Lecture Notes on:

Math 6710: Applied Linear Operators and Spectral Theory (Applied Mathematics)

Chris Miles
miles@math.utah.edu

Last Updated: June 18, 2015

Contents

Disclaimer: These notes are for my personal use. They come with no claim of correctness or usefulness.

Errors or typos can be reported to me via email and will be corrected as quickly as possible.

Acknowledgments: The content of these notes is based on lectures delivered (excellently) by David C. Dobson for Math 6710 in Fall 2013.

The notes themselves are derived from notes originally taken by Gregory Handy.

Many proofs have been omitted for brevity but can be found in either:

- E. Kreyszig, *Introductory Functional Analysis with Applications*, Wiley (1989).
- F. G. Friedlander and M. Joshi, *Introduction to the Theory of Distributions*, Cambridge (1999).

1

Metric spaces

1.1 Preliminaries

Definition 1.1. A **metric space** is a set X, together with a mapping $d : X \to \mathbb{R}$, with properties for all $x, y, z \in X$:

(D1) $d(x, y) \geq 0$ with $d(x, y) = 0$ iff $x = y$

(D2) $d(x, y) = d(y, x)$

(D3) $d(x, y) \leq d(x, z) + d(z, y)$

Example 1.1. $\mathbb{R}$ is a metric space with $d(x, y) \coloneqq |x - y|$.

Example 1.2. $\mathbb{R}^3$ is also a metric space with the Euclidean metric.

Example 1.3. Let $\ell^\infty \coloneqq \{x = (\xi_1, \xi_2, \dots) : \xi_j \leq c_x\}$, where c_x does not depend on j, just x. The distance defined on this metric space is: $d(x, y) \coloneqq \sup_j |\xi_i - \eta_i|$.

Example 1.4. Let $C[a, b]$ be the set of all continuous functions on $[a, b]$, then $d(x, y) \coloneqq \max_{t \in [a,b]} |x(t) - y(t)$ is a metric space. We know the maximum is attained due to the continuity of x, y.

Example 1.5. The following is known as the *discrete metric space*. X can be anything and the metric is:

$$d(x, y) \coloneqq \begin{cases} 1 & \text{if } x = y \\ 0 & \text{otherwise} \end{cases} .$$

This is typically used as a counter-example.

Example 1.6. Let $B(A)$ be the set of all bounded functions on A, i.e., $f : A \to \mathbb{R}$. Then $\exists c_f : \sup_{t \in A} |f(t)| \leq c_f$. We define this metric to be $d(x, y) \coloneqq \sup_{t \in A} |x(t) - y(t)|$.

Example 1.7. Let $\ell^p := \left\{x = (\xi_1, \xi_2, \ldots) : \sum_{j=1}^{\infty} |\xi_j|^p < \infty\right\}$. The distance defined on this metric space is: $d(x,y) := \left(\sum_{j=1}^{\infty} |\xi_j - \eta_j\right)^{1/p}$. Note that $p = 2$ is very similar to the Euclidean norm and will be a noteworthy case. It is worth definition two noteworthy inequalities:

Theorem 1.1 (Hölder's inequality)**.**

$$\sum_{j=1}^{\infty} |\xi_j \eta_j| \leq \left(\sum_{j=1}^{p} |\xi_j|^p\right)^{1/p} \cdot \left(\sum_{j=1}^{p} |\eta_j|^p\right)^{1/q},$$

where p, q are conjugate, i.e. $1/p + 1/q = 1$. Note, Cauchy-Schwarz inequality is a special case when $p = q = 2$.

Theorem 1.2 (Minkowski's inequality)**.**

$$\left(\sum_{j=1}^{p} |\xi_j + \eta_j|^p\right)^{1/p} \leq \left(\sum_{j=1}^{p} |\xi_j|^p\right)^{1/p} + \left(\sum_{j=1}^{p} |\eta_j|^p\right)^{1/p}$$

1.2 Open sets, closed sets, topology

For $x_0 \in X$, define the following:

$B(x_0, r) := \{x \in X : d(x_0, x) < r\}$ (open ball)
$\tilde{B}(x_0, r) := \{x \in X : d(x_0, x) \leq r\}$ (closed ball)
$s(x_0, r) := \{x \in X : d(x_0, x) < r\}$ (sphere)

Definition 1.2. A subset $M \subset X$ is called **open** if each point x_0 of M has a ball $B(x_0, \epsilon) \subset M$ for some $\epsilon > 0$. Also, $K \subset X$ is closed if $K^C := X - K$ is open. A **neighborhood** N of x_0 is any set which contains an open ball around x_0.

Definition 1.3. x_0 is an interior point of M if M is in a neighborhood of x_0. i.e. $\exists \epsilon > 0 :$ $B(x_0, \epsilon) \subset M$.

Definition 1.4. The collection of all interior points of M is called M^0 or $\mathrm{Int}(M)$.

The collection $\mathcal{J}$ of all open sets on (X, d) is called a **topology** for (X, d) if it satisfies:

(T1) $\varnothing \in \mathcal{J}$

(T2) The union of any collection (infinite or finite) of elements of $\mathcal{J}$ is in $\mathcal{J}$.

(T3) Any *finite* intersection is in $\mathcal{J}$.

Note, no metric is needed for a topology but a metric always induces a topology.

Definition 1.5. Let $\rho : X \to Y$ where (X, d) and $(Y, \bar{D})$ are metric spaces. T is called **continuous** at $x_0 \in X$ if: $\forall \epsilon > 0, \exists \delta > 0 : \bar{d}(Tx, Tx_0) < \epsilon$ whenever $d(x, x_0) < \delta$. Alternatively: $T(B(x_0, \delta)) \subset B(Tx_0, \epsilon)$.

Definition 1.6. Let $M \subset X$. A point $x_0 \in X$ is called an **accumulation point** (or a limit point), if every neighborhood of x_0 contains at least one point M, distinct from x_0.

Definition 1.7. X plus all of its accumulation points is called the **closure**, denoted $\overline{X}$. Interestingly, in a metric space $\tilde{B}(x_0, \epsilon)$ is not necessarily equal to $\overline{B(x_0, \epsilon)}$.

Definition 1.8. A subset $M \subset X$ is called **dense** if $\overline{M} = X$.

Definition 1.9. A metric space is called **separable** if it has a countable dense subset.

Example 1.8. $X = [0, 1], d(x, y) := |x - 1|$. $M = (0, 1)$ is dense in X.

Example 1.9. $\mathbb{Q} \subset \mathbb{R}$ is dense, thus $\mathbb{R}$ is separable.

Example 1.10. $\mathbb{C}$ is separable for the same reason.

Example 1.11. ℓ^∞ is *not* separable.

Proof. Choose $x \in [0, 1]$ and write as binary $x = \xi_1 + \xi_2/2 + \xi_3/2^2 + \cdots$ and let $y = (\xi_i) \in \ell^\infty$. There are uncountably many of these but the distance between them is always 1. Thus, consider a ball of radius $1/3$ around each point. No countable dense subset exists. Thus, not separable. □

Example 1.12. ℓ^p for $1 \leq p < \infty$ is separable.

Proof. Idea: $(\eta_1, \eta_2, \ldots, \eta_n, 0, \ldots)$ is close enough. □

1.3 Completeness

Definition 1.10. A sequence (x_n) in a metric space (X, d) is **convergent** if there is an $x \in X : \lim_{n \to \infty} d(x_n, x) \to 0$.

Definition 1.11. A sequence (x_n) is **Cauchy** if $\forall \epsilon > 0, \exists N : d(x_n, x_m) < \epsilon$ whenever $n, m > N$.

Fact: every convergent sequence is a Cauchy Sequence is convergent, but the reverse is not necessarily true.

Definition 1.12. A space is said to be **complete** if all Cauchy sequences convergence.

Definition 1.13. A sequence is **bounded** if the set of its elements is bounded.

Theorem 1.3. Convergent sequences are bounded. Furthermore, convergent sequences have unique limits and if $x_n \to x, y_n \to y$, then $d(x_n, y_n) \to d(x, y)$.

Theorem 1.4. Let $M \subset X$, then $x \in \overline{M}$ iff $\exists (x_n) \subset M : x_n \to x$. Also, M is closed iff every sequence in M which converges, converges to a point in M.

Theorem 1.5. A subspace M of X is complete iff M is closed as a subspace of X.

Example 1.13. ℓ^∞ is complete.

Proof. Idea: Consider $(x_n = \{\xi_1^{(n)}, \ldots\})$, a Cauchy sequence, then $\xi_1^{(j)}$ forms a Cauchy sequence of real numbers, thus converges and defines our limit. □

Example 1.14. Consider c, the subspace of ℓ^∞ consisting of the convergent sequences is complete by noting that it is closed.

Example 1.15. ℓ^p for $1 \leq p < \infty$ is complete.

Example 1.16. $C[a, b]$ is complete.

Proof. Let (x_n) be a Cauchy sequence in $C[a, b]$. For a fixed t_0, $x_n(t_0)$ is a Cauchy sequence of real numbers and therefore converges. Now, note that $x(t)$ is continuous by noting that $|x(t) - x(t+h)| < |x(t) - x_n(t)| + |x_n(t) - x_n(t+h)| + |x_n(t+h) - x(t+h)| < \epsilon$, thus $x(t)$ is continuous. □

Example 1.17. $\mathbb{Q}$, polynomials on $[a, b]$, and $C[a, b]$ with the metric $d(x, y) = \int_a^b |x(t) - y(t)|\, dt$ are all examples of *incomplete* spaces. Note the last example comes from approximating a step function.

Definition 1.14. The **completion** of a metric space (X, d), denoted $(\hat{X}, \hat{d})$ exists and is defined by an isometry (distance preserving) map to a dense subset of $\hat{X}$. The completion is unique up to isometry.

1.4 Banach Fixed-point Theorem

This theorem is also known as the *contraction mapping principle*. In general if we have a mapping $T : X \to X$ and we want to solve $Tx = x$, the solution is called a *fixed point*.

Definition 1.15. T is a **contraction** on a set S if $\exists a < 1 : d(Tx, Ty) < a\, d(x, y)$ for all $x, y \in S$.

Theorem 1.6 (Banach-fixed point theorem)**.** If (X, d) is a complete metric space and T is a *contraction*, then T has a unique fixed point $x \in X$.

We can define the iteration $x_{n+1} = Tx_n$ for $n = 0, 1, \ldots$. Using a geometric series argument, we see that $d(x_n, x) \leq \frac{a^n}{1-a} d(x_0, x)$, which gives us an estimate of the error.

Corollary 1.1. Let T be a mapping of a complete metric space (X, d) onto itself. Suppose T is a contraction on some closed ball Y with radius r around a point x_0 and also assume that $d(x_0, Tx_0) < (1-a)r$, then fixed point iteration converges to a unique fixed point in Y.

Thus, if we have a contraction on a smaller space, as long as it maps onto itself, we still have our contraction mapping.

Example 1.18. We can apply this to linear algebra. Consider $X = \mathbb{R}^n$ with $d(x, y) = \max_j |\xi_j - \eta_j|$. Consider trying to solve $x = Cx + B$, where $C = (c_{ij}), B = (b_1, b_2, \ldots b_n)$. We hope to solve this by $(I - C)x = B$, the form of our fixed point iteration. Thus, define $T : X \to X$ as $Tx := Cx + B$, then we again hope to solve $Tx = x$. Consider $|Tx - Ty| = |C(x - y)|$. Note, in particular:

$$\begin{aligned}|Tx_j - Ty_j| &= \left|\sum_{k=1}^{n} c_{jk}(\xi_k - \eta_k)\right| \\ &\leq \sum_{k=1}^{n} |c_{jk}| \max_k |\xi_k - \eta_k| \\ &= \sum_{k=1}^{n} |c_{jk}| d(x, y)\end{aligned}$$

and take the max over j on both sides, yielding: $K = \max_j \sum_{k=1}^{n} |c_{jk}|$ as our contraction parameter.

A similar result holds for *Jacobi* iteration, where $A = (A - D) + D$, where D are the diagonal entries of A. We now are solving the problem $(A - D)x + Dx = C$, meaning $x = -D^{-1}(A - D)x + D^{-1}C$. Using the same criteria we end up with $\sum_{k=1}^{n} \left|\frac{a_{jk}}{a_{jj}}\right| < 1$ for a contraction. Note this is equivalent to row diagonal dominance.

Gauss-Seidel iteration is also similar, where $A = -L + D - U$, where L, U are the lower and upper triangular parts of A respectively. We now are solving the problem $(D - L)x - Ux = C$, meaning $x = (D - L)^{-1}Ux + (D - L)^{-1}C$. The condition above is sufficient but a looser exists. Note this method converges in fewer iterations but is more difficult to parallelize than the *Jacobi* iteration.

We can also use the contraction mapping to prove the following theorem from ordinary differential equation theory:

Theorem 1.7 (Picard)**.** Given the initial value problem:

$$\begin{cases} x' = f(x, t) \\ x(t_0) = x_0, \end{cases} \tag{1.1}$$

where f is continuous on $R = \{(t, x_0) : |t - t_0| \leq \alpha, |x_0| \leq b\}$. Since f is continuous, it is bounded by $|f(x, t)| \leq c$ in this interval. Also, suppose f satisfies the Lipschitz condition:

$\exists K : |f(x,t) - f(y,t)| \le k|x-y|$, then the IVP defined by 1.1 has a unique solution on the interval $J = [t_0 - \beta, t_0 + \beta]$, where $\beta < \min\{a, b/c, 1/k\}$.

Proof. Integrating the IVP yields:

$$x(t) = x_0 + \int_{t_0}^{t} f(\tau, x(\tau))\, d\tau$$

Thus, define the integral operator: $(Tx)(t) \coloneqq x_0 + \int_{t_0}^{t} f(\tau, x(\tau))\, d\tau$, meaning we are solving $Tx = x$. Consider $X = C[J]$, where J is defined above and define $d(x,y) = \max_{t\in J} |x(t) - y(t)|$. Note this is $\max$ not $\sup$ because we have continuous functions on an interval. Let $\mathcal{J} \subset C[J]$ be the subspace consisting of functions that satisfy $|x(t) - x_0| \le c\beta$, which is closed and therefore complete. Now consider $T : \mathcal{J} \to \mathcal{J}$, which yields:

$$\begin{aligned} |Tx(t) - x_0| &= \left| \int_{t_0}^{t} f(\tau, x(\tau))\, d\tau \right| \\ &\le c|t - t_0| \qquad \text{since } x \in \mathcal{J}, \tau \in J \\ &\le c\beta. \end{aligned}$$

We must now show that T is a contraction:

$$\begin{aligned} |Tx(t) - Ty(t)| &\le \left| \int_{t_0}^{t} f(\tau, x(\tau)) - f(\tau, y(\tau))\, d\tau \right| \\ &\le |t - t_0| K \max_{t\in\mathcal{J}} |x(\tau) - y(\tau)| \\ &= K|t - t_0|\, d(x,y) \\ &\le K\beta\, d(x,y). \end{aligned}$$

This suggests that $d(Tx, Ty) \le K\beta\, d(x,y)$, thus if $K\beta < 1$, we have a contraction, which is true by the definition of β. Thus, we have the *Picard iteration*:

$$x_{n+1}(t) = x_0 + \int_{t_0}^{t} f(\tau, x_n(\tau))\, d\tau. \qquad \square$$

Example 1.19. We can also consider integral equations. Consider those of the *first kind* to be of the form: $\int_a^b k(t,\tau)x(\tau)\, d\tau$, which is the continuous extension of multiplying by a matrix. We assume k is continuous on this square and therefore is bounded.

We can also define integral equations of the *second kind* which take the form: $x(t) - \mu \int_a^b k(t,\tau)x(\tau)\, d\tau + v(t)$. Here, μ is a parameter and k is again assumed to be continuous. We typically hope to solve for x, which can be done by rearranging: $(I - \mu k)x = v$. Thus, define the integral operator: $Tx(t) \coloneqq \mu \int_a^b k(t,\tau)x(\tau)\, d\tau + v(t)$, meaning we hope to

$Tx(t) = x(t)$. We must show that T is a contraction:

$$\begin{aligned} |Tx(t) - Ty(t)| &= \left| \mu \int_a^b k(t,\tau) \left[x(\tau) y(\tau) \right] \, d\tau \right| \\ &\leq |\mu| \left| \int_a^b k(t,\tau) \left[x(\tau) y(\tau) \right] \, d\tau \right| \\ &\leq |\mu| c(b-a) \max_\tau |x(\tau) - y(\tau)| \\ &= |\mu| c(b-a) \, d(x,y) \end{aligned}$$

Thus, if $c|\mu|(b-a) < 1$, we have a contraction. Also note that if the integral upper bound is changed to t, we can redefine our kernel to be 0 unless $\tau \leq t$.

2

Normed Vector Spaces, Banach Spaces

2.1 Preliminaries

Definition 2.1. A **vector space** over a field K is a non-empty set of elements (vectors) together with two algebraic operations $(+, \times)$ that satisfy the usual vector space axioms not listed here.

Example 2.1. $\mathbb{R}^n, \mathbb{C}^n$ are both vector spaces.

Example 2.2. $C[a, b]$ is a vector space.

Example 2.3. ℓ^p are vector spaces for all p.

Definition 2.2. A **subspace** of a vector space X is a non-empty subset that is closed under addition and multiplication with scalars and the subset is a vector space itself.

As usual, we can write elements in a vector space as (finite) linear combinations of the other elements. A linearly independent set is one in which: $\alpha_1 x_1 + \cdots + \alpha_n x_n = 0 \implies \alpha_1 = \alpha_n = 0$, otherwise they are linearly dependent.

Definition 2.3. A vector space is called **finite-dimensional** if there is a non-negative integer n such that the space contains n linearly independent vectors but any subset of the space with $n + 1$ vectors is linearly dependent. In this case, we consider $\dim X = n$.

Example 2.4. $X = \{0\}, \dim X = 0$.

Example 2.5. $\dim \mathbb{R}^n = \dim \mathbb{C}^n = n$.

Example 2.6. $\dim \ell^p = \infty$.

If B is any linearly independent subset of X such that $\operatorname{span} B = X$ then B is called a (Hamel) basis of X. Every x has a unique representation as a linearly combination of elements in B.

Note, every vector space has a Hamel basis but it can't always be found. Every Hamel basis has the same cardinality. All proper subspaces have lesser dimension.

2.2 Norms

Definition 2.4. A norm $\|\cdot\|$ on a vector space X is a (continuous by Lipschitz, see reverse triangle inequality) function $X \to \mathbb{R}$ that satisfies the usual norm properties, including the triangle inequality.

The norm provides us with the notion of the size of an element. Equipping a space with a norm has a profound impact and provides complex structure.

Definition 2.5. A **normed vector space (NVS)** is a vector space with a norm defined. A norm on X automatically defines a metric: $d(x, y) = \|x - y\|$ and this is called the **induced norm**. Note, not every metric is induced by a norm (for instance, $d(x, y) = \frac{|x-y|}{1+|x-y|}$ is not).

Definition 2.6. A **Banach space** is a normed vector space which is also complete. Note, every normed vector space has a unique completion.

Example 2.7. $\mathbb{R}^n, \mathbb{C}^n$ where $\|x\| = (|x_1|^2 + \cdots |x_n|^2)^{1/2}$.

Example 2.8. ℓ^p with $\|x\| = \left(\sum_{j=0}^{\infty} |\xi_j\|^p\right)^{1/p}$.

Example 2.9. $C[a, b]$ with $\|x\| = \max_t |x(t)|$ is a Banach space, but inducing this same space with the norm $\|x\| = \left(\int_a^b |x(t)|^2 \, dt\right)^{1/2}$ provides a normed vector space, but not a complete one. The completion of this space is $L^2[a, b]$, a very important space.

Theorem 2.1. A subspace $Y \subset X$ is complete iff Y is closed in X. Note, needs to be a *subspace*.

Definition 2.7. A sequence $\{x_n\}$ is said to be **convergent** if there exists an $x \in X$ such that $\lim_{n\to\infty} \|x_n - x\| \to 0$.

Definition 2.8. If $\sum_{j=1}^{\infty} \|x_j\| \leq \infty$, the sum is considered **absolutely convergent** . Note: $\{\text{absolute convergence} \implies \text{convergence}\} \Leftrightarrow \{X \text{ is complete}\}$.

Definition 2.9. If a NVS X contains a sequence $(e_1, e_2, \ldots)$ with the property that every vector can be represented uniquely $x = \sum_{j=0}^{\infty} \alpha_j e_j$, then (e_n) is said to be a **Schauder basis** of X.

Example 2.10. In ℓ^p, the natural choice of basis vectors $(1, 0, 0, \ldots), (0, 1, 0, \ldots), \ldots$ form a Schauder basis.

Theorem 2.2. If X has a Schauder basis, it is separable.

2.3 Finite dimensional normed vector spaces

Lemma 2.1. Let $\{x_1, \ldots, x_n\}$ constitute a linearly independent set in a NVS X. Then there exists a number $c > 0$ such that for every selection of scalars $\{\alpha_1, \ldots, \alpha_n\}$:

$$\|\alpha_1 x_1 + \cdots \alpha_n x_n\| \geq c(|\alpha_1| + \cdots |\alpha_n|).$$

Proof. Follows from Bolzano-Weierstrauss. □

Theorem 2.3. Every finite dimensional subspace of a normed vector space is complete (regardless of if the original vector space is complete).

Theorem 2.4. Every finite dimensional subspace of a normed vector space is closed in the original space.

Definition 2.10. A norm $\|\cdot\|$ on a NVS X is said to be **equivalent** to another norm $\|\cdot\|_0$ if there are constants $a, b > 0$ such that $a\|x\|_0 \leq \|x\| \leq b\|x\|_0$ for all $x \in X$.

Theorem 2.5. On any finite dimensional NVS, all norms are equivalent.

2.4 Compactness and finite dimension

Definition 2.11. A metric space is (sequentially) compact if every sequence has a convergent subsequence. Note: a subset is compact if every sequence in the subset converges to an element of the subset.

Theorem 2.6. Every compact subset M of a metric space is closed and bounded. Note: the converse is not always true.

Example 2.11. Consider $(e_n) \in \ell^2$. The set is bounded and closed but every element is distance 1 apart and therefore has no convergent subsequence.

Theorem 2.7. If X is finite dimensional, then {closed and boundedness}$\Leftrightarrow$ compactness.

Proof. Uses Bolzano-Weierstrauss. □

Lemma 2.2 (F. Riesz)**.** Let Y and Z be a subspace of a NVS X, where Y is closed and a proper subset of Z. Then for every number $\theta \in (0, 1) : \exists z \in Z : \|z\| = 1$ and $\|z - y\| > 0$ for all $y \in Y$.

Theorem 2.8. If the closed unit ball $B = \{x : \|x\| \leq 1\}$ is compact then X is finite dimensional.

Proof. Uses Lemma 2.2. □

Theorem 2.9. Let $T : X \to Y$ be continuous, where X, Y are metric spaces, then the image $T(M)$, where M is compact, is compact.

Theorem 2.10. A continuous function $T : M \to \mathbb{R}$ over a compact M attains both a minimum and maximum over M.

3

Linear operators

3.1 Preliminaries

Definition 3.1. A **linear operator** T is a mapping with the following properties:

(T1) The domain $\mathcal{D}(T)$ is a vector space and the range, $\mathcal{R}(T)$ lies in a vector space over the same field.

(T2) $T(ax + by) = aTx + bTy$ for all $x, y \in \mathcal{D}(T)$ and scalars a, b.

Definition 3.2. Define the **null space** of an operator to be $\mathrm{N}(T) = \{x \in \mathcal{D}(T) : Tx = 0\}$.

Example 3.1. $I : X \to X$ where $Ix \coloneqq x$ is the *identity operator*.

Example 3.2. If $A \in \mathbb{R}^{m\times n}$ then Ax is defined by matrix operation over a finite dimensional space.

Example 3.3. Let $X = \{p[0,1] \to \mathbb{R} : p$ is polynomial$\}$ and define $Dp(t) \coloneqq \frac{\mathrm{d}p}{\mathrm{d}t}$. Note this is unbounded.

Example 3.4. $Q : L^1[0,1] \to L^1[0,1]$ defined by $Qf(t) \coloneqq= \int_0^1 f(s)\,ds$. Note in this case:

$$\begin{aligned}
\|Qf\|_1 &= \int_0^1 \|Qf(t)\|\,dt \\
&= \int_0^1 \| \int_0^t f(s)\,ds\|\,dt\| \\
&\leq \int_0^1 \int_0^t |f(s)|\,ds\,dt \\
&\leq \int_0^1 |f(s)|\,ds \\
&= \|f\|_1.
\end{aligned}$$

Thus, from this, we can conclude that Q maps $[0,1]$ onto itself.

Example 3.5. Let k be a continuous function on $[0,1]^2$, then define $K : C[0,1] \to C[0,1]$ to be $(Kf)(t) \coloneqq \int_0^1 k(s,t)f(s)\,ds$ for $t \in [0,1]$. This is again the continuous matrix extension.

Theorem 3.1. The operator T is **injective** (or one-to-one) if $Tx_1 = Tx_2 \implies x_1 = x_2$. If $T : \mathcal{D}(T) \to \mathcal{R}(T)$ is injective, then $\exists T^{-1} : \mathcal{R}(T) \to \mathcal{D}(T)$ defined by $T^{-1}y = x$ where $Tx = y$. For linear operators, note that T^{-1} exists if $\mathrm{N}(T) = \{0\}$. Furthermore, if T^{-1} exists and T is a linear operator, then T^{-1} is also a linear operator.

3.2 Bounded linear operators

Definition 3.3. Let X, Y be normed vector spaces and $T : \mathcal{D}(T) \subset X \to Y$ be a linear operator. T is considered a **bounded linear operator (BLO)** if $\exists c > 0 : \|Tx\|_Y \leq c\|x\|_X$ for all $x \in \mathcal{D}(T)$.

For a bounded linear operator, the smallest constant c defines the norm on the operator:

$$\|T\| \coloneqq \sup_{\substack{x\in\mathcal{D}(T)\\ x\neq 0}} \frac{\|Tx\|}{\|x\|} \sup_{\substack{x\in\mathcal{D}(T)\\ \|x\|=1}} \|Tx\|.$$

Example 3.6. The *identity* operator is clearly bounded with $\|I\| = 1$.

Example 3.7. For a matrix $A \in \mathbb{R}^{n\times m}$ then $\|A\| = \sup_{\|x\|=1} \|Ax\|$ is the spectral radius and is therefore a bounded linear operator.

Example 3.8. Note that the derivative operator is not bounded because $Dx_n(t) = n \to \infty$, where $x_n(t) \coloneqq nt^{n-1}$.

Example 3.9. Our standard integral operator $(Kf)(t) \coloneqq \int_0^1 k(s,t)f(s)\,ds$ is bounded by noting:

$$\begin{aligned}
\|Kf\| &= \max_{t\in[0,1]} \left| \int_0^1 k(s,t)f(s)\,ds \right| \\
&\leq \max_t \int_0^1 |k(s,t)|\,|f(s)| ds \\
&\leq \max_t \underbrace{\int_0^1 |k(s,t)|\,ds}_{=c} \int_0^1 |f(s)|\,ds \\
&\leq c\|f\|.
\end{aligned}$$

Thus, we can conclude $\|K\| \leq \|Kf\|/\|f\| \leq c$ and therefore K is a BLO.

Theorem 3.2. If a normed space X is finite dimensional, then every linear operator on X is bounded.

Theorem 3.3. Let $T : \mathcal{D}(T) \subset X \to Y$ be a linear operator, then T is bounded iff T is continuous (even at a single point).

Note that if T is bounded then $\mathrm{N}(T)$ is closed. Also, if T_1, T_2 satisfy the norm property: $\|T_1\, T_2\| \leq \|T_1\|\, \|T_2\|$.

3.3 Linear functionals

Consider the special case of a linear operator where $f : \mathcal{D}(F) \subset X \to K$, where K is the field associated with X, typically $\mathbb{R}$ or $\mathbb{C}$, then this operator is called a linear functional. As with linear operators, we can define the notion of boundedness:

Definition 3.4. A **bounded linear functional (BLF)** is a linear functional if there $\exists c$ such that $|f(x)| \leq c\|f\|$. From this, we can define the norm of the functional to be:

$$\|f\| := \sup_{\substack{x \in \mathcal{D}(f) \\ x \neq 0}} \frac{\|fx\|}{\|x\|}.$$

Example 3.10. Let $y_0 \in \mathbb{R}^n$ be fixed, then the dot product with y_0, that is $fx = x \cdot y_0$ defines a linear functional where $|f(x)| = |x \cdot y| \leq \|x\|\|y_0\|$ by Cauchy-Schwartz which suggests $\|f\| \leq \|y_0\|$.

Example 3.11. Consider $y_0 \in L^q$ and define $f : L^p \to \mathbb{R}$ by $f(x) := \int x(t) y_0(t)\, dt$. It's clear f is linear and we can also note $|f(x)| \leq \int |x(t)||y_0(t)|\, dt \leq \|x(t)\|_p \|y_0\|_q$ by Hölder's inequality. Thus, we have a BLF. Note we can obtain equality in the integral by choosing $x(t) := \operatorname{sgn} y(t) |y(t)|^{q-1}$.

Example 3.12. Consider $x \in C[0,1]$ and $t_0 \in [0,1]$, then $f(x) := x(t_0)$ is a BLF. This can be seen by noting $|f(x)| = |x(t_0)| \leq \max_t |x(t)| = \|x\|$, thus $\|f\| \leq 1$.

Example 3.13. Consider again the space of polynomials on $[0,1]$ and let $f(x) := \frac{\mathrm{d}f}{\mathrm{d}x}(1)$. Note that if we consider $x_n(t) = t^n$ this is again unbounded.

3.4 Algebraic dual spaces

Definition 3.5. The set of all linear functionals over a vector space X forms a vector space, which we will denote X^* and is called the **algebraic dual**.

Note, since X^* is itself a vector space, it also has an algebraic dual, denoted X^{**}. We can identify elements of this vector space: for $x \in X$, define $g \in X^{**}$ which maps $g : X^* \to K$ by $g(f) = f(x)$.

From this, we can define a mapping $c : X \to X^{**}$ by $c(x) = g$, where g is defined above. Note that c is linear. This mapping is called the **canonical embedding** of $X \to X^{**}$.

Definition 3.6. If c is bijective, X and X^{**} are isomorphic, in which case X is considered to be **algebraically reflexive**.

Theorem 3.4. Finite dimensional vector spaces are always algebraically reflexive.

3.5 Dual spaces

Definition 3.7. The set of all *bounded* linear functionals on a normed vector space X is said to be the **dual space** X', where the definition of bounded linear functional remains the same as defined in Definition 3.4.

Denote $B(X, Y)$ to be the space of all BLO from X to Y. This space is itself a vector space. In fact, it is also a normed vector space by taking $\|B\| = \sup_{\|x\|=1} \|Bx\|$. From this, we can rewrite the dual space as $X' = B(X, K)$.

Theorem 3.5. Let X be a NVS and Y be a Banach space, then $B(X, Y)$ is a Banach space.

Proof. We already know that we have a NVS. We need only show completeness, but this follows from linearity of the operator. □

Note we can think of this from a linear algebra perspective. Each $x \in X$ has a unique representation $x = \xi_1 e_1 + \cdots + \xi_n e_n$, but conversely every n-tuple of coefficients $\{\xi_1, \ldots, \xi_n\}$ defines a unique x under the same mapping. Thus, these spaces are isomorphic.

Denote $Tx := \xi_1 Te_1 + \cdots \xi_m Te_m$ and let $B = (b_1, \ldots, b_n)$ be a basis for Y. For each k there is a unique representation: $Te_k = \tau_{1k} b_1 + \cdots \tau_{nk} b_n$ since $Te_k = y$, which then suggests that $Tx = \sum_{k=1}^{m} \xi_k T(e_k) = \sum_{k=1}^{m} \xi_k \sum_{j=1}^{n} \tau_{jk} b_k = \sum_{j=1}^{n} \left(\sum_{k=1}^{m} \tau_{jk} \xi_k \right) b_j$. Thus, $n \times m$ matrices are just mappings from $\dim X = m \to \dim Y = n$. The elements of X can be thought of as m dimensional column vectors (i.e. a $m \times 1$ matrix) and the elements of X^* are effectively $1 \times m$ matrices, or row vectors.

From this, we can conclude that $X = X^*$ for finite dimensional X and since every BLO on a finite dimensional space is bounded, we know $X^* = X'$.

Example 3.14. $(\mathbb{R}^n)' = \mathbb{R}^n$. This follows from the above analysis. $|f(x)| = |\sum \xi_k \gamma_k| \leq \sum |\xi_k \gamma_k| \leq \left(\sum |\xi_k|\right)^{1/2} \left(\sum |\gamma_k|\right)^{1/2} = \|x\| \left(\sum |\gamma_k|\right)^{1/2}$. Note we can obtain equality by choosing $\xi_k = \gamma_k$, thus $\|f\| = \|g\|$, where $g = (\gamma_1, \ldots, \gamma_n)$.

Example 3.15. $(\ell^1)' = \ell^\infty$. Note that ℓ^1 has a standard Schauder basis and can be written as $x \in \ell^1 \implies \lim_{n\to\infty} \sum_{k=1}^n \xi_k e_k$, so $f \in (\ell^1)' \implies f(\sum \xi_k e_k) = \sum \xi_k f(e_k)$. Note nowe that $|f(x)| \leq \sum |\xi_k \gamma_k| \leq \sup_k |\gamma_k| \sum |\xi_k| = \sup_k |\gamma_k| \|x\|_1 \implies \|f\| \leq \sup_k \|\gamma_k\| = \|g\|_\infty$. We can again obtain equality in this case telling us that $\|f\| = \|g\|_\infty$.

4

Inner product spaces

4.1 Preliminaries

Definition 4.1. An **inner product space** is a vector space X together with an inner product.

Definition 4.2. An **inner product** is a mapping $\langle \cdot, \cdot \rangle : X \times X \to K$, where K is the field. Note the inner product satisfies the typical set of axioms.

Also note the inner product defines a norm: $\|x\| = \sqrt{\langle x, x \rangle}$. This is the norm induced by an inner product, thus every inner product space is a normed vector space. One key inequality for inner product spaces is the following:

Theorem 4.1 (Schwarz inequality)**.** For every $x, y \in X$:

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

or equivalently:

$$| \langle x, y \rangle |^2 \leq \langle x, x \rangle \langle y, y \rangle .$$

We can also provide two other properties about inner product spaces that are occasionally useful.

Theorem 4.2 (Parallelogram Law)**.**

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

Theorem 4.3 (Polarization Identities)**.** For a real inner product space:

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

and for a complex inner product space:

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

Definition 4.3. A complete inner product space is called a **Hilbert space**.

Example 4.1. $\mathbb{R}^n$ is a Hilbert space with the inner product defined to be the dot product.

Example 4.2. $\mathbb{C}^n$ is also a Hilbert space with $\langle x, y\rangle := x \cdot \bar{y}$.

Example 4.3. ℓ^2 is a Hilbert space with $\langle x, y\rangle := \sum \xi_j \bar{\eta}_j$.

Example 4.4. ℓ^p for $p \neq 2$ is *not* a Hilbert space.

Example 4.5. $L^2[a, b]$ is a Hilbert space where $\langle x, y\rangle := \int_a^b x(t)\bar{y}(t)\, dt$.

Example 4.6. Again,$L^p[a, b]$ is not a Hilbert space for $p \neq 2$.

Example 4.7. $C[a, b]$ is not a Hilbert space with the ∞ norm.

Just as the norm is continuous, as is the inner product. Thus $x_n \to x, y_n \to y \implies \langle x_n, y_n\rangle \to \langle x, y\rangle$.

4.2 Completions, Sobolev spaces

Theorem 4.4. Let X be an inner product space, then there is a Hilbert space H and an isomorphism $T : X \to W$, where $W \subset H$ and is dense. H is unique up to isomorphism (that is, preserves inner product). In other words, every inner product space has a completion.

Now, we can see that $C[a, b]$, along with $C^1[a, b]$, the subspace of $C[a, b]$ where derivatives are continuous and $C_0^1[a, b]$, the subspace of $C^1[a, b]$ where the endpoints are both 0 are all Banach spaces but not Hilbert spaces.

We know that $L^2[a, b]$ is the completion of $C[a, b]$, but what are the completion of the other spaces? Define $H_1[a, b]$ to be the completion of $C^1[a, b]$. It is called a **Sobolev space**. The inner product in this case is defined to be:

$$\langle x, y\rangle_{H^1[a,b]} := \int_a^b x(t)\bar{y}(t)\, dt + \int_a^b x'(t) + \bar{y}'(t)\, dt.$$

Similarly, the Sobolev space $H_0^1[a, b]$ can be defined, corresponding to elements such that $\lim_{t\to a^+} = \lim_{t\to a^-} = 0$.

Example 4.8. Using this notion of Sobolev spaces, we can now begin to form alternate formulations of problems, which we will revisit later. Consider the following ODE:

$$x'' - cx = f, \quad t \in (a, b), \quad x(a) = x(b) = 0.$$

Note, we can now multiply both sides of this equation by some function $y(t)$ and integrate from a to b to obtain:

$$\int_a^b x''(t)y(t)\,dt - c\int_a^b x(t)y(t)\,dt = \int_a^b f(t)y(t)\,dt$$

Integration by parts yields:

$$-\int_a^b x'(t)y'(t)\,dt - c\int_a^b x(t)y(t)\,dt = \int_a^b f(t)y(t)\,dt$$

and if we assume $c = 1$ then we have the following form:

$$-\langle x, y\rangle_{H^1} = \langle f, y\rangle_{L^2}.$$

Note that in general we can define $H^k(0,1)$ with the norm: $\|x\|_{H^k} := \int |x(t)|^2\,dt + \int |x'(t)|^2\,dt + \cdots \int |x^k(t)|^2\,dt$. We know that $H^0 = L^2$.

4.3 Orthogonal projections

We know x, y are said to be **orthogonal** if $\langle x, y\rangle = 0$ and is denoted $x \perp y$. Also, x can be orthogonal to a set M if $\langle x, y\rangle = 0$ for all $y \in M$. The set of all x that are orthogonal to a set Y is denoted the **orthogonal complement**, denoted $Y^\perp$.

Definition 4.4. A set $M \subset X$ is convex if for every $x, y \in M$, the element $\lambda x + (1-\lambda)y \in M$ for $\lambda \in [0,1]$.

Theorem 4.5. Let X be an inner product space and M be a non-empty, convex, complete subset of X. Then for $x \in X$ there is a unique $y \in M$ such that:

$$\|x - y\| = \inf_{\tilde{y}\in m} \|x - \tilde{y}\|.$$

Furthermore, if M is a subspace, then $x - y$ is orthogonal to M.

If X is a vector space with subspaces Y, Z, we can write X as a decomposition, that is: $X = Y \oplus Z$, meaning every element $x \in X$ can be written uniquely as $x = y + z$, where $y \in Y$ and $z \in Z$. In this case, X is called the **direct sum** of Y and Z. If Y, Z are orthogonal, this is called an **orthogonal decomposition**.

Example 4.9. Consider $\mathbb{R}^n$, then $Y = \{y = (\xi_1, \ldots, \xi_{n-1}, 0)\}$ and $Z = \{y = (0, \ldots, \eta_n)\}$, then $\mathbb{R}^n = Y \oplus Z$.

Theorem 4.6. Let H be a Hilbert space and Y be a closed, complete, convex subspace, then H can always be written as an orthogonal decomposition:

$$H = Y \oplus Y^\perp. \tag{4.1}$$

In the above case, we can define $P : X \to Y$ to be the **orthogonal projection** such that $Px = y$, where $x = y + z$. Note that $P^2 = P$, $\mathrm{N}(P) = Y^\perp$.

Definition 4.5. A set M is said to be orthogonal if $\langle x, y \rangle = 0$ for all $x, y \in M$. It is called **orthonormal** if $\|x\| = 1$ for all $x \in M$.

Example 4.10. In $\mathbb{R}^n$, the standard basis elements are orthonormal.

Example 4.11. In $L^2(0, 1)$, the set $M = \{\sin(2\pi nx)\} \bigcup \{\cos(2\pi nx)\}$ forms an orthogonal set. It can be rescaled to be orthonormal. Note, this is the Fourier basis.

Note, for orthogonal elements, the triangle inequality becomes the Pythagorean theorem, that is: $\|x + y\|^2 = \|x\|^2 + \|y\|^2$ and also $\| \sum \xi_i \|^2 = \sum \|\xi_i\|^2$.

If $M = \{e_1, \ldots, e_n\}$ is an orthogonal set and let $Y = \operatorname{span} M$, then any $y \in Y$ can be written as $y = \alpha_1 e_1 + \cdots + \alpha_n e_n \implies \langle y \rangle e_j = \alpha_j$.

Note, now the sum $\sum_{i=1}^n \langle x, e_i \rangle e_i$ is well defined for any $x \in X$ since we know $x = y + z$ for $y \in Y, z \in Y^\perp$. We then can define the projection of x onto y as:

$$y = \sum_{j=1}^n \langle x, e_j \rangle e_j = \sum_{j=1}^n \langle y + z, e_j \rangle e_j = \sum_{j=1}^n \langle y, e_j \rangle e_j$$

Note that we know $\|x\|^2 = \|y\|^2 + \|z\|^2$ which suggests that $\|y\|^2 \leq \|x\|^2$, which leads to the following inequality:

Theorem 4.7 (Bessel's inequality). Let (e_k) be an orthogonal sequence in X, then for every $x \in X$:

$$\sum_{j=1}^\infty |\langle x, e_j \rangle| \leq \|x\|^2$$

Note that $\langle x, e_k \rangle$ are called **Fourier coefficients**.

Also notice that if M is any orthonormal set in X then the number of non-zero Fourier coefficients must be countable, which follows from Bessel's inequality by arguing: how many coefficients can we have with $| \langle x, e_k \rangle | > 1/n$? If it was infinite, then Bessel's inequality would not hold, thus it must be finitely many.

4.4 Total orthonormal sets

Definition 4.6. A set M is considered **total** in a normed vector space X if the $\operatorname{span} M$ is dense in X.

A few facts worth noting are that for a Hilbert space H, H must contain a total orthonormal set. All total orthonormal sets for H have the same cardinality, denoted the **Hilbert dimension**. A given subset M of H contains a countable total orthonormal set. Also, we have the following variation of Bessel's Inequality:

Theorem 4.8 (Parseval's equality). An orthonormal set M is total if and only if for every $x \in H$:

$$\sum_{j=1}^{\infty} |\langle x, e_j \rangle| = \|x\|.$$

Note the sum is to ∞ but there are only countably non-zero coefficients.

Note that if two Hilbert spaces have the same Hilbert dimension, they are isomorphic, therefore "once you've seen one real Hilbert space of dimension n, you've seen them all". Besicovich spaces are "almost periodic" functions which form a non-separable Hilbert space but it rarely occurs in applications.

Note, this analysis provides the justification for the Fourier series expansion: $f(t) = \sum_{k=1}^{\infty} a_k \sin(2\pi kt) + \sum_{j=1}^{\infty} b_j \cos(2\pi jt) + c_0$. As mentioned before, these form a total orthonormal set in L^2.

What about the Fourier transform? $f(t) = \int \hat{f}(\omega) e^{i\omega t}\, d\omega$? Do the exponential functions correspond to a uncountable orthonormal set? No. They aren't even in L^2.

We can also consider the **Gram-Schmidt Process** which produces a total orthonormal set $\{y_1, \dots, y_n\}$ from a linearly independent set $\{x_1, \dots, x_n\}$.

The construction of the first element is easy: $y_1 = \frac{x_1}{\|x_1\|}$. Now, the projection theorem says that $x_2 = y + z$, where $y \in \text{span}\{y_1\}$ and $z \perp y$. Thus, $x_2 = \langle x_2, y_1 \rangle y_1 + v_2$, which suggests that $v_2 = x_2 - \langle x_2, y_1 \rangle y_1 \neq 0$ since x_1, x_2 are linearly independent. Finally, set $y_2 = \frac{v_2}{\|v_2\|}$ to get normality.

Similarly, for x_3 we know that $x_3 = \langle x_3, y_1 \rangle y_1 + \langle x_3, y_2 \rangle y_2 + v_3$ and we can repeat the process. Note, this is guaranteed to work if the x_j are linearly independent, but this method is numerically unstable due to dividing by a potentially arbitrarily small number.

4.5 Representation of functions on Hilbert spaces

Recall that X' was the set of all bounded linear functionals on X. We know if X is reflexive then $X\tilde{X}'$. To check the reflexivity of a Hilbert space we have the following theorem.

Theorem 4.9 (Riesz representation theorem). Let H be a Hilbert space. For every bound linear functional $f \in H'$, there is a unique element $z \in H$ such that for all $x \in H$:

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

Furhtermore, $\|f\| = \|z\|$.

Proof. First note that $\mathrm{N}(f)$ is closed since f is bounded. If $\mathrm{N}(f) = H$, then $z = 0$ and $f = 0$ and we are done, so assume that $H = \mathrm{N}(f) \oplus \mathrm{N}\, f^{\perp}$ and chose $z_0 \in \mathrm{N}(f)^{\perp} \setminus \{0\}$.

We know immediately that $f(z_0) \neq 0$, so note that

$$f\left(\underbrace{x - \frac{f(x)z_0}{f(z_0)}}_{\in \mathrm{N}(f)}\right) = f(x) - \frac{f(x)}{f(z_0)} f(z_0) = 0.$$

From this, we can conclude:

$$0 = \left\langle x - \frac{f(x)}{f(z_0)} z_0, z_0 \right\rangle = \langle x, z_0 \rangle - \frac{f(x)}{f(z_0)} \langle z_0, z_0 \rangle$$

and therefore:

$$f(x) = \langle x, z_0 \rangle \cdot \frac{f(z_0)}{\|z_0\|^2} = \left\langle x, \frac{\overline{f(z_0)}}{\|z_0\|} z_0 \right\rangle.$$

So, if we choose $z = \frac{\overline{f(z_0)}}{\|z_0\|} z_0$, we then have $f(x) = \langle x, z \rangle$ as desired. Uniqueness is obvious. To show the norm relation:

$$\|z\|^2 = |\langle z, z \rangle| = |f(z)| \leq \|f\| \|z\| \implies \|z\| \leq \|f\|,$$

and similarly:

$$|f(x)| = |\langle x \rangle z| \leq \|x\| \|z\| \implies \frac{|f(x)|}{\|x\|} \leq z \implies \|f\| \leq \|z\|.$$

Thus, we can conclude $\|f\| = \|z\|$. □

Thus, we know that there is a bijective mapping between X and X', meaning every Hilbert space is reflexive.

4.6 Lax-Milgram theorem

Definition 4.7. Let X and Y be real vector spaces, then a function $X \times Y \to \mathbb{R}$ is considered a **bilinear form** if it satisfies the typical linearity properties.

Example 4.12. $a(x, y) := \langle x, y \rangle$ is a bilinear form.

Note, if the field is the complex numbers and the second term becomes conjugated in linearity, the form is considered **sesquilinear**.

We say the form is bounded if there exists a constant c such that: $|a(x, y)| \leq c\|x\|_X\|y\|_Y$. As usual, the norm of this operator is defined to be the smallest c that satisfies this property.

Theorem 4.10. Let a be a bounded bilinear form over real Hilbert spaces $a : H_1 \times H_2 \to \mathbb{R}$, then there exists a uniquely determined linear operator $S : H_1 \to H_2$ such that:

$$a(x, y) = \langle Sx, y \rangle .$$

Furthermore, $\|a\| = \|S\|$.

Proof. For a fixed x, the bilinear form becomes a BLF and this follows immediately from the Riesz Representation Theorem. □

Definition 4.8. A bilinear form $b(x, y)$ is called **coercive** (or uniformly elliptic) if there exists a constant c such that:

$$b(x, x) \geq c\|x\|^2.$$

Theorem 4.11 (Lax-Milgram)**.** Let H be a real Hilbert space and let $b : H \times H \to \mathbb{R}$ be a bounded, coercive, bilinear form. Let $B : H \to H'$ be defined by $(Bx)(y) = b(x, y)$, then $B^{-1} : H' \to H$ exists and is bounded.

Proof. Let $S : H \to H$ be the operator such that $b(x, y) = \langle Sx, y \rangle$, then we know from coercivity:

$$c\|x\|^2 \leq |b(x, x)| \leq \|Sx\|\|x\| \implies c\|x\| \leq \|Sx\|$$

Therefore, we know S is injective since if $Sx = Sy \implies c\|x - y\| \leq \|Sx - Sy\| = 0 \implies x = y$. We now claim $\mathcal{R}(S)$ is closed. Let $y \in \overline{\mathcal{R}(S)}$ and therefore $\exists (y_n) \in S$ such that $y_n \to y$ and we know $y_n = Sx_n$. Since y_n is convergent, it must be Cauchy, which suggests x_n is Cauchy and therefore convergent. Thus, $Sx = y$ so $y \in \mathcal{R}(S)$, meaning that the range is closed. Suppose that $\mathcal{R}(S) \neq H$ then $\exists z \in \mathcal{R}(S)^\perp \setminus \{0\}$ such that $\langle Sx, z \rangle = 0$ and choose $x = z$ which suggests that $\langle Sz, z \rangle = 0$ but this violates the coercivity estimate, meaning $\mathcal{R}(S) = H$. Thus, S is injective and the range is the whole space meaning it is onto and therefore bijective, meaning S^{-1} exists and is bounded.

Next consider $c\|x\| \leq \|Sx\|$ by the injectivity of S, which suggests that $Sx = y \implies x = S^{-1}y$ and S is bounded. We can then define $B^{-1} : H' \to H$ by the following analysis: let $g \in H' \implies \exists f \in H : g(y) = \langle f, y \rangle$ and $Bx = g$ means $(Bx)(y) = g(y) = \langle f, y \rangle \implies B^{-1}g = S^{-1}y$ which uniquely determines B^{-1}. □

In general, we can now solve problems of the form $b(x, y) = g(y)$ for all y and turn them into the form $Bx = g$ in our functional space.

Example 4.13. Consider the following boundary value problem:

$$\begin{cases} -u'' + au = f \\ u(0) = u(1) = 0. \end{cases} \tag{4.2}$$

Assume $a(t) \in C[0, 1]$ and $a(t) \geq a_0 > 0$ in this interval. We can also assume $f \in L^2(0, 1)$ to proceed. Consider a test function $v \in C_0^1[0, 1]$ and multiply both sides and integrate:

$$\begin{aligned} -u''v + auv &= fv \\ -\int_0^1 u''v\,dt + \int_0^1 auv\,dt &= \int_0^1 fv\,dt \\ \int_0^1 u'v'\,dt + \int_0^1 auv\,dt &= \int_0^1 fv\,dt. \end{aligned} \tag{4.3}$$

Thus, we now have the *weak version* of our problem. That is, solving (4.2) and (4.3) is equivalent. We can now define our bilinear operator $b(u, v) := \int u'v' + \int auv'$ for $u, v \in C_0^1[0, 1]$. Then note:

$$\begin{aligned} |b(u, v)| &\leq |\int u'v'| + |\int auv| \\ &\leq \int |u'v'| + \max_t a(t) \int |uv| \\ &\leq \|u'\|_{L_2}\|v'\|_{L_2} + \max_t a(t)\|u\|_{L_2}\|v\|_{L_2} \\ &\leq \max\{1, \|a\|_\infty\}\left[\|u'\|_{L_2}\|v'\|_{L_2} + \|u\|_{L_2}\|v\|_{L_2}\right] \\ &\leq \underbrace{2\max\{1, \|a\|_\infty\}}_{=c}\|u\|_{H^1}\|v\|_{H^1} \end{aligned}$$

Thus, our bilinear form is bounded. We need now show coercivity as well:

$$\begin{aligned} b(u, u) &= \int (u')^2 + \int a(u)^2 \\ &\geq \int (u')^2 + a_0 \int u^2 \\ &\geq \min\{1, a_0\}\left[\int (u')^2 + \int u^2\right] \\ &= \min\{1, a_0\}\|u\|_{H_1}^2 \end{aligned}$$

Now, we have $b(u, v) = g(v) \Leftrightarrow Bu = g$. We know $g(v) = \int fv$ and also $|g(v)| \leq \|f\|_{L^2}\|g\|_{L^2} \leq \|f\|_{L^2}\|v\|_{H^1}$, thus g is bounded (and therefore continuous), meaning our weak problem has a unique solution $u \in H_0^1[0, 1]$ which depends continuously on f.

5

Linear operators II

5.1 Hilbert adjoint

Definition 5.1. Let H_1, H_2 be Hilbert spaces and $T : H_1 \to H_2$ be a BLO. The **adjoint** of T, denoted T^* is defined to be: $\langle TX, y\rangle_{H_2} \coloneqq \langle x, T^* y\rangle_{H_1}$.

Note, "Hilbert adjoint" and "adjoint" are equivalent when working in a Hilbert space. The adjoint always exists and is unique with $\|T^*\| = \|T\|$. If we have a sesquilinear form $h(x, y)$ then this can be written as $h(y, x) = \langle y, Tx\rangle = \langle T^* x, y\rangle$.

Example 5.1. Consider a matrix $A : \mathbb{R}^n \to \mathbb{R}^m$ as a linear operator. Then $A = (a_{ij}) \in \mathbb{R}^{m\times n}$. In this case:

$$\langle Ax, y\rangle = Ax \cdot y = (Ax)^T y = x^T A^T y = \left\langle x, A^T y\right\rangle \implies A^T = A^*.$$

Example 5.2. Consider $K : L^2[0, 1] \to L^2[0, 1]$ defined by the usual integral operator $(Ku)(t) = \int_0^1 k(t, s)u(s)\, ds$. We can interchange the integrals as we did the sums above and the resulting adjoint becomes $K^* v(s) = \int_0^1 k(t, s)v(t)\, ds$.

A few other notable facts include: $\|T^*T\| = \|TT^*\| = \|T\|^2$ and $(S + T)^* = S^* + T^*$. Also note that if a BLO satisfies $T = T^*$, it is considered **self-adjoint** . If it is not bounded, it is considered symmetric.

5.2 Boundedness

Theorem 5.1 (Hahn-Banach)**.** Let X be a vector space, Z be a subspace of X and consider $\rho : X \to \mathbb{R}$ satisfy $\rho(x + y) \leq \rho(x) + \rho(y)$ and $\rho(ax) \leq |a|\rho(x)$, that is, we have a sub-linear functional. Let f be a functional on Z such that $|f(x)| \leq \rho(x)$, then there exists a linear extension $\tilde{f}$ of f onto X satisfying $|\tilde{f}| \leq \rho$.

We can now present some consequences of the Hanh-Banach theorem.

Theorem 5.2. Let f be a BLF on a subspace Z of the NVS X, then there is a BLF $\tilde{f}$ which extends to X and has the same norm.

Corollary 5.1. Let X be a normed vector space. Choose $x_0 \neq 0 \in X$. Then there is a BLF $\tilde{f} \in X'$ such that $\|\tilde{f}\| = 1$ and $\tilde{f}(x_0) = \|x_0\|$.

Theorem 5.3. Let X be a NVS. For every $x \in X$, $\|x\| = \sup_{\substack{f \in X' \\ f \neq 0}} \frac{|f(x)|}{\|f\|}$. Also, if x_0 has the property that $f(x_0) = 0 \forall f \in X' \implies x_0 = 0$.

Theorem 5.4 (Uniform Boundedness Theorem)**.** Let X be a Banach space and let $\{\tau_n\}$ be a sequence of BLO such that $\tau_n : X \to Y$, where Y is a normed vector space. If for each fixed $x \in X : \{\tau_n(x)\}$ is bounded, then there is a c such that $\|\tau_n\| \leq c$.

5.3 Reflexive spaces

Recall the canonical embedding map $c : X \to X^{**}$, where given $x \in X$, $g \in X^{**}$, we can define $g(f) = f(x) \implies cx = g$. Also note c is an isometry. X is then isomorphic with $\mathcal{R}(c) \subset X''$ and recall a space is **reflexive** if $\mathcal{R}(c) = X''$, that is $X\tilde{X}''$.

Example 5.3. ℓ^p with $1 < p < \infty$ is reflexive by recalling $(\ell^p)' = \ell^q$.

Example 5.4. L^p with $1 < p < \infty$ follows from an identical argument.

Example 5.5. Note that $p = 1, \infty$ is not reflexive for either case.

Theorem 5.5. Let X be a normed vector space, then:

(1) If X is reflexive, then X is a Banach space. (since X' must be Banach)

(2) If X is finite-dimensional, then X is reflexive.

(3) If X is a Hilbert space, then X is reflexive. (by Riesz representation theorem)

Lemma 5.1. Let X be a normed vector space and $Y \subset X$ be a closed, proper subspace. Choose $x_0 \in X - Y$ and define: $\delta = \inf_{y \in Y} \|x_0 - y\|$, then there exists $f \in X'$ such that $f = 0$ on Y and that $\|f\| = 1$ and $f(x_0) = \delta$.

Theorem 5.6. Let X be a normed vector space. If X' is separable, then X' is also separable.

Proof. Follows from previous lemma. □

Thus, we can conclude that if X is reflexive, then if either X or X'' is separable, then the other must be separable as well. Also note that if X is separable and X' is not, X cannot be reflexive. For instance, note that ℓ^1 is separable but ℓ^∞ is not, thus ℓ^1 is not reflexive.

5.4 Weak convergence and weak* convergence

Definition 5.2. The sequence $\{x_n\}$ in a normed vector space X is said to **strongly converge** if $\|x_n - x\| \to 0$ as $n\ to \infty$.

Definition 5.3. The sequence $\{x_n\}$ in a normed vector space X is said to **weakly converge** if $\|f(x_n) - f(x)\| \to 0$ as $n\ to \infty$ for all $f \in X'$.

In this case, we denote this limit $x_n \rightharpoonup x$ and say x is the weak limit of $\{x_n\}$. It's clear that strong convergence $\implies$ weak convergence, but the converse is not necessarily true. Note, if we have a finite dimensional space, the converse *is* true.

Example 5.6. Consider $X = \ell^p$ and define $(x_n) := (e_n)$. Since $(\ell^p)' = \ell^q$, we can represent $f \in \ell^q$ as $f(x) := \sum_{j=1}^{\infty} \xi_j \overline{\eta_j}$. Note, $|\eta_j| \to 0$. Therefore $f(x_n) \to 0$ as well, but x_n is not Cauchy and doesn't converge, but $x_n \rightharpoonup 0$.

Note that the weak limit is always unique and that if a sequence is weakly convergent, then every subsequence is as well. Also, $\{\|x_n\|\}$ must also be bounded, but the proof is tricky.

Let $\{f_n\}$ be a sequence of bound linear functionals in X'. Then recall that we say $f_n \to f$, that is, converges strongly if $\|f_n - f\| \to 0$ and $f_n \rightharpoonup$, that is, converges weakly if $g(f_n - f) \to 0$ for all $g \in X''$. We can introduce yet another notion of convergence.

Definition 5.4. If the above sequence has the property that $f_n(x) - f(x) \to 0$ for all $x \in X$, then we say $f_n \overset{*}{\rightharpoonup} f$, or f_n converges to f in the sense of **weak*** in X^*.

Weak convergence is stronger than weak* convergence. That is, weak implies weak*. To see this, note that if $f_n \rightharpoonup$ then $g(f_n) - g(f) \to 0$ for all $g \in X''$, thus every $x \in X$ maps to a $g \in X''$ by the canonical map $g(f) = f(x)$. Thus, $g(f_n) = f_n(x)$ and $g(f) = f(x)$, thus $f_n(x) - f(x) \to 0$.

Also note that if X is reflexive, meaning $X = X''$, then weak and weak* are equivalent.

Suppose $\{u_n\}$ is a sequence in L^q, where $1 < q < \infty$. L^q is reflexive and its dual space is L^p, where $1/p + 1/q = 1$. $u_n \rightharpoonup u$ in L^q means that $\int u_n \overline{v} \to \int u \overline{v}$ for all $v \in L^p$ and in this case, weak* is the same.

Next consider $\{u_n\}$ as a sequence in L^∞. Here, $u_n \rightharpoonup$ means that for all $f \in (L^\infty)' : f(u_n) \to f$, but we don't know what $(L^\infty)'$ is. We can note that $L^\infty = (L^1)'$, which now means that $u_n \overset{*}{\rightharpoonup} u$ which suggests that $\int u_n \overline{v} \to \int u \overline{v}$ for all $v \in L^1$.

Example 5.7. Consider the following boundary value problem:

$$\begin{cases} u'' - a(x)u = f \\ u(0) = u(1) = 0, \end{cases}$$

where $a(x)$ corresponds to some periodic material coefficient, say a periodic square wave from 1 to 2 such that $a_n(x) = a(x/n)$. We can now construct the weak form by multiplying by a $v \in H_0^1(0,1)$:

$$\int u'v' + \int a_n uv = -\int fv$$

Note that $\{a_n\}$ is a sequence in $L^\infty(0,1)$ and that $\int |uv| \leq \|u\|_{L^2}\|v\|_{L^2}$ so $uv \in L^1$. This then means that if $\int a_n(uv) \to \int a(uv)$ for some a, then the solution converges.

Another possible a_n choice is $a_n(x) = 2 + \sin(2\pi nx)$. In this case, $\int a_n(x)1\,dx = 2$, and also, $\int_0^1 (1-\theta(x))\,dx = 1$, and finally $\int a_n v \to \int 2v\,dx$ by the Riemann-Lebesgue lemma, thus $\sin nx \overset{*}{\rightharpoonup}$ in L^∞.

5.5 Compact operators

Recall that a set K is compact if every sequence $\{x_n\}$ in K has a convergent subsequence $\{x_{n_k}\}$ and converges to an element in K.

Definition 5.5. Let X, Y be normed vector spaces. An operator $T : X \to Y$ is said to be **compact** if T is linear and if for every bounded subset M of X, then $\overline{T(M)}$ is compact.

Lemma 5.2. Every compact operator $T : X \to Y$ is bounded. If $\dim X = \infty$, then $I : X \to X$ is not compact.

To see this, note that $u = \{x \in X : \|x\| = 1\}$ is bounded and $\overline{T(u)}$ is compact and hence bounded, meaning $\sup_{\|x\|=1} \|Tx\| \leq \infty \implies T$ is bounded. To see the second statement, note that $I(u) \cdot \overline{u}$ is not compact when $\dim X = \infty$.

Theorem 5.7 (Compactness criterion I)**.** An equivalent criteria for compactness can be summarized as: T is compact if and only if it maps every bounded sequence $\{x_n\}$ in X onto a sequence $\{Tx_n\}$ in Y which has a convergent subsequence.

$C(X,Y)$ is the space of all compact operators $T : X \to Y$ and is a vector space since $C(X,Y) \subset B(X,Y)$, which is Banach if Y is Banach.

Theorem 5.8 (Compactness criterion II)**.** Yet another equivalent criteria is: If T is bounded and linear and either:

(1) $\dim(X) < \infty$

(2) $\dim(\mathcal{R}(T)) < \infty$

(3) $\dim(Y) < \infty$,

then T is compact. In other words, if we have a finite dimensional bounded operator, it is indeed compact.

Theorem 5.9. If (T_n) is a sequence of compact operators and $\|T_n - T\| \to 0$, then T is also compact.

Proof. Let (x_n) be a bounded sequence in X. We will construct a subsequence y_n such that Ty_n converges.

We know T_1 is compact, so $\exists$ a subsequence $x_{1,m}$ of x_m such that $T_1 x_{1,m}$ is Cauchy. We can repeat this process for $T_1, \ldots, T_n$. Thus we have a sequence $(T_k x_{j,k})$ that is Cauchy whenever $k \leq j$. We can then define the "diagonal sequence" $y_m = x_{m,m}$. For any fixed k, $(T_k y_m)$ is Cauchy because eventually $k \leq m$. Since $\{x_m\}$ is bounded, we have that $\{y_m\}$ is also bounded. We want to show that (Ty_m) is Cauchy. Given $\epsilon > 0$, choose p, N large enough such that $\|T_p - T\| < \epsilon/3c$ and $\|T_p y_k - T_p y_j\| \leq \epsilon/3$

$$\begin{aligned}\|Ty_k - Ty_j\| &\leq \|Ty_k - T_p y_k\| + \|T_p y_k - T_p y_j\| + \|T_p y_j - Ty_j\| \\ &\leq \|T - T_p\|\|y_k\| + \epsilon/3 + \|T - T_p\|\|y_j\| < \epsilon.\end{aligned}$$ □

Example 5.8. Consider $T : \ell^2 \to \ell^2$ defined by $Tx := (\xi_1/1, \xi_2/2, \xi_3/3, \ldots)$, which can be thought of as multiplying by an infinite diagonal matrix. We hope to show that T is compact. Define T_n to be T truncated to n terms. Note that since $\dim T_n = n$, it is compact. We also note that $\|T_n x - Tx\|^2 = \sum_{j=n+1}^{\infty} |\xi_j/j|^2 \leq \left(\frac{1}{n+1}\right)^2 \sum_{j=n+1}^{\infty} |\xi_j|^2 \leq \left(\frac{1}{n+1}\right)^2 \|x\|^2 \implies \|T_n x - Tx\| \leq \frac{1}{n+1}\|x\| \implies \|T_n - T\| \to 0$. Therefore, by our previous theorem, T is compact.

Theorem 5.10. Let X, Y be NVS and let T be compact, then if $x_n \rightharpoonup x$ then $Tx_n \to Tx$.

6

Spectral Theory

6.1 Preliminaries

Suppose $T : X \to X$ where $\dim(X) < \infty$ and X is a NVS, then we can represent T as a matrix $A = (a_{ij})$. Consider:

$$Ax = \lambda x. \tag{6.1}$$

Recall that $\lambda \in \mathbb{C}$ is an **eigenvalue** if (6.1) has a solution, in which case $x \neq 0$ is the corresponding eigenvector. The nullspace of $(A - \lambda I)$ is the eigenspace corresponding to the eigenvalue λ. The collection of eigenvalues of A is called the **spectrum** of A, denoted $\sigma(A)$.

Note that eigenvalues are the same for similar matrices, that is $A_2 = C^{-1}A_1C$, where C is non-singular, thus $\sigma(A)$ is unambiguously defined. Also, $\det(A - \lambda I)$ is a polynomial of degree n in λ. The Fundamental Theorem of Algebra says that there must be at least one root and no more than n including multiplicities.

This works for finite dimensional operator, but defining the spectrum of an infinite dimensional operator becomes less straightforward. Thus, consider $X \neq \{0\}$be a complex NVS and let $T : \mathcal{D}(T) \subset X \to X$ be a linear operator.

For $\lambda \in \mathbb{C}$, define $T_\lambda \coloneqq T - \lambda I$ and $R_\lambda(T) \coloneqq T_\lambda^{-1} = (T - \lambda I)^{-1}$ when it exists. $R_\lambda(T)$ is called the **resolvent operator** because it is solving something.

Definition 6.1. The **resolvent set** $\rho(T)$ is the set of all $\lambda \in \mathbb{C}$ such that:

(R1) $R_\lambda(T)$ exists and therefore T_λ^{-1} exists.

(R2) $R_\lambda(T)$ is bounded and therefore T_λ^{-1} is bounded.

(R3) The domain of $R_\lambda(T)$ is dense in X, that is $\overline{\mathcal{D}(R_\lambda(T))} = X$. Note $\mathcal{D}(R_\lambda(T)) = \mathcal{R}(T - \lambda I)$.

We can then define the **spectrum** $\sigma(T)$ to be $\mathbb{C} - \rho(T)$. Thus, one or more of the conditions must fail for $\lambda \in \sigma(T)$. We can then split $\sigma(T)$ into three disjoint sets depending on which property is violated:

(1) $\sigma_p(T)$, the **point spectrum**, consists of all λ such that $R_\lambda(T) = (T - \lambda I)^{-1}$ does not exists. Note that in this case $(T - \lambda I)x = 0$ has a non-zero solution, meaning that we can think of this as $Tx = \lambda x$, thus denoting λ as an **eigenvalue**. This x is also considered an **eigenvector**.

(2) If $R_\lambda(T) = (T - \lambda I)^{-1}$ exists but is not bounded, then λ exists in the **continuous spectrum**, denoted $\sigma_c(T)$.

(3) Again, if $R_\lambda(T) = (T - \lambda I)^{-1}$ exists but the domain is not dense in X, then $\lambda \in \sigma_r(T)$, the **residual spectrum**.

By definition, we know $\sigma(T) = \sigma_p(T) \cup \sigma_c(T) \cup \sigma_r(T)$, thus we can characterize an operator by its spectrum.

Example 6.1. Consider the identity operator I. Then $R_\lambda(I) = (I - \lambda I)^{-1} = [(1 - \lambda)I]^{-1} = \frac{1}{1-\lambda} I$. Note that $R_\lambda(I)$ is not defined for $\lambda = 1$ so $1 \in \sigma_p(I)$. If $\lambda \neq 1$, then $R_\lambda(I)$ exists and is bounded and is defined for all X, thus $\lambda \in \rho(I)$.

Example 6.2. Consider $T : \ell^2 \to \ell^2$ defined by $Tx := (\xi_1, \xi_2/2, \xi_3/3, \ldots)$. If $\lambda = 1/n$, then $(T - \lambda I) = (T - \frac{1}{n} I) \implies (T - \frac{1}{n} I)e_n = 0$, thus $R_\lambda(T)$ does not exist, meaning $\frac{1}{n}$ is an eigenvalue with eigenvector e_n. Note that when $\lambda = 0$, $(T - \lambda I) = T$ and $T^{-1} = (\xi_1, 2\xi_2, 3\xi_3, \ldots)$, which exists and is defined on a dense subset of ℓ^2, which are all sequences with finitely many non-zero empties. Note that $\|T^{-1}e_n\| = n$, thus this is unbounded and therefore $0 \in \sigma_c(T)$.

Example 6.3. Consider $T : \ell^2 \to \ell^2$ defined by $Tx := (0, \xi_1, \xi_2, \ldots)$, that is, we have the right shift operator. First consider $\lambda = 0$. Then $(T - \lambda I) = T$ and therefore $T^{-1}x = (\xi_2, \xi_3, \ldots)$. Note, R_λ exists and is bounded, but the range of T are only sequences with a 0 in the first term, which is not dense in ℓ^2, meaning $\lambda \in \sigma_r(T)$.

6.2 Spectral properties of bounded linear operators

Theorem 6.1. Let $T : X \to X$ be a BLo where X is a Banach space, then if $\|T\| < 1$ then $(I - T)^{-1}$ exists, is bounded, and $(I - T)^{-1} = I + \sum_{k=1}^{\infty} T^k$.

The above theorem is known as a **Neumann series** for a bounded linear operator. The proof is clear by considering a geometric series and that we know that in a Banach space, absolute convergence implies convergence.

Theorem 6.2. Let X be a complex Banach space and $T : X \to X$ be a BLO. The resolvent set, $\rho(T)$ is an open set in $\mathbb{C}$ and for every $\lambda_0 \in \rho(T)$, $R_\lambda(T) = \sum_{k=0}^{\infty} (\lambda - \lambda_0)^k R_{\lambda_0}^{k+1}(T)$ and the series is absolutely convergent for $|\lambda - \lambda_0| < 1/\|R_{\lambda_0}\|$.

Theorem 6.3. Let X be a complex Banach space and $T : X \to X$ be a BLO, then $\sigma(T)$ is compact and $\sigma(T) \subset \{\lambda \in \mathbb{C} : |\lambda| \leq \|T\|\}$.

Theorem 6.4. If $X \neq \{0\}$ is a Banach space then $\sigma(T) \neq \varnothing$.

Definition 6.2. The **spectral radius** $r(T)$ of an operator T is defined to be $r(T) = \sup_{\lambda \in \sigma(T)} |\lambda|$. Note if $\sigma(T)$ is compact, this $\sup$ is reached.

Although it is immediately clear $r(T) \leq \|T\|$, it is more difficult to show that $r(T) = \lim_{n\to\infty} \|T^n\|^{1/n}$.

Theorem 6.5. Let T be a LO on a complex NVS X. Eigenvectors corresponding to distinct eigenvalues are linearly independent.

6.3 Spectral properties of compact operators

Theorem 6.6. Let X be a NVS and $T : X \to X$ be *compact*. The set of all eigenvalues $\sigma_p(T)$ is countable with the only accumulation point possible for $\sigma_p(T)$ is 0.

From this, we can know that the set of all eigenvalues of a compact linear operator can be arranged into a sequence whose only *possible* point of accumulation is 0.

Theorem 6.7. Let X be a NVS. Let $T : X \to X$ be *compact*. Then the eigenspace corresponding to every non-zero eigenvalue $\lambda \in \sigma_p(T)$ is finite dimensional.

Theorem 6.8 (Fredholm alternative)**.** Let X be a NVS and let $T : X\ toX$ be a compact linear operator. Then either:

(1) The problem $(I - T)x = 0$ has a nontrivial solution $x \in X$.

(2) For each $y \in X$, $(I - T)x = y$ has a unique solution.

Note, the two cases are mutually exclusive. In case (2), the inverse operator $(I - T)^{-1}$ whose existence is asserted, is bounded: $\|x\| \leq \|(I - T)^{-1}\|\|y\|$.

Lemma 6.1. Let X be a NVS and Y be a proper closed subspace of X, then given $\theta < 1$, there is a vector $y \in X - Y$ with $\|y\| = 1$ and $\inf_{x\in Y} \|x - y\| \geq \theta$.

Lemma 6.2. Let X be a NVS, and $T : X \to X$ be compact, $S : X \to X$ be bounded, then TS and ST are both compact.

Both these lemmas combined can be used to prove the statement of the Fredholm alternative.

Theorem 6.9. Let $T : X \to X$ be compact and X be a NVS. Then every non-zero $\lambda \in \sigma(T)$ is an eigenvalue.

It can also be shown that $0 \in \sigma(T)$ for all compact linear operators with infinite dimension. This follows from the fact that if $0 \in \rho(T)$, then T^{-1} would eist and be bounded, meaning $T^{-1}T = I$ would be compact, but note that I can only be a compact operator under finite dimensions, thus we have a contradiction.

6.4 Bounded self-adjoint operators

Recall that if $T : H \to H$ is bounded on the Hilbert space H then the adjoint $T^* : H \to H$ defined by $\langle Tx, y\rangle := \langle x, T^*y\rangle$ exists and is bounded by $\|T^*\| = \|T\|$. Note that self-adjointness implies $T^* = T$.

Theorem 6.10. Let $T : H \to H$ be a bounded self-adjoint linear operator on the Hilbert space H. If T has eigenvalues, they must be real and the eigenvectors corresponding to distinct eigenvalues are orthogonal. Note, this is the same result that holds for symmetric matrices.

Theorem 6.11. Let $T : H \to H$ be a bounded self-adjoint linear operator, then $\sigma_r(T) = \varnothing$.

Theorem 6.12. Let $T : H \to H$ be a BSALO on the Hilbert space H. Then $\lambda \in \rho(T)$ iff $\exists c > 0$ such that $\|(T - \lambda I)x\| \geq c\|x\|$ for all $x \in H$.

Example 6.4. Define $(Tx)(t) = tx(t)$ for $t \in [0, 1]$ with $T : C[0, 1] \to C[0, 1]$, which immediately implies that $\sigma(t) = [0, 1]$. We argued previously that $\sigma = \sigma_r$ in this case as the domain of R_λ or the range of $T - \lambda I$ is not dense in $C[0, 1]$. If we consider this same function over L^2 the spectral properties change. It stays bounded, but we can first ask if there are any eigenvalues by considering: $(T - \lambda)x = 0 = (t - \lambda)x = 0$ for all $\lambda \in [0, 1] \implies x = 0$, thus there are no eigenvalues. It can also be shown that T is self adjoint, meaning that $\sigma_r(T) = \varnothing$. Now, consider $(t - \lambda)x(t) = y(t) \implies x(t) = \frac{y(t)}{t-\lambda} \implies \|x\| = \|(T - \lambda)^{-1}y\|$. For instance, consider $\lambda = 0$, then $\|x\| \to \infty$, thus $\sigma = \sigma_c$.

Theorem 6.13. Let $T : H \to H$ be a BSALO on the Hilbert space H, then $\sigma(T)$ is real.

Theorem 6.14. For $T : H\ toH$, a BSALO, then: $\|T\| = \max\{|m|, |M|\} = \sup_{x\neq 0} |q(x)|$ and furthermore m, M are both spectral values.

Theorem 6.15 (Spectral theorem)**.** Let $T : H \to H$ be a compact, self adjoint, linear operator on a Hilbert space H, then there exists a complete orthonormal basis for H consisting of eigenvectors for T.

We know some basic facts already:

(1) Eigenspaces E_λ corresponding to $\lambda \neq 0$ are finite dimensional.

(2) $E_{\lambda_1} \perp E_{\lambda_2}$ if $\lambda \neq \lambda_2$.

(3) There are countably many eigenvalues.

(4) The spectrum consists of only eigenvalues plus possibly $0 \in \sigma_p(T) \cup \sigma_c(T)$.

We can also consider some consequences of the Spectral theorem. Particularly we can write $x = \sum_{j=1}^\infty \langle x, u_j \rangle u_j$ where $x \in H$ and $\{u_j\}$ form an orthonormal basis for H, which suggests that $Tx = \sum_{j=1}^\infty \langle x, u_j \rangle Tu_j = \sum_{j=1}^\infty \lambda_j \langle x, u_j \rangle u_j$, which tells us everything about $\{u_j\}$.

So, to solve $(I-T)x = y$, we need only solve $\sum_{j=1}^\infty (1-\lambda_j) \langle x, u_j \rangle u_j = \sum_{j=1}^\infty \langle y, u_j \rangle u_j \implies (1-\lambda_j) \langle x, u_j \rangle = \langle y, u_j \rangle$. In other words, we diagonalize the problem. We can see from this that $\langle x, u_j \rangle = \langle y, u_j \rangle / (1-\lambda_j)$ and therefore $x = \sum_{j=1}^\infty \frac{\langle y, u_j \rangle}{1-\lambda_j} u_j$, which follows from $x = \sum_{j=1}^\infty \langle x, u_j \rangle u_j$.

Thus, if we want to solve $Tx = y$, we can simply solve $x = \sum_{j=1}^\infty \frac{\langle y, u_j \rangle}{\lambda_j} u_j$, but we know $\lambda_j \to 0$ since T is compact, thus T^{-1} must be unbounded.

Overall, the big take-away from this theorem is: compact operators can be approximated with finite matrices.

6.5 Return to Fredholm theory

We can now formulate a stronger version of the Fredholm alternative for Hilbert spaces.

We first need the lemma:

Lemma 6.3. If $T : H \to H$ is a compact linear operator, then T^* is also compact.

Theorem 6.16 (Fredholm alternative)**.** Let $T : H \to H$ be a BLO on the Hilbert space H, then $\overline{\mathcal{R}(T)} = \mathrm{N}(T^*)^\perp$. Furthermore, there are three exclusive alternatives:

(1) $(I-T)x = y$ has a solution $x \Leftrightarrow y \in \mathrm{N}(I-T^*)^\perp$.

(2) $(I-T^*x) = y$ has a solution $x \Leftrightarrow y \in \mathrm{N}(I-T)^\perp$.

(3) If neither $\mathrm{N}(I-T^*) = 0$ nor $\mathrm{N}(I-T) = 0$, then both equations have solutions x for all $y \in H$ and the inverse operators $(I-T)^{-1}$ and $(I-T^*)-1$ is bounded.

6.6 Sturm-Liouville problem

We can apply most of the theory covered in this course by considering a single example, the **Sturm-Liouville problem**.

In particular, we want to solve the 1 dimensional equation describing acoustic waves:

$$\left(\frac{1}{\rho}v_x\right)_x - \kappa v_{tt} = 0, \tag{6.2}$$

where $\rho(x)$ is the density of the medium and $\kappa(x)$ is the compressibility, which is equivalent to $1/\nu(x)$, where ν is the bulk modulus.

We'd like to separate variables $v(x,t) = u(x)g(t)$, which yields:

$$g(t)\left(\frac{1}{\rho}u'(x)\right)_x - \kappa u(x)g''(t) = 0 \implies \frac{\left(\frac{1}{\rho}u'(x)\right)_x}{\kappa u(x)} = \frac{g''(t)}{g(t)} = -\lambda.$$

Note, as with any separation of variables problem, both sides of this equation are constant so we obtain two ordinary differential equations:

$$\left(\frac{1}{\rho}u'\right)' = -\kappa\lambda u(x)$$
$$g''(t) = -\lambda g(t).$$

We can observe that $g(t) = e^{i\omega t}$, $\lambda = \omega^2$ so $\lambda \geq 0$. Other solutions for $g(t)$ exist but this will suffice. We can generalize the u equation to the form:

$$(\sigma u')' - qu + \lambda\kappa u = 0, \quad q(x) \geq 0, \quad \sigma(x) \geq \sigma_0 > 0, \kappa(x) \geq \kappa_0 > 0. \tag{6.3}$$

Note, $q(x)$ corresponds to absorption, and here we will assume we have Dirichlet boundaries, that is $u(0) = u(1) = 0$. Let $w \in C_0^1[0,1]$ be a test function, then we have:

$$0 = \int_0^1 \left[(\sigma u')'\, w - quw + \lambda\kappa uw\right]\, dx$$
$$= \int_0^1 \left[(\sigma u')w' + quw\right]\, dx - \int_0^1 \kappa uw\, dx.$$

Thus, if we let $a(u,w) \coloneqq \int_0^1 \left[\sigma u'w' + quw\right]\, dx$ and $b(u,w) \coloneqq \int_0^1 \kappa uw\, dx$, we have a problem of the form $a(u,w) = \lambda b(u,w)$, which is the weak form of our problem and true for any $w \in H_0^1(0,1)$. We hope to find $\lambda \in \mathbb{C}$ and $u \in H_0^1(0,1)$.

If a, b are bounded, the Riesz representation theorem implies there are bounded operators: $A : H_0^1 \to H_0^1$ and $B : H_0^1 \to H_0^1$ such that $a(u,w) = \langle Au, w\rangle$ and $b(u,w) = \langle Bu, w\rangle$, in which case our problem becomes: $\langle Au - \lambda Bu, w\rangle = 0$ for all $w \in H_0^1(\Omega)$ which suggests that $Au - \lambda Bu = 0$. Thus, we must check if a, b are bounded, that is if $|a(u,w)| \leq c\|u\|\|w\|$.

Recall that $\|u\|_{H_1}^2 = \int |u'|^2 + |u|^2 \, dx$ and that $\langle u, w \rangle = \int u'w' + uw \, dx$. Thus:

$$\begin{aligned}
|b(u,w)| &= \left| \int_0^1 \kappa u w \, dx \right| \\
&\leq \int_0^1 \kappa(x) |u(x)||w(x)| \, dx \\
&\leq \max_{x \in [0,1]} \kappa(x) \int_0^1 |u(x)||w(x)| \, dx \\
&\leq c\|u\|_{L^2} \|w\|_{L^2} \\
&\leq c\|u\|_{H_1} \|w\|_{H_1}.
\end{aligned}$$

Note, a similar argument holds for a:

$$\begin{aligned}
|a(u,w)| &= \left| \int_0^1 \sigma u'w' + quw \, dx \right| \\
&\leq c_1 \|u'\|_{L^2} \|w'\|_{L^2} + c_2 \|u\|_{L^2} \|w\|_{L^2} \\
&\leq c\|u\|_{H_1} \|w\|_{H_1}.
\end{aligned}$$

So we can now consider the form $Au - \lambda Bu = 0$ and we want $(T - \lambda I)u = 0$. Lax-Milgram says that if a is coercive then A^{-1} exists and is bounded. Coercivity, recall, means $a(u,u) \geq c_0 \|u\|^2$.

First, we prove the **Poincaré's inequality** in 1 dimension by first noting that $u(x) = \int_0^x u'(t) \, dt$, and therefore:

$$\begin{aligned}
\|u\|_{L^2}^2 &= \int_0^1 u(x)^2 \, dx \\
&= \int_0^1 \left(\int_0^x u'(t) \, dt \right)^2 \, dx \\
&\leq \int_0^1 \left(\int_0^x u'(t)^2 \, dt \right) , dx \qquad \text{by Jensen's inequality} \\
&\leq \int_0^1 \int_0^1 u'(t)^2 \, dt dx \\
&= \int_0^1 u'(t)^2 \, dt \\
&= \|u'\|_{L^2}^2,
\end{aligned}$$

which suggests that $\|u\|_{L^2}^2 \leq \|u'\|_{L^2}$ for $u \in H_0^1$, which is the statement of Poincaré's inequality in one dimension. This then allows us to continue:

$$\begin{aligned}
a(u,u) &= \int_0^1 \sigma u'^2 + qu^2 \, dx \\
&\geq \int_0^1 \sigma u'^2 \, dx \\
&\geq \inf_x \sigma(x) \int_0^1 u'(x)^2 \, dx \\
&= \sigma_0 \int_0^1 u'(x)^2 \, dx \\
&= \sigma_0 \int_0^1 \left[\frac{u'(x)^2}{2} + \frac{u'(x)^2}{2} \right] \, dx \\
&\geq \frac{\sigma_0}{2} \int_0^1 u'(x) + u(x)^2 \, dx \quad \text{Poincaré's inequality} \\
&= \frac{\sigma_0}{2} \|u\|_{H^1}^2
\end{aligned}$$

By Lax-Milgram $A^{-1} : H_0^1 \to H_0^1$ exists and is bounded. Thus, $(A - \lambda B)u = 0$, where A, B are both BLO now becomes $(I - \lambda A^{-1}B)u = 0$ and define $T := A^{-1}B$. Let $\nu = 1/\lambda$ then we now have the form $(T - \nu I)u = 0$.

Note, A^{-1} is bounded, so if B is compact, then T is also compact. We must show that B is compact. Define $\tilde{B}$ by $\left\langle \tilde{B}u, w \right\rangle_{H^1} := \int uw \, dx$ since $\langle Bu, w \rangle = \int \kappa uw$.

Consider $\{\sin(n\pi x)\}_{n=1}^{\infty}$, a complete orthogonal set in H_0^1. Let $\phi_n := c_n \sin(n\pi x)$ where $c_n := \frac{\sqrt{2}}{\sqrt{(\pi n)^2+1}}$ for normalization. Note that now $\langle \phi_n, \phi_n \rangle = \int_0^1 \phi_n'^2 + \phi_n^2 = 1$, meaning we have a complete orthonormal set (basis) for H_0^1. Therefore, $\tilde{B}$ becomes:

$$\left\langle \tilde{B}\phi_n, \phi_m \right\rangle = \begin{cases} \int_0^1 \phi_n \phi_m = 0 & \text{if } n \neq m \\ \int_0^1 c_n^2 \sin^2(n\pi x) = \frac{1}{(n\pi)^2+1} & \text{if } n = m. \end{cases}$$

But note, $\frac{1}{(n\pi)^2+1} \to 0$. Thus, $\tilde{B}$ is analogous to our prototypical compact operator on ℓ^2 and is compact by the same argument. Note that $Bu = \tilde{B}(\kappa u)$.

If (u_n) is a sequence in H_0^1 with $\|u_n\| \leq 1$, then assuming that κ is smooth, $\|\kappa u_n\| \leq c$, where $\|\kappa u_n\|^2 = \int (\kappa u_n)'^2 + (\kappa u_n)^2$. Since $\tilde{B}$ is compact, $\tilde{B}(\kappa u_n) = B(u_n)$ has a convergent subsequence, meaning B is also compact and therefore T is compact.

Thus, we can invoke our properties of compact operators. Specifically, we know that the set of eigenvalues is at most countable, the only point of accumulation is 0, and the eigenspaces for $\nu \neq 0$ are finite dimensional.

Also note that A, B are self-adjoint. It can be shown that if A^{-1} is bounded and A is self adjoint, then so is A^{-1}. We can define a new inner product $(\cdot,\cdot)$ by $(u,w) := \langle Bu, w\rangle = \int_0^1 \kappa uw\,dx$. T is then self adjoint with respect to this new inner product. To see this:

$$(Tu,w) = \langle Tu, Bw\rangle = \left\langle A^{-1}Bu, Bw\right\rangle = \left\langle Bu, A^{-1}Bw\right\rangle = \langle Bu, Tw\rangle = (u, Tw)\,.$$

Thus, T is self adjoint with respect to $(\cdot,\cdot)$, meaning the eigenvalues are real and the eigenvectors associated with distinct eigenvalues are orthogonal with respect to $(\cdot,\cdot)$.

Note, we can also show that λ is always positive by noting that:

$$\langle Au, u\rangle = \underbrace{a(u,u)}_{\geq 0} = \lambda\,\underbrace{b(u,u)}_{\geq 0} = \langle Bu, u\rangle \implies \lambda \geq 0$$

Consider $T = A^{-1}B$, then $Tu = 0 \implies A^{-1}Bu = 0 \implies Bu = 0$ so $\mathrm{N}(T) = \mathrm{N}(B)$. But note, $Bu = 0 \implies \tilde{B}(\kappa u) = 0$ but by construction, we know that $(\tilde{B}) = \{0\}$ and also $\kappa u = 0 \implies u = 0$, thus $\mathrm{N}(T) = \{0\}$, meaning 0 is not an eigenvalue of T so $\nu_n \to 0$ and therefore $\lambda_n \to \infty$.

7

Distribution Theory

7.1 Multi-index notation

A multi-index is an n-tuple of non-negative integers $\alpha = (\alpha_1, \dots \alpha_n)$. We define the order of $|\alpha| = \alpha_1 + \cdots \alpha_n$. For partial derivatives, we have the notation:

$$\partial^\alpha u \coloneqq \frac{\partial^{|\alpha|}}{\partial^{\alpha_1} x_1 \cdots \partial^{\alpha_n} x_n} u$$

Example 7.1.

$$\frac{\partial^3 u}{\partial^2 x_1 \partial x_3} = \partial^\alpha u \implies \alpha = (2, 0, 1).$$

7.2 Definitions

Let $\Omega \subset \mathbb{R}^n$ be open. We can define the following spaces:

(1) $C^k(\Omega) \coloneqq \{u \in C(\Omega) : \partial^\alpha u \in C(\Omega), \forall |\alpha| \leq k\}$

(2) $C^k(\overline{\Omega}) \coloneqq \{u \in C(\Omega) : \partial^\alpha u$ can be extended continuously to $\overline{\Omega}, \forall |\alpha| \leq k\}$

(3) $C^\infty(\Omega) \coloneqq \bigcap_{k=0}^\infty C^k(\Omega)$

The **support** of a function u is the complement of the largest open set on which $u = 0$. Denote this set by $\operatorname{supp} u$. Note, for continuous functions u, $\operatorname{supp} u = \overline{\{x : u(x) \neq 0\}}$.

Define yet another space: $C_0^\infty(\Omega) \coloneqq \{u \in C^\infty(\Omega) : \operatorname{supp} u \subset \Omega$ and is compact$\}$. Historically $C_0^\infty(\Omega) = D(\Omega)$, the space of distributions.

Finally, the last space we need is $D_k = \{u \in C_0^\infty(\mathbb{R}^n) : \operatorname{supp} u \subset K : K$ is compact$\}$.

Then if $K \subset \Omega$, that is, a compact subset of Ω and $p \geq 0$ be an integer, then we define:

$$\|\phi\|_{k,p} = \sup_{\substack{x \in k \\ |\alpha| < p}} |\partial^\alpha \phi(x)| = \max_{|\alpha| \leq p} \|\partial^\alpha \phi\|_{L^\infty(K)}$$

Note that $\|\phi\|_{k,p} = 0 \not\Longrightarrow \phi = 0$ since we are only evaluating it on the compact subset K.

Definition 7.1. Let $\{\phi_n\}$ be a sequence in $C_0^\infty(\Omega)$. We say that $\phi_n \to \phi \in C_0^\infty(\Omega)$ if $\exists K \subset \Omega : \operatorname{supp} \phi_j \subset K$ for all j and $\|\phi_n - \phi\|_{k,p}$. Note again, K must be compact.

Note, this notion of convergence has very strict requirements.

7.3 Distributions

Definition 7.2. A **distribution** $u \in D'(\Omega)$ is a continuous linear functional on $C_0^\infty(\Omega)$.

Note the action of $u \in D^1(\Omega)$ on $\phi \in C_0^\infty(\Omega)$ by $u(\phi) = \langle u, \phi \rangle$, but this is not an inner product, just notationally the same.

The continuity means that $\phi_j \to \phi \implies \langle u, \phi_j \rangle \to \langle u, \phi \rangle$ or equivalently $\phi_j \to 0 \implies \langle u, \phi_j \rangle \to 0$.

Theorem 7.1. A linear functional u on $C_0^\infty(\Omega)$ is continuous iff for every compact $K \subset \Omega$, $\exists c, p : |\langle u, \phi \rangle| \leq c\|\phi\|_{k,p}$ for all $\phi \in D_k$.

Definition 7.3. If there is an integer p such that for every compact $K \subset \Omega$, $\exists c : |\langle u, \phi \rangle| \leq c\|\phi\|_{k,p}$ for all $\phi \in D_k$, then u is said to have order $\leq p$.

Example 7.2. We can consider the δ ''function'' defined by $\langle \delta, \phi \rangle := \phi(0)$. Note, in this case, $|\langle \delta, \phi \rangle| = |\phi(0)| \leq \|\phi\|_{k,0}$ for every k, so δ is a distribution of order zero. The classical notation is $\int \delta(x)\phi(x)\,dx = \langle \delta, \phi \rangle$.

Example 7.3. Every $f \in L^(\Omega)$ defines a distribution u by $\langle u, \phi \rangle := \int_\Omega f(x)\phi(x)\,dx$. Here, $|\langle u, \phi \rangle| = \left|\int_\Omega f(x)\phi(x)\,dx\right| \leq \max|\phi| \int |f(x)|dx = c\|\phi\|_{k,0}$ for any k. We then just need f to be *locally* integrable for this to hold.

Example 7.4. If α is any multi-index, u defined by $\langle u, \phi \rangle := (\partial^\alpha \phi)(0)$ is a distribution. Note here $|\langle u, \phi \rangle| = |\partial^\alpha \phi(0)| \leq \|\phi\|_{k,p}$ where $p = |\alpha|$.

Example 7.5. Take $\Omega = \mathbb{R}^n$. Denote $\phi \in C_0^\infty(\Omega)$ by $\phi(x, y)$. Define $u \in D'(\Omega)$ by $\langle u, \phi \rangle := \int_{-\infty}^{\infty} \phi(0, y)\,dy$. That is, the mass is concentrated at the y-axis.

7.4 Operations on distributions

We can consider multiplication of a distribution by a smooth function. Let $a \in C^\infty(\Omega)$ and $u \in D'(\Omega)$. Define $au \in D'(\Omega)$ by $\langle au, \phi \rangle := \langle u, a\phi \rangle$. Note, if u is a function, then $\langle au, \phi \rangle = \int au\phi \, dx = \langle u, a\phi \rangle$.

We can also consider the derivative of a distribution. Define $\left\langle \frac{\partial u}{\partial x_j}, \phi \right\rangle := -\left\langle u, \frac{\partial \phi}{\partial x_j} \right\rangle$. Note that the minus sign comes from integration by parts. For any multi-index α, $\langle \partial^\alpha u, \phi \rangle :=$ $=(-1)^{|\alpha|} \langle u, \partial^\alpha \phi \rangle$.

In general, if $T : C_0^\infty(\Omega) \to C_0^\infty(\Omega)$ is linear and continuous with T' mapping the same spaces such that $\int_\Omega (T\phi)\psi = \int_\Omega \phi(T'\psi)$ for all $\phi, \psi \in C_0^\infty(\Omega)$, then given $u \in D'(\Omega)$, define Tu by $\langle Tu, \phi \rangle := \langle u, T'\phi \rangle$. Note that T only has to be defined over smooth functions.

$$\left\langle \overbrace{\underbrace{u}_{\text{distribution}}, \underbrace{\phi}_{\text{test function}}}^{\text{dual pairing}} \right\rangle$$

We force our test functions to be limited but our distributions can be very general in our example, we pair $[H_0^1(0,1)]'$ with $H_0^1(0,1)$.

We can define the *linear partial differential operator*, defined by $L := \sum_{|\alpha| \le k} a_\alpha \partial^\alpha$ wjere $a_\alpha \in C^\infty$, k is the order of the operator. Then, note: $\int (L\phi)\psi = \sum_{|\alpha| \le k} \int \phi(\partial^\alpha (a_\alpha \psi))(-1)^\alpha = \int \phi \sum_{|\alpha| \le k} \underbrace{(-1)^\alpha \partial^\alpha (a_\alpha \psi)}_{L'\psi}$, which suggests that $\langle Lu, \phi \rangle = \langle u, L'\psi \rangle$, which is defined for every distribution.

We can also define the *translation operator* τ_x defined by $(\tau_x \phi)(y) := \phi(x+y)$. Note: $\int (\tau_x \phi)\psi = \int \phi \tau_{-x} \psi$ by a change of variables, then $\langle \tau_x u, \phi \rangle = \langle u, \tau_{-x} \phi \rangle$.

Returning back to the δ function, we will try to solve $Lu = \delta$, where L is a partial differential operator. The solution u *fundamental solution* of a PDE.

For test functions ϕ, ψ, define the *convolution* to be: $(\phi * \psi)(x) := \int \phi(x-y)\psi(y) \, dy = \int \phi(y)\psi(x-y) \, dy := (\psi * \phi)(x)$. Also define the *reflection operator* to be: $(R\phi)(x) := \phi(-x)$. Note: $(\phi * \psi)(x) = \int \phi(y)(R\tau_{-x}\psi)(y) \, dy$ by noting that $R\tau_{-x} = \tau_x R$. Thus, we can define the convolution by $(u * \phi)(x) := \langle u, R\tau_{-x}\phi \rangle$.

This provides us with a point-wise definition for u or an alternate way to define distributions by considering the limit as ϕ becomes localized. In this case, $(\delta * \phi)(x) = \phi(x)$, that is, δ becomes the identity operator: $u = L^{-1}\delta$. Now, if we solve $Lu = f$, we can take $v = u * f$ so $Lu = \delta \implies Lu * f = \delta * f = f$.

Index