} = \|x\|,$$

using the Cauchy-Schwarz inequality, in particular with $x = y$. Hence $\|Jx\|_{X^*} = \|x\|_X$.

For bijectivity, every linear isometry is 1-1 ($\ker J = 0$ since the lengths of vectors is preserved); J is onto by Riesz' theorem.

□

This shows that X and X^* are essentially isomorphic as Hilbert spaces. In Banach spaces in general the most that can be said is that X is isomorphic to a closed subspace of X^{**}. Note further that Hilbert spaces are therefore *reflexive* Banach spaces i.e. $X^{**} \equiv X$.

3.1.1 Exercises

1. Use the Riesz map to show that the two equations $\|\phi\| = \sup_x(\phi(x)/\|x\|)$ and $\|x\| = \sup_\phi(\phi(x)/\|\phi\|)$ become the same for Hilbert spaces.

2. Show that the norm of an operator is given by $\|T\| = \sup_{x,y}(|\langle y, Tx\rangle|/\|x\|\|y\|)$.

3. Show that, under the Riesz bijection, the annihilator A° of a set corresponds to $A^\perp$.

4. Prove the Hahn-Banach theorem for Hilbert spaces as follows. Start with any functional ϕ on a closed subspace M. Show that it must correspond to a vector in M, and hence find an extension of ϕ on X. Prove the rest of the theorem.

3.2 Adjoint Map

Recall that for Banach spaces, for every continuous operator $T : X \to Y$ there is a continuous adjoint or dual map $T^\top : Y^* \to X^*$ defined by $T^\top \phi(x) = \phi \circ T(x)$. For Hilbert spaces, the duals X^* and Y^* are essentially isomorphic to X and Y respectively, so we should get an operator $T^* : Y \to X$. In fact, it is defined by $T^*(y) = J\langle y, T.\rangle$ where J is the Riesz map i.e.,

$$\langle T^* y, x\rangle = \langle y, Tx\rangle \qquad \forall x, y \in X.$$

T^* is well-defined by the Riesz correspondence.

3.2.1 Example

Consider the left-shift map $L : \ell^2 \to \ell^2$ defined by $L(x_n) = (x_{n+1})$. Let $L^*(y_n) = (a_n)$, then $\langle L^* y, x\rangle = \langle y, Lx\rangle$ means

$$\sum_n a_n x_n = \sum_n y_n x_{n+1} = \sum_n y_{n-1} x_n.$$

In particular, taking x to be a basis vector $e_m = (\delta_{nm})$, we get $a_m = y_{m-1}$ so that $L^*(y_n) = (y_{n-1})$ i.e. $L^* = R$ the right-shift operator.

Proposition 3.3 **T^* is unique, linear and continuous, and satisfies**

$$T^{**} = T, \quad (S+T)^* = S^*+T^*, \quad (\lambda T)^* = \bar{\lambda} T^*, \quad (ST)^* = T^* S^*, \quad \|T^*\| = \|T\|$$

Proof. Uniqueness: Suppose A and B are adjoints of T i.e. $\langle Ay, x\rangle = \langle x, Ty\rangle = \langle By, x\rangle$ for any x, y. Then $Ay = By$ for all y so that $A = B$. We are therefore justified in using the notation T^*.

Linearity:

$$\langle T^*(y_1 + y_2), x\rangle = \langle y_1 + y_2, Tx\rangle = \langle y_1, Tx\rangle + \langle y_2, Tx\rangle = \langle T^* y_1 + T^* y_2, x\rangle;$$

$$\langle T^*(\lambda y), x\rangle = \langle \lambda y, Tx\rangle = \bar{\lambda}\langle y, Tx\rangle = \bar{\lambda}\langle T^* y, x\rangle = \langle \lambda T^* y, x\rangle.$$

Moreover,

$$\langle T^{**} x, y\rangle = \langle x, T^* y\rangle = \overline{\langle T^* y, x\rangle} = \overline{\langle y, Tx\rangle} = \langle Tx, y\rangle;$$

$$\langle (S+T)^* y, x\rangle = \langle y, (S+T)x\rangle = \langle y, Sx\rangle + \langle y, Tx\rangle = \langle (S^* + T^*) y, x\rangle;$$

$$\langle (\lambda T)^* y, x\rangle = \langle y, \lambda Tx\rangle = \lambda\langle T^* y, x\rangle = \langle \bar{\lambda} T^* y, x\rangle;$$

$$\langle (ST)^* y, x \rangle = \langle y, (ST)x \rangle = \langle S^* y, Tx \rangle = \langle T^* S^* y, x \rangle.$$

Continuity:

$$\|T^*\| = \sup_{x,y} \frac{|\langle T^* y, x \rangle|}{\|x\|\|y\|} = \sup_x \frac{|\langle y, Tx \rangle|}{\|x\|\|y\|} = \|T\|.$$

□

Proposition 3.4 **Let T be a continuous operator and M a closed linear subspace. Then**

$$(i)\ \ker T^* = (\mathrm{im} T)^\perp; \quad (\mathrm{im} T^*)^\perp = \ker T;$$

$$(ii)\ M \textbf{ is } T\textbf{-invariant} \ \Leftrightarrow M^\perp \textbf{ is } T^*\textbf{-invariant}.$$

Proof. $y \in \ker T^* \Leftrightarrow \langle T^* y, x \rangle = 0\ \forall x \in X, \Leftrightarrow \langle y, Tx \rangle = 0 \Leftrightarrow y \perp \mathrm{im} T$. Moreover, $\ker T = \ker T^{**} = (\mathrm{im} T^*)^\perp$.

The second part is basically a translation of the fact that for Banach spaces, M is T-invariant if, and only if, M° is $T^\top$-invariant. In detail, for any $y \in M^\perp$ and $x \in M$, $\langle T^* y, x \rangle = \langle y, Tx \rangle = 0$ since $Tx \in M$, so that $T^* y \in M^\perp$.

Conversely, if $M^\perp$ is T^*-invariant then $M^{\perp\perp}$ is T^{**}-invariant; but $T^{**} = T$ and $M^{\perp\perp} = M$ for a closed subspace M.

□

3.2.2 Application

When an operator T does not have an inverse, the equation $Tx = b$ need not have a solution. The next best thing is to ask for a vector x which *minimizes* $\|Tx - b\|$.

Proposition 3.5 **A vector x minimizes $\|Tx - b\|$ if, and only if**

$$T^* Tx = T^* b.$$

Proof. Consider the closed linear subspace $M = \overline{\mathrm{im} T}$. The problem is equivalent to finding a vector $y = Tx \in M$ which is closest to b. A necessary and sufficient condition is $b - y \in M^\perp = \ker T^*$. Hence $T^*(b - y) = 0$ i.e. $T^* Tx = T^* b$.

$\square$

Note however that $\|Tx - b\|$ may have no minimum when $\text{im}T$ is not closed. Moreover such an x need not be unique, when $\ker T \neq 0$. However, there will be a unique such x with smallest norm. The mapping from b to this x is then well-defined and is denoted by $T^\dagger$, called the Moore-Penrose generalized inverse. In the simple case when T is 1-1, it is given by $T^\dagger = (T^*T)^{-1}T$ (even if T is not onto).

3.2.3 Exercises

1. Show (i) $(\lambda I)^* = \bar{\lambda} I$, (ii) if $T : H \to H$ is invertible then $(T^{-1})^* = (T^*)^{-1}$.

2. For $T : H \to H$, prove that $\|T^*T\| = \|T\|^2$, and deduce from this that $\|T^*\| = \|T\|$.

3. Find the adjoint map of the Fourier map $F : L^2[-\pi, \pi] \to \ell^2$.

4. Let X be a real Hilbert space and let T be a continuous operator, a a fixed vector and b a fixed real number. Show that the minimum of the quadratic function $q(x) = \langle x, Tx\rangle + \langle a, x\rangle + b$ occurs when $(T+T^*)x + a = 0$. (Hint: consider $q(x+h) \geqslant q(x)$ for all h).

5. Show that $\text{im}T$ is dense in X if, and only if, T^* is 1-1.

6. Show that if T and T^* are onto, then T has a continuous inverse (you need to use the open mapping theorem).

7. Let $a+M$ be a coset of a closed linear subspace M. Show that there is a unique vector $x \in a+M$ with smallest norm. (Hint: this is equivalent to finding the closest vector in M to $-a$.)

8. Find the best approximation for $\boldsymbol{x}$ that solves
$$\begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix} \boldsymbol{x} = \begin{pmatrix} 4 \\ -1 \\ 0 \end{pmatrix}.$$

4 Normal Operators

In this chapter and the next we will concentrate on operators from X to itself. These form a so-called C^*-algebra $B(X)$ i.e there are the usual binary operations of addition and multiplication on $B(X)$ and a unary operation of taking the adjoint.

Before we delve into the properties of these operators, let us note that an operator is identified by the numbers $\langle y, Tx\rangle$ for all $x, y \in X$. For complex Hilbert spaces, we can even say that they are identified by the numbers $\langle x, Tx\rangle$ for all $x \in X$. This is so because $\langle x, Tx\rangle = 0, \forall x \Leftrightarrow T = 0$ (Prove!)

Definition A **normal** operator T is a continuous operator such that $T^*T = TT^*$.

A special case is that of *self-adjoint* operators which satisfy $T^* = T$.

Proposition 4.1 **For T a normal operator,**

$$(i)\ \|T^*x\| = \|Tx\|, \quad \|T^2\| = \|T\|^2;$$

$$(ii)\ \ker T^* = \ker T; \quad \textbf{im}T \textbf{ is dense in } X \Leftrightarrow T \textbf{ is 1-1}.$$

Proof. (i) follows from

$$\|T^*x\|^2 = \langle T^*x, T^*x\rangle = \langle x, TT^*x\rangle = \langle x, T^*Tx\rangle = \langle Tx, Tx\rangle = \|Tx\|^2.$$

Hence $\|T^2x\| = \|T^*Tx\|$, and

$$\|Tx\|^2 = \langle x, T^*Tx\rangle \leqslant \|T^*T\|\|x\|^2,$$

from which follows $\|T\|^2 \leqslant \|T^2\|$.

(ii) follows from the fact that $T^*x = 0 \Leftrightarrow \|T^*x\| = 0 \Leftrightarrow \|Tx\| = 0 \Leftrightarrow Tx = 0$, using (i). Also, $(\mathrm{im}T)^\perp = \ker T^* = \ker T$ which is 0 if, and only if, T is 1-1.

□

4.0.4 Exercises

1. Prove that for complex Hilbert spaces $\langle x, Tx\rangle = 0, \forall x \Leftrightarrow T = 0$, by substituting the vectors $x + y$ and $x + iy$. Deduce that if $\langle x, Tx\rangle = \langle x, Sx\rangle$, $\forall x$ then $T = S$.

2. Show that a continuous operator which satisfies $\|T^*x\| = \|Tx\|$ is normal.

3. Prove that for a normal operator T, (i) T^*, $T - \lambda$, and $p(T)$ for a polynomial p, are also normal, (ii) $\ker T$ is T^*-invariant and $(\ker T)^\perp$ is T-invariant.

4. Use (ii) of the proposition to show that $\ker T = \ker T^2 = \ker T^k$. Deduce that the only normal nilpotent operator (i.e. $T^k = 0$ for some k) is 0.

5. Prove that for a separable Hilbert space, there can be at most a countable number of distinct eigenvalues.

6. Suppose T is diagonalisable i.e there is a Hilbert basis of eigenvectors. Show that T is normal. (Hint: first show that $T^* e_n = \bar{\lambda}_n e_n$.)

4.1 Spectrum

Definition The spectrum of an operator $T \in B(H)$ is the set $\sigma(T) := \{ \lambda \in \mathbb{C} : T - \lambda \text{ does not have an inverse} \}$.

(Note: The standard definition requires that $T - \lambda$ does not have a continuous inverse. But, by the open mapping theorem, every bijective continuous linear operator has a continuous inverse, so the two definitions are equivalent.)

From the theory of Banach spaces, it follows that the spectrum of any operator is a closed bounded non-empty subset of $\mathbb{C}$. But the proof requires complex analysis. For our purposes, we can only show that the spectrum of a self-adjoint operator is non-empty.

Proposition 4.2 **The spectrum of any (self-adjoint) operator is non-empty.**

Proof. First notice that for a self-adjoint operator A, $\langle x, Ax \rangle$ is a real number since

$$\langle x, Ax \rangle = \langle Ax, x \rangle = \overline{\langle x, Ax \rangle}.$$

Claim: If A is self-adjoint such that $\langle x, Ax \rangle \geqslant 0$, then

$$\forall x, y, \quad |\langle x, Ay \rangle|^2 \leqslant \langle x, Ax \rangle \langle y, Ay \rangle.$$

The proof is as follows: for any $z = te^{i\theta} \in \mathbb{C}$,

$$\begin{aligned} 0 &\leqslant \langle x + zy, A(x + zy) \rangle \\ &= \langle x, Ax \rangle + 2t \operatorname{Re} e^{i\theta} \langle x, Ay \rangle + t^2 \langle y, Ay \rangle \\ \therefore (\operatorname{Re} e^{i\theta} \langle x, Ay \rangle)^2 &\leqslant \langle x, Ax \rangle \langle y, Ay \rangle \\ \therefore |\langle x, Ay \rangle|^2 &\leqslant \langle x, Ax \rangle \langle y, Ay \rangle, \end{aligned}$$

since θ can be chosen to rotate $\langle x, Ay\rangle$ into $|\langle x, Ay\rangle|$.

Now, let $\lambda := \inf_{\|x\|=1} \langle x, Tx\rangle$. This implies, for all x,

$$\langle x, (T-\lambda)x\rangle = \langle x, Tx\rangle - \lambda\|x\|^2 \geqslant 0.$$

Using $y = (T-\lambda)x$ and $A = T-\lambda$,

$$\|Ax\|^4 = |\langle x, A^2x\rangle|^2 \leqslant \langle x, Ax\rangle\langle y, Ay\rangle \to 0$$

as $\langle x, Tx\rangle \to \lambda$. Thus, λ belongs to $\sigma(T)$.

□

Proposition 4.3 **For a normal operator T,**

(i) **eigenvalues of T^* are conjugates of eigenvalues of T,**

(ii) **eigenvectors of distinct eigenvalues of T are orthogonal.**

$$(iii) \quad 0 \notin \sigma(T) \Leftrightarrow \exists c > 0 \ \|x\| \leqslant c\|Tx\|.$$

Proof. (i) is a direct application of the previous proposition on $\ker(T-\lambda) = \ker(T^*-\bar{\lambda})$. For eigenvalues λ and μ with corresponding eigenvectors x and y, we have

$$\lambda\langle y, x\rangle = \langle y, Tx\rangle = \langle T^*y, x\rangle = \langle \bar{\mu}y, x\rangle = \mu\langle y, x\rangle,$$

implying that either $\lambda = \mu$ or $\langle y, x\rangle = 0$.

(ii) Suppose the spectrum $\sigma(T)$ does not contain 0. Then T has a continuous inverse T^{-1} and $\|T^{-1}y\| \leqslant c\|y\|$, implying $\|x\| \leqslant c\|Tx\|$.

Conversely, suppose the right inequality is true for all x. Then $Tx = 0$ implies $x = 0$ so that T is 1-1. The image of T is closed since suppose $y_n = Tx_n$ is a Cauchy sequence in the image of T. Then $\|x_n - x_m\| \leqslant c\|Tx_n - Tx_m\| = c\|y_n - y_m\| \to 0$ as $n, m \to \infty$. Hence $x_n \to x$ implying that $y_n = Tx_n \to Tx$ since T is continuous. These two facts imply that the image of T must be X itself since $\mathrm{im}T^\perp = \ker T^* = \ker T = 0$. Thus T is invertible, and continuous:

$$\|T^{-1}y\| \leqslant c\|TT^{-1}y\| = c\|y\|.$$

□

That is, a value λ is in the spectrum iff there are unit vectors x_n such that $(T-\lambda)x_n \to 0$.

Proposition 4.4 **The spectrum of T has spectral radius $\|T\|$, contains no residual spectrum, and is a subset of the closure of the set $\{\langle x,Tx\rangle/\langle x,x\rangle\}$.**

Proof. The spectral radius of $\sigma(T)$ is given by $r_T = \lim_{n\to\infty}\|T^n\|^{1/n} = \lim_{k\to\infty}\|T^{2^k}\|^{2^{-k}} = \|T\|$.

Either $T-\lambda$ is 1-1, in which case its image is dense in X and λ forms part of the continuous spectrum, or it is not 1-1, in which case there must be an eigenvector and λ is an eigenvalue (point spectrum).

Let $A := \overline{\{\langle x,Tx\rangle/\langle x,x\rangle\}}$. Then $\lambda \notin A$ implies that

$$\exists\alpha>0 \quad \forall x \quad |\lambda - \frac{\langle x,Tx\rangle}{\|x\|^2}| \geqslant \alpha,$$

$$\Rightarrow \quad |\frac{\langle x,(T-\lambda)x\rangle}{\|x\|^2}| \geqslant \alpha > 0,$$

$$\Rightarrow \quad \alpha\|x\|^2 \leqslant |\langle x,(T-\lambda)x\rangle| \leqslant \|x\|\|(T-\lambda)x\|,$$

$$\Rightarrow \quad \alpha\|x\| \leqslant \|(T-\lambda)x\|,$$

$$\Rightarrow \quad \lambda \notin \sigma(T).$$

Notice that since $|\langle x,Tx\rangle| \leqslant \|T\|\|x\|^2$ we have that the set A is contained in the ball of radius $\|T\|$. Moreover, since the spectral radius of T is $\|T\|$, the "radius" of A itself is $\|T\|$.

□

Theorem 4.5 **(Spectral Theorem in Finite Dimensions)**
A normal operator on a finite-dimensional vector space is diagonalisable.

Proof. In finite dimensions, $B(X)$ has dimension $(\dim X)^2$ and so every operator T must satisfy a polynomial $m_T(x) = (x-\lambda)^k \ldots$. This induces a decomposition of the vector space as $X = \ker(T-\lambda)^k + \ldots$. For normal operators $\ker(T-\lambda)^k = \ker(T-\lambda)$, so that the minimal polynomial consists of simple factors. This is equivalent to the diagonalizability of T.

4.1.1 Exercises

1. By taking $T = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$ and $x = \begin{pmatrix}1\\1\end{pmatrix}$, show that the spectrum of a normal operator need not equal the closure of the set $\{\,\langle x,Tx\rangle/\|x\|^2\,\}$.

2. Show that $\|p(T)\| = \max_{\lambda\in\sigma(T)}\|p(\lambda)\|$ when T is normal and p is a polynomial.

4.2 Unitary Operators

Definition A **unitary** operator is a Hilbert-space automorphism i.e. a linear bijection which preserves the inner-product,

$$\langle Tx, Ty\rangle = \langle x, y\rangle \quad \forall x, y \in X.$$

Proposition 4.6 **The following are equivalent**

$$\langle Tx, Ty\rangle = \langle x, y\rangle \quad \forall x, y \Leftrightarrow T^*T = I \Leftrightarrow \|Tx\| = \|x\| \quad \forall x.$$

Proof. $\langle x, y\rangle = \langle Tx, Ty\rangle = \langle x, T^*Ty\rangle\ \forall x, y \Leftrightarrow T^*T = I$. By the comment at the beginning of this chapter, we can also say that $\langle x, x\rangle = \langle Tx, Tx\rangle = \langle x, T^*Tx\rangle\ \forall x \Leftrightarrow T^*T = I$.

□

This proposition is basically saying that preserving the inner-product (lengths *and* angles) is equivalent to preserving lengths. Recall also that norm-preserving linear operators must be 1-1. Hence we conclude that a linear operator is unitary if, and only if, it is onto and $T^*T = I$.

Proposition 4.7 **Let T be a unitary operator, then**

1. **the adjoint $T^* = T^{-1}$ is also unitary;**
2. **T is continuous with norm $\|T\| = 1$;**
3. **T is normal;**
4. **the set of unitary operators form a group i.e. if S, T are unitary then so is ST.**

Proof. If T is unitary then it is invertible and $T^*T = I$; hence $T^* = T^{-1}$. T^* is itself unitary since $(T^*)^{-1} = (T^{-1})^* = T^{**}$. For continuity, $\|T\| = \sup_x(\|Tx\|/\|x\|) = 1$. T is normal since T always commutes with its inverse, which for the case of unitary operators, is its adjoint. Finally, $(ST)^*(ST) = T^*S^*ST = I$ and ST is invertible when S, T are.

□

Proposition 4.8 **The spectrum of a unitary operator lies in the unit circle,**

$$\sigma_T \subseteq e^{i\mathbb{R}}.$$

Proof. The spectrum must lie in the unit ball since $\|T\| = 1$. Take $|\lambda| < 1$, then $T - \lambda = T(1 - \lambda T^*)$ and $\|\lambda T^*\| = |\lambda|\|T^*\| = |\lambda| < 1$. Therefore $(1 - \lambda T^*)$, and so $T - \lambda$, are invertible, with continuous inverse.

□

4.2.1 Exercises

1. Let T be an invertible normal operator. Show that T^*T^{-1} is unitary.

2. Show that the right shift operator on ℓ^2 satisfies $R^*R = I$ but is not unitary.

3. Let $T : \ell^2 \to \ell^2$ be defined by $T(x_n) = (i^n x_n)$. Show that T is unitary with eigenvalues $\{\pm 1, \pm i\}$. Find the set $\{\langle x, Tx \rangle / \|x\|^2\}$ and deduce that the spectrum of T is precisely the set of eigenvalues.

4. Show that an operator is unitary if, and only if, it maps Hilbert bases to Hilbert bases.

5. Show directly that eigenvalues of unitary operators must satisfy $|\lambda| = 1$.

6. In the proof of the equivalence of the first proposition of this section, use was made that the Hilbert space is over the complex field. Show that for real Hilbert spaces, $T^*T = I \Leftrightarrow \|Tx\| = \|x\|$ is also valid.

7. Let T_g be a group of unitary operators for $g \in G$ a group. Suppose M is T_g-invariant for all $g \in G$. Show that $M^\perp$ is also invariant.

8. (*) Show that a function on a Hilbert space which preserves the inner-product must be linear. Deduce that isometries (i.e. preserve distances) on a Hilbert space must be of the type $P(x) = Ux + a$ where $U^*U = I$. (Hint: show $\langle f(x+y), f(z) \rangle = \langle f(x) + f(y), f(z) \rangle$, so $f(x+y) - f(x) - f(y) \in \langle \operatorname{im} f \rangle \cap \langle \operatorname{im} f \rangle^\perp$.)

4.3 Self-Adjoint Operators

Definition A **self-adjoint** operator is a continuous linear operator such that $T^* = T$.

Self-adjoint operators are obviously normal.

Proposition 4.9 **Every operator in $B(X)$ can be written uniquely as $T = S + iR$ for S, R self-adjoint.**

Proof. Simply check that $S = (T + T^*)/2$ and $R = (T - T^*)/2i$ are self-adjoint and add to T. Uniqueness follows from the fact that $A + iB = 0$ for A, B self-adjoint if, and only if, $A = B = 0$ (since then $A = A^* = (-iB)^* = iB = -A$.)

□

Note that in this case, $T^* = S - iR$; and T is normal if, and only if, $SR = RS$; T is unitary if, and only if, T is normal and $S^2 + R^@ = 1$; T is self-adjoint if, and only if, $R = 0$.

Proposition 4.10 **T is self-adjoint iff the set $\{\, \langle x, Tx\rangle/\|x\|^2 \,\}$ is real.**

Proof. Since T is self-adjoint, $\langle x, Tx\rangle = \langle T^*x, x\rangle = \langle Tx, x\rangle = \overline{\langle x, Tx\rangle}$ is a real number. Hence the set $\{\, \langle x, Tx\rangle/\|x\|^2 : x \in X \,\}$ is real.

Conversely, if the given set is real, then

$$\langle x + y, Tx + Ty\rangle = \langle x, Tx\rangle + \langle x, Ty\rangle + \langle y, Tx\rangle + \langle y, Ty\rangle$$

implies that $\langle x, Ty\rangle + \langle y, Tx\rangle$ is real. Similarly, starting with $x + iy$ gives that $\langle x, Ty\rangle - \langle y, Tx\rangle$ is purely imaginary. Hence $\langle x, Ty\rangle = \overline{\langle y, Tx\rangle} = \langle Tx, y\rangle$. Thus $T^* = T$ and T is self-adjoint.

Corollary **The spectrum of a self-adjoint operator is real.**

Note that the converse is also true i.e. a normal operator whose spectrum is real must be self-adjoint.

4.3.1 Exercises

1. Prove that a continuous projection P is normal iff it is orthogonal iff it is self-adjoint.

2. By taking $\lambda = \alpha + i\beta$ with $\beta \neq 0$ and expanding $\|(T - \lambda)x\|$, show that $\|(T - \lambda)x\| \geqslant |\beta|\|x\|$ and hence that $\lambda \notin \sigma(T)$.

3. Let $Ty = -\frac{1}{p}(py')' + qy$ on the Hilbert space $L^2[a,b]$. Show that $\langle Tu, v\rangle = \langle u, Tv\rangle + [-p(u'v - uv')]_a^b$. Deduce that when we restrict to the subspace of those functions which also satisfy (i) the periodic condition $y(a) = y(b)$, or (ii) the unmixed conditions $\alpha_1 u(a) + \alpha_2 u'(a) = 0$, $\beta_1 u(b) + \beta_2 u'(b) = 0$, then T is self-adjoint.

4. A self-adjoint operator is called *positive definite* when $\langle x, Tx\rangle > 0$ fora all $x \neq 0$. Show that in this case the map $\langle\langle x, y\rangle\rangle = \langle x, Ty\rangle$ defines an inner-product on X.

4.4 Compact Normal Operators

Definition T is called a **compact** operator when for every bounded set B, $\overline{TB}$ is compact.

In particular, if x_n are unit vectors then Tx_n has a convergent subsequence. Compact operators are obviously continuous.

For example, matrices $T : \mathbb{C}^N \to \mathbb{C}^M$ are compact since TB has to be finite-dimensional, so its closure must be closed and bounded, hence compact. But in infinite dimensions, such as ℓ^2, the identity operator I is not compact (take any orthonormal sequence).

Example. The map $T : \ell^2 \to \ell^2$ defined by $T(a_n) := (a_n/n)$ is compact.

Consider the linear maps $T_N : \ell^2 \to \ell^2$, $T_N(a_n) := (a_1, a_2/2, \ldots, a_N/N, 0, \ldots)$; T_N is compact since its image is a subspace of $\mathbb{C}^N$. Now let $\boldsymbol{x}_n$ be a bounded sequence, say $\|\boldsymbol{x}_n\|_{\ell^2} \leqslant M$. For each N, there is a subsequence $\boldsymbol{x}_{n,N}$ such that

$$\|T_N\boldsymbol{x}_{n,N} - T_N\boldsymbol{x}_{m,N}\| \leqslant \frac{1}{2^N}.$$

Hence

$$\begin{aligned}\|T\boldsymbol{x}_{n,n} - T\boldsymbol{x}_{m,n}\| &= \|(T-T_n)\boldsymbol{x}_{n,n} + (T_n - T_m)\boldsymbol{x}_{n,n} + T_m(\boldsymbol{x}_{n,n} - \boldsymbol{x}_{m,n}) + (T_m - T)\boldsymbol{x}_{m,n}\| \\ &\leqslant \|(T-T_n)\boldsymbol{x}_{n,n}\| + \|(T_n - T_m)\boldsymbol{x}_{n,n}\| + \|T_m(\boldsymbol{x}_{n,n} - \boldsymbol{x}_{m,n})\| + \|(T_m - T)\boldsymbol{x}_{m,n}\| \\ &\to 0\end{aligned}$$

Thus every bounded sequence x_n gives a convergent subsequence of Tx_n.

Theorem 4.11 **For T compact normal, the non-zero part of the spectrum consists of a countable number of eigenvalues whose only possible limit point is 0. Each non-zero eigenvalue has a finite-dimensional eigenspace.**

Proof. Let $\lambda \in \sigma(T)$, $\lambda \neq 0$. Then there are unit vectors x_n such that $(T - \lambda)x_n \to 0$. But when T is compact, there is a convergent subsequence

$Tx_{n_i} \to y$. This implies

$$\lambda x_{n_i} = Tx_{n_i} - (T-\lambda)x_{n_i} \to y$$

so $x_{n_i} \to y/\lambda$ and $Tx_{n_i} \to Ty/\lambda$. Since limits are unique, $Ty = \lambda y$, and λ is an eigenvalue. (Note that y is not zero since x_{n_i} are unit vectors.)

Suppose there are an infinite number of unit eigenvectors e_n with non-zero eigenvalues λ_n. The bounded sequence λ_n must have a subsequence converging to some λ. From $Te_n = \lambda_n e_n$ and the fact that $Te_n \to y$ for some subsequence, we get $\lambda_n e_n \to y$. Yet, as T is normal, its eigenvectors e_n of distinct eigenvalues are orthogonal, and these cannot have a convergent subsequence since $\|e_n - e_m\| = \sqrt{2}$.

Keeping this in mind, if $\lambda \neq 0$, and there are an infinite number of distinct eigenvalues λ_n, then their orthonormal eigenvectors would converge to $e_n \to y/\lambda$, a contradiction. The only limit point must then be 0, and there can only be a countable number of non-zero eigenvalues. Similarly, if λ has an infinite number of linearly independent eigenvectors, choose a countable number of them e_n, make them orthonormal by the Gram-Schmidt algorithm, and this furnishes another contradiction.

□

Corollary **Spectral Theorem for Compact Normal operators**
Every compact normal operator has a countable orthonormal set of eigenvectors e_n with non-zero eigenvalues λ_n, and

$$Tx = \sum_n \lambda_n \langle e_n, x\rangle e_n.$$

Proof. T restricted to the orthogonal space of $M := [\![e_n]\!]$ is still compact normal. Yet it cannot have any non-zero eigenvalues because these are already accounted for in e_n. Its spectrum must therefore be $\{0\}$, that is the T restricted to $M^\perp$ is the zero operator. Hence $M^\perp = \ker T$. Writing any vector $x = \sum_n \langle e_n, x\rangle e_n + v$ where $v \in \ker T$, we find

$$Tx = \sum_n \langle e_n, x\rangle Te_n = \sum_n \lambda_n \langle e_n, x\rangle e_n.$$

Thus compact normal operators are diagonalizable. In terms of their representation in ℓ^2, every vector $x \sim (a_n)$ is mapped to $Tx \sim (\lambda_n a_n)$.