The Book of Narcowich

His Holiness S. Gerberding

January 10, 2026

Contents

Chapter 1

Vector Spaces

Definition 1.1. A vector space V is a set with operations $+, \times$ and scalars S ($\mathbb{C}$ or $\mathbb{R}$) such that:

1. For all $u, v \in V$, $u + w \in V$. (Closure under addition)
2. For all $c \in S$ $u \in V$, $cu \in V$. (Closure under scalar multiplication)
3. For all $u, v, w \in V$, $(u + v) + w = u + (v + w)$. (Addition is commutative)
4. There exists $0 \in V$ such that $0 + v = v$ for all $v \in V$. (Identity elements)
5. For all $u \in V$, there exists $-u \in V$ such that $u + (-u) = 0$. (Inverses)
6. For all $c \in S$, $u, v \in V$, $c(u + v) = cu + cv$. (Scalar distribution)
7. For all $c, d \in S$,$v \in V$, $(c + d)v = cv + dv$. (Vector distribution)
8. For all $c, d \in S$, $u \in V$, $c(dv) = (cd)v$. (Scalar associativity)
9. For all $u \in V$, $1u = u$.

Definition 1.2. A subset U of V is a subspace if under $+, \cdot$ (from V), U is its own vector space.

Theorem 1.3. *Let V be a VS. $U \subset V$ is a subspace iff the following hold:*

1. $0 \in U$;
2. *U is closed under $+$;*
3. *U is closed under $\cdot$.*

Proof. Suppose that U is a subspace. Then all three items hold from the definition of a vector space.

Now suppose the three conditions hold. Since these operations are inherited from V, and $U \subset V$, we have that all of the axioms hold. Hence U is a vector space, and thus a subspace. □

Definition 1.4. Let $S = \{v_1, ..., v_n\}$ be a subset of a VS V. The *span of S* is the set of all linear combinations of S, ie, $Span(S) = \{c_1v_1 + ...c_nv_n\}$, where $c_j \in \mathbb{C}$.

Proposition 1.5. Let V be a VS, and $S \subset V$. Then $Span(S)$ is a subspace.

Proof. Let $c_j = 0$ for all j. Then, $0 \in Span(S)$. Let $u, v \in Span(S)$. Then, $u = c_1x_1 + ... + c_nx_n$, $v = d_1x_1 + ... + d_nx_n$. Then, for $a \in \mathbb{C}$, $au + v = \sum_1^n (ac_j + d_j)x_n \in Span(S)$. Thus $Span(S)$ is a subspace. □

Definition 1.6. Let $S = \{v_1, ...v_n\} \subset V$ be a set. Then S is *linearly independent* iff:

$$a_1v_n + ... + a_nv_n = 0 \implies a_1 = a_2 = ... = a_n = 0.$$

There there exists $a_j \neq= 0$ such that the equation holds, then S is *linearly dependent.*

Definition 1.7. A subset $B = \{v_1, ..., v_n\}$ of a VS V is a basis for V if B spans V and B is linearly independent. Equivalently, B is a basis if it *maxiamlly linearly indepedent*; that is, B is not a proper subset of some other linearly independent set.

Theorem 1.8. *Every basis for V has the same number of vectors.*

Proof. Omitted. $\square$

Remark 1.9. If V has arbitrarily large linearly independent sets, it is *infinite dimensional.*

Definition 1.10. Let V, W be vector spaces. $T : V \to W$ is *linear transformation* (*linear map*, or just linear) if, for all $u, v \in V$, $c \in \mathbb{C}$, $T(cu + v) = cT(u) + T(v)$.

Definition 1.11. A bijective linear map is called an isomorphism.

Proposition 1.12. Let U, V, W be VS, $S : U \to V$, $T : V \to W$, $R : V \to W$ be linear. Then $T \circ S := TS$ is linear. If S is an isomorphism, S^{-1} is an isomorphism. Furthemore, for all $c \in \mathbb{C}$, $(aS + R) : U \to V$ is linear.

Proof. Let $au + v \in V$. Then, $TS(au + v) = T(S(au + v)) = T(aS(u) + S(v)) = aTS(u) + TS(v)$. Thus, TS is linear. Suppose S is an isomorphism, and let $x, y \in V$. Then, there exists $u, v \in U$ such that $Su = x$, $Sv = y$. Thus, $S(au + v) = aSu + Sv = ax + y$. Therefore, $au + v = aS^{-1}x + S^{-1}v = S^{-1}(ax + y)$. $\square$

Theorem 1.13. *Let V be a finite dimensional vector space and let $B = \{v_1, ..., v_n\}$ be a basis for V. Then every $v^{-1}V$ can be uniquely written as $v = c_1v_1 + ... + c_nv_n$.*

Remark 1.14. Because this association is unique, we may represent v as $[v]_B := \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$. Note that is dependent upon the basis. We may therefore define a *coordinate operator* ϕ_B such that $\phi_B(x) = [x]_B$. Observe that this is a bijection, and therefore, $\phi_B^{-1}([x]_B) = x$ is also well-defined.

Thus, we have A, B as bases in a VS V, we can therefore have a change of basis as follows: let $[x]_B$ be the coordinate vector of x in terms of B and ϕ_B, ψ_A the coordinate operators for their respective bases. Then: $[x]_A = \psi_A(\phi_B^{-1}([x]_B))$. Of course, it this does not give us a matrix representation for this operator. But we may do so as follows.

Let $\alpha = \{v_1, ..., v_n\}$ and $\beta = \{w_1, ..., w_n\}$ be bases for VS V. Then, for $x \in V$, $x = \sum_{j=1}^n a_jv_j$ and $x = \sum_{i=1}^n b_kw_k$. Now, for each v_j, $v_j = \sum_{k=1}^n c_{jk}w_k$. Thus, $x = \sum_{j=1}^n a_j(\sum_{k=1}^n c_{jk}w_k)v_j = \sum_{k=1}^n(\sum_{j=1}^n a_jc_{jk})w_k$. Since the representation of a vector with respect to a basis is unique, we have:

$$b_i = \sum_{j=1}^n a_jc_{jk}.$$

Thus:

$$[x]_\beta = \begin{bmatrix} \sum_{j=1}^n a_jc_{1j} \\ \vdots \\ \sum_{j=1}^n a_jc_{nj} \end{bmatrix} = \begin{bmatrix} c_{11} & \dots & c_{1n} \\ \vdots & & \vdots \\ c_{1n} & \dots & c_{nn} \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} c_{11} & \dots & c_{1n} \\ \vdots & & \vdots \\ c_{1n} & \dots & c_{nn} \end{bmatrix} [x]_\alpha = [[v_1]_\beta, ...[v_n]_\beta][x]_\alpha.$$

Observe there that we therefore have $[x]_\beta = [T]_\alpha^\beta[x]_\alpha = [Tx]_\beta$, and so $T = I$.

If we want to know the matrix representation of a linear transformation T, we only need to consider how it changes the coordinates with respect to the bases. Let $T : V \to W$, T, W FDVSs, and let $\alpha = \{v_1, ..., v_n\}$ and $\beta = \{w_1, ..., w_m\}$ be their respective basis. Then,

$$T(x) = a_1T(v_1) + ...a_nT(v_n).$$

We also have that $T(v_j) = c_{1j}w_1 + ... + c_{mj}w_m = \sum_{i=1}^m c_{ij}w_i$ for each j. Thus, $T(x) = \sum_{j=1}^n a_j(\sum_{i=1}^m c_{ij}w_i) = \sum_{i=1}^m(\sum_{j=1}^n a_j c_{ij})w_i$. Thus,

$$[Tx]_\beta = \begin{bmatrix} \sum_{j=1}^n a_j c_{1j}) \\ \vdots \\ \sum_{j=1}^n a_j c_{nj}) \end{bmatrix} = \begin{bmatrix} c_{11} & \dots & c_{1n} \\ \vdots & & \vdots \\ c_{1n} & \dots & c_{nn} \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} c_{11} & \dots & c_{1n} \\ \vdots & & \vdots \\ c_{1n} & \dots & c_{nn} \end{bmatrix} [x]_\alpha = [[T(v_1)]_\beta, ...[T(v_n)]_\beta][x]_\alpha.$$

Thus we we may define the matrix representation (with respect to the appropriate bases) as $[[T(v_1)]_\beta, ..., [T(v_n)]_\beta] =: [T]_\alpha^\beta$.

Definition 1.15. A *linear functional* is a linear transformation $\phi : V \to$ scalars of V.

Definition 1.16. The set V^* of linear functionals on V is the (algebraic) dual of V.

Proposition 1.17. V^* is a vector space under the operations of addition and scalar multiplication of a function.

Chapter 2

Inner Product Spaces

Definition 2.1. Let V be a vector space over $\mathbb{C}$. Then, an inner product is a function $\langle ; \rangle : V \times V \to \mathbb{C}$ such that the following hold:

1. For all $v \in V$, $\langle v, v\rangle \geq 0$ (note that this means they are real valued);
2. $\langle u, v\rangle = \overline{\langle v, u\rangle}$;
3. $\langle cu, v\rangle = c\langle u, v\rangle$;
4. $\langle u + x, v\rangle = \langle u, v\rangle + \langle x, v\rangle$.

Observe that this all implies $\langle u, x + v\rangle = \langle u, x\rangle + \langle u, v\rangle$ and $\langle u, cv\rangle = \bar{c}\langle u, v\rangle$.

Definition 2.2. Let $u \in V$ from above. Then, $||u|| := \sqrt{\langle u, u\rangle}$.

Proposition 2.3. $||u + e^{i\alpha}v||^2 = ||u||^2 + 2|\langle u, v\rangle| + ||v||^2$.

Proof.

$$\begin{aligned} ||u + e^{i\alpha}v||^2 =& \langle u, u\rangle + \langle u, e^{i\alpha}v\rangle + \langle e^{i\alpha}v, u\rangle + \langle e^{i\alpha}v, e^{i\alpha}v\rangle \\ &= ||u||^2 + e^{-i\alpha}\langle u, v\rangle + \overline{\langle u, e^{i\alpha}v\rangle} + ||v||^2 \\ &= ||u||^2 + e^{-i\alpha}\langle u, v\rangle + \overline{e^{-i\alpha}\langle u, v\rangle} + ||v||^2. \end{aligned}$$

The above holds for all α. Using the polar form of a complex number, we have $\langle u, v\rangle = e^{i\theta}|\langle u, v\rangle|$ for some θ. Thus if we set $\alpha = \theta$, we deduce:

$$\begin{aligned} ||u + e^{i\alpha}v||^2 &= ||u||^2 + e^{-i\alpha}\langle u, v\rangle + \overline{e^{-i\alpha}\langle u, e^{v}\rangle} + ||v||^2 \\ &= ||u||^2 + e^{-i\theta}e^{i\theta}|\langle u, v\rangle| + \overline{e^{-\theta}e^{\theta}|\langle u, v\rangle|} + ||v||^2 \\ &= ||u||^2 + 2|\langle u, v\rangle| + ||v||^2. \end{aligned}$$

□

Theorem 2.4 (Cauchy-Schwarz Inequality)**.** $|\langle u, v\rangle| \leq ||u||||v||$.

Proof. If u or 0, then the above follows immediately. So suppose $u, v \neq 0$. Define $\alpha = sgn(\langle u, v\rangle)$, $z = \alpha v$ and let $t \in \mathbb{R}$. Then, we have:

$$\langle u, z\rangle = \langle u, \alpha v\rangle = \bar{\alpha}\langle u, v\rangle = \frac{\overline{\langle u, v\rangle}\langle u, v\rangle}{|\langle u, v\rangle|} = |\langle u, v\rangle|.$$

As well as $||z|| = ||\alpha v||^2 = |\alpha|^2||v||^2 = ||v||^2$. Therefore, we deduce:

$$\begin{aligned} 0 \le \langle u - tz, u - tz\rangle &= \langle u, u - tz\rangle - \langle tz, u - tz\rangle \\ &= ||u||^2 - t\langle u, z\rangle - t\langle z, u\rangle + t^2||z||^2 \\ &= ||u||^2 - t(|\langle u, v\rangle| + |\langle u, v\rangle|) + t^2||v||^2 \\ &= ||u||^2 - 2t|\langle u, v\rangle| + t^2||v||^2. \end{aligned}$$

This is a positive polynomial, and therefore the discriminant is such that $4|\langle u, v\rangle|^2 - 4||u||^2||v||^2 \le 0$. Thus, $|\langle u, v\rangle|^2 \le ||u||^2||v||^2$. □

Corollary 2.5. Equality holds in the Scharwz inequality iff $u,\ v$ are linearly dependent.

Proof. First suppose equality holds. That means, from the above, for some t, we have $0 = \langle u - tz, u - tz\rangle$, so $u - tz = 0$, thus $u = tz = t\alpha v$. Thus, we have linear dependence. If $u = cv$, then $|\langle u, v\rangle| = |\langle cv, v\rangle| = |c|||v||^2 = |c|||v|| \cdot |v||$. Thus, we have equality. □

Theorem 2.6 (Triangle Inequality)**.** *For $u, v \in V$, $||u + v|| \le ||u|| + ||v||$.*

Proof. Let $u,\ v \in V$, then:

$$\begin{aligned} ||u + v||^2 &= \langle u + v, u + v\rangle \\ &= ||u||^2 + \langle u, v\rangle + \langle v, u\rangle + ||v||^2 \\ &= ||u||^2 + 2Re(\langle u, v\rangle) + ||v||^2 \\ &\le ||u||^2 + 2|\langle u, v\rangle| + ||v||^2 \\ &\le ||u||^2 + 2||u||||v|| + ||v||^2 \\ &= (||u|| + ||v||)^2. \end{aligned}$$

Therefore, $||u + v|| \le ||u|| + ||v||$. □

Remark 2.7. If V is a real vector space, CS gives us that:

$$-1 \le \frac{\langle u, v\rangle}{||u||||v||} \le 1.$$

Therefore, we may define the angle θ be u and v to be

$$\theta := \arccos(\frac{\langle u, v\rangle}{||u||||v||}.$$

Definition 2.8. Let V be an inner product space, $u,\ v \in V$. Then, u and v are *orthogonal* if $\langle u, v\rangle = 0$.

The following are some examples of inner product spaces with their inner products:

1. $\mathbb{R}^2$: $\langle x, y\rangle = \sum_1^n x_j y_j = y^T x$.
2. $\mathbb{C}^n$: $\langle x, y\rangle = \sum_1^n x_j \overline{y_j} = y^* x$.
3. Real $L^2[a, b]$: $\langle f, g\rangle = \int_a^b f(x)g(x)dx$.
4. Complex $L^2[a, b]$: $\langle f, g\rangle = \int_a^b f(x)\overline{g(x)}dx$.
5. 2π-periodic L^2 functions: $\langle f, g\rangle = \int_0^{2\pi} f(x)\overline{g(x)}dx$.
6. Weighted L^2 inner products (real): $\langle f, g\rangle = \int_a^b f(x)g(x)w(x)dx$, where $w(x) \ge 0$.

7. Let A be an $n \times n$ matrix with real entries. Suppose that A is self-adjoint and that $x^T Ax > 0$, unless $x = 0$: $\langle x, y\rangle = y^T Ax$.

8. Biinfinite complex sequences, l^2: $\langle x, y\rangle = \sum_{-\infty}^{\infty} x_n \overline{y_n}$.

Example 2.9. Let f, g real valued function in $C[0,1]$ and define $\langle f, g\rangle := \int_0^1 f(x)g(x)dx$. Show that $\langle f, g\rangle$ is a real inner product for $C[0,1]$.

Solution 2.10. Observe that additivity, symmetry, and scalars all follow from properties of the integral. Thus all that remains is positivity. Since $f^2 \geq 0$, $\int_a^b f^2 dx \geq \int_a^b 0dx = 0$. Thus, $\langle f, f\rangle \geq 0$. Clearly, if $f = 0$, then $\langle f, f\rangle = 0$. Now suppose $\langle f, f\rangle = 0$. That is, $\int_a^b f^2 dx = 0$. Since $f^2 \geq 0$, this implies that $f = 0$. Thus we have an inner product.

Note: This is also easily seen by the fact that $C[0,1] \subset L^2([0,1])$, and so inherits the inner product.

2.1 Orthogonality

Definition 2.11. Let V be an inner product space, and $S = \{v_1, ..., v_n, ...\} \subset V$. Then, S is orthogonal if none of the vectors are 0 (this is more of a formality that saves needless wordings) and $\langle v_j, v_k\rangle = 0$ if $j \neq k$.

Proposition 2.12. Let $S = \{v_1, ..., v_n, ...\}$. If S is orthogonal, then S is linearly independent.

Proof. Suppose S is linearly dependent, so $v_j \neq 0$ for all j. Wlog, suppose $v_1 = \sum_2^k c_j v_j$. The,

$$\langle \sum_2^k c_j v_j, v_1\rangle > 0 \text{since } v_1 \neq 0$$

Thus, $\sum_2^k \langle v_j, v_1\rangle > 0$. Thus, there exists some p such that $\langle v_p, v_1\rangle \neq 0$. Thus, S is not orthogonal. □

Alternative Proof. Suppose $S = \{v_1, ..., v_n, ...\}$ is orthogonal. Consider

$$a_1 v_1 + ...a_n v_n = 0$$

For, for each v_j, $j = 1, ..., n$, we have:

$$\langle v_j, \sum_1^n a_i v_i\rangle = a_j \langle v_j, v_j\rangle = 0$$

Since $v_j \neq 0$, we have that $a_j = 0$. This holds for all j, and so S is linearly independent. □

Definition 2.13. A set S is *orthonormal* if it is orthogonal and each vector has norm 1. That is, $\langle v_j, v_k\rangle = \delta_{j,k}$.

Definition 2.14. Let $U, W \subset V$. Then, U and V are orthogonal if every vector in U is orthogonal to every vector in W, and we denote $U \perp W$.

Definition 2.15. Let $U \subset V$. The orthogonal complement is $U^\perp := \{v \in V : \langle u, v\rangle = 0 \text{ for all } u \in U\}$.

2.2 Minimization Problems

The basic problem (from high school for example) is as follows. Suppose we have a data set $\{y_1, ..., y_n\}$ collected at times $\{t_1, ...t_n\}$, or (t_j, y_j). To get a straight line, choose a, b such that the line $y = a + bt$ minimizes the sum of squares $y_j - y = y_j - a - bt_j$. That is, we minimize $D^2 = \sum_1^n (y - a - bt_j)^2$. Put into linear algebra terms, let $y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$, $1 = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$ $t = \begin{bmatrix} t_1 \\ \vdots \\ t_n \end{bmatrix}$. we are looking for a, b such that $D^2 = ||y - a1 - bt||^2$ is minimzied. If we let $U = span(\{1, t\})$, we are looking for $p \in U$ such that $||y - p = \min_{u \in U} ||y - u||$.

The general problem is as follows: Given an inner product space V, a vector $v \in V$, and a subspace $U \subset V$, find $p \in U$ such that $||v - p|| = \min_{u \in U} ||v - u||$. This occurs iff there exists a $p \in U$ such that $v - p \in U^{\perp}$. When, this occurs, p is unique and $p = Pv$, where P is the orthogonal projection of v onto U. Note that in finite dimensions, this is always true. If we also have an o.n. basis for U, then Pv is given by:

$$Pv = \sum_{1}^{n} \langle v, u_j \rangle u_j.$$

This follows from Parseval's equation.

2.3 Gram-Schmidt Process

The main question is: can we find an o.n. basis for a inner product space?

In the finite dimensional case, begin with a basis $B = \{v_1, ..., v_n\}$. We will build up a orthogonal space, then normalization will give the result. Define $U_k = span(\{v_1, ..., v_k\})$. Let $w_1 = v_1$. Let $w_2 = v_2 - P_1 v_2 = v_2 - \frac{\langle v_2, w_1 \rangle}{||w_1||^2} w_1$, where P_1 is the orthogonal projection to U_1. Then, $v_2 - P_1 v_2 \in U_1^{\perp}$ (by construction), and so $w_1 \perp w_2$. We may continue this as follows, using $w_k := v_k - P_k v_k$. This give us an orthogonal basis, and we are done.

2.4 QR-Factorization

Let A be an $n \times n$ matrix with real entries and linearly independent columns $\{v_1, ..., v_n\}$. Then there exists a matrix Q whose columns form an o.n. basis for R^n and an upper triangular matrix R with positive diagonal entries such that $A = QR$.

We are working with three different bases: the standard one $E = \{e_1, ..., e_n\}$, the columns of A (they are linearly independent, there are n of them, and so are a basis) $F = \{v_1, ..., v_n\}$, and the o.n. set generated from $F : G = \{q_1, ..., q_n\}$. That is, $A = [[v_1]_E, ..., [v_n]_E]$ or the change of basis from F to E. Likewise, $Q = [[q_1]_E, ..., [q_n]_E]$, or the change of basis from G to E. Then, $Q^{-1}A = R$ is the change of basis from F to G. The properties of R follow from the Gram-Schmidt properties. In particular, we have:

$$w_k = v_k - \sum_{1}^{k-1} \frac{\langle v_k, w_j \rangle}{||w_j||^2} w_j \implies v_k = w_k + \sum_{1}^{k-1} \frac{\langle v_k, w_j \rangle}{||w_j||^2} w_j.$$

If we normalize, and wet $q_j = w_j / ||w_j||$, we arrive at:

$$v_k = ||w_k|| q_k + \sum_{1}^{k-1} \langle v_k, q_j \rangle q_j.$$

Thus, $R_{j,k} = \langle v_k, q_j \rangle$ for $j < k$, $R_k k = ||w_k||$, and $R_{j,k} = 0$ when $j > k$. This matrix is thus upper triangular.

2.5 Normed Linear Spaces

Definition 2.16. Let V be a vector space. A mapping $|| \cdot || : V \to [0, \infty)$ is a *norm* on V if:

1. Positivity: $||v|| \geq 0$, $v = 0$ iff $||v|| = 0$.
2. Positive homogenity: $||cv|| = |c| ||v||$.
3. Triangle Inequality: $||u + v|| \leq ||u|| + ||v||$.

Examples of normed linear spaces

- Any inner product space: $||v|| = \sqrt{\langle v, v \rangle}$.

- $\mathbb{R}^2$ or $\mathbb{C}^n$, with $||x||_p = (\sum_1^n |x_j|^p)^{1/p}$, where $1 \leq p < \infty$. When $p = \infty$, $||x||_\infty = \max_j |x_j|$.
- Continuous function on $[a, b]$: $||f||_C = \max_{x \in [a,b]} |f(x)|$.
- k-times continuously differentiable functions $||f||_{C^k} = \sum_{j=0}^k ||f^j||_C$.
- L^p spaces with their norms.

2.6 Norms on dual spaces

Definition 2.17. Let V be a vector space. Then, for $\phi \in V^*$:

$$||\phi||_* = \sup_{||u||_V = 1} |\phi(v)|.$$

Chapter 3

Coordinates and Bases

3.1 Coordinate Maps

(This is a bit of rehash of some of the above material.) Suppose we have a vector space with a (ordered) basis $B = \{v_1, ..., v_n\}$. If v is a vector in V, then we can uniquely represent v as a linear combination of the vectors in B. That is, there exists scalars $c_1, ..., c_n$ such that

$$v = c_1 v_1 + ... + c_n v_n.$$

The c_j's are the coordinates of v with respect to the basis B. We can collect them into a coordinate vector:

$$[v]_B = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}.$$

Since these coordinates are unique, we can define a map $K_B : V \to \mathbb{C}^n$ via $K_B(v) = [v]_B$. K_B is the coordinate map for B. It's easy to see that this map is a bijection, and so an inverse exists:

$$K_B^{-1}([v]_B) = K_B^{-1}\left(\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}\right) = v = c_1 v_1 + ... + c_n v_n.$$

Example 3.1. Let $V = P^2$ (set of polynomials of at most degree 2) and $B = \{1, x, x^2\}$. What is the coordinate vector of $[5 + 3x - x^2]_B$? What is the vector p such that $[p]_B = [3, 0, -4]^T$?

Solution 3.2.

$$[10 + x - 12x^2]_B = \begin{bmatrix} 10 \\ 1 \\ -12 \end{bmatrix}$$

$$p = K_b^{-1}([p_b]) = 3 - 4x^2.$$

3.2 Matrices for linear transformations

Consider a linear transformation $T : V \to W$, where V and dimension m and W n, and V and W have bases α and β respectively. Then we have the coordinate maps K_α and K_β respectively. Then, we have that that the matrix representation, A_T, is given by $A_T = K_\beta \circ T \circ K_\alpha^{-1}$. Then, for e_k, we have $A_T = [T(e_k)]_\beta$. Therefore, $A^T = [[T(e_1)]_\beta, ..., [T(e_m)]_\beta]$.

3.3 Changing Bases and Coordinates

Let V be a finite dimensional vector space with two bases α and β. Then, for each $v_j \in \alpha$, $v_j = c_1 w_1 + ... + c_n w_n$. . Then, we have that $[v_j]_\beta = K_\beta \circ K_\alpha^{-1}([v_j]_\alpha)$ (observe that this corresponds to the identity transformation). Thus, have that the basis transformation is of the form $A = [[v_1]_\beta, ..., [v_n]_\beta]$.

Chapter 4

Diagonalization

Definition 4.1. Let $L : V \to W$ be a linear transformation (V and W have the same dimension). A scalar λ is an *eigenvalue of* L if there exists some $x \neq 0$ such that $Lx = \lambda x$, and x is called an eigenvector associated with λ. The span of all eigenvectors corresponding to λ is called the *eigenspace* of λ and denoted by $\mathcal{E}_\lambda$.

We can write $Lv = \lambda$ as $(L - \lambda I) = 0$ and vice versa. That is, if $x \in \mathcal{E}_\lambda$, then $x \in Ker(L - \lambda I)$. Thus, the dimension of $\mathcal{E}_\lambda$ is the same as $Ker(L - \lambda I)$. This dimension is the *geometric multiplicity* of λ, denoted by γ_λ.

Definition 4.2. A matrix A is called diagonalizable if there exists a basis of V composed of eigenvectors of A.

We can diagonalize a linear transformation T by the following:

1. Fix a basis B. Find the characteristic polynomial for T, $p_T(\lambda) = \det(T - \lambda I)$ and factor it:

$$p_T(\lambda) = (\lambda_1 - \lambda)^{\alpha_1} ... (\lambda_r - \lambda)^{\alpha_r}$$

.

Here, $\lambda_1, ..., \lambda_r$ are the *distinct* roots of p_T and are the eigenvalues of T and α_j is the *algebraic multiplicity* of λ_j.

2. For each λ_j, find the eigenvectors that form a basis for the eigenspace $\mathcal{E}_{\lambda_j}$. Do so via standard row reductions. (Recall here that the geometric multiplicity is the dimension of E_{λ_j}, so there will be γ_j is the number of eigenvectors to be found.)

Then we have:

Theorem 4.3. *The linear transformation T will be diagonizable iff $\gamma_j = \alpha_j$ for all j. If so, then the matrix:*

$$S = [\mathcal{E}_{\lambda_1} basis, ..., \mathcal{E}_{\lambda_r} basis]$$

is the matrix that changes from coordintes to D, the basis of eigenvectors, to the coordinates relative to B. We also have $\lambda = S^{-1}TS$.

We want a few lemmas here:

Lemma 4.4. *Let x_i be eigenvectors corresponding to disctinct eigenvalues λ_i. Then, $\{x_1, ..., x_k\}$ is linearly independent.*

Proof. We proceed via induction. If $k = 1$, then we are done because there is one eigenvector, which is a basis. Now suppose $k = m$ works, and consider:

$$v := a_1x_1 + ... + a_{m+1}x_{m+1} = 0$$

$$T(v) = a_1\lambda_1 x_1 + ... + a_{m+1}\lambda_{m+1}x_{m+1} = 0$$

If we subtract $a_1\lambda_{m+1}x_1 + ... + a_{m+1}\lambda_{m+1}x_{m+1} = 0$, we have:

$$a_1(\lambda_1 - \lambda_{m+1}x_1 + ... + a_m(\lambda_m - \lambda_{m+1})x_m = 0$$

By assumption, we therefore have $a_j(\lambda_j - \lambda_{m+1})$, and since $\lambda_j \neq \lambda_{m+1}$, we have $a_j = 0$ for $1 \leq j \leq m$. Thus, $a_{m+1}x_{m+1} = 0$, and so $a_{m+1} = 0$. Thus we have linear independence. □

Lemma 4.5. *For each i $(1 \leq i \leq k)$, let $\{x_i, ..., x_{i,n_i}\}$ be a linearly independent set of eigenvectors corresponding to λ_i, where these are distinct eigenvalues. Then, $S = \{x_{1,1}, ..., x_{1,n_1}\} \cup ... \cup \{x_{k,1}, ..., x_{k,n_k}\}$ is linearly independent.*

Proof. Consider:

$$a_{1,1}x_{1,1} + ... + a_{1,n}x_{1,n} + ... + a_{k,1}x_{k,1} + ... + a_{k,n_k}x_{k,n_k} = 0$$

Now, each of these individual sums is itself an eigenvector (say $\hat{x_i}$, so we have $x_1 + ... + x_k = 0$. By our above lemma, this is not possible. Thus each $x_j = 0$, or

$$a_{1,1}x_{1,1} + ... + a_{1,n}x_{1,n} = 0$$

These are linearly independent, and thus $a_{i,n_i} = 0$ for all i. Thus we linear independence. □

We are now ready to prove our theorem.

Proof. Suppose T is diagonizable. Then there exists a basis consisting of eigenvectors. These are all linearly independent. If there are n disctinct eigenvalues, then we have $\gamma_j = \alpha_j$. If the eigenvalues have multiplicities greater than 1, then we must have that $\alpha_j = \gamma_j$, otherwise we would violate the above lemma.

Now we suppose that T is diagonalizable. Let E be the eigenvector B and E any other basis. Then, clearly, S satisfies the above (see earlier comments about change of basis). Then, $[T]_E = S^{-1}[T]_B^B S$. □

Note 4.6. What the above theorem says is that if an eigenvalue λ has multiplicity m, but its eigenspace has only $n < m$ basis elements, then T is *not* diagonalizable. If all eigenvalues' multiplicity equals its eigenspace dimension, then it is diagonalizable. BUT THIS MUST BE CHECKED MANUALLY.

Chapter 5

Adjoints and Self-Adjoint Operators: Fintie Dimensional Case

5.1 Adjoints

Definition 5.1. Let V, W be real or complex finite dimensional vector spaces with inner producet $\langle \cdot, \cdot \rangle_V$ and $\langle \cdot, \cdot \rangle_W$ respectively. Let $L : V \to W$ be linear. Then, the *adjoint* of L is a linear transformation $L^* : W \to V$ such that:

$$\langle Lv, w \rangle_V = \langle v, L^* w \rangle_V$$

for all $v \in V$, $w \in W$.

Proposition 5.2. Let $L : V \to W$ be linear. Then L^* exists, is unique, and is linear.

Proof. Let $\alpha = \{v_1, ..., v_n\}$ be an o.n. basis for V and $\beta = \{w_1, ..., w_,\}$ for W. We make the following observation: let $u = c_1 v_1 + ... + c_n v_n \in V$, $v = b_1 v_1 + ... + c_n v_n \in V$. Then:

$$\langle u, v \rangle = \langle \sum_1^n c_j v_j, \sum_1^n b_j v_j \rangle = \bar{b_1} c_1 + ... + \bar{b_n} c_n = \langle [u]_\alpha, [v]_\alpha \rangle_{\mathbb{C}} = [v]_\alpha^* [u]_\alpha.$$

Note that this holds for W as well. Thus we have that:

$$\langle Lu, w \rangle_W = [w]_\beta^* A_L [u]_\beta = (A_L^* [w]_\beta)^* [u]_\beta.$$

Clearly, this $A_L^* \in \mathbb{C}^{m \times n}$ exists, is unique, and sends $\mathbb{C}^m \to \mathbb{C}^n$.

Let $y = [w]_\beta$ and let $x = A_L^* y$. We can define $v = \sum_1^n x_j v_j$ so that $x = [v]_\alpha$. Thus, $[v]_\beta = A_L^* [w]_\beta$. This is thus a linear map $L^* : W \to V$. Finally, we have:

$$\langle L^* w, v \rangle_V = [v]_\alpha^* A_L^* [w]_\beta = (A_L [v]_\alpha)^* [w]_\beta = \langle w, Lv \rangle_W.$$

Then taking conjugates (and using symmetry) gives the result. □

Corollary 5.3. Let V be a finite dimensional vector space and α an orthonormal basis for V. If $L : V \to V$ is a linear transformation with matrix representation A_L, then the matrix representation of L^* is A_L^*.

5.2 Spectral Theory for Self-Adjoint Operators

Proposition 5.4. Let V be a complex vector space with an inner product. If $L : V \to V$ is a self-adjoint linear transformation, then the eigenvalues of L are real numbers and the eigenvectors of L corresponding to distinction eigenvalues are orthogonal.

Proof. Let λ be an eigenvalue of L and let x be it's corresponding eigenvector. Then, we have:

$$\lambda\langle x, x\rangle = \langle Lx, x\rangle = \langle x, Lx\rangle = \langle x, \lambda x\rangle = \bar{\lambda}\langle x, x\rangle.$$

Therefore, $(\lambda - \bar{\lambda})\langle x, x\rangle = 0$. Since $x \neq 0$ (as it is an eigenvector), we have that $\lambda = \bar{\lambda}$, and thus λ is real.

Now suppose $\lambda_1 \neq \lambda_2$ are distinct eigenvalues with x_1 and x_2 their corresponding eigenvectors. Then we have:

$$\lambda_1\langle x_1, x_2\rangle = \langle Lx_1, x_2\rangle = \langle x_1, Lx_2\rangle = \langle x_1, \lambda_2 x_2\rangle = \lambda_2\langle x_1, x_2\rangle.$$

□

Therefore, $(\lambda_1 - \lambda_2)\langle x_1, x_2\rangle = 0$. Since $\lambda_1 \neq \lambda_2$, we have that $\langle x_1, x_2\rangle = 0$. Therefore, they are orthogonal.

Lemma 5.5. *Let V be a finite dimensional vector space with dimension $n \geq 1$. If $L : v\ toV$ is linear, then L has at least one eigenvector.*

Proof. Let β be a basis for V and let A_L be the matrix representation of L with respect to that basis. Then, $\det(A_L - \lambda I) = 0$ is a polynomial in λ, it has at most n distinct roots. Thus, there is at least one eigenvalue. □

Theorem 5.6. *Let V be a complex, finite dimensional vector space. If $L : V \to V$ is a self-adjoint linear transformation, then there is an orthonormal basis for V consisting of eigenvectors of L. The matrix for L with respect to this basis is diagonal.*

Proof 1. Let $\alpha = x_1, ..., x_k$ be the eigenvectors of L and let $S = span(\alpha)$. Then $V = S \bigoplus S^{\perp}$. We claim that that $S^{\perp} = \{0\}$. To do this, we show that $S^{\perp}$ is invariant under L. Once we do that, if $S \neq \{0\}$, it has an (nonzero) eigenvector, and so we have that $S \cap S \perp \neq \{0\}$, which is a contraction. Thus, $V = S$, and so S gives rise to a basis for V. We can apply Gram-Schmidt to α to achieve an orthonormal basis.

We now prove our claim. Let $u \in S^{\perp}$, and let v_j be an eigenvector of L. Then:

$$\langle Lu, v_j\rangle = \langle u, Lv_j\rangle = \langle u, \lambda_j v_j\rangle = \bar{\lambda_j}\langle u, v_j\rangle = 0$$

Therefore, $Lu \in S^{\perp}$. Thus, we have that $L : S^{\perp} \to S^{\perp}$ is a self-adjoint operator, and so has an eigenvector. The rest follows as above. □

Proof 2 (I believe this is a better proof: the above has some details missing). We proceed by induction. Suppose $n = 1$. Then L has one eigenvalue, call it x_1, and so $V = span(\{x_1\})$. Now suppose the result holds for $k-$dimensions. Consider a self-adjoint operator $L : V \to V$, where $dim(V) = k + 1$. Again, L has at least one eigenvector, x_1. Let $S = span(\{x_1\})$. Then $V = S \bigoplus S^{\perp}$, where $dim(S^{\perp}) = k$. Now let $s \in S^{\perp}$. Then,

$$\langle Ls, x_1\rangle = \langle x, Lx_1\rangle = \langle s, \lambda_1 x_1\rangle = \bar{\lambda_1}\langle s, x_1\rangle = 0.$$

Therefore, $L : S^{\perp} \to S^{\perp}$ is a well defined self-adjiont linear operator. Since $dim(S^{\perp}) = k$, by our assumption, $S^{\perp}$ has an o.n. basis consisting of eigenvectors $\{y_1, ..., y_k\}$. Then, $\{x_1, y_1, ..., y_k\}$ is an orthogonal set, and so is linearly indepdent, and thus it spans V. Thus it is an o.n. basis consisting of eigenvectors of L.

Observe that therefore the matrix representation of L with respect to the o.n. basis is:

$$A_L = \begin{bmatrix} \lambda_1 & & 0 \\ \vdots & \ddots & 0 \\ 0 & & \lambda_n \end{bmatrix}$$

Therefore, if A_L^* is given with respect to a different basis, all we need is a change of basis matrix S from the o.n. to the given basis and we get: $\Lambda := A_L = S^{-1}A_L^*S^{-1}$. That is, $A_L^* = S\Lambda S^{-1}$. □

Note 5.7. Observe, therefore, that S has columns $[x_1, ..., x_n]$, where x_j is an eigenvector of L. It's just that S has x_j in terms of the other basis. Thus, we can make it so that S has orthogonal column vectors.

Definition 5.8. A matrix A is *orthogonal* if the columns are orthonormal.

Theorem 5.9. *Let A be an $n \times n$ complex matrix. The following are equivalent:*

1. *A is orthogonal;*
2. *A is invertible and $A^{-1} = A^*$;*
3. *$\langle Au, Av\rangle = \langle u, v\rangle$ for all u, v.*
4. *$||Au|| = ||u||$.*

Proof. Suppose 1: Then we have:

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix}$$

$$A^* = \begin{bmatrix} \overline{a_{11}} & \dots & \overline{a_{n1}} \\ \vdots & & \vdots \\ \overline{a_{1n}} & \dots & \overline{a_{nn}} \end{bmatrix}$$

Then, the ij^{th} entry of AA^* is $x_{ij} = \sum_{k=1}^{n} a_{ik}\overline{a_{kj}}$, which is the inner product of the j^{th} column of A and the i column of A. Thus, $x_{ij} = \delta_{ij}$. Thus, $AA^* = I$, and so $A^* = A^{-1}$.

Now suppose 2. Then, $\langle Au, Av\rangle = (Av)^*Au = v^*A^*Au = v^*u = \langle u, v\rangle$.

Now suppose 3. Then, $||Au||^2 = \langle Au, Au\rangle = \langle u, u\rangle = ||u||^2$. Thus, $||Au|| = ||u||$. (Later, this implies that $||A|| = 1$.) Now suppose 4. Using the first proposition from the inner product section, we have that $||Ae_i + Ae_je^{i\alpha}|| = ||e_i|| + 2|\langle Ae_i, Ae_j\rangle| + ||e_j||^2 = 2 + 2|\langle Ae_i, Ae_j\rangle|$. We also have:

$$||Ae_i + Ae_je^{i\alpha}||^2 = ||A(e_i + e_je^{i\alpha})||^2 = ||e_i + e_je^{i\alpha}||^2 = ||e_i||^2 + 2|\langle e_i, e_j\rangle| + ||e_j||^2 = 2.$$

Combining these two equations, we get that $|\langle Ae_i, Ae_j\rangle| = 0$. Thus, $\langle Ae_i, Ae_j\rangle = 0$. Thus the column vectors of A are orthogonal, and we can normalize if necessary. Thus A is orthogonal. $\square$

Note 5.10. Note, then, that S from above is orthogonal and thus satisfies the above theorem.

5.3 Courant-Fischer Theorem

Theorem 5.11. *Let A be a real $n \times n$ self-adjoint matrix having eigenvalues $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_n$. Then,*

$$\lambda_k = \min_{C \in \mathbb{R}^{(k-1)\times n}} \max_{||x||=1;\ Cy=0} x^T Ax.$$

Proof. By diagonalization, we get $x^T Ax = x^T S\Lambda S^T x$ ($S^* = S^T$ when S is real). Define $S^T x = y$. Since S is orthogonal, $||x|| = ||Sy|| = ||y||$. Furthermore, since S is invertible, CS runs over all matrices as C does. Then, we can rewrite the above as:

$$\lambda_k = \min_{C \in \mathbb{R}^{(k-1)\times n}} \max_{||y||=1,\ Cy=0} y^T \Lambda y.$$

Define $q(y) := y^T\Lambda y$. Then, $q(y) = \sum_1^n y_j\lambda_j y_j = \sum_1^n \lambda_j y_j^2$. Now consider:

$$C_0 = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & \\ 0 & 1 & 0 & & 0 & \\ \vdots & & & & & \\ 0 & \dots & 1 & \dots & 0 & \end{bmatrix}$$

where the 1 in the bottom row is in the $k-1$ spot. So C_0 looks like the identity matrix for the first $k-1$ columns then 0's. Thus, to get $C_0 y = 0$, we only need to consider $y = \sum_{j=k}^{n} y_j e_j$ such that $||y|| = 1$. Thus, $C_0 y = 0$, and

$$q(y) = \sum_{j=k}^{n} \lambda_j y_j^2 = \leq \lambda_k \sum_{1}^{n} y_j^2 = \lambda_k.$$

Now take $y = (0, ..., 1, ...0)$ where 1 is in the k-spot. Then $||y|| = 1$, $C_0 y = 0$ and $q(y) = \lambda_k$.

$$\max_{||y||=1 \ C_0 y=1} q(y) = \lambda_k.$$

Now we consider any C. If we can find some y' such that $q(y') \geq \lambda_k$, we would have $\max_C q(y) \geq \lambda_k$, then we will be done. (Will explain more on this.) So take $C \in \mathbb{R}^{(k-1)\times n}$. Now augment C to $\tilde{C}$ as follows:

$$\tilde{C} = \begin{bmatrix} C \\ e_{k+1}^T \\ \vdots \\ e_n^T \end{bmatrix}$$

Observe, therefore, that $\tilde{C}$ is an $(n-1) \times n$. Since $rank(\tilde{C}) \leq n-1$, $null(\tilde{C}) \geq 1$. Therefore, there exists a $y' \neq 0$ such that $\tilde{C} y' = 0$. Then, we have that $Cy' = 0$ and $e_j^T y' = 0 = y_j$. We can normalize y'. Then, $q(y') = \sum_1^k \lambda_j y_j' \geq \lambda_k$. Thus, $\max_{Cy=0||y||=1} q(y) \geq \lambda_k$.

To put it all together, we have a set with these maxes. One of them is λ_k that we got from C_0. But, for all C, there max is greater than λ_k. Thus the minimum over all C is λ_k, and this completes the proof. □

Theorem 5.12 (Fredholm Alternative)**.** *Let $L : V \to W$, finite dimensional vector spaces, be linear. The equation $Lv = w$ has a solution iff $w \in Null(L^*)^{\perp}$. That is $Range(L) = Null(L^*)^{\perp}$.*

Proof. Suppose there exists a $v \in V$ such that $Lv = w$. Let $u \in Null(L^*)$. Then:

$$\langle w, u\rangle_W = \langle Lv, u\rangle_W = \langle v, L^* u\rangle_V = \langle v, 0\rangle_V = 0.$$

Thus, $w \in Null(L^*)^{\perp}$, and $Range(L) \subset Null(L^*)$. Thus, if $Lv = w$ has a solution, $w \in Null(L^*)^{\perp}$.

Now suppose that $w \in Null(L^*)^{\perp}$ but $w \notin Range(L)$ (note then that $w \neq 0$). Since W is finite dimensional, we have $W = Range(L) \bigoplus Range(L)^{\perp}$. And so $w \in Range(L)^{\perp}$. Then, if $v \in V$, $Lv \in Range(L)$, so $\langle Lv, w\rangle = \langle v, L^* w\rangle$. This holds for all $v \in V$. Thus, pick v such that $L^* w = v$. Then,

$$\langle Lv, w\rangle = \langle v, L^* w\rangle = \langle L^* w, L^* w\rangle = 0.$$

Thus, $L^* w = 0$, and so $w \in Null(L^*)$. But then $w \in Null(L^*) \cap Null(L^*)^{\perp}$, so $w = 0$, a contradiction. Thus, $w \in Range(L)$, and $Lv = w$ has a soluction. Also, we have that $Null(L^*)^{\perp} \subset Range(L)$ and so $Range(L) = Null(L^*)^{\perp}$. □

Chapter 6

Banach and Hilbert Spaces

6.1 Complete Spaces

Definition 6.1. Let V be a normed vector space over either the real or complex numbers. A sequence of vectors $\{v_j\}_1^\infty$ *converges* to $v \in V$ if
$$\lim_{j\to\infty} ||v_j - v|| = 0.$$

Definition 6.2. Let V be a normed vector space over either the real or complex numbers. A sequence of vectors $\{v_j\}_1^\infty$ is *Cauchy* if for every $\epsilon > 0$, there exists and $N \in \mathbb{N}$ such that for $n, m \geq N$
$$||v_n - v_m|| < \epsilon.$$

Proposition 6.3. Every convergent sequence is Cauchy.

Proof. Let $\{v_j\}_1^\infty$ be a convergent sequence, say $v_j \to v$. Let $\epsilon > 0$. Then, there exists an N such that for all $n \geq N$, $||v_n - v|| < \epsilon/2$. So choose $n, m \geq N$. Then, by the triangle inequality:
$$||v_n - v_m|| = ||v_n + v - v - v_m|| \leq ||v_n - v|| + ||v_m - v|| < \epsilon/2 + \epsilon/2 = \epsilon.$$

Thus the sequence is Cauchy. $\square$

Example 6.4. It is not always the case that a Cauchy sequence is convergent. Consider $\mathbb{Q}$ and then let $(1.4, 1.41, 1.414, ..)$ be a sequence. This sequence goes to $\sqrt{2}$, but $\sqrt{2} \notin \mathbb{Q}$. (We know that this sequences is Cauchy because it converges in $\mathbb{R}$.) However, this sequence does not converge in $\mathbb{Q}$.

Definition 6.5. A normed vector space V is a *Banach Space* if it is complete.

Definition 6.6. A inner product space V is a *Hilbert Space* if it is complete.

Let $x = (x_1, x_2, ...)$ be a sequence. The following are examples of Banach Spaces:

1. $l^p := \{x : \ \sum_{j=1}^\infty |x_j|^p := ||x||_{l^p} < \infty\},\ 1 \leq p < \infty$;
2. $l^\infty := \{x : \ \sup_j |x_j| < \infty\}$, $||x||_{l^\infty} = \sup_j |x_j|$;
3. $c = \{x : \ \lim_{j\to\infty} x_j \text{ exists}\},\ ||x||_c = ||x||_\infty$;
4. $c_0 = \{x : \ \lim_{j\to\infty} x_j = 0\},\ ||x||_{c_0} = ||x||_\infty$;

Note that, expect for l^∞, the above spaces are separable. (Just consider Q as possible options.)
The following *function* spaces are also Banach Spaces:

1. $C[0,1],\ ||f||_C = \max_{x\in[0,1]} |f(x)|$;
2. $C^k[a,b],\ ||f||_{C^k} = \sum_{j=0}^k \sup_{x\in[0,1]} |f^j(x)|$;

3. $L^P(I),\ ||f||_p := \left(\int_I |f(x)|^p\right)^{1/p}$; item $L^\infty(I),\ ||f||_\infty(I) = \text{ess-sup}_{x\in I}|f(x)|$.

Note that $L^2,\ l^2$ are Hilbert spaces.

Proposition 6.7. The space l^∞ is a Banach Space.

Proof. Let $\{x_n\}_{n=1}^\infty$ be a Cauchy sequence in l^∞ (so we have a sequence of sequences). Let $x_n(j)$ denote the j^{th} term in the n^{th} sequence. Then, for all $\epsilon > 0$, there exists an N such that for all $n, m \geq N$ and all j,

$$|x_n(j) - x_m(j)| < \epsilon.$$

Then, $\{x_n(j)\}_{n=1}^\infty$ is Cauchy sequence of real or complex numbers, and therefore converges, say to $x(j)$. From the above inequality, fix m and let $\epsilon = 1$. Then we have $|x_n(j)| < 1 + |x_m(j)|$. Let $n \to \infty$, and we have $|x(j)| < 1 + |x_m(j)| < 1 + ||x_m||_\infty < \infty$. This holds for all j and therefore $x \in l^\infty$. All we need to show is that $x_n \to x$ in norm. From above, we have for all j:

$$|x_n(j) - x_m(j)| < \epsilon \text{ for } n, m \text{ big enough.}$$

Let $n \to \infty$. Then, we have $|x(j) - x_m(j)| < \epsilon$, which holds for all j. Thus, $||x - x_m|| < \epsilon$ for $m \geq N$. Thus, $x_n \to x$. $\square$

6.2 Continuous Functions

Proposition 6.8. $C[0,1]$ with the sup norm is complete.

Proof. Let $\{f_n\}_1^\infty$ be a Cauchy sequence in $C[0,1]$, so for every $\epsilon > 0$, there exists some $N \in \mathbb{N}$ such that for all $n, m \geq N$, $||f_n - f_m|| < \epsilon$. That is, for all x, $||f_n(x) - f_m(x)|| < \epsilon$. Thus, if fix x, $\{f_n(x)\}_1^\infty$ is a Cauchy sequence of real numbers and therefore converges. Therefore, we may define a new function:

$$f(x) := \lim_{n\to\infty} f_n(x).$$

Now we need to show that f is continuous and f_n converges to f in the sup norm. First, we show continuity.

Let $\epsilon > 0$. Then, we have

$$|f(x) - f(y)| \leq |f(x) - f_n(x)| + |f_n(x) - f_n(y)| + |f_n(y) - f(y)|.$$

From the definition of f, there exists a N_1 such that $|f(x) - f_n(x)| < \epsilon/3$ and an N_2 such that $|f_n(y) - f(y)| < \epsilon/3$. Let $N = \max\{N_1, N_2\}$. Then, since f_N is continuous, we have that there exists and $\delta > 0$ such that $|f_N(x) - f_N(y)| < \epsilon/3$ when $|x - y| < \delta$. There, we have:

$$|f(x) - f(y)| \leq |f(x) - f_N(y)| + |f_N(x) - f_N(y)| + |f_N(y) - f(y)| < \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon.$$

Thus f is continuous.

Now we show that $f_n \to f$ in the sup norm. Let $\epsilon > 0$. Then, from earlier we have that, for all $x \in [0,1]$, there exists an N such that for all $n, m \geq N$,

$$|f_n(x) - f_m(x)| < \epsilon$$

Let $m \to \infty$, and we have $|f_n(x) - f(x)| < \epsilon$. Since this holds for all x, have $||f_n - f|| < \epsilon$, and so $f_n \to f$ in the sup norm. This completes the proof. $\square$

Chapter 7

Notes on the Lebesgue Integral

7.1 Brief Introduction

Consider the function:

$$\chi(x) = \begin{cases} 1 & x \in \mathbb{R} \setminus \mathbb{Q} \cap [0,1] \\ 0 & x \in \mathbb{Q} \cap [0,1] \end{cases}.$$

This function does not have a Riemann integral because partitioning the x-axis does not work for the purposes of integration. However, The Lebesgue integral *can*. It does so by partitioning the y-axis instead of the x-axis.

Here are some of the technical details (for Narcowich– 607 does not give this treatment). Let $P = \{c \leq y_0 < y_1 < ... < y_n \leq d\}$ be a sequence of points such that the range of f is contained in $[c, d]$. Let $E_j = \{x \in [a, b] : \ y_j \leq f(x) < y_{j+1}\} = f^{-1}([y_j, y_{j+1}))$ and choose a y_j^* from each interval $[y_j, y_{j+1}]$. We then define the Lebesgue sum to be:

$$L_{P,Y^*}(f) = \sum_{j=0}^{n-1} y_j^* \mu(E_j).$$

where $\mu(E_j)$ denotes the measure or length of the set E_j and $Y^* = \{y_j^*\}_{j=0}^{n-1}$. For this sum to make sense, we need a concept of measure for more than just intervals. For example, for the first function above, $\chi^{-1}(1/2, 3/2)$ is the set of all irrationals in the interval. So we need to figure out how to measure this set (and others).

7.2 Measureable

Definition 7.1. A σ-algebra is a set $\mathcal{M} \subset \mathcal{P}(X)$ such that:

- $\mathcal{M}$ is nonempty;
- $\mathcal{M}$ is closed under complements;
- $\mathcal{M}$ is closed under countable unions.

Definition 7.2. A *measure* $\mu : \mathcal{M} \to [0, \infty)$ is a set function such that:

- $\mu([a, b]) = b - a$;
- $0 \leq \mu(A)$ for all $A \in \mathcal{M}$; (Positivity)
- If $\{A_j\}_1^\infty$ is pairwise disjoint, then $\mu(\bigsqcup_1^\infty A_j) = \sum_1^\infty \mu(A_j)$. (Countable additivity)

- $\mu(\emptyset) = 0$.

Proposition 7.3. Every countable set has measure 0.

Proof. We begin by showing that singletons have measure 0. (We will assume here the regularity of the Lebesgue measure– that is, $\mu(A) = \inf \mu(O) : \ A \subset O, \ O$ is open). Then, for every n:

$\mu(\{x\}) \le \mu((x - 1/n, x + 1/n)) = 2/n$.

Let $n \to \infty$, and the result follows for singletons. Then, if $A = x_1, x_2, ...$ we have

$$\mu(A) = \sum_1^\infty \mu(\{x_j\}) = 0.$$

$\square$

7.3 Measurable Functions

Definition 7.4. A function $f : [a, b] \to \mathbb{R}$ is *measurable* if $f^{-1}([c, d))$ is measurable. More general, $f : A \to \mathbb{R}$ is measurable if for every measurable $E \subset \mathbb{R}$, $f^{-1}(E) \subset$ is a measurable subset of A.

Note 7.5. Know that if f, g are measurable, then $af + g$ is measurable, fg and f/g are all measurable. Every continuous function is measurable. If f is continuous and the range g is in the domain of f, then $f \circ g$ is measurable, but the reserve may not hold.

Definition 7.6. Let f and g be functions. Then, $f = g$ *almost everywhere* if $\{x : g(x) \ne f(x)\}$ has measure 0.

Proposition 7.7. Suppose that A is a measurable set and $f_n : A \to \mathbb{R}$ is a sequence of measurable functions such that $f_n(x) \to f(x)$ for all x. Then f measurable.

Definition 7.8. Let A be a measurable set. A function $\phi : A \to \mathbb{R}$ is simple if the range of s has finite values.

Definition 7.9. The *characteristic function* χ_A is:

$$\chi_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}.$$

Note then that χ_A is simple.

Proposition 7.10. Let $s : A \to \mathbb{R}$ be a simple function with range $\{d_j\}_{j=1}^n$ and $n < \infty$. Denote $E_j = s^{-1}(d_j)$. Then s has the form

$$s = \sum_1^n d_j \chi_{E_j}.$$

Conversely, if s has the from above, then s is simple.

Proof. Suppose $s : A \to \mathbb{R}$ is simple, and let $t = \sum_1^n d_j \chi_{E_j}$. Since $s^{-1}(\{d_j\}) = E_j$, $A = s^{-1}(Rang(s)) = s^{-1}(\bigsqcup \{d_j\}) = \bigsqcup_1^n E_j$. Let $t : A \to \mathbb{R}$ is well defined. Now let $x \ inA$. Without loss of generality, let $x \in E_j$. Then, $s(x) = d_j = \sum_1^n d_j \chi_{E_j}(x) = t(x)$. This hold for all $x \in A$, so s is the sum.

For the other direction, simply observe that since s is well defined, $E_j \cap E_k = \emptyset$ for $j \ne k$. Thus, $A = \bigsqcup_1^n E_j$. Finally, since $n < \infty$, $|Range(s)| < \infty$, and so is simple. $\square$

7.4 The Lebesgue Integral

Definition 7.11. Let s be a simple function, $E_j = s^{-1}(d_j)$.. Then the integral of s is:

$$\int_A s d\mu = \sum_{j=1}^{n} d_j \mu(E_j).$$

Definition 7.12. Let f be a measurable function. Then, for s a simple function:

$$\int_A^+ f d\mu = \inf\{\int_A s d\mu : \ s(x) \geq f(x)\}$$

$$\int_A^- f d\mu = \inf\{\int_A s d\mu : \ s(x) \leq f(x)\}$$

Define $\int_A f d\mu = \int_A^+ f d\mu = \int_A^- f d\mu$, when they exists and are equal.

Definition 7.13. Let f be a nonnegative measurable function. Then:

$$\int_A f d\mu = \sup_s \int_A s d\mu : s(x) \leq f(x).$$

Definition 7.14. A measurable function f is *integrable* if $\int_A |f| d\mu < \infty$. When f is not nonnegative, then we define $f^+ := 1/2(f + |f|)$, $f^- = 1/2(f - |f|)$, and so $\int_A f d\mu = \int_A f^+ d\mu - \int_A f^- d\mu$.

Note 7.15. Know the following:

- If f is a bounded, measurable function, then the integral exists.
- Integreation is a linear operation.
- If $\inf_A f^2 dx$ and $\int_A g^2 dx$ both exists, then $\int_A fg dx$ and $\int_A (f+g)^2 dx$ exist.
- $\int_{A \bigsqcup B} f d\mu = \int_A f d\mu + \int_B f d\mu$.
- If $f = g$ almost everywhere, then $\int_A f d\mu = \int_B f d\mu$.
- If the Riemann integral of x exists, then the Lebesgue integral exists and the integrals are equal.

We give the following without definition.

Theorem 7.16 (Monotone Convergence Theorem)**.** *Let $\{f_j\}$ be a collection of measurable functions on A such that $0 \leq f_1 \leq f_2 \leq ... \leq f_n \leq ...$ almost everywhere. Define $f(x) := \lim_{n\to\infty} f_n(x)$. Then, f is measurable and*

$$\lim_{n\to\infty} \int_A f_n d\mu = \int_A \lim_{n\to\infty} f d\mu = \int_A f d\mu.$$

Theorem 7.17 (Dominated Convergence Theorem)**.** *Let $\{f_j\}$ be a set of measurable functions that converges pointwise to a function f. Assume there exists an integrable function g such that $|f_j(x)| \leq g(x)$ almost every where. Then, f is integrable and*

$$\lim_{n\to\infty} \int_A f_j d\mu = \int_A \lim_{n\to\infty} f_j d\mu = \int_A f d\mu.$$

Theorem 7.18 (Fubini's Theorem)**.** *Let f be a measurable function on $A \times B$. If $\int_{A\times B} |f(x,y)| d\mu(x,y) < \infty$, then $\int_A \int_B f(x,y) d\mu(x) d\mu(y)$ exists and the order of integration can be switched.*

Definition 7.19.

$$L^p([a,b]) = \{f : \ \left(\int_a^b |f|^p d\mu\right)^{\frac{1}{p}} < \infty\}.$$

Note 7.20. $L^p([a,b])$ is a normed space with the norm $||f||_p = \left(\int_a^b |f|^p d\mu\right)^{\frac{1}{p}}$.

Definition 7.21. When $p=\infty$, we define $||f||_\infty = ess\ sup|f| = \inf\{a \in \mathbb{R} : \ \mu(\{x : |f|(x) > a\}) = 0\}$.

Theorem 7.22. *The space $L^p([a,b])$ is a Banach Space. When $p=2$, it is a Hilbert Space.*

Example 7.23. Let $f_n(x) = \dfrac{1}{\sqrt{1+\frac{1}{n}}}$ for $x \in [0,1]$, $n \geq 1$. Let f be the pointwise limit of f_n. Show that f is integrable.

Proof. Obsevrse that $\dfrac{1}{\sqrt{x+\frac{1}{n}}} \leq \dfrac{1}{\sqrt{1+\frac{1}{n+1}}}$, so that $f_n(x) \leq f_n(x)$ for all x. Thus, by MCT:

$$\lim_{n\to\infty} \int_0^1 f_n(x)dx = \int_0^1 \lim_{n\to\infty} \frac{1}{\sqrt{x+\frac{1}{n}}} dx = \int_0^1 f dx = \int_0^1 \frac{1}{\sqrt{x}} dx < \infty.$$

□

Example 7.24. Let $f \in L^1(\mathbb{R})$, define $\hat{f}(w0 = \int_{-\infty}^{\infty} f(t)e^{iwt}dt$. Show that $\hat{f}$ is continuous.

Proof. By computation, we get:

$$\begin{aligned}
\hat{f}(x+1/n) - f(x) &= \int_{-\infty}^{\infty} f(t)e^{-i(w+1/n)t}dt - \int_{-\infty}^{\infty} f(t)e^{iwt}dt \\
&= \int_{-\infty}^{\infty} f(t)\left(e^{-iwt-i(1/n)t} - e^{-iwt}\right)dt \\
&= \int_{-\infty}^{\infty} f(t)e^{-iwt}(e^{-i(1/n)t} - 1)dt.
\end{aligned}$$

Let $g_n(t) = f(t)e^{-iwt}(e^{-i(1/n)t} - 1)$. Observe that $g_n(t) \to 0$ pointwise.
Then, $|g_n(t)| \leq |f(t)e^{-iwt}(e^{-i(1/n)t} - 1)| \leq |f(t)||e^{-i(1/n)t} - 1| \leq 2|f(t)|$. Since $2|f(t)| \in L^1(\mathbb{R})$, we can apply DCT:

$$\lim_{n\to\infty} \hat{f}(w+1/n) - \hat{f}(w) = \lim_{n\to\infty} \int_{-\infty}^{\infty} g_n(t)dt = \int_{-\infty}^{\infty} 0dt = 0.$$

Therefore $\lim_{h\to 0} \hat{f}(w+h) = \hat{f}(w)$, and $\hat{f}$ is continuous. □

Chapter 8

Orthonormal Sets

In this section, let $\mathcal{H}$ be a Hilbert space and $V \subset \mathcal{H}$ be a subspace. Also, let $\{u_j\}_1^\infty$ be an orthonormal basis for $\mathcal{H}$. We first give a brief lemma.

Lemma 8.1. *Let $\{u_j\}_1^\infty$ be an o.n. set. Then $||\sum_1^\infty \alpha_j u_j||^2 = \sum_1^\infty |\alpha_j|^2$.*

Proof.

$$\begin{aligned} ||\sum_1^\infty \alpha_j u_j||^2 =& \langle \sum_1^\infty \alpha_j u_j, \sum_1^\infty \alpha_k u_k \rangle \\ &= \sum_1^\infty \sum_1^\infty \alpha_j \overline{\alpha_k} \langle u_j, u_k \rangle \\ &= \sum_1^\infty \sum_1^\infty \alpha_j \overline{\alpha_k} \delta_{j,k} \\ &= \sum_1^\infty \alpha_j \overline{\alpha_j} \\ &= \sum_1^\infty |\alpha_j|^2. \end{aligned}$$

□

Recall from earlier, that $p \in V$ minimizes $||f - v||$ for $v \in V$ iff $f - p \in V^\perp$. When such a p exists, there is a projection P_V such that $P_V f = p$ (since p is unique). When V is finite dimensional, such a p always exsits, and so $P_V : \mathcal{H} \to V$ is well defined.

We want to verify that the projection does in fact minimize over V. So, if we let $\{u_1, ..., u_n\}$ be an o.n. basis for V, then we have that $P_V f = \sum_1^n \langle f, u_j \rangle u_j$. Then, by the Pythagorean theorem, we have:

$$||f - P_v f||^2 = ||f||^2 - ||P_V f||^2 = ||f||^2 - \langle \sum_1^n \langle f, u_j \rangle u_j, \sum_1^n \langle f, u_i \rangle u_i \rangle = ||f||^2 - \sum_1^n |\langle f, u_j \rangle|.$$

Now suppose $v \in V$ is given by $v = \sum_1^n \alpha_j u_j$, where α_j are arbitrary. By the Pythagorean Theorem, we know that for any $u \in \mathcal{H}$, we have: $||u||^2 = ||P_V u||^2 + ||(I - P_v)u||^2$. Thus, for $v \in v$, we have $f - v \in \mathcal{H}$, $||f - v||^2 = ||P_V(f - v)||^2 + ||(I - P_V)(f - v)||^2 = ||P_v f - v||^2 + ||f - P_v f||^2$. Then, from earlier, we deduce $||f - v||^2 = ||f||^2 - \sum_1^n |\langle f, u_j \rangle|^2 + \sum_1^n |\alpha_j - |\langle f, u_j \rangle|^2$. Clearly, this is minimized when $\alpha_j = |\langle f, u_j \rangle$, that is, when v is the projection.

Naturally, if we have an o.n. set $\{u_j\}_1^\infty$ would want to know if we can represent any $f \in \mathcal{H}$ as $\sum_1^\infty \langle f, u_j \rangle u_j$. Unfortunately, this is not always the case. When it does occur, we have the following:

Definition 8.2. Let $\{u_j\}_1^\infty := U$ be a o.n. set. If every vector in $\mathcal{H}$ can be written as $\sum_1^\infty \langle f, u_j \rangle u_j$, then U is *complete*. (Note this is different than Cauchy completeness.)

Proposition 8.3. The following are equivalent:

1. Every vector in $\mathcal{H}$ may be uniquely represented as the series $f = \sum_1^\infty \langle f, u_j \rangle u_j$.
2. U is maximal in the sense that there is no non-zero vector in $\mathcal{H}$ that is orthogonal to U. (Equivalently, U is not a proper subset of any other o.n. set in $\mathcal{H}$.

Proof. For 1) implies 2), we proceed by contradiction. Suppose every vector can be represented as it's projection, and suppose there exists and o.n. set W such that $U \subsetneq W$. Let $w \neq 0 \in W$. Then, $w = \sum_1^\infty \langle w, u_j \rangle u_j$. But then $w = \sum_1^\infty \langle w, u_j \rangle u_j = 0$. For 2) implies 1 we proceed via contrapositive. Suppose f cannot be represented by it's projection. Then, $f - Pf \neq 0$. Computation shows:

$$\langle f - Pf, u_j \rangle = \langle f, u_j \rangle - \langle Pf, u_j \rangle = \langle f, u_j \rangle - \sum_1^\infty \langle f, u_k \rangle u_k u_j = \langle f, u_j \rangle - \langle f, u_j \rangle = 0.$$

This holds for all j, so $g := f - Pf \in U^\perp$. Thus, $U \cup \{g\}$ is an o.n. basis (we can normalize g if necessary), and so U is no longer maximal. □

Theorem 8.4 (Bessel's Inequality)**.** *Let $\{u_j\}_1^\infty$ be an o.n. set. Then, for all $f \in \mathcal{H}$,*

$$\sum_1^\infty |\langle f, u_j \rangle|^2 \leq ||f||^2.$$

Proof. For each n, $U_n := \{u_1, ..., u_n\}$ is a finite set, so a projection exists, and thus from the above (Pythagorean theorem), $0 \leq ||f - P_{U_n} f|| = ||f||^2 - \sum_1^n |\langle f, u_j \rangle|^2$, thus $\sum_1^n |\langle f, u_j \rangle|^2 \leq ||f||^2$. This holds for all n, and so the partial sums form a bounded monotone increasing sequence, and thus converges. Therefore, $\lim_{n\to\infty} \sum_1^n |\langle f, u_j \rangle|^2 = \sum_1^\infty |\langle f, u_j \rangle|^2 \leq ||f||^2$. □

Theorem 8.5 (Parseval's Equation)**.** *An o.n. set U is complete iff $||f||^2 = \sum_1^\infty |\langle f, u_j \rangle|^2$ for all $f \in \mathcal{H}$.*

Proof. If U is complete, then $f = \sum_1^\infty \langle f, u_j \rangle u_j$. Then, $||f - \sum_1^\infty \langle f, u_j \rangle u_j|| = 0$ implies that $||f|| \leq ||\sum_1^\infty \langle f, u_j \rangle u_j||$. Thus, $||f||^2 \leq ||\sum_1^\infty \langle f, u_j \rangle u_j||^2 = \sum_1^\infty |\langle f, u_j \rangle|^2$. This coupled with Bessel's Inequality gives Parseval's Inequality.

Now suppose Parseval's Equality holds. Then, we have:

$$\lim_{n\to\infty} ||f - \sum_1^n \langle f, u_j \rangle u_j||^2 = ||f||^2 - \lim_{n\to\infty} \sum_1^n |\langle f, u_j \rangle|^2 = ||f||^2 - \sum_1^\infty |\langle f, u_j \rangle|^2 = 0$$

Thus, $f = \sum_1^\infty \langle f, u_j \rangle u_j$, and U is complete. □

Definition 8.6. Let $\{u_j\}_1^\infty := U$ be an o.n. set. Then, $\mathcal{H}_U = \{g \in \mathcal{H} : \ g = \sum_1^\infty \alpha_j u_j\}$.

Note 8.7. Observe that for $g \in U$, $\langle g, u_j \rangle = \langle \sum_1^\infty \alpha_k u_k, u_j \rangle = \alpha_j$, so the α's are uniquely given (as exactly what we would expect).

Proposition 8.8. $\mathcal{H}_U$ is a closed subspace of $\mathcal{H}$.

Proof. Let $\{g_k\}_1^\infty$ be a sequence in $\mathcal{H}_U$ such that $g_k \to g$. Let P_n denote the projection to $span(\{u_1, ..., u_n\})$. Observe that $g \in \mathcal{H}_U$ iff $g = \lim_{n\to\infty} P_n g$. Thus we show the second portion. Let $\epsilon > 0$. Then, there exists an L such that $||g - g_L|| < \epsilon/2$ Then, for g_L, (since $g_L \in \mathcal{H}_U$, there exists an N such that $||g - P_N g|| < \epsilon_2$. Then, by properties of projections:

$$||g - P_N g|| \leq ||g - P_N g_L|| \leq ||g - g_L|| + ||g_L - P_N g_L|| < \epsilon/2 + \epsilon/2 = \epsilon.$$

Therefore $\lim_{n\to\infty} P_n g = g$, and we are done. □

Corollary 8.9. An o.n. set U is complete iff $\mathcal{H}_U = \mathcal{H}$.

Proof. Suppose U is complete, so every $f \in \mathcal{H}$ can be given by its projection. Thus $\mathcal{H} = \mathcal{H}_U$. Now suppose $\mathcal{H}_U = \mathcal{H}$. Then, every $f \in \mathcal{H}$ can be given by its projection, and so U is complete. □

Proposition 8.10. Let D be a dense subset of $\mathcal{H}$. An o.n. set U is complete iff $D \subset \mathcal{H}_U$.

Proof. Suppose U is complete. Then $\mathcal{H}_U = \mathcal{H}$ and so $D \subset \mathcal{H}_U$. Now suppose $D \subset \mathcal{H}$. Since D is dense, every $f \in \mathcal{H}$ is a limit point of a sequence in D. Since $\mathcal{H}_U$ is closed, this implies that the limit is in $\mathcal{H}_U$. Therefore, $\mathcal{H} = \mathcal{H}_U$. □

Corollary 8.11. Let D be a dense subset of $\mathcal{H}$. An o.n. set U is complete iff every vector $f \in D$ can be written as $f = \sum_1^\infty \alpha_j u_j$.

Proof. If U is complete, then the consequence follows immediately. If f can be represented as the sum, then $D \subset \mathcal{H}_U$, and so the result follows from the above proposition. □

8.1 Orthogonal Polynomials

In this section, we focus primarily on $L^2[a, b]$. Also note that we are given that the continuous function are dense in L^2.

Proposition 8.12. The polynomials are dense in L^2.

Proof. Let $f \in L^2$, $\epsilon > 0$. Then, there exists a continuous g such that $||f - g||_2 < \epsilon/2$. Then, since the polynomials are dense in the continuous functions, there exists a polynomial p such that $||g - p||_u < \epsilon/(2\sqrt{b-a})$. Then, we have that:

$$||g - p||_2 = \left(\int_a^b |g - p|^2 dx \right)^{\frac{1}{2}} \leq \left(\int_a^b \left(\frac{\epsilon}{2\sqrt{b-a}} \right) dx \right)^{\frac{1}{2}} = \frac{\epsilon}{2\sqrt{b-a}} \sqrt{b-a} = \epsilon/2.$$

Therefore, $||f - p||_2 \leq ||f - g||_2 + ||g - p||_2 < \epsilon/2 + \epsilon/2 = \epsilon$.

This completes the proof. □

Definition 8.13. The *Legendre Polynomials* are the polynomials generated by using Gram-Schimdt on $\{1, x, x^2, ...\}$ in $L^2[-1, 1]$.

Corollary 8.14. The Legendre Polynomials form a complete set in $L^2[-1, 1]$.

Proof. Let $D = span(\{1, x, x^2, ...\})$ or the polynomials. This set is dense by the above proposition. Then, we can $P = \{p_1, p_2, ...\}$, the Legendre Polynomials. This set is orthonormal by construction. Since this set is generated by orthonomalizing the basis for polynomials, it preserves the span of that basis. Therefore, $D \subset \mathcal{H}_P$. Thus, by the above proposition, P is complete. □

Chapter 9

Approximation of Continuous Functions

9.1 Modulus of Continuity

Consider $f \in C[0,1]$. Since $[0,1]$ is compact, we have that f is uniformly continuous. Thus, for every $\epsilon > 0$, there exists a $\delta > 0$ such that $|x-y| < \delta$ implies $|f(x)-f(y)| < \epsilon$. One way to think about it is, if f is uniformly continuous, if I give you a $\epsilon > 0$, you give me $\delta > 0$ that satisfies the definition. But what we switched it; what if I gave you a $\delta > 0$ and asked for an $\epsilon > 0$? This is the modulus of continuity.

Definition 9.1. Let $f \in C[0,1]$ and $\delta > 0$. The the *modulus of continuity* is:

$$\omega(f,\delta) = \sup\{|f(x)-f(y)| : \ |x-y| < \delta, \ x,y \in [0,1]\}.$$

Example 9.2. Let $f(x) = \sqrt{x}$, $0 \le x \le 1$. Then $\omega)f,\delta) = \sqrt{\delta}$.

Proof. Let $\delta > 0$. Then, wlog, let $x < y$. Then:

$$\begin{aligned}
0 < |f(y)-f(x)| = |\sqrt{y}-\sqrt{x}| &= \sqrt{y}-\sqrt{x} \\
&= \frac{y-x}{\sqrt{y}+\sqrt{x}} \\
&= \frac{\sqrt{y-x}\sqrt{y-x}}{\sqrt{y}+\sqrt{x}} \\
&= \sqrt{y-x}\frac{\sqrt{1-\frac{x}{y}}}{1+\sqrt{\frac{x}{y}}} \\
&= \sqrt{y-x}\frac{1-\sqrt{\frac{x}{y}}}{1+\sqrt{\frac{x}{y}}} \\
&\le \sqrt{y-x} \le \sqrt{\delta}.
\end{aligned}$$

Therefore, $\omega(f,\delta) \le \sqrt{\delta}$. If we take $x = 0$ and $y = \delta$, we hget $|f(y)-f(x)| = \dfrac{\delta}{\sqrt{\delta}}\sqrt{\delta} = \sqrt{\delta}$. Therefore, $\omega(f,\delta) = \sqrt{\delta}$. $\square$

Example 9.3. Suppose $f \in C^1[0,1]$. Then $\omega(f,\delta) \le \delta||f||_\infty$.

Proof. By the Fundamental Theorem of Calculus, $f(y) - f(x) = \int_x^y f'(t)dt$. Therefore:

$$\begin{aligned}
|f(y) - f(x)| = \left| \int_x^y f'(t)dt \right| &\leq \int_x^y |f'(t)|dt \\
&\leq ||f'||_\infty \int_x^y 1dt \\
&= ||f'||_\infty |y - x| \\
&< ||f'||_\infty \delta.
\end{aligned}$$

□

9.2 Approximation with Linear Splines

Suppose we have a continuous function f in $[0, 1]$. One way to approximate it would be to take a set of sample point $\{x_j\}_1^n$ and plot f at those points (so we get $\{f(x_j)\}_1^n$). Then we simply connect the dots with straight lines. This gives us a piecewise linear function (call it g for now) where $g(x_j) = f(x_j)$ for all j.

Formally, to construct a linear spline we begin with a *knot sequence*.

Definition 9.4. A *knot sequence* is a finite partition of $[0, 1]$ such that $\{x_0 = 0 < x_1 < ... < x_n = 1\} := \Delta$.

Definition 9.5. A *linear spline* on a knot sequence Δ is the set of all piecewise linear functions that are continuous on $[0, 1]$ that may have corner at the knots, but no where else. That is, is linear between knot points.

Proposition 9.6. Let $f \in C[0, 1]$ and let $\Delta = \{x_0 = 0 < x_1 < .... < x_n = 1\}$ be a knot sequence with norm $||\Delta|| = \max |x_j - x_{j+1}|$. If s_f is the linear spline that interpolates f, then $||f - s_f|| \leq \omega(f, ||\Delta||)$.

Proof. We will consider the interval $I_j = [x_j, x_{j+1}]$ (the rest follows easily). It can easily be found (draw a graph and note what constants must be used to ensure $s_f(x_j) = f(x_j)$) that, on I_j

$$s_f(x) = \frac{f(x_j)(x - x_{j+1}}{x_j - x_{j+1}} + \frac{f(xj + 1}{x_{j+1} - x_j}.$$

Also observe (flip the sign in the second term)

$$\frac{x - x_{j+1}}{x_j - x_{j+1}} = \frac{x - x_j}{x_{j+1} - x_j} = 1.$$

Finally, note that $\omega(f, \delta)$ is an increasing function (of δ). Then, we have (setting $\delta_j = x_{j+1} - x_j$):

$$\begin{aligned}
f(x) - s_f(x) =& f(x)1 - s_f(x) \\
&= f(x)\left(\frac{x - x_{j+1}}{x_j - x_{j+1}} + \frac{x - x_j}{x_{j+1} - x_j}\right) - \frac{f(x_j)(x - x_{j+1}}{x_j - x_{j+1}} - \frac{f(x_{j+1}(x - x_j}{x_{j+1} - x_j} \\
&= (f(x) - f(x_j))\left(\frac{x - x_{j+1}}{x_j - x_{j+1}}\right) + (f(x) - f(x_{j+1}))\left(\frac{x - x_j}{x_{j+1} - x_j}\right) \qquad =
\end{aligned}$$

Therefore:

$$\begin{aligned}
|f(x) - s_f(x)| &\leq |f(x) - f(x_j)| \left|\frac{x - x_{j+1}}{x_j - x_{j+1}}\right| + |f(x) - f(x_{j+1}| \left|\frac{x - x_j}{x_{j+1} - x_j}\right| \\
&< \omega(f, \delta_j) \left|\frac{x - x_{j+1}}{x_j - x_{j+1}}\right| + \omega(f, \delta_j) \left|\frac{x - x_j}{x_{j+1} - x_j}\right| \\
&\leq \omega(f, ||\Delta||) \left(\frac{x - x_{j+1}}{x - x_{j+1}} + \frac{x - x_j}{x_{j+1} - x_j}\right) \\
&= \omega(f, ||\Delta||).
\end{aligned}$$

This completes the proof. $\square$

9.3 The Weierstrass Approximation Theorem

When it comes to analysis, the Weierstrass Approximation Theorem is a pillar. It says that, for any continuous function, we get find a polynomial that "hugs" that function as close as we want. Take the coast of the United States. That certainly is a continuous function (pick a coast). But we can find a polynomial that resembles the coast of the US.

There are several ways to prove this theorem. One way is via the Stone-Weierstrass Theorem, which is the big brother version which makes this one a mere corollary. But our approach is constructive, whereas the other one merely gives existence. Our construction consists of the *Bernstein Polynomials.*

Definition 9.7. Let $n \in \mathbb{N}$. Then, the j^{th} *Bernstein Polynomial* is:

$$\beta_{j,n}(x) := \binom{n}{j} x^j (1-x)^{n-j}.$$

Observe that the Binomial theorem $(x+y)^n = \sum_1^n \binom{n}{j} x^j y^{n-j}$ gives the terms for the Bernstein polynomials when $y = 1-x$. These polynomials indeed form a basis for regular polynomials.

Proposition 9.8. The Bernstein polynomials $\Lambda := \{\beta_{j,n}(x)\}_{j=0}^n$ form a basis for $\mathcal{P}_n$.

Proof. The dimension of $\mathcal{P}_n$ is $n+1$, and we have $n+1$ Bernstein polynomials. We only need to show that Λ spans the polynomials. We do so by showing that each of the standard basis elements can be given by a linear combination of Bernstein polynomials. That, is we show x^k for $k = 0, ..., n$ can be given as linear combination of Bernstein polynomials. So let $0 \leq k \leq n$. By the Biomial Theorem, we have:

$$(x+y)^n = \sum_{j=-}^{n} \binom{j}{j} x^j y^{n-j}.$$

Taking the k^{th} x-partial derivative, we have:

$$n(n-1)...(n-k+1)(x+y)^n = \sum_{j=1}^{n} \binom{n}{j} j(j-1)...(j-k+1) x^{j-k} y^{n-j}.$$

Setting $y = 1-x$ gives and multiplying by x^K gives:

$$x^k n(n-1)...(n-k+1) = \sum_{j=0}^{n} \binom{n}{j} j(j-1)...(j-k+1) x^j (1-x)^{n-j} = \sum_{j=0}^{n} j(j-1)...(j-k+1) \beta_{j,n}(x).$$

Therefore:

$$x^k = \sum_{j=0}^{n} \frac{(n-k)!j!}{(j-k)!n!} \beta_{j,n}(x).$$

This completes the proof. $\square$

Before we proceed, we make the following observations. First, Bernstein polynomials are positive except at $x = 1, 0$ where it zeros out. Second, a simple min/max argument shows that $x = \frac{j}{n}$ gives the max value for $\beta_{j,n}(x)$. Finally, we have:

$$1 = \sum_{j=0}^{n} \beta_{j,n}(x)$$
$$x = \sum_{j=0}^{n} \frac{j}{n}\beta_{j,n}(x)$$
$$\frac{1}{n}x + (1-\frac{1}{n})x^2 = \sum_{j=0}^{n} \frac{j^2}{n^2}\beta_{j,n}(x).$$

The first two come from direct calculation, simpler than the one in the proof above. The third is more tricky. Now we are ready for Weierstrass.

Theorem 9.9 (Weierstrass Approximation Theorem)**.** *Let $f \in C[0,1]$. Then, for every $\epsilon > 0$, there exists a polynomial p such that $||f-p||_u < \epsilon$.*

Proof. Let $f \in C[0,1]$ and define:

$$f_n(x) := \sum_{j=0}^{n} f(j/n)\beta_{j,n}(x).$$

Observe that :

$$f(x) = f(x)\cdot 1 = \sum_{j=0}^{n} f(x)\beta_{j,n}(x).$$

Now let $\delta > 0$ and n be large. Then:

$$E_n(x) := \sum_{j=0}^{n} \left(f(x) - f(j/n)\right)\beta_{j,n}(x).$$

We now want to split this sum by j, whether or not $|x-j/n| \le \delta$ for $|x-j/n|| > \delta$. To do so, define:

$$F_n(x) = \sum_{|x-j/n|\le\delta} \left(f(x) - f(j/n)\right)\beta_{j,n}(x);$$
$$G_n(x) = \sum_{|x-j/n|>\delta} \left(f(x) - f(j/n)\right)\beta_{j,n}(x).$$

Naturally, since this runs over j, we have $E_n(x) = F_n(x) + G_n(x)$. The idea, then, is to bound both of these terms by some version of n and them make the bound as small as we please by making n large. We focus on $F_n(x)$ first. Since $\beta_{j,n}(x) \ge 0$, we have:

$$\begin{aligned}
|F_n(x)| &\le \sum_{|x-j/n|\le\delta} |f(x) - f(j/n)|\beta_{j,n}(x) \\
&\le \sum_{|x-j/n|\le\delta} \omega(f,\delta)\beta_{j,n}(x) \\
&\le \omega(f,\delta)\left(\sum_{j=0}^{n} \beta_{j,n}(x)\right) \\
&= \omega(f,\delta).
\end{aligned}$$

We not turn to $G(x)$. Suppose $x - j/n > \delta$ (the other case follows similarly). Then, there is a $k \in \mathbb{N}$ such that $k\delta < x - j/n < (k+1)\delta$. We can telescope $f(x) - f(n/j)$ into:

$$f(x) - f(n/j) = [f(x) - f(n/j + k\delta)] + [f(j/n + k\delta) - f(j/n + (k-1)\delta)] + ... + [f(j/n + \delta) - f(j/n)].$$

We therefore have:

$$f(x) - f(j/n) = [f(x) - f(j/n + k\delta)] + \sum_{m=0}^{k-1} [f(j/n + (k-m)\delta) - f(j/n + (k-m-1)\delta)].$$

Since $j/n + (k-m)\delta - (j/n + (k-m-1)\delta) = \delta$ and $x - j/n - k\delta < \delta$ (see bounds), we have that

$$|f(x) - f(n/j)| \leq (k+1)\omega(f, \delta).$$

Again, using the bounds, we get that $k + 1 < 1 + |x - j/n|\delta^{-1}$.. Thus:

$$|f(x) - f(j/n)| \leq (1 + |x - j/n|\delta^{-1})\omega(f, \delta).$$

Now, since $k \geq 1$, we get $\delta < k\delta < |x - j/n|$, and so $1 < |x - j/n|\delta^{-1}$. Thus, $|x - j/n|\delta^{-1} < |x - j/n|^2\delta^{-2}$. Thus,

$$|f(x) - f(j/n)| < \left(1 + \frac{|x - j/n|^2}{\delta^2}\right)\omega(f, \delta).$$

And:

$$\begin{aligned} |G_n(x)| &\leq \sum_{|x-j/n|>\delta} |f(x) - f(j/n)|\beta_{n,j}(x) \\ &\leq \sum_{j=0}^{n} \left(1 + \frac{|x - j/n|^2}{\delta^2}\beta_{j,n}(x)\right)\omega(f, \delta) \\ &\leq \sum_{j=0}^{n} \left(\left(1 + \frac{x^2}{\delta^2} - \frac{2xj}{\delta^2 n} + \frac{j^2}{n^2\delta^2}\right)\beta_{j,n}(x)\right)\omega(f, \delta). \end{aligned}$$

Distributing the $\beta_{j,n}(x)$ over the summand and then breaking the sum over the parts and using the identifies, we get that:

$$|G_n(x)| < (1 + \frac{x - x^2}{\delta_n^2}\omega(f, \delta).$$

Since $x - x^2 < 1/4$ on $[0, 1]$, we have (independent of x!):J

$$|E_n(x)| \leq |F_n(x)| + |G_n(x)| = (2 + \frac{1}{4}\frac{1}{\delta^2 n})\omega(f, \delta).$$

If we take $\delta = \frac{1}{\sqrt{n}}$, (and $2 + 1/4 \cdot 1/n^2 \leq 1$) we get:

$$|E_n(x)| \leq (2 + \frac{1}{4}\frac{1}{n^2})\omega(f, n^{-1/2}) \leq (2 + 1/4)\omega(f, n^{-1/2}) = 9/4\omega(f, n^{-1/2}).$$

Since modulus of continuity is increasing, we can choose n so large tht $9/4\omega(f, n^{-1/2}) < \epsilon$. This completes the proof.

$\square$

Chapter 10

Pointwise Convergence of Fourier Series

Definition 10.1. Let $f \in L^1[-\pi, \pi]$. The the *Fourier Series* of f is:

$$\sum_{-\infty}^{\infty} a_n e^{inx}, \text{ where } a_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-int}dt.$$

Our goal of this section to show that if f is piecewise continuous, then its Fourier Series equals f at points of continuity and is half way between the points of discontinuity (jump dicontinuities). To do so, we begin by massaging the partial sums of the series. So, denote:

$$\begin{aligned} S_N(x) &= \sum_{-N}^{N} a_n e^{inx} \\ &= \sum_{N}^{N} \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-int}dte^{inx} \\ &= \int_{-\pi}^{\pi} \frac{1}{2\pi}\sum_{-N}^{N} f(t)e^{in(x-t)}dt \end{aligned}$$

Make a change of variable: $u = t - x$, and get:

$$S_N(x) = \int_{-\pi-x}^{\pi-x} f(u+x)D_N(u)du.$$

Next, we define $D_N(u) := \frac{1}{2\pi}\sum_{-N}^{N} e^{inu}$, so we have $S_N(x) = \int_{-\pi}^{\pi} D_N(x-t)f(t)dt$. ($D_N(u)$ is called the *Dirichlet Kernel*). We have some properties about $D_N(u)$.

Proposition 10.2. The Dirichlet Kernel Satisfies:

1. $D_N(u) = \frac{1}{2\pi} + \frac{1}{\pi}\sum_{n=1}^{N}\cos(nu)$;
2. $D_N(u)$ is even and 2π-periodic;
3. $\int_{-\pi}^{\pi} D_N(u)du = 1$ and $\int_0^{\pi} D_N(u)du = 1/2$.

Proof. 1) $D_N(u) = \frac{1}{2\pi}\sum_{-N}^{N} e^{inu}$. Recall that $e^{inu} = \cos(nu) + i\sin(nu)$. Then:

$$\sum_{-N}^{N} e^{inu} = \sum_{-N}^{N} \cos(nu) + i\sin(nu)$$
$$= \cos(-Nu) + i\sin(-Nu) + ... + \cos(-u) + i\sin(-u) + 1 + + \cos(u) + i\sin(u) + ... + \cos(Nu) + i\sin(Nu)$$
$$= 1 + \cos(u) + ... + 2\cos(Nu).$$

Thus, $D_N(u) = \frac{1}{2\pi}\sum_{-N}^{N} e^{inu} = \frac{1}{2\pi}\left(1 + \sum_1^N 2\cos(nu)\right) = \frac{1}{2\pi} + \frac{1}{\pi}\sum_1^N \cos(nu)$.

2) Since $\cos(u)$ is even and 2π periodic, we have, by 1, that $D_N(u)$ is even and 2π periodic.

3) $\int_{-\pi}^{\pi} D_N(u)du = 2\int_0^{\pi} D_N(u)du$ by eveness. Then:

$$\int_0^{\pi} D_N(u)du = \int_0^{\pi} \frac{1}{2\pi} + \frac{1}{\pi}\sum_1^N \cos(nu) = \int_0^{\pi} \frac{1}{2\pi}du + \frac{1}{\pi}\sum_1^N \int_0^{\pi} \cos(nu) = \frac{1}{2} + \frac{1}{\pi}\sum_1^N (\frac{1}{n}\sin(n\pi) - \frac{1}{n}\sin(0) = \frac{1}{2}$$

.

4)

$$2\pi D_N(u) = \sum_{-N}^{N} e^{inu}$$
$$= e^{-iNu} + ... + e^{-iu} + 1 + e^{iu} + ... + e^{iNu}$$
$$= e^{-iNu}\left(1 + e^{iu} + e^{2iu} + ... + e^{2iNu}\right)$$
$$= e^{-iNu}\frac{e^{(2N+1)u-1}}{e^{iu}-1}$$

Multiply top and bottom by $\frac{e^{iu/2}}{2}$. The numerator is:

$$\frac{e^{(N+1/2)ui} - e^{-i(N+1/2)u}}{2} = \sin((N+1/2)u).$$

The denominator is:

$$\frac{e^{iu/2} - e^{-iu/2}}{2} = \sin(iu/2).$$

Thus:

$$2\pi = \frac{\sin((N+1/2)u}{\sin(iu/2)}.$$

□

Lemma 10.3. *Let g be a 2π periodic function that is integrable on each bounded interval in $\mathbb{R}$. Then $\int_{-\pi+c}^{\pi+c} g(u)du$ is independent of c. In particular, $\int_{-\pi}^{g}(u)du = \int_{-\pi+c}^{\pi+c} g(u)du$.*

Proof. We will show that $\int_{-\pi+c}^{\pi+c} g(u)du = \int_{-\pi}^{\pi} g(u)du$. Since g is 2π periodic, $g(-\pi+c) = g(\pi+c)$ and $\mu([-\pi, -\pi+c]) = \mu([\pi, \pi+c])$:

$$\int_{-\pi}^{-\pi+c} gdu = \int_{\pi}^{\pi+c} gdu.$$

Thus:

$$\int_{-\pi+c}^{\pi+c} gdu = \int_{-\pi+c}^{\pi} gdu + \int_{\pi}^{\pi+c} gdu = \int_{-\pi+c}^{\pi} gdu + \int_{\pi}^{-\pi+c} gdu = \int_{-\pi}^{\pi} gdu.$$

Thus, for constants $c,\ d$:

$$\int_{-\pi+d}^{\pi+d} gdu = \int_{-\pi}^{\pi} gdu = \int_{-\pi+c}^{\pi+c} gdu.$$

□

Now we can finished our form of $S_N(x)$. First, since $D_N(u)$ and $f(x+u)$ are 2π periodic, the lemma implies:

$$S_N(x) = \int_{-\pi-x}^{\pi-x} f(u+x)D_N(u)du = \int_{-\pi}^{\pi} f(u+x)D_N(u)du.$$

We make another change of variable: u to $-u$. Changing the bounds of integration and the variable of integration, we have:

$$S_N(x) = -\int_{\pi}^{-\pi} f(x-u)D_N(-u)du = \int_{-\pi}^{\pi} f(x-u)D_N(u)du.$$

At these this equation and the one above, divide by 2, and we get:

$$2S_N(x) = \int_{-\pi}^{\pi} f(u+x)D_N(u) + f(x-u)D_N(u)du = \int_{-\pi}^{\pi} \left(f(u+x) + f(x-u)\right)\frac{D_N(u)}{2}du.$$

Observe that the function $f(x+u)+f(x-u)$ is even in u (one replaces the other). Thus, since we are over a symmetric interval:

$$S_N(x) = \int_{0}^{\pi} \left(f(u+x) + f(x-u)\right) D_N(u)du.$$

The final lemma we need has a name:

Lemma 10.4 (Reimann Lebesgue Lemma). *If $f \in L^1[a,b]$, then:*

$$\lim_{\lambda\to\infty}\int_a^b f(x)\cos(\lambda x)dx = \lim_{\lambda\to\infty}\int_a^b f(x)\sin(\lambda x)dx - \lim_{\lambda\to\infty}\int_a^b f(x)e^{i\lambda x}dx = 0.$$

Proof. We will do the cosine case and the sine case will follow in a similar manner. With both sine and cosine, the $e^{\lambda i x}$ case follows immediately. Recall that the continuous functions are dense in $L^1[a,b]$. Therefore, we will show that the result holds for continuous functions. Also recall that, for a continuous function g, if s is the interpolating spline, then $||g-s||_u \le \omega(g, ||\Delta||)$. So, to generalize for splines, let f be a continuous piecewise differentiable function. Then:

$$\int_a^b f(x)\cos(\lambda x)dx = \frac{1}{\lambda}f(x)\sin(\lambda x)|_a^b - \frac{1}{\lambda}\int_a^b f(x)\sin(\lambda x)dx.$$

Then,

$$\begin{aligned}\left|\int_a^b f(x)\cos(\lambda x)dx\right| &\le \frac{1}{|\lambda|}|f(b)\sin(\lambda b)| + \frac{1}{|\lambda|}|f(a)\sin(\lambda a)| + \frac{1}{|\lambda|}\int_a^b |f'(x)||\sin(\lambda x)dx \\ &\le \frac{1}{|\lambda|}|f(b)| + \frac{1}{|\lambda|}|f(a)| + \frac{1}{|\lambda|}\int_a^b |f'(x)|dx.\end{aligned}$$

Then, $\int_a^b f(x)\cos(\lambda x)dx \to 0$ as $\lambda \to \pm\infty$.

Now let $f \in L^1[a,b]$ and $\epsilon > 0$. Then there exists a continuous g such that $||f-g||_1 < \epsilon/3$. Then let s be a linear spline interpolant. Then:

$$||g-s||_1 = \int_a^b |g-s|dx \leq \int_a^b \omega(g, ||\Delta||)dx = \omega(g, ||\Delta||)(b-a).$$

Thus, we can choose a spline such that $||g-s||_1 < \epsilon/3$. Finally, our earlier work, since s is continuous, piecewise smooth function, we can make λ large enough so that $\left|\int_a^b s\cos(\lambda x)dx\right| < \epsilon.3$.

Then:

$$\begin{aligned}
\left|\int_a^b f(x)\cos(\lambda x)dx\right| &= \left|\int_a^b (f(x)+g-g+s-s)\cos(\lambda x)dx\right| \\
&\leq \int_a^b |(f-g)\cos(\lambda x)|dx + \int_a^b |(g-s)\cos(\lambda x)|dx \left|\int_a^b s\cos(\lambda x)dx\right| \\
&\leq \int_a^b |f-g|dx + \int_a^b |g-x| + \left|\in_a^b s\cos(\lambda x)dx\right| \\
&= ||f-g||_1 + ||g-s||_1 + \left|\int_a^b s\cos(\lambda x)dx\right| \\
&< \epsilon/3 + \epsilon/3 + \epsilon/3 \\
&= \epsilon.
\end{aligned}$$

Thus, $\int_a^b f(x)\cos(\lambda x)dx \to 0$ as $\lambda \to \pm\infty$.

□

Now we can give the major theorem of the section.

Theorem 10.5 (Pointwise Convergence Theorem)**.** *If f is a 2π periodic piecewise continuous function that has a right-hand derivative $f'(x+)$ and a left-hand derivative $f'(x-)$ at x, then:*

$$\lim_{N\to\infty} S_N(x) = \begin{cases} f(x) & f \text{ is continuous at } x \\ \dfrac{f(x+)+f(x-)}{2} & \text{if } x \text{ is a jump discontinuities} \end{cases}$$

Proof. Since f is piecewise continuous, we only need to consider $\dfrac{f(x+)+f(x-)}{2}$. Define the error:

$$E_n(x) := S_N(x) - \frac{f(x+)+f(x-)}{2}.$$

Then:

$$\begin{aligned}
E_n(x) &= \int_0^\pi [f(u+x)+f(x-u)]D_N(u)du - \frac{f(x+)+f(x-)}{2} \\
&= \int_0^\pi [f(u+x)+f(x-u)]dx - \frac{f(x+)+f(x-)}{2} 2\int_0^\pi D_N(u)du \\
&= \int_0^\pi [f(u+x)+f(x-u)-f(x+)-f(x-)]D_N(u)du \\
&= \int_0^\pi [f(u+x)+f(x-u)-f(x+)-f(x-)]\frac{\sin((N+1)u)}{2\pi\sin(u/2)}du.
\end{aligned}$$

Now define $F(u) := \dfrac{f(u+x)+f(x-u)-f(x+)-f(x-)}{2\pi\sin(u/2)}$. Using L'Hopital's Rule (with respect to u!):

$$\lim_{u \to 0} \frac{f'(u+x) + f'(u-x) - f'(x+) - f'(x-)}{2/2\pi \cos(u/2)} = \frac{f'(x+) + f'(x)}{\pi}.$$

Thus, F has a right hand limit at 0. Since f is piecewise continuous and $\frac{1}{2\pi \sin(u/2)}$ is piecewise continuous on the interval, we have that $F(u)$ is piecewise continuous. Then we have:

$$E_n(x) = \int_0^{\pi} F(u) \sin((N+1)u) du.$$

Since F is piecewise continuous on $[0, \pi]$, $F \in L'(0, \pi)$. Thus, $\lim_{N \to \infty} E_N(x) = 0$. $\square$

Chapter 11

Discrete Fourier Transform

11.1 Introduction

Recall that, for a 2π-periodic function f, its Fourier Series is:

$$f(x) \backsim \sum_{-\infty}^{\infty} a_k e^{ikx} \text{ where } a_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-ikt}.$$

So, explicitly, $f(x) \backsim \sum_{k=-\infty}^{\infty} \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{ikt}dte^{ikx}$.

However, $\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-ikt}dt$ may be difficult to find, so we approximate it.

To do so, we first suppose that $F(t)$ is any continuous function. Recall the trapezoid rule: Let Δx be the k-th sub-interval of the partition of the domain $[a, b]$. Then, where x_k is the end point of an interval:

$$\int_a^b f(x)dx \approx \sum_{k=1}^{N} \frac{f(x_k) + f(x_{k+1}}{2}\Delta x.$$

For our purposes, we suppose $\Delta x_k = \frac{2\pi}{n}$. Then,:

$$\frac{1}{2\pi}\int_0^{2\pi} F(t)dt \approx \frac{1}{n}\sum_{j=1}^{n} F\left(\frac{2\pi j}{n}\right) = \frac{1}{n}\sum_{j=0}^{n-1} F\left(\frac{2\pi j}{n}\right).$$

Now since $f(t)e^{-int}$ is 2π-periodic, we have:

$$a_k = \frac{1}{2\pi}\int_0^{2\pi} f(t)e^{-ikt}dt \approx= \frac{1}{n}f\left(\frac{2\pi j}{n}\right)e^{-ik\frac{2\pi j}{n}}dt.$$

We then define $y_j := f\left(\frac{2\pi j}{n}\right)$, $\omega := e^{\frac{2\pi i}{n}}$, so that

$$a_k \approx \frac{1}{n}\sum_{j=0}^{n-1} y_j \overline{\omega}^{jk}.$$

Suppose now that we consider $a_{k+n} \approx \frac{1}{n}\sum_{j=0}^{n-1} y_j \overline{\omega}^{j(k+n)}$. Focus on $\overline{\omega}^{j(k+n)}$:

$$\overline{\omega}^{j(k+n)} = \overline{\omega}^{jk+jn} = \overline{\omega}^{jk}\overline{\omega}^{jn} = \overline{\omega}^{jk}(\overline{\omega}^{n})^j.$$

Since $\overline{\omega}^n = e^{-2\pi i} = 1$, we have:

$$a_{k+n} \approx \frac{1}{n}\sum_{j=0}^{n-1} y_j \overline{\omega}^{jk}.$$

Thus, when we are approximating the Fourier coefficients, we need to compute a_k for $k = 0, ..., n-1$. Now, suppose we have the approximation. Can we get y_j back? As it turns out, yes. Let

$$\hat{y}_k = \sum_{j=0}^{n-1} y_j \overline{\omega}^{jk},$$

so $na_k \approx \hat{y}_k$. Then, $\hat{y}_k w^{kl} = \sum_{j=0}^{n-1} y_k \overline{\omega}^{jk} \omega^{kl}$. Sum over k:

$$\begin{aligned}\sum_{j=0}^{n-1} \hat{y}_k \omega^{kl} &= \sum_{k=0}^{n-1}\sum_{j=0}^{n-1} y_k \overline{\omega}^{jk} \omega_{kl} \\ &= \sum_{k=0}^{n-1}\sum_{j=0}^{n-1} y_k \omega^{(l-j)k}.\end{aligned}$$

Also recall:

$$\sum_{j=0}^{n-1} z^k = \begin{cases} \dfrac{z^n - 1}{z - 1} & z \neq 1 \\ n & z = 1 \end{cases}.$$

Thus:

$$\sum_{j=0}^{n} \omega^{(l-j)k} = \begin{cases} \dfrac{w^{(l-j)kn} - 1}{\omega^{l-j} - 1} & \omega^{l-j} \neq 1 \\ n & \omega^{l-j} = 1 \end{cases} = \begin{cases} 0 & l \neq j \\ n & l = j \end{cases}.$$

Hence $\sum_{k=0}^{n-1} \hat{y}_k w^{kl} = \sum_{k=0}^{n-1}\sum_{j=0}^{n-1} y_k \omega^{(l-j)k} = y_l n$, and so $y_l = \frac{1}{n}\sum_{k=0}^{n-1} \hat{y}_k \omega^{kl}$. And so we can recover y_l from approximation data.

11.2 Formal Definitions

Definition 11.1. $S_n := \{(..., x_{n-1}, x_0, x_1, ...) : \ x_j \in \mathbb{C}, \ x_{n+j} = x_j\}$. That is, all sequences of complex numbers that are n-periodic.

Let $y = \{y_j\}_{-\infty}^{\infty} \in S_n$. Then, we can define a new sequence:

$$\hat{y}_k = \sum_{j=0}^{n-1} y_j \overline{\omega}^{jk}.$$

(This is the exact formula from above, but the y_js are not necessarily from a function.) Following an argument we alluded to earlier, we show that this new sequence is back in S_n:

$$\hat{y}_{k+n} = \sum_{j=0}^{n-1} y_j \overline{\omega}^{j(k+n)} = \sum_{j=0}^{n-1} y_j \overline{\omega}^{jk} (\overline{\omega}^n)^j = \sum_{j=0}^{n-1} y_j \overline{\omega}^{jk} = \hat{y}_k.$$

So, we may define $\mathcal{F} : S_n \to S_n$ by $\mathcal{F}[y] = \hat{y}$. Observe that the other "inverse" formula also holds: $y_j = \frac{1}{n}\sum_{k=0}^{n-1} \hat{y}_k \omega^{kj}$ is also a transformation on S_n, so that $\hat{y} \to y$, and so we define $\mathcal{F} : S_n \to S_n$ by $\mathcal{F}^{-1}[y_k] = y$.

Definition 11.2. Convolution: If $y, z \in S_n$ then $[y * z]_j := \sum_{m=0}^{n-1} y_m z_{j-m}$ is the convolution.

Proposition 11.3. If $y, z \in S_n$, then $[y * z] \in S_n$.

Proof. $[y * z]_{j+n} = \sum_{m=0}^{n-1} y_m z_{j+n-m} = \sum_{m=0}^{n-1} y_m z_{j-m} = [y * z]_j$. $\square$

Proposition 11.4. If z is the periodic sequence formed from $y \in S_n$ by $z_j = z_{j+1}$ (or the left shift), then $\mathcal{F}[z]_k = \omega^k \mathcal{F}[y]_k$.

Proof. By definition, $\mathcal{F}[y]_k = \sum_{j=0}^{n-1} y_j \overline{\omega}^{jk}$. Then:

$$\omega^k \mathcal{F}[y]_k = \sum_{j=0}^{n-1} y_j \overline{\omega}^{jk} \omega^{jk} = \sum_{j=0}^{n-1} y_j \overline{\omega}^{k(j-1)}.$$

Shift the index: $j \to j+1$. (Observe that $y_n = y_0$, $\omega^{jn} = \omega^0$.) Then:

$$\omega^k \mathcal{F}[y]_k = \sum_{j=0}^{n-1} y_{j+1} \overline{\omega}^{kj} = \sum_{j=0}^{n-1} z_j \overline{\omega}^{kj} = \mathcal{F}[z]_k.$$

□

Theorem 11.5 (The Convolution Theorem)**.** $\mathcal{F}[y * z]_k = \mathcal{F}[y]_k \mathcal{F}[z]_k$.

Proof. Let $p = [y * z]$. Then, $p_j = \sum_{m=0}^{n-1} y_m z_{j-m}$. Then:

$$\begin{aligned}
\mathcal{F}[p]_k &= \sum_{j=0}^{n-1} p_j \overline{\omega}^{jk} = \sum_{j=0}^{n-1} \left(\sum_{m=0}^{n-1} y_m z_{j-m} \right) \overline{\omega}^{jk} \\
&= \sum_{m=0}^{n-1} y_m \sum_{j=0}^{n-1} z_{j-m} \overline{\omega}^{jk} \\
&= \sum_{m=0}^{n-1} y_m \overline{\omega}^{mk} \sum_{j=0}^{n-1} z_{j-m} \overline{\omega}^{(j-m)k}.
\end{aligned}$$

Shifting $z_{j-m} \to z_j$ implies, by the shift proposition above, that $\mathcal{F}[p]_k = \sum_{m=0}^{n-1} y_m \overline{\omega}^{mk} \sum_{j=0}^{n-1} z_j \overline{\omega}^{jk} = \mathcal{F}[y]_k \mathcal{F}[z]_k$. □

Chapter 12

Contraction Mapping Theorem

Theorem 12.1 (Contraction Mapping Theorem)**.** *Let $\mathcal{H}$ be a Hilbert space and let $B \subset \mathcal{H}$ be a closed subspace. Let $f : \mathcal{H} \to \mathcal{H}$ be a continuous function such that:*

- *$f(B) \subset B$ (that is, $f(b) \in B$ for $b \in C$);*
- *f is Lipschitz with constant $0 \leq \alpha < 1$.*

Let $u_0 \in B$ and define the sequence $\{u_n\}_1^\infty$ by $u_{n+1} = f(u_n)$. Then, $\{u_n\}_1^\infty$ converges, say to u, and $f(u) = u$.

Proof. We show that $\{u_n\}_1^\infty$ is Cauchy. Let $m > n$, and consider:

$$\begin{aligned}||u_m - u_n|| &= ||u_m - u_{m-1} + u_{m+1} - ... - u_{n+1} + u_{n+1} - u_n|| \\ &\leq ||u_m - u_{m+1}|| + ... + ||u_{n+1} - u_n||.\end{aligned}$$

For each j:

$$||u_{j+1} - u_j|| = ||f(u_j) - f(u_{j-1})|| \leq \delta||u_j - u_{j-1}|| = \delta||f(u_{j-1} - f(u_{j-2})|| \leq \delta^2||u_{j-1} - u_{j-2}|| \leq ... \leq \delta^j||u_1 - u_0||.$$

Thus, from our sum:

$$||u_m - u_n|| \leq \delta^m||u_1 - u_0|| + ... + \delta^n||u_1 - u_0|| = \sum_{j=n}^{m} \delta^j||u_1 - u_0|| = \frac{\delta^n - \delta^{m+1}}{1 - \delta}||u_1 - u_0||.$$

Since $\delta < 1$, $\lim_{n,m\to\infty} \dfrac{\delta^n - \delta^{m+1}}{1 - \delta} = 0$. Thus, $\{u_n\}_1^\infty$ is Cauchy, and thus converges. Say, $u_n \to u$. Now, since f is continuous, $f(u) = \lim_{n\to\infty} f(u_n) = \lim_{n\to\infty} u_{n-1} = u$. $\square$

Chapter 13

Splines and Finite Element Spaces

We have already dealt with linear splines in a limited sense, but we want to develop a more useful general theory. For example, we may want splines that have certain differentiability conditions. We need three things to specfiy a spline:

1. Knot sequence: $\Delta = \{0 = x_0 < x_1 < ... < x_n = 1\}$;
2. The degree of polynomial k;
3. the Level of differntiability of the whole spline, r.

Definition 13.1. $S^\Delta(k, r)$ is the set of splines on a knot sequence Δ, polynomials of degree k, and smoothness C^r.

Proposition 13.2. $S^\Delta(k, r)$ is a vector space.

Proof. Assume $r \geq 0$. Then, $S^\Delta(k, r) \subset C[0, 1]$. Let $s, t \in S^\Delta(k, r)$, $\alpha \in \mathbb{R}$. Let $I_i = [x_i, i_{i+1}]$ be an interval. Then, on I_i, s and t are both degree k polynomials, and since degree k polynomials form a vector space, $\alpha s + t$ is a degree k polynomial on I_i. Furthermore, if s, t have degree r differentiability, then, for $\alpha s + t$ also has degree r differentiability. Thus, since $C[0, 1]$ is a vector space, $S^\Delta(k, r)$ is a vector space. $\square$

13.1 Basis Splines

Since $S^\Delta(r, k)$ is a vector space (of finite dimension), we want to know if it has a basis. Right now, we focus on linear splines. To do so, we consider the following "plus" function:

$$(x)_+ = \begin{cases} x & 0 \leq x \\ 0 & 0 > x \end{cases}.$$

(This function is the "tent" function, 0 at 0, then 1 at 1, then back to 0 at 2, and linear in between. We use the tent function to define a "sub-basis" function that we will build our basis from.

$$N_2(x) = (x)_+ - 2(x - 1)_+ + (x - 2)_+.$$

This is the tent function, For $x \leq 0$, $N_2(x) = 0 - 0 + 0 = 0$; for $0 \leq x \leq 1$, $N_2(x) = x + 0 + 0 = x$; for $1 \leq x \leq 2$, $N_2(x) = x - 2x + 0 = -x$; for $2 \leq x$, $N_2(x) = x - 2x + x = 0$.

Proposition 13.3. Let Δ be an equally spaced knot sequence with $x_j = j/n$, $j = 0, ..., n$. Then $B := \{N_x(nx - j + 1) : \ j = 0, ..., n\}$ is a basis for $S^\Delta(1, 0)$ (the space of linear splines).

Proof. Observe that $dim(B) = n$. Since any spline is uniquely determined by its value at the knots, and we have n knots $dim(S^\Delta(1, 0)) = n$. Thus $dim(B) = dim(S^\Delta(1, 0))$. All that remains is to show that B is linearly independent. Note that $N_2(k - j + 1) = \delta_{jk}$. Suppose

$$c_0 N_2(nx+1) + ... + c_n N_2(nx-n+1) = 0.$$

In particular, this holds for all $x \in [0,1]$. Then, for $x_j = j/n$

$$c_0 N_2(j/n+1) + ... + c_j N_2(1) + ... + c_n N_2(j/nn+1) = c_j = 0.$$

This holds for all j, and so B is linearly independent.

□

13.2 Finite Element Spaces

First, we give a better definition for splines.

Definition 13.4. Let $\Delta = \{x_0 = 0 < x_1 < ... < x_n = 1\}$ be a knot sequence on $[0,1]$. Define $I_j = [x_{j-1}x_j)$, $I_n = [x_{n-1}, x_n]$. Let $\mathbb{P}^k$ be the space of degree k polynomials. Then:

$$S^{\Delta}(k,r) := \{\phi : [0,1] \to \mathbb{R} \ : \ \phi|_{I_j}, \ \phi \in C^r[0,1]\}.$$

$S^{1/n}(k,r)$ is the set where Δ has evenly spaced intervals, so $x_j = j/n$.

Definition 13.5. A *finite element space* $S^{1/n}(k,r)$ degree k polynomials on each intervals have $r \leq k-1$ derivatives that match on the knots.

Proposition 13.6. *$dim(S^{1/n}(k,r) = n(k-r) + r + 1$.*

Proof. We have n intervals. On each interval, we have a polynomial of degree k, and so is specified by $k+1$ parameters. Thus, we have $n(k+1)$. But this is an overestimate, as we have restrictions on the derivatives.

On each of the $n-1$ intervals, there are $r+1$ equations that must match: r derivatives and the original polynomials. These remove degree of freedom from each node, so $dim(S^{1/n}(k,r)) = n(k+1) - (r+1)(n-1) = n(k-n) + r + 1$. □

13.3 Construction of Cubic Splines

We now want a spline that matches a given function not just on the nodes, but also the functions derivative. We will use (and analyze) $S^{1/n}(3,1)$, or the cubic splines. That is, for f, where $f(x_j)$ and $f'(x_j)$ are known, there exists a unique cubic spline $s(x)$ such that $s(x_j) = f(x_j)$ and $s'(x_j) = f'(x_j)$. Note that, by the above proposition, $dim(S^{1/n}(3,1) = 2n+2$. Observe that, for a given f, we have $n+1$ $f(x_j)$ values and $n+1$ $f'(x_j)$ values, and so $2n+2$ data points to fit.

Now we want to construct a basis for $S^{1/n}(3,1)$. So we start with interpolating functions. We want a $\phi(x)$ such that $\phi(0) = 1$, $\phi(1) = \phi'(1) = \phi'(0) = 0$. Consider the following function:

$$\phi(x) = A(x-1)^3 + B(x-1)^2.$$

This clearly satisfies the $\phi(1) = 0$. Now we find the coefficients that satisfy the remaining requirements. Now, $\phi(0) = A + B = 0$, $\phi'(x) = 3A(x-1)^3 + 2B(x-1)^2$, so $\phi'(1) = 0$ and $\phi'(0) = 3A + 2B = 0$. Solve this sytem of equations gives $A = 2$ $B = 3$, so

$$\phi(x) = 2(x-1)^2 + 3(x-1)^2 = (x-1)^2(2x = 1).$$

Now define (using the same notation, admittedly an abuse of it)

$$\phi(x) = \begin{cases} (|x|-1)^2(2|x|+1) & |x| \leq 1 \\ 0 & |x| > 1 \end{cases}.$$

That is, on $[0,1]$ $\phi(x)$ looks like $2(x-1)^3 + 3(x-1)^2$, and on $[-1,0]$ ϕ also looks like $2(x-1)^3 + 3(x-1)^2$ but reflected over the y-axis. (Plot to see the idea.) So ϕ is not a polynomial on $[-1,1]$, but is piecewise cubic. Form the construction, $\phi'(0) = \phi'(1) = \phi'(-1) = 0$ and $\phi(0) = 1$. Clearly outside of $[-1, 1]$, $\phi = 0$. From these observations, it is obvious $\phi \in S^{\mathbb{Z}}(1, 3)$. This function is useful for approximating the values of a given function. Now we want to approximate a given functions derivative.

So, we want a function ψ such that $\phi(1) = \phi(0) = \phi'(1) = 0$, $\psi'(0) = 1$. Again, consider ψ of the form:

$$\psi(x) = A(x-1)^3 + B(x-1)^2.$$

Then $\phi(1) = 0$, $\phi'(x) = 3A(x-1)^2 + 2B(x-1)$, so $\phi'(1) = 0$. Since $\phi(0) = -A + B = 0$, $A = B$, and $\psi'(0) = 3A - 2B = 1$, we have $A = B = 1$. So:

$$\psi(x) = (x-1)^3 + (x-1)^2 = x(x-1)^2.$$

We do the same as before:

$$\psi(x) = \begin{cases} |x|(|x|+1)^2 & |x| \le 1 \\ 0 & |x| > 1 \end{cases}.$$

Note, by a similar argument, that $\phi(x) \in S^{\mathbb{N}}(3, 1)$ and $\phi(0) = 0$, $\phi'(0) = 1$.

Next, we use these functions to build a basis. For $\phi_j(x)$, define:

$$\phi_j(x) := \phi(nx - j).$$

Then, $\phi_0(x) = \phi(nx)$, and $\phi_j(x) = \phi(n(x - j/n)) = \phi_0(x - j/n)$.Thus, ϕ_j is simply ϕ_0 translated by j/n. Clearly, ϕ_j is 0 outside of $[\frac{j-1}{n}, \frac{j+1}{n}]$, and $\phi(j/n) = 1$, $\phi'(j/n) = \phi'(j/n) = \phi'(\frac{j+1}{n}) = 0$. Thus, $\phi_j(k/n) = \delta_{j,k}$, $\phi_j'(k/n) = 0$.

As for ψ_j, consider that $(\psi(nx - j))' = n\psi'(nx - j)$, so $n\psi'(n(j/n) - j) = n$. Thus, for $\psi_j'(j/n)$ to equal 1, we must scale. So take

$$\psi_j(x) = \frac{1}{n}\psi'(nx - j).$$

Again, as before, the support of $\psi_j(x)$ is $[\frac{j-1}{n}, \frac{j+1}{n}]$ and $\psi_j(k/n) = 0$, $\psi'(k/n) = \delta_{j,k}$.

13.4 Interpolation with Cubic Splines

Proposition 13.7. The set $B := \{\phi_j, \psi_j\}_0^n$ (where ϕ_j and ψ are defined above) is a basis for $S^{1/n}(1, 3)$.

Proof. As noted before, $dim(S^{1/n}(1, 3) = 2n + 2$. Since we have $2n + 2$ elements in B, we only need to show that B is linearly independent. Suppose

$$\sum_{j=0}^{n} \alpha_j \phi_j(x) + \sum_{j=0}^{n} \beta_j \psi_j(x) = 0.$$

By assumption, this holds for all x. So in particular, for $x = k/n$, we have:

$$\sum_{j=0}^{n} \alpha_j \phi_j(k/n) + \sum_{j=0}^{n} \beta_j \psi(k/n) = \alpha_k = 0.$$

This holds for all k. Furthermore, if we differentiate (and apply $\alpha_j = 0$), we get:

$$\sum_{j=0}^{n} \beta_j \psi_j'(k/n) = \beta_k = 0.$$

Again, this holds for all k. Thus, $\alpha_1 = \alpha_2 = ... = \alpha_n = \beta_1 = \beta_2 = ... = \beta_n = 0$, and we have linear independence.

□

If we want a projection, we can take

$$s(x) := P_S f = \sum_{j=0}^{n} f(j/n)\phi_j(x) + \sum_{j=0}^{n} f'(j/n)\psi(j/n).$$

Then $s(j/n) = f(j/n)$ and $s'(j/n) = f'(j/n)$ for all j, as we set out to do at the beginning of the section.

13.5 Finite Element Methods and Galerkin Methods

Suppose we have some function f where $f(x_j) =: f_j$ and we want to find the "nicest" function in $S^{1/n}(1,3)$ such that $s(x_j) = f_j$ for all j. So, we want $s \in S^{1/n}(3,1)$ that minimizes

$$||s||^2 = \int_0^1 (s''(x))^2 dx$$

for s such that $s(x_j) = f_j$. Any function that satisfies this is given by

$$s(x) = \sum_{j=0}^{n} f_j \phi_j(x) + \sum_{j=0}^{n} \alpha_j \psi_j(x).$$

Let $f = \sum_{j=0}^{n} f_j \phi_j(x)$. So, we want to find the coefficients that minimize $||f - \sum_{j=0}^{n} \alpha_j \psi_j||$. This is a least squares problem, and so can be solved using the associated normal equations. That, if we set $g = \sum_{j=0}^{n} \alpha_j \psi_j$ we want to α_j such that:

$$\langle f - g, \psi_k \rangle = 0.$$

That is, finding

$$\sum_{j=0}^{n} \alpha_j \langle \psi_j, \psi_k \rangle = \langle f, \psi_k \rangle$$

where $\langle \psi_k, \psi_j \rangle = G_{j,k}$. Note G is invertible and $\langle \psi_j, \psi_k \rangle = \int_0^1 \psi_j'' \psi_k''$. We may deduce from the support of ψ_k and ψ_j that $\langle \psi_j, \psi_k \rangle$ is nonzero only when $k = j-1$, $k = j$ and, $k = j+1$. These systems are easy to solve numerically.

Chapter 14

Bounded Operators and Closed Subspaces

Definition 14.1. Let $V,\ W$ be Banach Spaces and let $L : V \to W$ be a linear operator. Then:

$$||L||_{op} = \sup_{||v||_V=1} ||Lv||_W = \sup_{v\neq 0} \frac{||Lv||_W}{||v||_V}.$$

(These two are equivalent, which can be shown).

Definition 14.2. Let $L : V \to W$ be an operator. Then, L is *continuous at* u if for every $\epsilon > 0$, there exists a $\delta > 0$ such that $||u - v||_V < \delta$ implies $||Lu - Lv||_W < \epsilon$. (Or the standard definition.)

Note 14.3. When L is linear, we need only to check $||x||_V < \delta$ implies $||Lx||_W < \epsilon$ since $Lx - Ly = L(x-y)$.

Proposition 14.4. Let $L : V \to W$ be linear. If L is continuous at 0, then L is continuous at v for all $v \in V$.

Proof. Let $\epsilon > 0,\ u,\ v \in V$. Then, if we let $u - v = z$, we have $Lu - Lv = Lz$. Since L is continuous at 0, there exists a $\delta > 0$ such that $||z - 0|| = ||z|| = ||u - v|| < \delta$ implies $||Lz - L0|| = ||Lz|| = ||Lu - Lv||\epsilon$. Thus we have continuity at u. This holds for all $u \in V$, and so we are done. $\square$

Proposition 14.5. A linear transformation $L : V \to W$ is countinuous iff it is bounded.

Proof. First suppose L is bounded. Let $\epsilon > 0, u, v \in V$. Then, $||Lu-Lv||_W = ||L(u-v)||_W \leq ||L||_{op}||u-v||_V$. Thus if we choose $\delta = \epsilon/||L||_{op}$, we have $||Lu - Lv||_W \leq ||L||_{op}||u - v||_V < \epsilon$, and so we have continuity.

Now suppose L is continuous. Then for $\epsilon = 1$, there exists a $\delta > 0$ such that $||u||_V \leq \delta$ implies that $||Lu||_W \leq 1$. Further suppose that $||u|| = 1$. Then, $||\delta u|| = \delta$, and so $||L(\delta u)||_W \leq 1$, and so $||Lu|| \leq 1/\delta$. This holds for all u such that $||u|| = 1$. Thus, $||L||_{op} \leq 1/\delta < \infty$, and so is bounded. $\square$

Definition 14.6. Let $L : C[0, 1] \to C[0, 1]$ be given as

$$Lu(x) = \int_0^1 k(x, y)u(y)dy$$

for some $k(x, y)$, which is called the *kernel*. If $k(x, y) \in L^2([0, 1] \times [0, 1])$, then k is called a *Hilbert-Schmidt Kernel*, and Lu is called a *Hilbert Schmidt Operator*.

Proposition 14.7. Let K be a Hilbert-Schmidt kernel. Then the operator $Lu(x) = \int_0^1 k(x, y)u(y)dy$ is bounded and $||L||_{op} \leq ||k||_{L^2(R)}$. ($R$ is the unit rectangle.)

Proof.

$$\begin{aligned}
||Lu||_{L^2}^2 &= \int_0^1 \left(\left| \int_0^1 k(x,y)u(y)dy \right| \right)^2 dx \\
&\leq \int_0^1 \left(\int_0^1 |k(x,y)||u(y)|dy \right)^2 dx \\
&\leq \int_0^1 \left(\int_0^1 |k(x,y)|^2 dy \right) \left(\int_0^1 u^2(y)dy \right) dx \\
&\leq \int_0^1 \int_0^1 |k(x,y)|dydx \int_0^1 u^2(y)dy \\
&= ||k(x,y)||_{L^2(R)} ||u||_{L^2[0,1]}.
\end{aligned}$$

(The second inequality follows from Cauchy Schwarz on the L^2 inner product.) Since u and $k(x,y)$ are both in their respective L^2 spaces, we have that L is a bounded operator. If we assume $||u||_{L^2} = 1$, we have the desired inequality. □

14.1 Closed Subspaces

Note that since our spaces have norms, we use the norm topology (which is a metric topology), and thus a set is closed if it contains all its limit points.

Definition 14.8. Let V be a subspace of a Hilbert space $\mathcal{H}$. The *orthogonal complement* $V^\perp$ is:

$$V^\perp := \{f \in \mathcal{H} : \langle f, g \rangle = 0 \ \forall g \in V\}.$$

Proposition 14.9. $V^\perp$ is a closed subspace of $\mathcal{H}$.

Proof. Suppose $\{f_n\}_1^\infty \subset V^\perp$ such that $f_n \to f$. Let $g \in V$. Since inner products are continuous,

$$0 = \lim_{n\to\infty} \langle f_n, g \rangle = \langle \lim_{n\to\infty} f_n, g \rangle = \langle f, g \rangle.$$

Thus $f \in V^\perp$. □

Definition 14.10. Let $L : V \to W$ be a bounded linear operator. Then:

- Domain of L: $D(L)$;
- Range of L: $R(L) = \{w \in W : \ \exists v \in V \ s.t. \ Lv = w\}$;
- Kernel of L: $\{v \in V : \ Lv = 0\}$.

Proposition 14.11. If $L : V \to W$ is a bounded linear operator, then $N(L)$ is a subspace of V.

Proof. Let $\{v_j\}_1^\infty \subset N(L)$ such that $v_j \to v$. Since L is bounded, it is continuous, and so:

$$0 = \lim_{j\to\infty} L(v_j) = L(\lim_{j\to\infty} v_j) = L(v).$$

Thus, $v \in N(L)$, and so is closed. □

Chapter 15

Several Important Chapters

Let $\mathcal{H}$ be a Hilbert space. When $V \subset \mathcal{H}$ is finite dimensional, we know an orthogonal projection always exists. That is, we can always find a $p \in V$ such that $||f - p|| = \min_{v \in V} ||f - v||$. But what happens when V is infinite dimensional? As it turns out, this is only possible iff V is closed. But first, a lemma.

Lemma 15.1. *Let $\mathcal{H}$ be a Hilbert space. For every $f,\ g \in \mathcal{H}$, we have:*

$$||f + g||^2 + ||f - g||^2 = 2(||f||^2 + ||g||^2).$$

Proof. Take the two equalities:

$$||f + g||^2 = ||f||^2 + \langle f, g\rangle + \langle g, f\rangle + ||g||^2$$

$$||f - g||^2 = ||f||^2 - \langle f, g\rangle - \langle g, f\rangle + ||g||^2.$$

Add them and the result follows. □

Theorem 15.2 (The Projection Theorem)**.** *Let $\mathcal{H}$ be a Hilbert space and $V \subset \mathcal{H}$ be a subspace. For every $f \in \mathcal{H}$, there exists a unique $p \in V$ such that $||f - p|| = \min_{v \in V} ||f - v||$ iff V is closed.*

Proof. Suppose a minimizer exiss, and $\{v_n\}_1^\infty \subset V$ is a sequence such that $v_n \to v$. By assumption, there exists a $p \in V$ such that $||p - v|| = \min_{w \in V} ||v - w||$. Since $v_n \to v$, for every $\epsilon > 0$, there exists a v_n such that $||v - v_n|| < \epsilon$. But since $v_n \in V$, we have that $||v - p|| \leq ||v - v_n|| < \epsilon$. Since this holds for all $\epsilon > 0$, we have that $v = p \in V$.

Now assume V is closed. For $f \in \mathcal{H}$, let $\alpha^2 := \inf_{v \in V} ||v - f||^2$. Thus, for $\epsilon > 0$, there exists $v_\epsilon \in V$ such that $\alpha^2 \leq ||v_\epsilon - f|| < \alpha^2 + \epsilon$. And so we may find a sequence $\{v_n\}_1^\infty$ such that for every n, $||f - v_n|| < \alpha^2 + 1/n$. For differen $n,\ m$ we have

$$0 \leq ||f - v_m||^2 + ||f - v_n||^2 - 2\alpha^2 < 1/n + 1/m.$$

Now, by our lemma:

$$\begin{aligned}
2(||f - v_m||^2 + ||f - v_m||^2) &= ||f - v_m - (f - v_n)||^2 + ||f - v_m + f - v_n||^2 \\
&= ||v_n - v_m||^2 + ||2f - (v_n + v_m)||^2 \\
&= ||v_n - v_m||^2 + ||2\left(f - \frac{v_n + v_m}{2}\right)||^2 \\
&= ||v_n - v_m||^2 + 4||f - \frac{v_n + v_m}{2}||^2.
\end{aligned}$$

.

From earlier, we get

$$0 \leq 2\left(||f - v_n||^2 + ||f - v_m||^2\right) - 4\alpha^2 < 2/n + 2/m.$$

Therefore

$$0 \leq ||v_n - v_m||^2 + 4||f - \frac{v_n + v_m}{2}||^2 - 4\alpha^2 < 2/n + 2/m.$$

Since V is a subspace, $\frac{v_n+v_m}{2} \in V$, and so $||f - \frac{v_n+v_m}{2}||^2 > \alpha^2$. Thus,

$$0 \leq ||v_n - v_m||^2 + 4\alpha^2 - 4\alpha^2 = ||v_n - v_m||^2 < 2/n + 2/m.$$

Thus, $\{v_n\}_1^\infty$ is Cauchy, and therefore we have that $\{v_n\}_1^\infty$ converges, say to v. Since V is closed, $v \in V$. Then, $\lim_{n\to\infty} 0 \leq ||v_n - f||^2 - \alpha^2 < 1/n$ implies $||v_n - f|| = \alpha$.

To show uniquness, suppose p, p' both minimize. Let $\epsilon > 0$. Then, since $||p - f|| = ||p' - f|| =: \delta$ and using the lemma:

$$\begin{aligned} ||p - p'||^2 &= 2||f - p||^2 + ||f - p'||^2 - ||2\left(f - \frac{p + p'}{2}\right)|| \\ &= 4\delta^2 - 4||f - \frac{p + p'}{2}||. \end{aligned}$$

Since $\frac{p+p'}{2} \in V$, $||f - \frac{p+p'}{2}|| < \delta$,

$$0 \leq ||p - p'||^2 \leq 4\delta^2 - 4\delta^2 = 0.$$

Thus $p = p'$. □

Corollary 15.3. Let V be a subspace of $\mathcal{H}$. There exists an orthogonal projection $P : \mathcal{H} \to V$ such that $||f - Pf|| = \min_{v\in V} ||f - v||$ iff V is closed.

Proof. Suppose that P exists. Then a minimizer exists and so by the theorem, V is closed. Now suppose V is closed. Then, by the theorem, a minimizer exists, so define $Pf := p$, where p is the minimizer for f. We now show that $f - Pf \in V^\perp$. Let $w \in V$, $t \in \mathbb{R}$. (This only shows the case for a real Hilbert Space.) Then, $||f- = +tw||^2$ is minimized when $t = 0$. Thus:

$$||f - p + tw||^2 = \langle f - p, f - p\rangle + \langle f - p, tw\rangle + \langle tw, f - p\rangle + \langle tw, tw\rangle = ||f - p||^2 + 2t\langle f - p, w\rangle + t^2||w||^2.$$

Taking a derivative, $d'(t) = 2\langle f - p, 2\rangle + 2t||w||^2$. Since the original equation was minimized by $t = 0$, we have $0 = \langle f - p, w\rangle$ and so $f - p \in V^\perp$. □

Corollary 15.4. Let V be a closed subspace of $\mathcal{H}$. Then, $\mathcal{H} = V \bigoplus V^\perp$ and $(V^\perp)^\perp = V$.

Proof. Let $f \in \mathcal{H}$. Then, $Pf \in V$, $f - Pf \in V^\perp$, thus $f = Pf + f - Pf$, so $\mathcal{H} = V \bigoplus V^\perp$. Now let $v \in V$, $w \in V^\perp$. Then, $\langle w, v\rangle = 0$, so $v \in (V\ \perp)^\perp$. Thus $V \subset (V^\perp)^\perp$. Now let $w \in (V^\perp)^\perp$. For earlier, $w = v + \hat{v}$, where $v \in V$ and $\hat{v} \in V^\perp$. Now let $z \in V^\perp$. Then, $0 = \langle v + \hat{v}, z\rangle = \langle v, z\rangle + \langle \hat{v}, z\rangle = \langle \hat{v}, z\rangle$. Since this holds for all $z \in V^\perp$, $\hat{v} = 0$, and so $w = v \in V$. Thus, $(V\ \perp)^\perp \subset V$, and so $V = (V^\perp)^\perp$. □

Definition 15.5. Let V be a Banach Space. Then, a linear bounded map $\phi : V \to \mathbb{R}$ or $\mathbb{C}$ is called a *linear functional.*

Note 15.6. The space of linear functionals is called the *dual space* of V, and denoted V^*. V^* is a Banach space under the norm:

$$||\phi||_{V^*} = \sup_{v\neq 0} \frac{|\phi(v)|}{||v||_V}.$$

Theorem 15.7 (Reisz Representation Theorem)**.** *Let $\mathcal{H}$ be a Hilbert space and $\phi : \mathcal{H} \to \mathbb{C}$ be a bounded linear functional on $\mathcal{H}$. Then, there is a unique $g \in \mathcal{H}$ such that for all $f \in \mathcal{H}$, $\phi(f) = \langle f, g\rangle$.*

Proof. Since ϕ is bounded, $N(\phi)$ is closed. If $N(\phi) = \mathcal{H}$, then $\phi(f) = 0$ for all $f \in \mathcal{H}$, and so simply take $g = 0$. Now suppose $N(\phi) \neq \mathcal{H}$. Since $N(\phi)$ is closed, by the above corollary, we have $\mathcal{H} = N(\phi) \bigoplus N(\phi)^{\perp}$. Since $N(\phi) \neq \mathcal{H}$, there exists a $g \in \mathcal{H}$ but $g \notin N(\phi)$. By our decomposition, we conclude that $g \in N(\phi)^{\perp}$. Note that $\phi(g) \neq 0$. Then, for $f \in \mathcal{H}$, define $w := \phi(g)f - \phi(f)g$. (Recall that $\phi(a)$ is a scalar.) Then, we have

$$\phi(w) = (\phi(\phi(g)f) - \phi(\phi(f)g) = \phi(g)\phi(f) - \phi(f)\phi(g) = 0.$$

Therefore, $w \in N(\phi)$, $\langle w, g\rangle = 0$. Now we simply solve for $\phi(f)$:

$$\begin{aligned} 0 = \langle \phi(g)f - \phi(f)g, g\rangle \qquad &= \phi(g)\langle f, g\rangle - \phi(f)\langle g, g\rangle \implies \\ \phi(f) = \frac{\phi(g)\langle f, g\rangle}{||g||^2} = \langle f, \frac{\overline{\phi(g)}g}{||g||^2}\rangle. & \end{aligned}$$

This holds for all f, and so if we take $h := \dfrac{\overline{\phi(f)}f}{||g||^2}$, we get $\phi(f) = \langle f, g\rangle$. For uniqueness, suppose $\phi(f) = \langle f, g_1\rangle = \langle f, g_2\rangle$. Then $\langle f, g_1 - g_2\rangle = 0$ holds for all $f \in \mathcal{H}$. Thus, $g_1 - g_2 = 0$, and so $g_1 = g_2$. $\square$

15.1 Adjoints of Bounded Linear Operators

Corollary 15.8. Let $L : \mathcal{H} \to \mathcal{H}$ be a bounded lienar operator. Then, there exists a bounded linear operator $L^* : \mathcal{H} \to \mathcal{H}$ such that $\langle Lf, g\rangle = \langle f, L^*g\rangle$ for all $f, g \in \mathcal{H}$. L^* is called the *adjoint of* L.

Proof. Let $h \in \mathcal{H}$ and define $\phi_h(f) = \langle Lf, h\rangle$. Clearly, ϕ_h is linear, and

$$|\phi(f)| = |\langle Lf, h\rangle| \leq ||Lf||||h|| \leq ||L||||f||||h||$$

implies that ϕ is bounded. Therefore, by Reisz Representation, there exists a g_h such that $\phi_h(f) = \langle f, g_h\rangle = \langle Lf, h\rangle$. We denote g_h as such because it depends on h. Indeed, by the uniqueness of Reisz Representation, we have that g_h is a function of h. Indeed, take $\psi(h) = g_h$. We claim that ψ is linear. That is, we need to show that adding vectors in the subscript of ϕ_h leads to addition of vectors in the Reisz Representation. So, let $p = \alpha h_1 + \beta h_2$. Then:

$$\phi_{\alpha h_1 + \beta h_2}(f) = \langle Lf, \alpha h_1 + \beta h_2\rangle = \overline{\alpha}\langle Lf, h_1\rangle + \overline{\beta}\langle Lf, h_2\rangle = \overline{\alpha}\langle f, g_1\rangle + \overline{\beta}\langle f, g_2\rangle = \langle f, \alpha g_1 + \beta g_2\rangle.$$

Thus, ψ is linear. (Spoiler alert: this is going to be our adjoint.) To show this correspondence is bounded, since $\phi_h(f) = \langle f, g_h\rangle$ for all $g_h \in \mathcal{H}$, we have that $\phi_h(g_h) = \langle g_h, g_h\rangle = ||g_h||^2$. Thus, $||g_h||^2 \leq ||\phi||$, and so is bounded. Thus, ψ is a bounded linear functional, and $\langle Lf, h\rangle = \langle f, g_h\rangle = \langle f, \psi(h)\rangle = \langle f, L^*(h)\rangle$. $\square$

Corollary 15.9. $||L|| = ||L^*||$

Proof. We have (from a previous homework) that $||L|| = \sup_{f,h} |\langle Lf, h\rangle|$, where $||f|| = ||h|| = 1$. On the other hand, $||L^*|| = \sup_{f,h} |\langle L^*h, f\rangle|$. And since $\langle Lf, h\rangle = \overline{\langle L^*h, f\rangle}$, they are equal in magnitude, and so $||L^*|| = ||L||$. $\square$

Theorem 15.10. *Let $R = [0, 1] \times [0, 1]$ and suppose that $k(x, y)$ is Hilbert-Schmidt kernel. If $Lu(x) = \int_0^1 k(x, y)u(y)dy$, then $L^*v(x) = \int_0^1 \overline{k(y, xv(y)}dy$.*

Proof.

$$\begin{aligned}
\langle Lu, v\rangle &= \int_0^1 \left(\int_0^1 k(x,y)u(y)dy\right)\overline{v(x)}dx \\
&= \int_0^1 \left(\int_0^1 k(x,y)v(x)dx\right)u(y)dy \\
&= \int_0^1 \overline{\int_0^1 \overline{k(x,y)v(x)dx}} u(y)dy \\
&= \langle u, L^*v\rangle.
\end{aligned}$$

Changing the appropriate variables will give the result. □

Theorem 15.11 (Fredholm Alternative)**.** *Let $L : \mathcal{H} \to \mathcal{H}$ be a bounded linear operator whose range is closed. Then, the equation $Lf = g$ has a solution iff $\langle g, v\rangle = 0$ for all $v \in N(L^*)$. That is, $R(L) = N(L^*)^{\perp}$.*

Proof. Let $g \in R(L)$, so there exists a $h \in \mathcal{H}$ such that $Lh = g$. Then, if $v \in N(L^*)$, $\langle g, v\rangle = \langle Lh, v\rangle = \langle h, L^*v\rangle = 0$, so $g \in N(L^*)^{\perp}$. Thus $R(L) \subset N(L^*)^{\perp}$.

Now let $f \in N(L^*)^{\perp}$. Since $R(L)$ exists, there exists an orthogonal projection $P : \mathcal{H} \to R(L)$ such that $Pf \in R(L)$, $f' := f - Pf \in R(L)^{\perp}$. Since $f \in N(L^*)^{\perp}$, $Pf \in R(L) \subset N(L^*)^{\perp}$, $f' \in N(L^*)^{\perp}$. Thus, $f' \in N(L^*)^{\perp} \cap R(L)^{\perp}$. Thus, $\langle Lh, f'\rangle = 0 = \langle h, L^*f'\rangle$ for all $h \in \mathcal{H}$. Thus, for $h = L^*f'$, $\langle L^*f', L^*f'\rangle = ||L^*f'||^2 = 0$. Thus, $L^*f' = 0$, and so $f' \in N(L^*)$. But then $f' \in N(L^*) \cap N(L^*)^{\perp}$, and so $f' = 0$. Thus, $f = Pf \in R(L)$, and so $R(L) = N(L^*)^{\perp}$. □

Corollary 15.12. Let $L : \mathcal{H} \to \mathcal{H}$ be a bounded linear operator such that $R(L)$ is closed. Then, either $Lf = g$ has a solution, or there exists a $v \in N(L^*)$ such that $\langle g, v\rangle \neq 0$.

15.2 A Resolvent Example

This purpose of this example is to see the power of the Fredholm Alternative: it allows us to deduce when a solution exists (or, put another way, when L is surjective). So consider the following problem:

Let $K(x,y) = xy^2$, $Ku(x) = \int_0^1 k(x,y)u(y)dy$ and $Lu = I - \lambda ku$, $\lambda \in \mathbb{C}$. Assume L has a closed range.

1. Find the values of λ for which $Lu = f$ has a solution for all $f \in \mathcal{H}$ (that is, what values of λ make L surjective). Solve $Lu = f$ for these values.
2. For the remaining values of λ, find a condition on f that gurantees a solution to $Lu = f$. When f satisfies this condition, solve $Lu = f$.

Solution 15.13. Since $R(L)$ is closed, the Fredholm Alternative applies. So, $R(L) = N(L^*)^{\perp}$. We want $R(L) = \mathcal{H}$, so do to so, we show that $N(L^*) = \{0\}$. (This follows from the decomposition of $\mathcal{H}$.) Then, $L^* = I - \overline{\lambda}K^*$. We want to find what values of λ make the statement "$L^*u = 0$ implies $u = 0$" true. So suppose $L^*u = 0$. Then:

$$\begin{aligned}
L^*u = 0 = u - \overline{\lambda}\int_0^1 k(y,x)u(y)dy = 0 \\
= u - \overline{\lambda}\int_0^1 yx^2u(y)dy \\
= u - \overline{\lambda}x^2\int_0^1 yu(y)dy.
\end{aligned}$$

$\int_0^1 yu(y)dy$ is just a constant, so we have $u = Cx^2$. Replace this in u both on the left and in the integral:

$$0 = Cx^2 - \overline{\lambda}x^2 \int_0^1 u(y)dy = Cx^2 - \overline{\lambda}x^2 \int_0^1 Cy^2 dy = Cx^2 - C\overline{\lambda}x^2/4.$$

Then $Cx^2 = C\overline{\lambda}x^2/4$, so $C = \overline{\lambda}C/4$. So when $\lambda \neq 4$, we have that $C = 0$, and this in turn implies that $u = 0$. Thus, when $\lambda \neq 4$, $L^*u = 0$ implies $u = 0$. Thus, $N(L)^* = \{0\}$, and so $R(L) = \mathcal{H}$.

Now suppose $\lambda \neq 4$, and $u - \lambda x \int_0^1 y^2 u(y)dy = f$. We want to find u. To do so, take:

$$ux^2 - \lambda x^3 \int_0^1 y^2 u(y)dy = fx^2$$

$$\int_0^1 ux^2 dx - \lambda \int_0^1 x^3 \int_0^1 y^2 u(y)dydx = \int_0^1 fx^2 dx$$

Thus, $\int_0^1 fy^2 dy = \left(1 - \frac{\lambda}{4}\right) \int_0^1 uy^2 dy$, and so

$$\frac{\int_0^1 fy^2 dy}{1 - \frac{\lambda}{4}} = \int_0^1 u(y)y^2 dy$$

.

Therefore

$$u(x) = f(x) + \frac{\lambda x}{1 - \frac{\lambda}{4}} \int_0^1 fy^2 dy = f(x) + \frac{4\lambda}{4 - \lambda} Kf(x).$$

That is, if $Lu = f$, then $L^{-1} = (I + \frac{4\lambda}{4-\lambda}K)f$, so $u = L^{-1}f$. L^{-1} is called the *resolvent* of K.

b) When $\lambda = 4$, $L^*u = 0$ implies that $u = Cx^2$, so $N(L^*) = \mathit{span}\{x^2\}$. By Fredhom Alternative, $Lu = f$ when $f \in N(L^*)^{\perp}$, or $\int_0^1 x^2 f dx = 0$. (This is the condition we sought to satisfy.) Thus, so solve $u - 4x \int_0^1 y^2 u(y)dy = f$, note that $\int_0^1 y^2 u(y)dy$ cannot be determined by our techniques used above, for $\int_0^1 uy^2 dy - 4/4 \int_0^1 y^2 u(y)dy = \int_0^1 y^2 f dy = 0$. That is, we have consistency, so $C = \int_0^1 y^2 u(y)dy$ is arbitrary. Thus $u(x) = f(x) + Cx$.

Chapter 16

Compact Operators

First, a brief note on notation:

$$\mathcal{B}(\mathcal{H}) = \text{set of all bounded operators on } \mathcal{H}. \text{ Know that this is also a Banach Space.}$$

Definition 16.1. A subset $S \subset \mathcal{H}$ is *compact* if every sequence has a convergent subsequence. S is *precompact* if its closure if compact.

Proposition 16.2. The following hold:

1. Every compact set is bounded.
2. Let S be bounded. Then S is precompact iff every sequence has a convergent subsequence.
3. Let $\mathcal{H}$ be fintie dimensional. Every closed and bounded set is compact.
4. In an infinite dimensional space, closed and bounded is not enough.

Proof. 1) Suppose S was compact but not bounded. Then, there exists a sequence $\{x_n\}_1^\infty \subset S$ such that $||x_n|| \to \infty$. Since S is compact, there exists a convergent subsequence of $\{x_n\}_1^\infty$, say $\{x_{n_j}\}_1^\infty$. But every convergent sequence is bounded, and thus we have a contradiction.

2) Suppose S is bounded. Let S be precompact, and let $\{x_n\}_1^\infty \subset S$. Then, $\{x_n\}_1^\infty \subset \overline{S}$, which is compact. Thus, $\{x_n\}_1^\infty$ has a convergent subsequence.

Now suppose every sequence has a convergent subsequence. We claim that $\overline{S}$ is compact. $\overline{S}$ is closed, and suppose $\{x_n\}_1^\infty \subset \overline{S}$. Then, every x_j is the limit of a sequence in S. Thus, for each n, there exists $\{y_{n,m}\}_1^\infty \subset S$ such that $y_{n,m} \to x_n$ (limit of m). Then, for each n there exists M_n such that for all $m \geq M_n$, $||y_{n,m} - x_n|| < 1/n$. Then, the sequence $\{y_{n,M_n}\}_1^\infty$ has a convergent subsequence, $\{y_{n_k,M_{n_k}}\}_1^\infty$ such that $y_{n_k,M_{n_k}} \to y$. We claim that $x_{n_k} \to y$. Let $\epsilon > 0$. Then, for large enough k, $||y_{n_k,M_{n_k}} - y|| < \epsilon/2$. Then:

$$||x_{n_k} - y|| \leq ||x_{n_k} - y_{n_k,M_{n_k}}|| + ||y_{n_k,M_{n_k}} - y|| < 1/n_k + \epsilon/2.$$

Chose K large if necessary so that $1_{n_k} < \epsilon/2$, so $||x_{n_k} - y|| < \epsilon$.

3) This is Heine-Borel.

4) Let $S = \{f \in \mathcal{H} : \ ||f|| \leq 1\}$. Let $\{\phi\}_1^\infty$ be an o.n. basis for $\mathcal{H}$, so $\{\phi_n\}_1^\infty \subset \mathcal{H}$. But: $||\phi_n - \phi_m||^2 = ||\phi_n||^2 - 2\langle \phi_n, \phi_m \rangle + ||\phi_m||^2 = 2$, so $||\phi_n - \phi_m|| = \sqrt{2}$ for $n \neq m$. So $\{\phi_n\}_1^\infty$ does not have a Cauchy subsequence and thus does not have a convergent subsequence. $\square$

16.1 Compact Operators

Definition 16.3. Let $K : \mathcal{H} \to \mathcal{H}$ be linear. If K maps bounded sets into precompact sets, then K is a *compact operator*. Equivalently, K is compact if, for every bounded sequence $\{v_n\}_1^\infty$, there exists a subsequence $\{v_{j_k}\}_1^\infty$ such that $\{Kv_{j_k}\}_1^\infty$ converges. Dentoe $\mathcal{C}(\mathcal{H})$ as the space of compact operators.

Proposition 16.4. If $K \in \mathcal{C}(\mathcal{H})$, then K is bounded.

Proof. Suppose not. Let $\{u_n\}_1^\infty$ be such that $||u_n|| = 1$ for all n but $||Ku_n|| \to \infty$. Then, since K is compact, there exists a subseqenence $\{u_{n_k}\}_1^\infty$ such that $\{Ku_{n_k}\}_1^\infty$ converges. But convergent subsequences are bounded. This is a contradiction. $\square$

Proposition 16.5. Ever finite rank operator K is compact.

Proof. Since the range of K is finite dimensional, every bounded set is precompact. Let $\{f \in \mathcal{H} : ||f|| \leq C\} = S$. Then, $K(S)$ is bounded, as $||Kf|| \leq ||K||_{op}||f|| \leq C||K||_{op}$. Thus, K maps bounded sets into bounded sets, which are precompact. Thus K is compact. $\square$

Lemma 16.6. *Let $\{\phi_j\}_1^\infty$ be an o.n. set in $\mathcal{H}$ and $K \in \mathcal{C}(\mathcal{H})$. Then, $\lim_{j\to\infty} kK\phi_j = 0$.*

Proof. Suppose not, then there exists a subsequence such that $||K\phi_{j_n}|| > \alpha > 0$ for all n. Since K is compact, there exists a further subsequence $\{\phi_{j_{n_k}}\} := \{\phi_k\}_1^\infty$ such that $\{K\phi_k\}1^\infty$ converges, say to ψ. Since $\{\phi_k\} - 1^\infty$ is a subsequence of $\{\phi_{j_n}\}$, $||K\phi_k|| > \alpha > 0$ for all k. Then, $\lim_{k\to\infty} ||K\phi_k|| = ||\psi|| > 0$. But, by Bessel's Inequality, $\sum_{k=1}^\infty |\langle K\phi_k, \psi\rangle| = \sum_{k=1}^\infty |\langle \phi_k, K^*\psi\rangle| \leq ||K^*\psi||^2$. Thus, $\lim |\langle K\phi_j, \psi\rangle| \to 0$, and thus $\langle K\phi_j, \psi\rangle \to 0$. Then, $||\psi||^2 = 0$, a contradiction. $\square$

Definition 16.7. A sequence $\{f_n\}_1^\infty$ is *weakly convergent* to $f \in \mathcal{H}$ if for all $g \in \mathcal{H}$, $\lim_{n\to\infty}\langle f_n, g\rangle = \langle f, g\rangle$.

Lemma 16.8. *If $f_n \to f$ weakly, and K is bounded linear operator, then $Kf_n \to Kf$ weakly.*

Proof. Lt $g \in \mathcal{H}$. Then, $\langle Kf_n, g\rangle - \langle Kf, g\rangle = \langle Kf_n - Kf, g\rangle = \langle K(f_n - f), g\rangle = \langle f_n - f, K^*g\rangle$. But this goes to 0 as $n \to \infty$. Thus $Kf_n \to Kf$ weakly. $\square$

Proposition 16.9. Let $\{f_n\}$ be weakly convergent to $f \in \mathcal{H}$. If $K \in \mathcal{C}(\mathcal{H})$, then $\lim_{n\to\infty} Kf_n = Kf$. That is, K maps weakly convergent sequences to strongly convergent ones.

Proof. Suppose not. Then, there exists a subsequence $\{f_{n_k}\}_1^\infty$ and an $\alpha > 0$ such that $||Kf_{n_k} - f|| > \alpha > 0$. Since K is compact, we may find a subsequence $\{f_{n_{k_j}}\}_1^\infty := \{f_j\}_1^\infty$ such that $Kf_j \to \psi$. Thus, $\lim_{j\to\infty} ||Kf_j - Kf|| = ||\psi - Kf|| > \alpha > 0$.

Then, by the lemma $Kf_j \to Kf$ weakly. So, $\lim\{Kf_j, g\} = \langle \psi, g\rangle = \langle Kf, g\rangle$ for all $g \in \mathcal{H}$. But then $||\psi - Kf||62 = \langle \psi - Kf, \psi - Kf\rangle = \langle \psi, \psi - Kf\rangle - \langle Kf, \psi - Kf\rangle = 0$. But this is a contradiction. $\square$

Theorem 16.10. *$\mathcal{C}(\mathcal{H})$ is a closed subspace of $\mathcal{B}(\mathcal{H})$.*

Proof. It is obvious that $\mathcal{C}(\mathcal{H})$ is a subspace. We only need to show that $\mathcal{C}(\mathcal{H}) \subset \mathcal{B}(\mathcal{H})$ and closure. We already know that first, since we showed that compact operators are bounded. Thus all the remains is closure. So let $\{K_n\}_1^\infty \subset \mathcal{C}(\mathcal{H})$ such that $K_n \to K$ in operator norm. We want to show that K is compact. So let $\{u_n\}_1^\infty$ be a bounded sequence.

Now we want to construct our sequence very carefully (this is the trick for this proof). For K_1, we can select a convergent subseqence. Denote it by $\{u_n^1\}_1^\infty$. Now this subsequence is bounded, and so we may take a further subsequence for K_2: $\{u_n^2\}_1^\infty$. Note that since $\{u_n^2\}_1^\infty$ is a subsequence of $\{u_n^1\}_1^\infty$, it is convergent under K_1. We continue this process: find a subsequence $\{u_n^j\}_1^\infty$ that is convergent under K_m for $1 \leq m \leq j$. Now we define the "diagonal" sequence: $\{u_j\}_1^\infty$, where $u_j = u_j^j$. That is, the j term comes from the j^{th} term of the the sequence $\{u_n^j\}_1^\infty$. By construction, $\{u_j\}$ is a subsequence of $\{u_n^j\}_1^\infty$ at least by the j term. (That is, after the j term, u_j becomes a subsequence of u_n^j. Thus, for any n, $\{K_nu_j\}_1^\infty$ is convergent. (Since $\{u_j\}_1^\infty$ is a subsequence of $\{u_n^j\}_1^\infty$ after the j term, $\{u_j\}$ is convergent after the j term.)

Now we are ready for the fireworks. We show that $\{Ku_j\}_1^\infty$ is Cauchy, and thus convergent. Let $\epsilon > 0$, and $C \leq ||u_m||$, since these subsequneces are all bounded. :

$$\begin{aligned}
||Ku_n - Ku_m|| &\leq ||Ku_n - K_pu_n|| + ||K_pu_n - K_pu_m|| + ||K_pu_m - Ku_m|| \\
&\leq ||K - K_p||||u_n|| + ||K_pu_n - K_pu_m|| + ||K_p - K||||u_m|| \\
&\leq C||K - K_p|| + ||K_pu_n - K_pu_m|| + C||K_p - K||||u_m|| \\
&= 2C||K - K_p|| + ||K_pu_n - K_pu_m||.
\end{aligned}$$

We can choose p so large such that $||K - K_p|| < \epsilon/(4C)$. Since $\{u_j\}_1^\infty$ is convergent under K_p, $\{K_p u_j\}_1^\infty$ is Cauchy, and so we can choose m, n so large so that $||K_p u_n - K_p u_m|| < \epsilon/2$. Thus, we $||Ku_n - Ku_m|| < \epsilon$, and thus is Cauchy. Therefore $\{Ku_j\}_1^\infty$ converges, and so K is compact. $\square$

Corollary 16.11. Hilbert-Schmidt Operators are compact.

Proof. Let $\mathcal{H} = L^2[0,1]$ and suppose $k(x,y) \in L^2(R)$. Then the Hilbert-Schmidt operator is: $Ku = \int_0^1 k(x,y)u(y)dy$. Let $\{\phi_n\}_1^\infty$ be an o.n. basis for L^2. Then, $\{\phi_n(x)\phi_m(y)\}_{n,m=1}^\infty$ is a basis for $L^2(R)$. Note that $||K||_{op} \leq ||k||^{L^2}$. Put $k(x,y)$ in its basis representation:

$$k(x,y) = \sum_{n,m=1}^{\infty} a_{n,m}\phi_n(x)\phi_m(y).$$

Define $k_N(x,y) := \sum_{n,m=1}^{N} a_{n,m}\phi_n(x)\phi_m(y)$, and $K_N u := \int_0^1 k_N(x,y)u(y)dy$, so K_N is a finite rank operator, and so is compact. Then:

$$\begin{aligned}
||K - K_n||_{op}^2 &\leq ||k - k_n||_{L^2}^2 \\
&= ||\sum_{n,m=1}^{\infty} a_{n,m}\phi_n(x)\phi_m(y) - \sum_{n,m=1}^{N} a_{n,m}\phi_n(x)\phi_m(y)||^2 \\
&= ||\sum_{n,m=N+1}^{\infty} a_{n,m}\phi_n(x)\phi_m(y)||^2 \\
&= \sum_{n,m=N+1}^{\infty} |a_{n,m}|^2.
\end{aligned}$$

Since this is finite, as $N \to \infty$, $||K - K_N||_{op}^2 \to 0$, and so $K_N \to K$. Thus, K is the limit of compact operators, and thus is compact. $\square$

Proposition 16.12. Let $K \in \mathcal{C}(\mathcal{H})$ and let $L \in \mathcal{B}(\mathcal{H})$. Then both KL and LK are compact.

Proof. Let $\{u_j\}_1^\infty$ be a bounded sequence. Then, since L is bounded, $\{Lu_j\}_1^\infty$ is a bounded sequence. Therefore, there exists a subsequence $\{Lu_{j_k}\}_1^\infty$ such that $\{KLu_{j_k}\}_1^\infty$ converges. Thus, KL is compact. For LK, since K is compact, there exists a subsequence $\{u_{j_k}\}_1^i = \infty$ such that $\{Ku_{j_k}\}_1^\infty$ converges. Since L is bounded, it is continuous, and so $\lim_{k\to\infty} L(Ku_{j_k}) = L(\lim_{k\to\infty} Ku_{j_k})$, which converges. Thus LK is compact. $\square$

Proposition 16.13. *K is compact iff K^* is compact.*

Proof. Suppose K is compact. Then it is bounded, and thus so is K^*. Therefore, KK^* is compact. If $\{u_n\}_1^\infty$ is a bounded sequence, then there exists a subsequence $\{u_{n_j} := u_j\}_1^\infty$ that is convergent under KK^*. Thus, $\{KK^*u_j\}_1^\infty$ is Cauchy. Observe

$$\langle KK^*(u_j - u_k), u_j - u_k\rangle = \langle K^*(u_j - u_k), K^*(u_j - u_k)\rangle = ||K^*(u_j - u_k)||^2.$$

Since $\{u_j\}_1^\infty$ is bounded, we thus have $||K^*(u_j - u_k)||^2 \leq C||KK^*(u_j - u_k)||$. Since $\{KK^*(u_j)\}_1^\infty$ is convergent, it is also Cauchy. Thus, we may the right hand side as small as we want by sending $j,k \to \infty$. Therefore, $\{K^*u_j\}_1^\infty$ is Cauchy, and thus converges. Therefore, K^* is compact. Since $(K^*)^* = K$, the other direction follows immediately. $\square$

Chapter 17

Closed Range Theorem

Theorem 17.1. *If $K \in \mathcal{H}(\mathcal{H})$, $\lambda \in \mathbb{C}$, then the range of the operator $L := I - \lambda K$ is closed.*

Proof. If $N(L) =\neq \{0\}$, then we may not have a unique solution to $Lf = g$. In particular, if $h \neq 0 \in N(L)$, and f is a solution, then $L(f+g) = Lf + Lh = Lf$. So our solution may not be unique. To alleviate this issue, we utilize the decomposition of $\mathcal{H}$ into $N(L) \bigoplus N(L)^{\perp}$. In particular, if $Lf = g$, then $f = f_1 + f_2$ where $f_1 \in N(L)$ and $f_2 \in N(L)^{\perp}$, so we simply redefine $L : N(L)^{\perp} \to R(L)$ to ensure uniqueness. (That is, we mod out the null space.)

Now that we have made this adjustment, we prove that there exists a $c > 0$ such that for $f \in N(L)^{\perp}$, $||Lf|| > c||f||$.

Suppose not. Then, we may find a sequence $\{f_n\}_1^{\infty} \subset N(L)^{\perp}$ such that $||f_n|| = 1$ but $||Lf_n|| \to 0$. However, $Lf_n = f_n - \lambda K f_n$, so $f_n = Lf_n + \lambda K f_n$. Since the f_ns are bounded and K is compact, there exists a subsequence $\{f_{n_k}\}_1^{\infty} := \{f_k\}_1^{\infty}$ such that $\{Kf_k\}_1^{\infty}$ converges. By our choice of sequences, $Lf_K \to 0$. Thus, both $\lim k \to \infty L f_k$ and $\lim_{k\to\infty} \lambda K f_k$ exist. Then define:

$$\hat{f} = \lim_{k\to\infty} f_k = \lim_{k\to\infty} \lambda K f_k + \lim_{k\to\infty} Lf_k.$$

However, since L is continuous, $L\hat{f} = \lim_{k\to\infty} Lf_k = 0$. Thus, $\hat{f} \in N(L)$. However, $\{f_K\}_1^{\infty} \subset N(L)^{\perp}$, and $N(L)^{\perp}$ being a closed subspace implies that $\hat{f} \in N(L)^{\perp}$. Therefore, $\hat{f} = 0$. But $0 = ||\hat{f}|| = \lim_{k\to\infty} ||f_k|| = 1$. Thus a contradiction.

Now with this bound, we proceed to the main part of the proof. Let $\{g_n\}_1^{\infty} \subset R(L)$ be such that $g_n \to g$. Our goal is to find an $f \in \mathcal{H}$ such that $Lf = g$. This will show that the range is closed. As before, we limit $f \in N(L)^{\perp}$ to ensure uniqueness. Now:

$$||g_n - g_m|| = ||L(f_n - f_M)|| \geq c||f_n - f_m||.$$

But since $\{g_n\}_1^{\infty}$ is convergent, it is Cauchy, and thus $\{f_n\}_1^{\infty}$ is Cauchy. Thus $f_n \to f$ for some f. Thus, $g = \lim_{n\to\infty} Lf_n = Lf$. So $g \in R(L)$. □

Chapter 18

Spectral Theory for Compact Self-Adjoint Operators

First, let us consider the finite dimensional case. Let A be an $n \times n$ matrix. We say that λ is an eigenvalue of A is $Ax = \lambda x$ for some $x \neq 0$. When we want to find λ, we find $(A - I\lambda)x = 0$. That is we want to find what values of λ make $N(A - \lambda I) \neq \{0\}$, so that $x \neq\in N(A - \lambda I)$. In particular we find λ such that $\det(A - \lambda I) = 0$, so that $A - I\lambda$ is singular, and thus *not* invertible. So $(A - I\lambda)^{-1}$ does *not* exist. What we may conclude, then, is that if $(A - I\lambda)^{-1}$ does exist, then λ is not an eigenvalue.

To further elaborate, if $(A - I\lambda)^{-1}$ exists, it is injective, so $N(A - I\lambda)^{-1}) = \{0\}$. But then the only x that satisfies $(A - I\lambda)x = 0$ is $x = 0$. But then λ is not an eigenvector. We can formalize this.

Definition 18.1. The spectrum of a finite dimensional linear operator is $\{\lambda \in \mathbb{C} : \ Ax = \lambda x \text{for some } x \neq 0\} = \{\lambda \in \mathbb{C} : \det(A - I\lambda) = 0\} = \{\lambda \in \mathbb{C} : \ (A - \lambda I)^{-1} \textbf{ does not exist}\} = \{\lambda \in \mathbb{C} : \ (A - \lambda I)^{-1} \text{ exists}\}^c$.

This is a bit overkill for finite dimensions, but we can use this definition to move into infinite dimensions.

Definition 18.2. Let $L \in \mathcal{B}(\mathcal{H})$. The *resolvent set of* L is $\rho(L) := \{\lambda \in \mathbb{C} : \ (L - \lambda I)^{-1} \in \mathcal{B}(\mathcal{H})\}$. The operator $R_L(\lambda) := (L - \lambda I)^{-1}$ is called the *resolvent of* L. The *spectrum of* L, $\sigma(L)$ is defined as the complement of the resovlent set: $\sigma(L) = \rho(L)^c$.

Note 18.3. There is a subtle difference between what we have for finite dimensions. To be in the spectrum is to say $(L - \lambda I)^{-1} \notin \mathcal{B}(\mathcal{H})$, so it may exists, but not be over all of $\mathcal{H}$.

Lemma 18.4. *Let $L = L^* \in \mathcal{B}(\mathcal{H})$. Then the eigenvalues of L are real and the eigenvectors corresponding to distinct eigenvalues are orthogonal.*

Proof. First we show that eigenvalues are real:

$$\lambda\langle x, x\rangle = \langle \lambda x, x\rangle = \langle Lx, x\rangle = \langle x, L^* x\rangle = \langle x, Lx\rangle = \langle x, \lambda x\rangle = \overline{\lambda}\langle x, x\rangle.$$

Thus, $(\lambda - \overline{\lambda})\langle x, x\rangle = 0$. Since $x \neq 0$, we have that $\lambda = \overline{\lambda}$, and so λ is real. Now we use this to show orthogonalithy. Let $\lambda_1 \neq \lambda_2$ be eigenvalues (and so are real) and let x_1 and x_2 be eigenvectors corresponding to the eigenvalues. Then:

$$\lambda_1\langle x_1, x_2\rangle = \langle Lx_1, x_2\rangle = \langle x_1, L^* x_2\rangle = \langle x_1, Lx_2\rangle = \overline{\lambda_2}\langle x_1, x_1\rangle.$$

Thus, $(\lambda_1 - \lambda_2)\langle x_1, x_2\rangle = 0$. Since $\lambda_1 \neq \lambda_2$, $\langle x_1, x_2 = 0\rangle$. □

Lemma 18.5. *Let $\{\phi\}_1^\infty$ be an o.n. set in $\mathcal{H}$ and let $K \in \mathcal{C}(\mathcal{H})$. Then* $\lim_{n\to\infty} K\phi_n = 0$.

Proof. Suppose not. Then, there exists a $c > 0$ such that $||K\phi_n|| > c$ for all n. Now $\{\phi_n\}_1^\infty$ is bounded, and so since K is compact, there exists a subsequence $\phi_{n_K} := \phi_k$ such that $\{K\phi_k\}_1^\infty$ converges. Since this is a subsequence $||\psi|| \lim ||K\phi_K|| > 0$, so $\psi \neq= 0$. Now, by Bessel's Inequality,

$$\sum_1^\infty |\langle K\phi_k, \psi\rangle|^2 = \sum_1^\infty |\langle \phi_k, K^*\psi\rangle|^2 \leq ||K^*\psi||^2 < \infty.$$

Then, the sum on the left is convergent, and so $|\langle K\phi_k, \psi\rangle| \to 0$, so $\langle K\phi_k, \psi\rangle \to 0$. But then

$$0 = \lim\langle K\phi_k, \psi\rangle = \langle \psi, \psi\rangle = ||\psi||^2.$$

But this implies that $\psi = 0$, which is a contradiction. □

Proposition 18.6. If $K \in \mathcal{K}(\mathcal{H})$, then $\sigma(K)$ consists only of eigenvalues together with 0.

Proof. While true in general, we will only show for when $K = K^*$. Suppose $\lambda \in \sigma(K)$, $\lambda \neq 0$. We want to show λ is in fact an eigenvalue. So, since $\lambda \in \sigma(K)$, $K - \lambda I \notin \mathcal{B}(\mathcal{H})$. This is possible either because $N(K - \lambda I) \neq \{0\}$ or $(K - \lambda I)^{-1}$ is not bounded. (Note that if the inverse is unbounded, then $(K - \lambda I)$ is mapped to a dense subset of $\mathcal{H}$, but not the whole space.)

If $N(K - \lambda I) \neq \{0\}$, for some $u \neq 0$ $(K - \lambda I)u = Ku - \lambda u = 0$, so $Ku = \lambda u$, and thus λ is an eigenvalue.

Now suppose the range of $K - \lambda I$ is not all of $\mathcal{H}$. Since K is compact, it has closed range (a simple adjustment to the Closed Range Theorem shows this), and so the Fredholm Alternative applies to $L := K - \lambda I$. That is, $R(L) = N(L^*)^\perp$. Recall that we have $N(L^*) \bigoplus N(L^*)^\perp = N(L^*) \bigoplus R(L)$. By assumption, $R(L) \neq \mathcal{H}$, thus there exists a $g \in N(L^*)$ such that $g \neq 0$. Then $L^*g = 0$, so $K^*g - \overline{\lambda}g = 0$, so $K^* = \overline{\lambda}g$, and $\overline{\lambda}$ is an eigenvalue of K^*. But $K = K^*$, so $\overline{\lambda}$ is an eigenvalue of K. Since all eigenvalues of a self-adjoint operator are real, $\overline{\lambda} = \lambda$. So λ is an eigenvalue.

Now we show $0 \in \sigma(K)$. Suppose not. Then, $(K - 0\lambda)^{-1} = K^{-1} \in \mathcal{B}(\mathcal{H})$. Let $\{\phi_n\}_1^\infty$ be an o.n. set and define $\psi_k = K\phi_n$. Then, $K^{-1}\psi_k = \phi_k$, so $||\phi_n|| = 1 = ||K^{-1}\phi_k|| \leq ||K^{-1}||||\psi_k||$. However, $\lim_{k\to\infty} K\phi_k = 0$, so $\psi = 0$. But this is a contradiction. Thus, $0 \in \sigma(K)$. □

Proposition 18.7. Let $K \in \mathcal{K}(\mathcal{H})$. If $\lambda \neq 0$ is an eigenvalue of K with eigenspace $\mathcal{E}_\lambda$, then $\mathcal{E}_\lambda$ is finite dimensional.

Proof. Since λ is an eigenvalue of K, if $x \in \mathcal{E}_\lambda$, then $Kx = \lambda x$ implies $(K - \lambda I)x = 0$, so $x \in N(K - \lambda I)$ and $\mathcal{E}_\lambda \subset N(K - \lambda I)$. Likewise, if $x \in N(K - \lambda I)$, then $Kx = \lambda x$, so $x \in \mathcal{E}_\lambda$. Thus, $\mathcal{E}_\lambda = N(K - \lambda I)$. Since K is bounded, $N(K - \lambda I) = \mathcal{E}_\lambda$ is closed. Thus, choose an o.n. basis $\{\phi_n\}_1^N$. If $N = \infty$, then $K\phi_n = \lambda\phi_n$, and $||K\phi_n|| = |\lambda|||\phi_n|| = |\lambda|$. But $||K\phi_n|| \to 0$, and thus a contradiction. Thus $N < \infty$, and we are done. □

Proposition 18.8. Let $K \in \mathcal{C}(\mathcal{H})$ be self-adjoint. Then 0 is the only possible accumulation point of the eigenvalues of K.

Proof. Suppose not. Then we may choose a sequence $\{\lambda\}_1^\infty$ such that $\lim_{n\to\infty} \lambda_n \neq 0$. Let $\{\phi_n\}_1^\infty$ be the set of eigenvalues corresponding to the respective eigenvalues such that $||\phi_n|| = 1$. Note that this implies that $\{\phi_n\}_1^\infty$ is an orthogonal set, and thus an o.n. set. Then, $||K\phi_n|| = |\lambda_n|$, but then $\lim ||K\phi_n|| = 0 = \lim |\lambda_n|$, so $\lambda_n \to 0$. This is a contradiction. □

18.1 Spectral Theory for Self-Adjoint Compact Operators

Lemma 18.9. *Let $L = L^* \in \mathcal{B}(\mathcal{H})$. Then, $||L|| = \sup_{||u||=1} |\langle Lu, u\rangle|$.*

Proof. Omitted. □

Lemma 18.10. *Let $K = 0 \in \mathcal{C}(\mathcal{H})$ be self-adjoint. Then either $||K||$ or $-||K||$ or both are eigenvalues.*

Proof. By the lemma, $||K|| = \sup_{||u||=1} |\langle Ku, u\rangle|$. Thus, we may choose a sequence $\{u_n\}$ such that $||u_n|| = 1$ and $||K|| = \lim_{n\to\infty} |\langle Ku_n, u_n\rangle|$. Thus, $\langle Ku_n, u_n\rangle \to \pm||K||$. Suppose we have $+||K||$. (The other case follows identically.) Then:

$$
\begin{aligned}
||Ku_n - ||K||u_n||^2 &= ||Ku_n||^2 - 2||K||\langle Ku_n, u_n\rangle + ||K||^2||u_n||^2 \\
&\leq ||K||^2 - 2||K||\langle Ku_n, u_n\rangle + ||K||^2 \\
&= 2||K||^2 - 2||K||\langle Ku_n, u_n\rangle \\
&= 2||K||(||K|| - \langle Ku_n, u_n\rangle).
\end{aligned}
$$

Since $\langle Ku_n, u_n\rangle \to ||K||$, we have $||Ku_n - ||K||u_n|| \to 0$. So, $Ku_n - ||K||u_n$ converges to 0. Now, since $||u_n|| = 1$ and K being compact implies that $\{Ku_{n_k}\}_1^\infty$ converges for some $\{u_{n_k}\}_1^\infty := \{u_k\}_1^\infty$. Say $Ku_k \to u$ for some u. Then

$$
\begin{aligned}
||K||u_k &= Ku_k - (Ku_k - ||K||u_k) \implies \\
||K|| \lim_{k\to\infty} u_k &= \lim Ku_k - \lim(Ku_k - ||K||u_k) \\
||K|| \lim_{k\to\infty} u_k &= \lim Ku_k - 0 \\
&= u.
\end{aligned}
$$

Now since K is continuous $||K||u = \lim ||K||Ku_k = Ku$. Then, $||K||u = Ku$. Since $||u|| = ||\lim_{k\to\infty} u_k|| = 1$, $u \neq 0$, and so $||K||$ is an eigenvalue. □

Definition 18.11. A self-adjoint operator K is *positive* if $\langle Ku, u\rangle \geq 0$ for all $u \in \mathcal{H}$.

Definition 18.12. Let $U \subset \mathcal{H}$ be a subspace, $L \in \mathcal{B}(\mathcal{H})$. Then U is *invariant under* L if $L(U) \subset U$.

Lemma 18.13. *Let M_n be the span of eigenvalues of $||K|| = \lambda_1 > \lambda_2 > ... > \lambda_n > 0$ and let $M_n^\perp$ be the orthogonal compliment. Then M_n and $M_n^\perp$ are both invariant, with $K = K^*$.*

Proof. Let $v \in M_n$, so $v = \sum_1^n \alpha_j u_j$ where u_j is an eigenvector. Then, $Kv = \sum_1^n \alpha_j K(u_j) = \sum_1^n \alpha_j \lambda_j u_j \in M_n$. Thus M_n is invariant.

Now let $u \in M_n^\perp$, $p \in M_n$. Then,

$$
\langle Ku, p\rangle = \sum_1^n \overline{\alpha_j}\langle Ku, u_j\rangle = \sum_1^n \overline{\alpha_j}\langle u, K^* u_j\rangle = \sum_1^n \overline{\alpha_j}\langle u, Ku_j\rangle = \sum_1^n \overline{\alpha_j}\langle u, \lambda_j u_j\rangle = \sum_1^n \overline{\alpha_j}\overline{\lambda_j}\langle u, u_j\rangle = 0.
$$

Thus, $K(M_n^\perp) \subset M_n^\perp$. □

Lemma 18.14. *Let $K \neq 0 \in \mathcal{C}(\mathcal{H})$ be self-adjoint and nonnegative. If K has n positive integers, then*

$$
\lambda_{n+1} = \sup\{\langle Ku, u\rangle : \ u \in M_n^\perp, \ ||u|| = 1\} < \lambda_n.
$$

Proof. Define $K_{n+1} := K|_{M_n^\perp}$, so that for $w \in M_n^\perp$, $K_{n+1}w = Kw$. Since K is compact on $\mathcal{H}$, K_{n+1} is compact on $M_n^\perp$. Thus:

$$
K_{n+1} = \sup\{\langle K_{n+1w}, w\rangle, \ w \in M_n^\perp, \ ||w|| = 1\}
$$

is an eigenvalue for K_{n+1} with $w \neq 0$ being the eigenvector. Therefore, $||K_{n+1}||$ is an eigenvalue for K as well. Define $\lambda_{n+1} = ||K_{n+1}||$.

To show the inequality, we observe that by construction $\lambda_j = ||K_j|| = \sup\{\langle Ku, u\rangle : \ u \in M_{j-1}^\perp, \ ||u|| = 1\}$. Note that since $M_n \subset M_{n+1}$, $M_{n+1}^\perp \subset M_n^\perp$. Then by properties of supremum, $\lambda_{j+1} < \lambda_j$. □

Proposition 18.15. From among the eigenvectors of K corresponding to the nonzero eigenvalues of K, one may select an orthonormal basis for $R(K)$. Moreover, if $R(K)$ is dense in $\mathcal{H}$, then that set forms an orthonormal basis for $\mathcal{H}$.

Proof. Let P_n be the orthogonal projection of $\mathcal{H}$ onto M_n and $P_n^\perp$ for $M_N^\perp$. Let $g = Ku \in R(K)$. We may write $u = u_n + u_n^\perp$, $P_n u = u_n$, $P_n^\perp u = u_n^\perp$. Since M_n and $M_n^\perp$ are invariant,

$$g = Ku = K(u_n + u_n^\perp) = Ku_n + Ku_n^\perp = g_n + g_n^\perp.$$

Thus, $g - g_n = g_n^\perp = K(u - u_n) = Ku_n^\perp = K_{n+1}u_n^\perp$. Then:

$$||g - g_n|| = ||K_{n+1}u_n^\perp|| \leq ||K_{n+1}||||u_n^\perp|| = \lambda_{k+1}||u_n^\perp|| \leq \lambda_{k=1}||u||.$$

Thus, $||g - g_n|| \leq \lambda_{k+1}||u||$. Now there are two possibilities to consider. If there are only n nonzero eigenvalues, $\lambda_{n+1} = 0$, so then $||g - g_n|| = 0$, so $g = g_n$ and $R(K) = M_n$.

Suppose there are infinitely many nonzero values. Then λ_n is a decreasing sequence bounded below, and so converges. More importantly, $\lambda_n \to 0$, as 0 is the only limit point of eigenvalues. Thus, $\lim_{n\to\infty} ||g_n - g|| \leq \lim_{n\to\infty} \lambda_{n+1}||u|| = 0$. So $g = g_n$ and

$$g = \lim_{n\to\infty} \sum_{k=1}^{\infty} \sum_{j=1}^{dim(\mathcal{E}_\lambda)} \langle g, \phi_{j,k} \rangle \phi_{j,k}$$

and we have a basis. □

Theorem 18.16 (Spectral Theorem)**.** *Let $K =\neq 0 \in \mathcal{C}(\mathcal{H})$ be self-adjoint. Then, from among the eigenvectors of K, including those for $\lambda = 0$, we may select an orthonormal basis for $\mathcal{H}$.*

Proof. Since $K = K^*$ and $N(K)$ is closed, we have $\mathcal{H} = \overline{R(K)} \bigoplus N(K^*)$. The basis from the above proposition gives us a basis for $R(K)$. We claim that this basis works for $\overline{R(K)}$. Let $\{g_n\}_1^\infty \subset R(K)$ and $g_n \to g$. Let $\{\phi_j\}_1^\infty$ be a basis for $R(K)$. Extend this basis to $\overline{R(K)}$ with $\{\psi_j\}_1^\infty$. Then

$$g = \sum_1^\infty \beta_j \phi_j + \sum_{k=1}^\infty \gamma_k \psi_k.$$

Then, $||g - g_n|| = ||\sum_{j=1}^\infty (\beta_j - \alpha_j)\phi_j + \sum_{k=1}^\infty \gamma_k \psi_k|| \to 0$.

Now the second sum is not in the span of the ϕ_js, and thus $\psi_k = 0$ for all k. Thus, $g \in span(\{\phi_j\}_1^\infty)$, and so $\{\phi_j\}_1^\infty$ is a basis for $\overline{R(K)}$. Now, for $N(K^*)^\perp = N(K)$, $N(K)^\perp = \mathcal{E}_{\lambda=0}$ and we can find a basis for this space. Take the two, and we are done. □