# Functional Analysis and Elliptic Partial Differential Equations

Jarosław Mederski

06.05.2026

## Table of Contents

Functional analysis is a branch of mathematical analysis that focuses on the study of vector spaces and operators acting upon them. It has a wide range of applications in both pure and applied mathematics, as well as in various scientific fields. Some of the key areas where functional analysis is used include:

- **Partial Differential Equations (PDEs):** Many techniques in functional analysis are used to study and solve PDEs, which model physical phenomena such as heat, fluid flow, and waves.
- **Quantum Mechanics:** Functional analysis provides the mathematical foundation for quantum theory, especially through Hilbert spaces and linear operators.
- **Signal Processing:** Tools from functional analysis, such as Fourier transforms and wavelets, are essential in signal and image processing.
- **Control Theory:** Functional analytic methods help in modeling and solving problems related to dynamic systems and feedback control.
- **Economics and Optimization:** Functional analysis is used in the study of infinite-dimensional optimization problems and economic equilibria.
- **Numerical Analysis:** It underpins the theoretical framework for various numerical methods, including finite element methods.

In summary, functional analysis plays a critical role in both theoretical and applied contexts, bridging abstract mathematical theory with practical applications.

The *goal* of this lecture is to provide an introduction to functional analysis, which will enable further interest and research in the aforementioned areas. In the final lectures, we will aim to build the framework of functional analysis and explore variational methods that allow us to solve elliptic equations, such as the famous nonlinear Schrödinger equation

$$-\Delta u + V(x)u = f(u),$$

or the Dirichlet problem, which is important from the perspective of physics, applications, and is interesting in terms of functional analytic tools.

# 1 Banach and Hilbert Spaces

## 1.1 Normed spaces and complete norms, examples

**Definition 1.1.** A *normed space* $(X, \|\cdot\|)$ is a vector space $X$ over $\mathbb{R}$ or $\mathbb{C}$ equipped with a function $\|\cdot\| : X \to \mathbb{R}$, called a *norm*, satisfying the following properties for all $x, y \in X$ and $\alpha \in \mathbb{R}$ (or $\mathbb{C}$):

- **Positivity:** $\|x\| \geq 0$, and $\|x\| = 0$ if and only if $x = 0$.
- **Homogeneity (absolute scalability):** $\|\alpha x\| = |\alpha| \cdot \|x\|$.
- **Triangle inequality:** $\|x + y\| \leq \|x\| + \|y\|$.

**Definition 1.2.** A sequence $(x_n)$ in a normed space $(X, \|\cdot\|)$ is called a *Cauchy sequence* if for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that for all $m, n \geq N$, we have $\|x_n - x_m\| < \varepsilon$.

A normed space is said to be *complete* if every Cauchy sequence in $X$ converges to a limit in $X$. A complete normed space is called a *Banach space*.

**Example 1.1** (Classical Banach spaces)**.**

- The sequence space $\ell^p$ for $1 \leq p \leq \infty$ consists of all sequences $x = (x_n)_{n=1}^{\infty}$ of scalars such that:
$$\|x\|_p = \begin{cases} \left(\sum_{n=1}^{\infty} |x_n|^p\right)^{1/p} & \text{if } 1 \leq p < \infty, \\ \sup_{n \in \mathbb{N}} |x_n| & \text{if } p = \infty, \end{cases}$$
is finite. Each $\ell^p$ space is a Banach space.

- The space $L^p([a, b])$ for $1 \leq p \leq \infty$ consists of (equivalence classes of) measurable functions $f : [a, b] \to \mathbb{R}$ (or $\mathbb{C}$) such that the $p$-th power of the absolute value is integrable:
$$\|f\|_p = \begin{cases} \left(\int_a^b |f(x)|^p \, dx\right)^{1/p} & \text{if } 1 \leq p < \infty, \\ \operatorname{ess\,sup}_{x \in [a,b]} |f(x)| & \text{if } p = \infty. \end{cases}$$
These spaces are Banach spaces.

- The space $C([a, b])$ of continuous real (or complex) functions on $[a, b]$ equipped with the *supremum norm*
$$\|f\|_\infty = \sup_{x \in [a,b]} |f(x)|$$
is also a Banach space.

**Example 1.2** (Non-complete normed space)**.** Let $c_{00}$ denote the space of sequences with only finitely many nonzero terms, equipped with the $\ell^p$ norm for some $1 \leq p < \infty$. Then $(c_{00}, \|\cdot\|_p)$ is a normed space, but it is not complete — its completion is $\ell^p$.

**Remark 1.1.** Every norm induces a metric $d(x, y) = \|x - y\|$, so every normed space is a metric space. However, not every metric space arises from a norm.

### 1.1.1 Direct sum of normed spaces

Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed vector spaces over the same field (e.g., $\mathbb{R}$ or $\mathbb{C}$). The *direct sum* of $X$ and $Y$, denoted by $X \oplus Y$, is the Cartesian product $X \times Y$ equipped with the following operations:

$$(x_1, y_1) + (x_2, y_2) := (x_1 + x_2, y_1 + y_2), \quad \lambda \cdot (x, y) := (\lambda x, \lambda y)$$

and a norm defined, for example, by:

$$\|(x, y)\|_1 := \|x\|_X + \|y\|_Y$$

Alternatively, one can use:

$$\|(x, y)\|_\infty := \max\{\|x\|_X, \|y\|_Y\}$$

or:

$$\|(x, y)\|_2 := \left(\|x\|_X^2 + \|y\|_Y^2\right)^{1/2}$$

Each of these norms turns $X \oplus Y$ into a normed space. The specific choice of norm depends on the context and desired properties. All the above norms are *equivalent*, i.e. there are constants $0 < a < b$ such that

$$a\|(x, y)\|_i \leq \|(x, y)\|_\infty \leq b\|(x, y)\|_i$$

for any $(x, y) \in X \times Y$ and $i = 1, 2$.

### 1.1.2 Quotient space

Let $(X, \|\cdot\|)$ be a normed vector space and let $X_0 \subset X$ be a closed linear subspace. The *quotient space* $X/X_0$ is the set of equivalence classes

$$X/X_0 := \{[x] : x \in X\}, \quad \text{where } [x] := x + X_0 = \{x + x_0 : x_0 \in X_0\}.$$

Two elements $x, y \in X$ belong to the same equivalence class if and only if $x - y \in X_0$. The space $X/X_0$ becomes a normed vector space when equipped with the norm

$$\|[x]\|_{X/X_0} := \inf_{x_0 \in X_0} \|x + x_0\| = \inf_{z \in [x]} \|z\|.$$

With this norm, $X/X_0$ is a normed space.

**Theorem 1.1.** Let $(X, \|\cdot\|)$ be a Banach space and let $X_0 \subset X$ be a closed linear subspace. Then the quotient space $X/X_0$, equipped with the quotient norm

$$\|[x]\|_{X/X_0} := \inf_{x_0 \in X_0} \|x + x_0\|,$$

is a Banach space.

*Proof.* Let $([x_n])_{n\in\mathbb{N}}$ be a Cauchy sequence in the normed space $(X/X_0, \|\cdot\|_{X/X_0})$. We will show that $([x_n])$ converges in $X/X_0$.

Since $([x_n])$ is Cauchy, we can pick an increasing sequence of indices $(n_k)_{k\in\mathbb{N}}$ such that

$$\|[x_{n_{k+1}}] - [x_{n_k}]\|_{X/X_0} < 2^{-k} \qquad \text{for all } k \in \mathbb{N}.$$

Note that

$$[x_{n_{k+1}}] - [x_{n_k}] = [x_{n_{k+1}} - x_{n_k}].$$

By the definition of the quotient norm, for each $k$ there exists $u_k \in X_0$ such that

$$\|(x_{n_{k+1}} - x_{n_k}) + u_k\| < 2^{-k}.$$

Now, we define a sequence $(y_k)$ in $X$ by

$$y_1 := x_{n_1}, \qquad y_{k+1} := x_{n_{k+1}} + \sum_{j=1}^{k} u_j \quad (k \geq 1).$$

Since each $u_j \in X_0$, we have $[y_k] = [x_{n_k}]$ for all $k$. Moreover,

$$y_{k+1} - y_k = \Big(x_{n_{k+1}} - x_{n_k}\Big) + u_k,$$

hence

$$\|y_{k+1} - y_k\| < 2^{-k} \qquad (k \in \mathbb{N}).$$

It follows that $(y_k)$ is Cauchy in $X$, because for $m > k$,

$$\|y_m - y_k\| \leq \sum_{j=k}^{m-1} \|y_{j+1} - y_j\| < \sum_{j=k}^{m-1} 2^{-j} \leq \sum_{j=k}^{\infty} 2^{-j} = 2^{-(k-1)} \xrightarrow[k\to\infty]{} 0.$$

Since $X$ is Banach, there exists $y \in X$ such that $y_k \to y$ in $X$. We claim that $[x_n] \to [y]$ in $X/X_0$. First, for the subsequence we have

$$\|[x_{n_k}] - [y]\|_{X/X_0} = \|[y_k - y]\|_{X/X_0} = \inf_{x_0\in X_0} \|(y_k - y) + x_0\| \leq \|(y_k - y) + 0\| = \|y_k - y\| \to 0.$$

Thus $[x_{n_k}] \to [y]$ in $X/X_0$. Since $([x_n])$ is Cauchy in $X/X_0$, $[x_n] \to [y]$ in $X/X_0$. $\square$

### 1.1.3 Completion of a normed space

Let $(X, \|\cdot\|)$ be a normed vector space over $\mathbb{R}$ or $\mathbb{C}$. A *completion* of $X$ is a Banach space $(\widehat{X}, \|\cdot\|_{\widehat{X}})$ together with a linear isometry

$$J : X \to \widehat{X}$$

such that $J(X)$ is dense in $\widehat{X}$, i.e. $\overline{J(X)} = \widehat{X}$. In this situation we identify $X$ with its image $J(X)$ and writes $X \subset \widehat{X}$. We denote also

$$\overline{X}^{\,\|\cdot\|} = \widehat{X}.$$

**Abstract construction:** Consider the set $\mathcal{C}(X)$ of all Cauchy sequences in $X$:

$$\mathcal{C}(X) := \Big\{(x_n)_{n\in\mathbb{N}} \subset X : \ (x_n) \text{ is Cauchy in } \|\cdot\|\Big\}.$$

Define an equivalence relation on $\mathcal{C}(X)$ by

$$(x_n) \sim (y_n) \quad \iff \quad \|x_n - y_n\| \xrightarrow[n\to\infty]{} 0.$$

Let $\widehat{X} := \mathcal{C}(X)/\sim$ be the set of equivalence classes and denote the class of $(x_n)$ by $[(x_n)]$. Define addition and scalar multiplication by

$$[(x_n)] + [(y_n)] := [(x_n + y_n)], \qquad \lambda[(x_n)] := [(\lambda x_n)].$$

These operations are well-defined and make $\widehat{X}$ a vector space. Now we define

$$\|[(x_n)]\|_{\widehat{X}} := \lim_{n\to\infty} \|x_n\|.$$

This limit exists because $(x_n)$ Cauchy implies $(\|x_n\|)$ is Cauchy in $\mathbb{R}$, hence convergent. Moreover, if $(x_n) \sim (y_n)$ then $\|x_n\| - \|y_n\| \to 0$, so the limit does not depend on the chosen representative. Thus $\|\cdot\|_{\widehat{X}}$ is well-defined and is a norm on $\widehat{X}$.

Finally, define the canonical embedding (isometry)

$$J : X \to \widehat{X}, \qquad J(x) := [(x, x, x, \dots)].$$

Then $\|J(x)\|_{\widehat{X}} = \|x\|$ for all $x \in X$, so $J$ is an isometry. Moreover, $J(X)$ is dense in $\widehat{X}$: given $[(x_n)] \in \widehat{X}$, we have $J(x_n) \to [(x_n)]$ in $\widehat{X}$.

**Completeness and uniqueness:** The space $(\widehat{X}, \|\cdot\|_{\widehat{X}})$ is complete, hence a Banach space. Moreover, if $(Y, \|\cdot\|_Y)$ is another completion of $X$ with an isometric embedding $I : X \to Y$ whose range is dense, then there exists a unique linear isometry $T : \widehat{X} \to Y$ such that $T \circ J = I$. In particular, completions are unique up to a unique isometric isomorphism.

**Example 1.3.** The completion of $(\mathbb{Q}, |\cdot|)$ is $(\mathbb{R}, |\cdot|)$.

**Example 1.4.** For $1 \le p < \infty$, we have the inclusion $C([a,b]) \subset L^p([a,b])$ and $C([a,b])$ is dense in $L^p([a,b])$. Hence the completion of $C([a,b])$ with respect to the norm $\|\cdot\|_{L^p}$ is $L^p([a,b])$.

**Example 1.5.** Let $\Omega \subset \mathbb{R}^n$ be open and $1 \le p < \infty$. Then $C_0^\infty(\Omega)$ is dense in $L^p(\Omega)$, i.e.

$$\overline{C_0^\infty(\Omega)}^{\|\cdot\|_{L^p}} = L^p(\Omega).$$

Therefore the completion of the normed space $\Big(C_0^\infty(\Omega), \|\cdot\|_{L^p(\Omega)}\Big)$ is $L^p(\Omega)$.

Later, we will construct Sobolev spaces as the completion of $C_0^\infty(\Omega)$ with respect to norms involving weak derivatives (differential operators), and we will use these spaces in the study of partial differential equations.

## 1.2 Bounded Linear Operators

**Definition 1.3.** A linear operator $T : X \to Y$ between normed spaces is *bounded* if there exists $C \geq 0$ such that $\|Tx\|_Y \leq C\|x\|_X$ for all $x \in X$. The least possible constant $C$ such that the above inequality holds is denoted by $\|T\|$.

Let $X$ and $Y$ be normed vector spaces over the same field (usually $\mathbb{R}$ or $\mathbb{C}$). We define

$$\alpha(X, Y) := \{T : X \to Y \mid T \text{ is linear and continuous}\}$$

as the set of all continuous linear operators from $X$ to $Y$.

**Theorem 1.2.** $T \in \alpha(X, Y)$ if and only if one of the following condition holds

- $T$ is continuous,
- $T$ is continuous at 0,
- $T$ is bounded.

In this case,

$$\|T\| = \inf\{C \geq 0 : \|Tx\|_Y \leq C\|x\|_X \ \forall x \in X\} = \sup_{\|x\|_X \leq 1} \|Tx\|_Y = \sup_{\|x\|_X = 1} \|Tx\|_Y.$$

*Proof.* (1) ⇒ (2) is obvious.

(2) ⇒ (3). Assume $T$ is continuous at 0. Then there exists $\delta > 0$ such that

$$\|x\|_X < \delta \quad \Longrightarrow \quad \|Tx\|_Y < 1.$$

Let $x \in X$, $x \neq 0$, and set $\lambda := \delta/(2\|x\|_X)$. Then $\|\lambda x\|_X = \delta/2 < \delta$, hence $\|T(\lambda x)\|_Y < 1$. By linearity,

$$\|Tx\|_Y = \frac{1}{\lambda} \|T(\lambda x)\|_Y < \frac{1}{\lambda} = \frac{2}{\delta} \|x\|_X.$$

For $x = 0$ the inequality is trivial, so $T$ is bounded with $C = 2/\delta$.

(3) ⇒ (1). Assume $\|Tx\|_Y \leq C\|x\|_X$ for all $x \in X$. Then for any $x, y \in X$,

$$\|Tx - Ty\|_Y = \|T(x - y)\|_Y \leq C\|x - y\|_X.$$

Thus $T$ is Lipschitz, hence continuous on $X$.

Finally, for a bounded linear $T$ define

$$M := \sup_{\|x\|_X \leq 1} \|Tx\|_Y.$$

Then for any $x \neq 0$, writing $x = \|x\|_X \cdot \frac{x}{\|x\|_X}$ gives

$$\|Tx\|_Y = \|x\|_X \left\| T\left(\frac{x}{\|x\|_X}\right) \right\|_Y \leq \|x\|_X \, M,$$

so $M$ is an admissible constant. Conversely, if $\|Tx\|_Y \le C\|x\|_X$ for all $x$ then $\|Tx\|_Y \le C$ for all $\|x\|_X \le 1$, hence $M \le C$. Therefore $M$ is the least such constant and equals $\|T\|$. Moreover,

$$\sup_{\|x\|_X \le 1} \|Tx\|_Y = \sup_{\|x\|_X = 1} \|Tx\|_Y,$$

since $\|Tx\|_Y \le \|T\| \, \|x\|_X \le \|T\|$ on the unit ball and, for any $0 < \|x\|_X < 1$,

$$\|Tx\|_Y = \|x\|_X \left\| T\left( \frac{x}{\|x\|_X} \right) \right\|_Y \le \sup_{\|u\|_X = 1} \|Tu\|_Y.$$

□

The set $\alpha(X, Y)$ forms a vector space itself, and can be equipped with the *operator norm*

$$\|T\| = \sup_{\|x\|_X \le 1} \|T(x)\|_Y = \sup_{\|x\|_X = 1} \|T(x)\|_Y.$$

**Theorem 1.3.** $(\alpha(X, Y), \|\cdot\|)$ is a normed vector space, and a Banach space if $(Y, \|\cdot\|_Y)$ is a Banach space.

*Proof.* It is clear that $\alpha(X, Y)$ is a vector space, indeed if $S, T \in \alpha(X, Y)$ and $\lambda \in \mathbb{F}$, then $S + T$ and $\lambda T$ are linear. They are also continuous because

$$\|(S + T)x\|_Y \le \|Sx\|_Y + \|Tx\|_Y, \qquad \|\lambda Tx\|_Y = |\lambda| \, \|Tx\|_Y.$$

Hence $S + T, \lambda T \in \alpha(X, Y)$. Observe that the operator norm is a norm. For $T \in \alpha(X, Y)$ we have $\|T\| \ge 0$ by definition and

$$\|\lambda T\| = \sup_{\|x\|_X \le 1} \|\lambda Tx\|_Y = |\lambda| \sup_{\|x\|_X \le 1} \|Tx\|_Y = |\lambda| \, \|T\|.$$

Also,

$$\|S + T\| = \sup_{\|x\|_X \le 1} \|(S + T)x\|_Y \le \sup_{\|x\|_X \le 1} \Big( \|Sx\|_Y + \|Tx\|_Y \Big) \le \|S\| + \|T\|.$$

Finally, if $\|T\| = 0$, then $\|Tx\|_Y = 0$ for all $\|x\|_X \le 1$, hence by homogeneity $\|Tx\|_Y = 0$ for all $x \in X$, so $T = 0$. Thus $\|\cdot\|$ is a norm on $\alpha(X, Y)$.

We show the of completeness $\alpha(X, Y)$ provided that $Y$ is complete. Let $(T_n)$ be a Cauchy sequence in $(\alpha(X, Y), \|\cdot\|)$. Fix $x \in X$. Then

$$\|T_n x - T_m x\|_Y = \|(T_n - T_m)x\|_Y \le \|T_n - T_m\| \, \|x\|_X,$$

so $(T_n x)$ is Cauchy in $Y$ and hence converges (since $Y$ is complete). Define

$$Tx := \lim_{n \to \infty} T_n x \qquad (x \in X).$$

We show that $T \in \alpha(X, Y)$ and $T_n \to T$ in operator norm.

*Linearity of $T$.* For $x, y \in X$ and scalars $\lambda$,

$$T(x + y) = \lim_{n \to \infty} T_n(x + y) = \lim_{n \to \infty} (T_n x + T_n y) = Tx + Ty,$$

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

*Boundedness of $T$ and convergence in norm.* Since $(T_n)$ is Cauchy, there exists $N$ such that for all $n \geq N$,

$$\|T_n - T_N\| \leq 1.$$

Hence for all $n \geq N$,

$$\|T_n\| \leq \|T_N\| + \|T_n - T_N\| \leq \|T_N\| + 1.$$

Let $M := \|T_N\| + 1$. Then $\|T_n\| \leq M$ for all $n \geq N$. For any $x \in X$,

$$\|Tx\|_Y = \lim_{n\to\infty} \|T_n x\|_Y \leq \limsup_{n\to\infty} \|T_n\| \, \|x\|_X \leq M\|x\|_X,$$

so $T$ is bounded, hence continuous, i.e. $T \in \alpha(X, Y)$.

Now fix $\varepsilon > 0$. Since $(T_n)$ is Cauchy, there exists $N_\varepsilon$ such that $\|T_n - T_m\| < \varepsilon$ for all $n, m \geq N_\varepsilon$. Fix $n \geq N_\varepsilon$ and let $m \to \infty$. For any $\|x\|_X \leq 1$,

$$\|(T_n - T)x\|_Y = \lim_{m\to\infty} \|(T_n - T_m)x\|_Y \leq \limsup_{m\to\infty} \|T_n - T_m\| \leq \varepsilon.$$

Taking the supremum over $\|x\|_X \leq 1$ yields $\|T_n - T\| \leq \varepsilon$ for all $n \geq N_\varepsilon$. Thus $T_n \to T$ in operator norm.

Therefore every Cauchy sequence in $(\alpha(X, Y), \|\cdot\|)$ converges, so $\alpha(X, Y)$ is Banach when $Y$ is Banach. □

## 1.3 Hahn-Banach Theorem

**Theorem 1.4** (Hahn-Banach)**.** Let $p : X \to \mathbb{R}$ be a sublinear function, i.e.

- $p(\lambda x) = \lambda p(x)$,
- $p(x + y) \leq p(x) + p(y)$, for all $x, y \in X$, $\lambda > 0$,

and let $f$ be a linear functional defined on a subspace $Y \subseteq X$ such that $f(y) \leq p(y)$ for all $y \in Y$. Then $f$ can be extended to a linear functional $F$ on $X$ such that $F(x) \leq p(x)$ for all $x \in X$.

*Proof.* The proof uses Zorn's Lemma. Consider the set of all linear extensions of $f$ to subspaces of $X$ dominated by $p$. Order this set by extension. Zorn's Lemma ensures a maximal element, which can be shown to be defined on all of $X$. □

**Theorem 1.5** (Supporting functional / norm-attaining at a point)**.** Let $(X, \|\cdot\|)$ be a normed space over $\mathbb{R}$ or $\mathbb{C}$ and let $x_0 \in X$, $x_0 \neq 0$. Then there exists a bounded linear functional $f \in X^*$ such that

$$\|f\| = 1 \qquad \text{and} \qquad f(x_0) = \|x_0\|.$$

In particular,

$$\|x_0\| = \sup\{|f(x_0)| : \; f \in X^*, \; \|f\| \leq 1\}.$$

*Proof.* Let $Y := \operatorname{span}\{x_0\}$ and define $f_0 : Y \to \mathbb{F}$ by

$$f_0(\lambda x_0) := \lambda \|x_0\| \qquad (\lambda \in \mathbb{F}),$$

where $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$.

Then $f_0$ is linear and for $y = \lambda x_0$ we have

$$|f_0(y)| = |\lambda| \, \|x_0\| = \|\lambda x_0\| = \|y\|.$$

Hence $|f_0(y)| \le \|y\|$ for all $y \in Y$.

Apply Hahn–Banach in the normed form (domination by the sublinear functional $p(x) := \|x\|$) to extend $f_0$ to a linear functional $f : X \to \mathbb{F}$ such that

$$|f(x)| \le \|x\| \qquad \forall x \in X.$$

This implies $\|f\| \le 1$. Moreover, $f$ extends $f_0$, so

$$f(x_0) = f_0(x_0) = \|x_0\|.$$

Therefore $\|f\| \ge \dfrac{|f(x_0)|}{\|x_0\|} = 1$, hence $\|f\| = 1$ and $f(x_0) = \|x_0\|$. □

## 1.4 Duals of Normed Spaces

**Definition 1.4.** The *dual space* $X^* = \alpha(X, \mathbb{R})$ of a normed space $X$ is the set of all bounded linear functionals on $X$. In view of Theorem 1.3, $X^*$ is a Banach space.

**Theorem 1.6.** Let $1 < p < \infty$ and let $q$ be the *Hölder conjugate exponent* of $p$, i.e.,

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

Then the dual space $(\ell^p)^*$ is isometrically isomorphic to $\ell^q$.

*Proof.* Let $e_n = (0, \cdots, 0, 1, 0, \cdots)$ with 1 at the $n$-th place. For every continuous linear functional $\varphi \in (\ell^p)^*$, there exists a unique sequence $y = (y_n)$ such that

$$\varphi(x) = \sum_{n=1}^{\infty} x_n y_n \quad \text{for all } x = (x_n) \in \ell^p,$$

that is $y_n := \phi(e_n)$. We will show that $y = (y_n) \in \ell^q$. Let $z_n := |y_n|^{q-2} y_n$ if $y_n \ne 0$, otherwise $z_n = 0$. Observe that

$$\|\phi\| \ge \frac{|f(\sum_{i=1}^n z_i e_i)|}{\| \sum_{i=1}^n z_i e_i \|_p} = \frac{\sum_{i=1}^n |y_i|^q}{(\sum_{i=1}^n |y_i|^{(q-1)p)})^{1/p}} = \| \sum_{i=1}^n y_i e_i \|_q$$

for each $n$. Letting $n \to \infty$, we get $y \in \ell^q$ and $\|\phi\| \ge \|y\|_q$, Moreover,

$$\phi(x) \le \|x\|_p \|y\|_q,$$

so $\|\phi\| \leq \|y\|_q$. This correspondence defines an isometric isomorphism

$$\Phi : (\ell^p)^* \to \ell^q, \quad \Phi(x) = y$$

This mapping $\Phi$ is linear, bijective, and satisfies

$$\|\Phi(\phi)\|_q = \|\phi\|.$$

$\square$

**Theorem 1.7.** The dual space of $\ell^1$, denoted $(\ell^1)^*$, is isometrically isomorphic to $\ell^\infty$. That is, for every continuous linear functional $\varphi \in (\ell^1)^*$, there exists a unique sequence $y = (y_n) \in \ell^\infty$ such that

$$\varphi(x) = \sum_{n=1}^{\infty} x_n y_n \quad \text{for all } x = (x_n) \in \ell^1.$$

This correspondence defines an isometric isomorphism.

**Remark 1.2.** $\ell^1 \subsetneq (\ell^\infty)^*$

**Theorem 1.8.** Let $(X, \mathcal{A}, \mu)$ be a measure space.
If $1 < p < \infty$:

$$(L^p(\mu))^* \cong L^q(\mu), \quad \text{where } \frac{1}{p} + \frac{1}{q} = 1$$

That is, the dual space of $L^p$ is isometrically isomorphic to $L^q$, and each bounded linear functional $\phi \in (L^p)^*$ can be represented as:

$$\phi(f) = \int_X f(x) g(x) \, d\mu(x), \quad \text{for some } g \in L^q(\mu)$$

If $p = 1$:

$$(L^1(\mu))^* \cong L^\infty(\mu)$$

Each bounded linear functional on $L^1$ is given by integration against a function in $L^\infty$.
If $p = \infty$:

$$L^1(\mu) \subsetneq (L^\infty(\mu))^*$$

In this case, the dual of $L^\infty$ is strictly larger than $L^1$. Specifically:

$$(L^\infty(\mu))^* \cong \mathrm{ba}(\mu)$$

where $\mathrm{ba}(\mu)$ denotes the space of bounded finitely additive set functions (not necessarily $\sigma$-additive).

**Summary:**

$$\begin{aligned}
&\text{For } 1 < p < \infty: && (L^p)^* \cong L^q, \quad \text{where } \frac{1}{p} + \frac{1}{q} = 1 \\
&\text{For } p = 1: && (L^1)^* \cong L^\infty \\
&\text{For } p = \infty: && L^1 \subsetneq (L^\infty)^*
\end{aligned}$$

### 1.4.1 Application of Hahn-Banach theorem

As in Theorem 1.5, for a given point $x_0 \in X$, we know that there is $f \in X^*$ such that $f(x_0) = \|x_0\|$ and $\|f\|_{X^*} = 1$.

**Corollary 1.1.** Let $(X, \|\cdot\|)$ be a normed space. Then

$$\|x\| = \sup_{f \in X^*,\, \|f\| \le 1} |f(x)| = \max_{f \in X^*,\, \|f\| \le 1} |f(x)|.$$

*Proof.* If $f \in X^*$ and $\|f\| \le 1$, then $|f(x)| \le \|x\|$ for any $x \in X$. On the other hand there is $g \in X^*$ such that $g(x) = \|x\|$ and $\|g\|_{X^*} = 1$. □

Given two disjoint convex subsets $A$ and $B$ of a normed space $X$, can we find a continuous linear functional $f \in X^*$ that separates them – that is, such that the images $f(A)$ and $f(B)$ do not overlap?

We will show that, under certain mild conditions on the sets $A$ and $B$, such a functional always exists.

**Theorem 1.9** (Separation of Convex Sets)**.** Let $(X, \|\cdot\|)$ be a normed vector space and let $A, B \subset X$ be two disjoint convex subsets.

(i) If $A$ is open, then there exists $f \in X^*$ and $c \in \mathbb{R}$ such that

$$f(a) < c \le f(b) \quad \text{for all } a \in A,\ b \in B.$$

(ii) If $A$ is compact and $B$ is closed, then there exists $f \in X^*$ and $c_1, c_2 \in \mathbb{R}$, with $c_1 < c_2$, such that

$$f(a) \le c_1 < c_2 \le f(b) \quad \text{for all } a \in A,\ b \in B.$$

In order to prove theorem we need the following lemma.

**Lemma 1.1** (Minkowski functional)**.** Let $(X, \|\cdot\|)$ be a normed vector space, and let $C \subset X$ be an open convex set with $0 \in C$. Define, for each $x \in X$,

$$p(x) := \inf \left\{ \alpha > 0 : \alpha^{-1} x \in C \right\}.$$

Then the function $p : X \to [0, \infty)$ satisfies:

1. $p(\lambda x) = \lambda p(x)$ for all $\lambda > 0$ (positive homogeneity),
2. $p(x + y) \le p(x) + p(y)$ for all $x, y \in X$ (subadditivity),
3. There exists a constant $M > 0$ such that $p(x) \le M\|x\|$ for all $x \in X$,
4. $C = \{x \in X : p(x) < 1\}$.

*Proof.* 1. Let $\lambda > 0$. Then

$$p(\lambda x) = \inf \left\{ \alpha > 0 : \alpha^{-1} \lambda x \in C \right\} = \inf \left\{ \lambda\beta > 0 : \beta^{-1} x \in C \right\} = \lambda p(x).$$

2. Let $x, y \in X$ and $\varepsilon > 0$. Choose $\alpha < p(x) + \varepsilon/2$, $\beta < p(y) + \varepsilon/2$, such that

$$\frac{x}{\alpha} \in C, \quad \frac{y}{\beta} \in C.$$

Define $\lambda = \frac{\alpha}{\alpha+\beta}$, so that

$$\lambda \cdot \frac{x}{\alpha} + (1 - \lambda) \cdot \frac{y}{\beta} = \frac{x + y}{\alpha + \beta} \in C$$

by convexity. Hence,

$$p(x + y) \leq \alpha + \beta < p(x) + p(y) + \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, we conclude that $p(x + y) \leq p(x) + p(y)$.
3. Since $C$ is open and contains 0, there exists $r > 0$ such that $B(0, r) \subset C$. Then for any $x \in X$, we have

$$\left\| \frac{r}{\|x\|} x \right\| = r \Rightarrow \frac{r}{\|x\|} x \in C \quad \text{if } x \neq 0.$$

Thus $p(x) \leq \frac{\|x\|}{r}$. Let $M := \frac{1}{r}$. Then $p(x) \leq M\|x\|$ for all $x \in X$.
4. ($\subseteq$) Let $x \in C$. Since $C$ is open, there exists $\varepsilon > 0$ such that $(1 + \varepsilon)x \in C$. Hence $p(x) \leq (1 + \varepsilon)^{-1} < 1$.

($\supseteq$) Suppose $p(x) < 1$. Then there exists $\alpha < 1$ such that $\alpha^{-1}x \in C$. Since $C$ is convex and $0 \in C$, we have:

$$x = \alpha \cdot (\alpha^{-1}x) + (1 - \alpha) \cdot 0 \in C.$$

Hence, $x \in C$. Therefore,

$$C = \{x \in X : p(x) < 1\}.$$

□

**Lemma 1.2** (Separation of a point and an open convex set)**.** Let $(X, \|\cdot\|)$ be a normed space, let $C \subset X$ be a nonempty open convex set, and let $x_0 \notin C$. Then there exists a nonzero $f \in X^*$ such that

$$f(x) < f(x_0) \quad \text{for all } x \in C.$$

In other words, the *affine hyperplane* $\{x : f(x) = f(x_0)\}$ strictly separates the point $x_0$ and the convex set $C$.

*Proof.* Since $C$ is convex and open and does not contain $x_0$, we can translate everything by picking some $c_0 \in C$. Define

$$D := C - c_0 \quad \text{and} \quad y_0 := x_0 - c_0.$$

Then $D$ is still open, convex, nonempty, and $0 \in D$, while $y_0 \notin D$. Define the gauge (Minkowski) functional $p : X \to [0, \infty)$ by

$$p(x) := \inf\{\alpha > 0 : \alpha^{-1}x \in D\}.$$

By Lemma 1.1, $p$ is sublinear and satisfies $p(x) < 1 \Leftrightarrow x \in D$. In particular, $p(y_0) \geq 1$.

Consider the one-dimensional subspace $Y = \text{span}\{y_0\}$. Define a linear functional $g : Y \to \mathbb{R}$ by

$$g(\lambda y_0) = \lambda.$$

Then for any $\lambda \in \mathbb{R}$,

$$g(\lambda y_0) = \lambda \leq |\lambda| \leq p(\lambda y_0),$$

where the last inequality uses sublinearity of $p$ and $p(y_0) \geq 1$. So $g \leq p$ on $Y$.

By the Hahn–Banach Theorem 1.4, $g$ can be extended to some $f \in X^*$ such that $f \leq p$ everywhere and $f(y_0) = g(y_0) = 1$. In particular, $f$ is nonzero.

For any $x \in C$, write $x = d + c_0$ with $d \in D$. Then $p(d) < 1$, and hence

$$f(x) = f(d + c_0) = f(d) + f(c_0) < 1 + f(c_0).$$

On the other hand,

$$f(x_0) = f(y_0 + c_0) = f(y_0) + f(c_0) = 1 + f(c_0).$$

Thus $f(x) < f(x_0)$ for all $x \in C$, completing the proof. □

*Proof of Theorem 1.9.* (i) Assume $A$ is open, convex, and disjoint from convex set $B$. Fix $a_0 \in A$, $b_0 \in B$ and set

$$x_0 := a_0 - b_0.$$

Define the set

$$C := \{a - b : a \in A,\ b \in B\}.$$

Then $C \subset X$ is convex, and since $A \cap B = \emptyset$, we have $0 \notin C$. Moreover, because $A$ is open, $C$ is also open in $X$. .

Now apply Lemma 1.2 and there exists a continuous linear functional $f \in X^*$ and $\alpha > 0$ such that

$$f(x) < f(0) = 0 \quad \text{for all } x \in C.$$

In particular, for all $a \in A,\ b \in B$,

$$f(a - b) = f(a) - f(b) < 0 \Rightarrow f(a) < f(b).$$

Since $A$ is open, then $f(A)$ is open and

$$f(a) < \sup_{x \in A} f(x) \leq f(b)$$

(ii) Now suppose that $A$ is compact and $B$ is closed. Since the two sets are disjoint and $A$ is compact, the distance between them is strictly positive:

$$\rho := \inf\{\|a - b\| : a \in A,\ b \in B\} > 0.$$

We define the open $\rho$-neighborhood of $A$ as

$$A_\rho := \{x \in X : \text{dist}(x, A) < \rho\}.$$

This set is open, contains $A$, and is still disjoint from $B$, because every point in $A_\rho$ lies at a distance strictly less than $\rho$ from $A$, while all points in $B$ are at least $\rho$ away.

Now we can apply the result from part (i). Since $A_\rho$ is open and convex, and disjoint from the convex set $B$, there exists a continuous linear functional $g \in X^*$ and a scalar $c_2 \in \mathbb{R}$ such that

$$g(a) < c_2 \leq g(b) \quad \text{for all } a \in A_\rho,\ b \in B.$$

Because $A \subset A_\rho$ and $A$ is compact, the image $g(A)$ is compact in $\mathbb{R}$, and the supremum

$$c_1 := \sup_{a \in A} g(a)$$

is attained. Hence, we obtain

$$g(a) \leq c_1 < c_2 \leq g(b) \quad \text{for all } a \in A,\ b \in B.$$

This establishes the strict separation between the sets $A$ and $B$.

□

## 1.5 Open Mapping Theorem

Let $(X, d)$ be a metric space. A subset $D \subset X$ is called *dense* in $X$ if every nonempty open set $G \subset X$ intersects $D$; i.e. for every $x \in X$ and any open neighborhood $U \ni x$, $D \cap U \neq \emptyset$.

**Theorem 1.10** (Baire's Theorem)**.** Let $(X, d)$ be a complete metric space, and let $\{D_n\}_{n \in \mathbb{N}}$ be a sequence of subsets of $X$, where each $D_n$ is open and dense in $X$. Then the intersection

$$D := \bigcap_{n=1}^{\infty} D_n$$

is also dense in $X$.

**Lemma 1.3.** Let $(X, \|\cdot\|)$ be a Banach space and let $K \subset X$ be a closed, convex, and symmetric set such that

$$X = \bigcup_{n=1}^{\infty} nK, \quad \text{where } nK := \{nx : x \in K\}.$$

Then $K$ must contain a neighborhood of the origin.

*Proof.* Since $X = \bigcup_{n=1}^{\infty} nK$, we obtain

$$\bigcap_{n=1}^{\infty} (nK)^c = \emptyset$$

Because $K$ is closed, each $(nK)^c$ is open in $X$ and in view of Theorem 1.10, $(nK)^c$ cannot be dense for some $n$. Hence $K^c$ is not dense as well. This means that $K$ has nonempty interior. So for some $z$, there exists a ball $B(z, r) \subset K$, i.e., $K$ contains a ball centered at some point

$z$. Now, since $K$ is symmetric, it contains both $z$ and $-z$, and since $K$ is convex, it must also contain the midpoint between them, which is the origin:

$$0 = \frac{1}{2}(-z) + \frac{1}{2}(z) \in K.$$

Using convexity again,

$$B\Big(0, \frac{1}{2}r\Big) = \frac{1}{2}(-z) + \frac{1}{2}B(z,r) \subset K.$$

□

**Theorem 1.11** (Open Mapping Theorem)**.** Let $X$ and $Y$ be Banach spaces, and let $T \in \alpha(X,Y)$ be surjective. Then $T$ is an open mapping; that is, for every open set $U \subset X$, the image $T(U) \subset Y$ is also open.

*Proof.* We start by considering the image of the unit ball in $X$, defined as

$$K := \overline{T(B_X(0,1))}.$$

This set $K \subset Y$ is convex, symmetric, and closed. Because $T$ is surjective, we have:

$$Y = \bigcup_{n=1}^{\infty} nK.$$

Now we use Lemma 1.3, which tells us that if a symmetric, convex, closed set $K$ satisfies $Y = \bigcup_{n=1}^{\infty} nK$, then $K$ must contain a neighborhood of the origin. In particular, there exists a constant $c > 0$ such that:

$$B_Y(0, 4c) \subset K.$$

Observe that, for every $n \geq 0$

$$B_Y\Big(0, \frac{1}{2^{n-2}}c\Big) \subset \overline{T\Big(B_X\Big(0, \frac{1}{2^n}\Big)\Big)}.$$

Now we show that $B_Y(0,c) \subset T(B_X(0,1))$. Let $y \in B_Y(0,c)$. Then, since $B_Y(0,c) \subset \overline{T(B_X(0,1/4))}$, there exists a sequence of points in $X$ that approximate $y$ by $T$. Indeed, take $\varepsilon = \frac{c}{2}$. Then there exists $x_1 \in X$ such that

$$\|x_1\| < \frac{1}{4}, \quad \text{and} \quad \|y - Tx_1\| < \frac{c}{2}.$$

This implies that the remainder $y - Tx_1 \in B_Y(0, c/2) \subset \overline{T(B_X(0,1/8))}$, so we can find $x_2 \in X$ with

$$\|x_2\| < \frac{1}{8}, \quad \text{and} \quad \|y - Tx_1 - Tx_2\| < \frac{c}{4}.$$

Proceeding inductively, for each $n \in \mathbb{N}$, we choose $x_n \in X$ such that

$$\|x_n\| < \frac{1}{2^{n+1}}, \quad \text{and} \quad \Big\|y - \sum_{k=1}^{n} Tx_k\Big\| < \frac{c}{2^n}.$$

Define the partial sums

$$z_n := \sum_{k=1}^{n} x_k.$$

Then the sequence $\{z_n\}$ is Cauchy, because for $n > m$,

$$\|z_n - z_m\| = \left\| \sum_{k=m+1}^{n} x_k \right\| \leq \sum_{k=m+1}^{n} \|x_k\| < \sum_{k=m+1}^{\infty} \frac{1}{2^{k+1}} = \frac{1}{2^{m+1}}.$$

So $\{z_n\}$ converges in the Banach space $X$ to some element $z \in X$. By continuity of $T$, we have

$$Tz = \lim_{n\to\infty} Tz_n = y.$$

Furthermore,

$$\|z\| \leq \sum_{n=1}^{\infty} \|x_n\| < \sum_{n=1}^{\infty} \frac{1}{2^{n+1}} = \frac{1}{2},$$

which shows that $z \in B_X(0,1)$, and thus $y = Tz \in T(B_X(0,1))$. This proves that

$$B_Y(0,c) \subset T(B_X(0,1)).$$

So we have shown that the image of the unit ball under $T$ contains an open ball around zero in $Y$. Now we use this to show that $T$ maps any open set in $X$ to an open set in $Y$. Let $U \subset X$ be any open set, and take any point $y \in T(U)$. Then there exists $x \in U$ such that $T(x) = y$. Since $U$ is open, we can find a radius $r > 0$ such that:

$$B_X(x,r) \subset U.$$

Then:

$$T(U) \supset T(B_X(x,r)) = T(x + B_X(0,r)) = y + T(B_X(0,r)).$$

But since $T(B_X(0,r)) = r \cdot T(B_X(0,1)) \supset r \cdot B_Y(0,c) = B_Y(0,rc)$, we conclude that:

$$T(U) \supset B_Y(y,rc).$$

This means that $T(U)$ contains an open ball around every point $y \in T(U)$, so $T(U)$ is open in $Y$. □

**Remark 1.3.** Observe that a normed space does not contain any proper and open subspace. Hence if $T : X \to Y$ is a linear and open map between two normed spaces, then $T$ is surjective. Indeed, $T(X)$ is an open subspace of $Y$, hence $T(X) = Y$.

**Theorem 1.12** (Bounded Inverse Theorem)**.** Let $X$ and $Y$ be Banach spaces, and let $T \in \mathcal{L}(X,Y)$ be a bijective bounded linear operator. Then its inverse $T^{-1} : Y \to X$ is also bounded and linear; that is, $T^{-1} \in \mathcal{L}(Y,X)$.

*Proof.* Since $T$ is a bounded bijective linear operator between Banach spaces, the Open Mapping Theorem 1.11 applies. Thus, $T$ is an open mapping. In particular, the image of the unit ball in $X$, defined by

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

contains a ball around the origin in $Y$; that is, there exists $r > 0$ such that:

$$B_Y(0,r) \subset T(B_X(0,1)).$$

Now take any $y \in B_Y(0,r)$. Then $y = T(x)$ for some $x \in B_X(0,1)$, hence:

$$\|T^{-1}(y)\|_X = \|x\|_X < 1.$$

Therefore,

$$T^{-1}(B_Y(0,r)) \subset B_X(0,1).$$

This implies that $T^{-1}$ is bounded and

$$\|T^{-1}\| \leq \frac{1}{r}.$$

□

**Corollary 1.2.** Let $\|\cdot\|_1$ and $\|\cdot\|_2$ be two norms on a vector space $X$, and suppose both norms are complete (i.e., define Banach spaces), and there exists $c > 0$ such that

$$\|x\|_1 \leq c\|x\|_2 \quad \text{for all } x \in X.$$

Then the norms are equivalent; that is, there exists $c' > 0$ such that

$$\|x\|_2 \leq c'\|x\|_1 \quad \text{for all } x \in X.$$

*Proof.* Consider the identity map:

$$\text{Id} : (X, \|\cdot\|_2) \to (X, \|\cdot\|_1).$$

This map is clearly linear, and the assumption $\|x\|_1 \leq c\|x\|_2$ means that it is bounded.

Since both normed spaces are Banach (complete), and the identity map is bijective and bounded, we can apply Theorem 1.12, the inverse map

$$\text{Id}^{-1} : (X, \|\cdot\|_1) \to (X, \|\cdot\|_2)$$

is also bounded. That means:

$$\|x\|_2 \leq c'\|x\|_1 \quad \text{for some } c' > 0.$$

Hence, the norms $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent. □

## 1.6 Banach-Steinhaus Theorem

**Theorem 1.13.** Let $X$ be a Banach space and $Y$ be a normed vector space. Let $\mathcal{T} \subseteq \alpha(X, Y)$ be a family of bounded linear operators such that:

$$\sup_{T \in \mathcal{T}} \|T(x)\|_Y < \infty \quad \text{for every } x \in X.$$

Then:

$$\sup_{T \in \mathcal{T}} \|T\|_{\alpha(X,Y)} < \infty.$$

*Proof.* Define the set

$$K = \left\{ x \in X : \sup_{T \in \mathcal{T}} \|T(x)\|_Y \leq 1 \right\}.$$

It is easy to verify that $K$ is a closed, convex, and symmetric subset of $X$. For any $x \in X$, define:

$$M_x = \sup_{T \in \mathcal{T}} \|T(x)\|_Y < \infty.$$

Then

$$\left\| T\left(\frac{x}{M_x}\right) \right\|_Y = \frac{\|T(x)\|_Y}{M_x} \leq 1 \quad \text{for all } T \in \mathcal{T},$$

which implies

$$\frac{x}{M_x} \in K$$

and $x \in M_x K$. Therefore

$$X = \bigcup_{n=1}^{\infty} nK.$$

By Lemma 1.3, it follows that $K$ contains a neighborhood of the origin. Thus, there exists $r > 0$ such that the open ball $B(0, r) \subseteq K$. Hence $\sup_{T \in \mathcal{T}} \|T\| \leq r^{-1}$. $\square$

## 1.7 Closed Graph Theorem

**Theorem 1.14.** If $T : X \to Y$ is a linear operator between Banach spaces, and its graph is closed in $X \times Y$, then $T$ is bounded, i.e. $T \in \alpha(X, Y)$.

*Proof.* Let $X$ and $Y$ be Banach spaces, and consider their direct sum $X \oplus Y$ equipped with the norm

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

Define the canonical projections:

$$P_X : X \oplus Y \to X, \quad P_X(x, y) := x,$$

$$P_Y : X \oplus Y \to Y, \quad P_Y(x, y) := y.$$

Both projections $P_X$ and $P_Y$ are bounded linear operators, since they act as coordinate projections in the Banach space $X \oplus Y$. Observe that $Z := \mathit{Graph}(T) \subset X \oplus Y$ is a closed subspace such that the restriction

$$P_X|_Z : Z \to X$$

is bijective. Then, by Theorem 1.12, its inverse

$$(P_X|_Z)^{-1} : X \to Z$$

is also a bounded linear operator. Define the operator $T : X \to Y$ by

$$T := P_Y \circ (P_X|_Z)^{-1}.$$

In other words, for each $x \in X$, we find the unique $(x, y) \in Z$ with first coordinate $x$, and define $T(x) := y$. Since both $P_Y$ and $(P_X|_Z)^{-1}$ are bounded, the composition $T$ is also bounded. Therefore,

$$T \in \mathcal{L}(X, Y).$$

□

### 1.7.1 Neuman series

**Theorem 1.15** (Neumann Series)**.** Let $X$ be a Banach space and let $A \in \alpha(X, X)$ be such that $\|A\| < 1$. Then the series

$$\sum_{n=0}^{\infty} A^n = I + A + A^2 + A^3 + \dots$$

converges in the operator norm to a bounded operator $S \in \alpha(X, X)$, and this operator satisfies

$$S = (I - A)^{-1}.$$

*Proof. Step 1: Convergence of the series.* Since $\|A\| < 1$, the sequence of partial sums

$$S_N = \sum_{n=0}^{N} A^n$$

is a Cauchy sequence in the Banach space $\alpha(X, X)$. Indeed, for $M > N$ we have

$$\|S_M - S_N\| = \left\| \sum_{n=N+1}^{M} A^n \right\| \leq \sum_{n=N+1}^{M} \|A\|^n \leq \frac{\|A\|^{N+1}}{1 - \|A\|}.$$

Since the right-hand side tends to 0 as $N \to \infty$, $(S_N)$ is Cauchy. Because $\alpha(X, X)$ is Banach, $S_N$ converges in norm to some $S \in \alpha(X, X)$.
*Step 2: Computation of the limit.* For each $N \geq 0$,

$$(I - A)S_N = (I - A) \sum_{n=0}^{N} A^n = \sum_{n=0}^{N} A^n - \sum_{n=0}^{N} A^{n+1} = I - A^{N+1}.$$

Taking the limit as $N \to \infty$ gives

$$(I - A)S = \lim_{N \to \infty} (I - A)S_N = I - \lim_{N \to \infty} A^{N+1}.$$

Since $\|A^{N+1}\| \leq \|A\|^{N+1} \to 0$ as $N \to \infty$, we obtain

$$(I - A)S = I.$$

Similarly, one shows $S(I - A) = I$. Thus $S$ is the inverse of $(I - A)$, i.e.

$$S = (I - A)^{-1}.$$

We conclude that the Neumann series converges in $\alpha(X, X)$ and its sum is the inverse of $I - A$. $\square$

## 1.8 Strong and Weak Convergence

**Definition 1.5.** A sequence $\{x_n\}$ converges *strongly* to $x$ if $\|x_n - x\| \to 0$.

**Definition 1.6.** A sequence $\{x_n\}$ converges *weakly* to $x$ if $f(x_n) \to f(x)$ for all $f \in X^*$.

Let $(X, \|\cdot\|)$ be a normed vector space. We define the so-called *canonical embedding*

$$J = J_X : X \to X^{**} = (X^*)^*, \quad \text{by } J_X(x)(f) := f(x) \quad \text{for all } x \in X,\ f \in X^*.$$

**Theorem 1.16.** $J_X$ is a linear and injective operator. Moreover

$$\|J_X(x)\|_{X^{**}} = \|x\|_X.$$

for every $x \in X$.

*Proof.* Let $x, y \in X$ with $x \neq y$. Then $x - y \neq 0$. By the Hahn–Banach theorem 1.4, there exists a functional $g \in X^*$ such that

$$\|g\|_{X^*} = 1 \quad \text{and} \quad g(x - y) = \|x - y\|_X > 0.$$

Thus,

$$J_X(x)(g) = g(x) \neq g(y) = J_X(y)(g),$$

so $J_X(x) \neq J_X(y)$, which shows that $J_X$ is injective. Next, we compute the norm of $J_X(x) \in X^{**}$:

$$\|J_X(x)\|_{X^{**}} = \sup_{f \in X^*,\, \|f\|_{X^*} \leq 1} |J_X(x)(f)| = \sup_{\|f\|_{X^*} \leq 1} |f(x)| = \|x\|_X.$$

The inequality $\|J_X(x)\|_{X^{**}} \leq \|x\|_X$ follows directly from the dual norm inequality:

$$|f(x)| \leq \|f\|_{X^*} \cdot \|x\|_X.$$

To obtain equality, again a the Hahn–Banach theorem 1.4, there exists $f \in X^*$ with $\|f\|_{X^*} = 1$ and $f(x) = \|x\|_X$. Then:

$$\|J_X(x)\|_{X^{**}} \geq |J_X(x)(f)| = |f(x)| = \|x\|_X.$$

Combining both bounds gives:

$$\|J_X(x)\|_{X^{**}} = \|x\|_X,$$

which completes the proof. $\square$

**Theorem 1.17.** Let $X$ and $Y$ be normed vector spaces, and let $\{x_n\} \subset X$. Then

(i) If $x_n \to x$ strongly in $X$, then $x_n \rightharpoonup x$ weakly in $X$.

(ii) If $x_n \rightharpoonup x$ weakly in $X$, then the sequence $\{\|x_n\|_X\}$ is bounded and

$$\liminf_{n\to\infty} \|x_n\|_X \geq \|x\|_X.$$

(iii) If $x_n \rightharpoonup x$ weakly in $X$ and $T \in \mathcal{L}(X,Y)$, then $Tx_n \rightharpoonup Tx$ weakly in $Y$.

*Proof.* (i) Let $f \in X^*$. Since

$$|f(x_n) - f(x)| = |f(x_n - x)| \leq \|f\|_{X^*} \cdot \|x_n - x\|_X \to 0 \quad \text{as } n \to \infty,$$

we conclude that $f(x_n) \to f(x)$, which means $x_n \rightharpoonup x$ weakly in $X$.
(ii) By the Hahn–Banach Theorem 1.4, there exists a functional $f \in X^*$ such that $\|f\|_{X^*} = 1$ and $f(x) = \|x\|_X$. Since $x_n \rightharpoonup x$, we have

$$f(x_n) \to f(x) = \|x\|_X.$$

Also, since $|f(x_n)| \leq \|f\|_{X^*} \cdot \|x_n\|_X$, it follows that

$$\|x\|_X = \lim_{n\to\infty} f(x_n) \leq \liminf_{n\to\infty} \|x_n\|_X.$$

Assume that $x_n \rightharpoonup x$. Then, by definition of weak convergence, we have:

$$J(x_n)(f) = f(x_n) \to f(x) = J(x)(f), \quad \text{for all } f \in X^*.$$

This means that the sequence $\{J(x_n)\} \subset X^{**}$ converges pointwise on $X^*$ to $J(x)$. Hence

$$\sup_{n\geq 1} |J(x_n)(f)| < \infty$$

for every $f \in X^*$. Hence, for each $f \in X^*$, the sequence $\{J_X(x_n)(f)\}$ is bounded. By the Banach-Steinhaus Theorem 1.13, it follows that the sequence $\{J_X(x_n)\}$ is uniformly bounded in $X^{**}$, i.e.

$$\sup_{n\geq 1} \|J_X(x_n)\|_{X^{**}} < \infty.$$

Now, by Theorem 1.16 we conclude:

$$\sup_{n\geq 1} \|x_n\|_X = \sup_{n\geq 1} \|J_X(x_n)\|_{X^{**}} < \infty.$$

This proves that any weakly convergent sequence in $X$ is norm-bounded.
(iii) Let $g \in Y^*$. Then $g \circ T \in X^*$, so

$$g(Tx_n) = (g \circ T)(x_n) \to (g \circ T)(x) = g(Tx),$$

which proves that $Tx_n \rightharpoonup Tx$ weakly in $Y$. □

**Example 1.6.** Let $1 < p < \infty$, and define a sequence $\{u_n\} \subset L^p(\mathbb{R})$ by

$$u_n(t) = \begin{cases} 1 & \text{if } t \in [n, n+1), \\ 0 & \text{otherwise.} \end{cases}$$

This sequence converges weakly to zero in $L^p(\mathbb{R})$, but not strongly.
Indeed, observe that for any $n \neq m$,

$$\|u_n - u_m\|_{L^p} = \left( \int_{\mathbb{R}} |u_n(t) - u_m(t)|^p \, dt \right)^{1/p} = 2^{1/p},$$

because their supports are disjoint. So the sequence is not Cauchy, hence not strongly convergent. To prove weak convergence, recall that any $f \in (L^p)^*$ corresponds to a function $v \in L^q(\mathbb{R})$ (where $\frac{1}{p} + \frac{1}{q} = 1$) such that

$$f(u) = \int_{\mathbb{R}} u(t)v(t) \, dt.$$

Then, by the Hölder inequality and Lebesgue's dominated convergence theorem

$$f(u_n) = \int_{\mathbb{R}} u_n(t)v(t) \, dt = \int_n^{n+1} v(t) \, dt \leq \left( \int_n^{n+1} |v(t)|^q \, dt \right)^{1/q} \to 0 \quad \text{as } n \to \infty,$$

because $v \in L^q(\mathbb{R})$. Hence, $u_n \rightharpoonup 0$ weakly in $L^p(\mathbb{R})$.

### 1.8.1 Eberlein–Šmulian theorem

The *weak topology* on $X$ is the coarsest topology on $X$ such that all elements of $X^*$ remain continuous.

In general topological spaces (especially non-metrizable ones), compactness does not coincide with sequential compactness. The weak topology on an infinite-dimensional Banach space is not metrizable, so sequential compactness does not automatically follow from compactness. However the following theorem shows that weak compactness and weak sequential compactness are equivalent in Banach spaces.

**Theorem 1.18** (Eberlein–Šmulian Theorem)**.** Let $X$ be a Banach space and $A \subset X$. Then the following are equivalent:

1. $A$ is weakly compact.

2. $A$ is *weakly sequentially compact* (every sequence in $A$ has a weakly convergent subsequence).

### 1.8.2 Remarks on integrability and Vitali Convergence Theorem

For a family $\{f_n\} \subset L^1(X, \mu)$ on a (possibly infinite) measure space, we say it is *tight* if

$$\forall \varepsilon > 0 \; \exists \text{ measurable } E_0 \subset X, \; \mu(E_0) < \infty \text{ such that } \sup_n \int_{X \setminus E_0} |f_n| \, d\mu < \varepsilon.$$

**Theorem 1.19** (Vitali Convergence Theorem)**.** Suppose that

1. The family $\{f_n\} \subset L^1(X, \mu)$ is tight,
2. and $\{f_n\}$ is *uniformly integrable*, i.e. for every $\epsilon > 0$ there exists $\delta > 0$ such that for all measurable $A \subset X$ with $\mu(A) < \delta$,

$$\sup_n \int_A |f_n| \, d\mu < \epsilon.$$

If $f_n \to f$ pointwise a.e. on $X$, then $f$ is integrable on $X$ and

$$\lim_{n\to\infty} \int_X |f_n - f| \, d\mu = 0,$$

that is, $f_n \to f$ in $L^1(X, \mu)$.

The theorem generalizes the Dominated Convergence Theorem, where domination by an integrable function implies tightness and uniform integrability.

## 1.9 Weak* Topology and Banach-Alaoglu Theorem

**Definition 1.7.** If $\{f_n\} \subset X^*$ is a sequence, then we say that $f_n$ converges to $f \in X^*$ in the *weak* topology* $f_n \overset{*}{\rightharpoonup} f$ provided that

$$\lim_{n\to\infty} f_n(x) = f(x) \quad \text{for every } x \in X.$$

**Theorem 1.20** (Banach-Alaoglu)**.** The closed unit ball

$$D_{X^*} = \{f \in X^* : \|f\| \leq 1\}$$

in the dual of a normed space is compact in the weak* topology.

*Proof.* For each $x \in X$, consider the closed interval:

$$[-\|x\|, \|x\|] \subset \mathbb{R}.$$

Now define the product space:

$$K := \prod_{x \in X} [-\|x\|, \|x\|].$$

Each coordinate space is compact in $\mathbb{R}$, and so by Tychonoff's Theorem (in its elementary form for products of compact metric spaces), the product $K$ is compact with the product topology. Define a map

$$\Phi : D_{X^*} \to K, \quad \Phi(f) := (f(x))_{x\in X}.$$

This is well-defined since for all $f \in B_{X^*}$, we have:

$$|f(x)| \leq \|f\| \cdot \|x\| \leq \|x\|, \quad \text{so } f(x) \in [-\|x\|, \|x\|].$$

The mapping $\Phi$ is injective and continuous with respect to the weak* topology in $D_{X^*}$ Therefore, $\Phi(D_{X^*})$ is a subset of a compact space $K$, and since it is closed in the weak-* topology, it is compact. Thus, the unit ball $D_{X^*}$ is compact in the weak-* topology. □

## 1.10 Reflexive Spaces

**Definition 1.8.** A Banach space $X$ is *reflexive* if the natural map $J : X \to X^{**}$ defined by $J(x)(f) = f(x)$ is surjective.

**Theorem 1.21.** Any reflexive normed space is a Banach space.

*Proof.* Let $(X, \|\cdot\|)$ be a reflexive normed space. By definition the canonical map

$$J_X : X \to X^{**}, \quad J_X(x)(f) = f(x) \text{ for all } f \in X^*,$$

is an isometric isomorphism onto its image, and moreover, $J_X(X) = X^{**}$. That is, $X \cong X^{**}$ isometrically and surjectively. Since the dual space $X^*$ is a Banach space, then bidual $X^{**}$ is also a Banach space. Therefore $X$ inherits completeness from $X^{**}$, and hence $X$ is also a Banach space. □

**Corollary 1.3.** Let $X$ be a reflexive Banach space. Then every bounded sequence $\{x_n\} \subset X$ has a subsequence that converges weakly in $X$.

*Proof.* Since $\{x_n\}$ is bounded, there exists $R > 0$ such that $x_n \in RD_X$ for all $n$, where $D_X = \{x \in X : \|x\| \leq 1\}$ is the closed unit ball. Since $X$ is reflexive, the canonical embedding $J : X \to X^{**}$ is surjective. By the Banach–Alaoglu theorem, the closed unit ball $D_{X^{**}} \subset X^{**}$ is weak* compact. Reflexivity implies that $J(D_X) = D_{X^{**}}$ is weakly compact in $X$.

Therefore, the bounded closed set $RD_X$ is weakly compact in $X$. By the Eberlein–Šmulian theorem, weak compactness implies weak sequential compactness, so $(x_n)$ has a weakly convergent subsequence. □

**Example 1.7** (Reflexive Banach Spaces)**.** The following are examples of reflexive Banach spaces:

- Every finite-dimensional normed space, e.g. $\mathbb{R}^n$.
- Every Hilbert space $H$ (by the Riesz Representation Theorem) – we will see later.
- $L^p(\mu)$ spaces for $1 < p < \infty$ (with dual $L^q(\mu)$, $1/p + 1/q = 1$).
- $\ell^p$ spaces for $1 < p < \infty$ (with dual $\ell^q$).

**Example 1.8** (Non-Reflexive Banach Spaces)**.** The following are not reflexive:

- $L^1(\mu)$ and $L^\infty(\mu)$.
- $\ell^1$ and $\ell^\infty$.
- $c_0$, the space of sequences converging to 0.

**Theorem 1.22.** Any closed subspace of a reflexive Banach space is reflexive.

*Proof.* Let $X$ be a reflexive Banach space and $Y \subset X$ a closed subspace. We want to show that $Y$ is reflexive. Since $Y$ is a Banach space with the induced norm, consider its dual $Y^*$. By the Hahn–Banach theorem, each $f \in Y^*$ can be extended to an element $F \in X^*$ with the same norm. Thus, $Y^*$ can be identified with the quotient space

$$X^*/Y^\perp,$$

where

$$Y^\perp = \{F \in X^* : F(y) = 0 \text{ for all } y \in Y\}.$$

Taking duals again, we have

$$(Y^*)^* \cong (X^*/Y^\perp)^*.$$

From basic duality theory, the dual of a quotient space is isometrically isomorphic to the annihilator of the quotient, that is,

$$(X^*/Y^\perp)^* \cong (Y^\perp)^\perp \subset X^{**}.$$

Because $X$ is reflexive, we have $X^{**} \cong X$. Hence

$$(Y^*)^* \cong (Y^\perp)^\perp \subset X.$$

But $(Y^\perp)^\perp$ is exactly the weak-* closure of $Y$ in $X^{**}$, which equals $Y$ since $Y$ is closed in $X$. Thus

$$(Y^*)^* \cong Y,$$

so $Y$ is reflexive. □

**Lemma 1.4** (Riesz's Lemma)**.** Let $X$ be a normed space and $Y \subset X$ be a proper closed subspace ($Y \neq X$). Then for every $0 < \alpha < 1$, there exists $x \in X$ such that

$$\|x\| = 1 \quad \text{and} \quad \operatorname{dist}(x, Y) > \alpha,$$

where

$$\operatorname{dist}(x, Y) := \inf_{y \in Y} \|x - y\|.$$

*Proof.* Since $Y \neq X$, there exists $z \in X \setminus Y$. Because $Y$ is closed, the distance

$$d := \operatorname{dist}(z, Y) = \inf_{y \in Y} \|z - y\| > 0.$$

Choose $y_0 \in Y$ such that

$$\|z - y_0\| < \frac{d}{1 - \alpha}.$$

Define

$$x := \frac{z - y_0}{\|z - y_0\|}.$$

Then $\|x\| = 1$. For any $y \in Y$, we have $y + y_0 \in Y$, so

$$\|x - y\| = \frac{\|z - y_0 - \|z - y_0\| y\|}{\|z - y_0\|}.$$

By the choice of $y_0$, one shows that $\operatorname{dist}(x, Y) > \alpha$. Hence, such an $x$ exists, which completes the proof. □

Using Riesz's lemma, in an infinite-dimensional normed space $X$ we can find a sequence $(x_n) \subset B_X$ such that

$$\|x_n\| = 1 \quad \text{and} \quad \|x_n - x_m\| \geq \frac{1}{2} \text{ for } n \neq m.$$

This sequence has no convergent subsequence. Hence $B_X$ is never compact (in the strong topology) provided that $X$ is infinite-dimensional.

## 1.11 Compact Operators

**Definition 1.9.** $T : X \to Y$ is *compact* if it maps bounded sets to relatively compact sets.

The following are equivalent ways to characterize when a linear operator $T$ is compact:

1. $T$ is a compact operator.
2. For every bounded sequence $(x_n) \subset X$, there exists a subsequence $(x_{n_k})$ such that $(Tx_{n_k})$ converges in $Y$.

We denote by $\mathcal{K}(X,Y)$ the set of all *compact operators*.

**Theorem 1.23.** Let $X$ be a normed space, $Y$ a Banach space, and $(T_n)$ a sequence of compact operators $T_n : X \to Y$ such that $T_n \to T$ in the operator norm. Then $T$ is also a compact operator, i.e. $T \in \mathcal{K}(X,Y)$. In other words, $\mathcal{K}(X,Y)$ is a closed linear subspace of $\alpha(X,Y)$ .

*Proof.* Take any $\varepsilon > 0$. By the compactness properties, it is enough to prove that the set for any $\varepsilon$, the set $T(B_X(0,1))$ can be covered by finitely many balls in $Y$ of radius $\varepsilon$ (i.e. it is totally bounded). Since $T_n \to T$ in the operator norm, we can choose an index $m_1$ such that

$$\|T_{m_1} - T\| < \frac{\varepsilon}{2}. \tag{1}$$

Since $T_{m_1}$ is compact, there exist points $z_1, \dots, z_k \in Y$ such that

$$T_{m_1}(B_X(0,1)) \subset \bigcup_{j=1}^{k} B_Y(z_j, \varepsilon/2). \tag{2}$$

Now, for every $x \in B_X(0,1)$, from (1) and (2) we can find an index $j$ with

$$\|T(x) - z_j\|_Y \leq \|T(x) - T_{m_1}(x)\|_Y + \|T_{m_1}(x) - z_j\|_Y < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Hence

$$T(B_X(0,1)) \subset \bigcup_{j=1}^{k} B_Y(z_j, \varepsilon).$$

This shows that $T(B_X(0,1))$ is totally bounded, and therefore $T$ is compact. $\square$

**Definition 1.10.** A linear operator $T : X \to Y$ is said to have *finite rank* if its range $R(T) := T(X)$ is a finite-dimensional subspace of $Y$.

If we take vectors $y_1, \dots, y_m \in Y$ and bounded linear functionals $\lambda_1, \dots, \lambda_m \in X^*$, then the operator

$$Tx = \sum_{i=1}^{m} \lambda_i(x)\, y_i \tag{$*$}$$

has finite rank because its range is contained in the finite-dimensional subspace of $Y$ spanned by $y_1, \dots, y_m$.

**Lemma 1.5.** Every finite rank operator can be written in the form ($*$).

*Proof.* Let $y_1, \dots, y_m \in Y$ be a basis of $R(T)$. Any element $y \in R(T)$ can be uniquely expressed as

$$y = c_1 y_1 + \cdots + c_m y_m,$$

where the coefficients $c_i$ depend linearly and continuously on $y$. These coefficients define bounded linear functionals $c_i : R(T) \to \mathbb{R}$. Thus, for any $y \in R(T)$ we have

$$y = c_1(y) y_1 + \cdots + c_m(y) y_m.$$

In particular, for all $x \in X$,

$$Tx = c_1(Tx) y_1 + \cdots + c_m(Tx) y_m.$$

Hence $T$ has the form ($*$) if we define $\lambda_i(x) = c_i(Tx)$. □

**Theorem 1.24.** Every finite rank operator $T : X \to Y$ is compact.

*Proof.* Take any bounded sequence $\{x_n\} \subset X$. Then the sequence $\{Tx_n\} \subset Y$ is bounded and lies in $R(T)$. Since $R(T)$ is finite-dimensional, every bounded sequence in $R(T)$ has a convergent subsequence. Therefore $T$ is compact. □

Let $X$ and $Y$ be Banach spaces. A bounded linear operator

$$T : X \to Y$$

is called a *Fredholm operator* if the following conditions hold:

1. The kernel $\ker(T) = \{x \in X : Tx = 0\}$ is finite-dimensional.
2. The range $R(T) = \{Tx : x \in X\}$ is closed in $Y$.
3. The cokernel $Y/R(T)$ is finite-dimensional, i.e.

$$\operatorname{codim}(R(T)) = \dim(Y/R(T)) < \infty.$$

The *index* of a Fredholm operator $T$ is defined as

$$\operatorname{index}(T) = \dim(\ker T) - \operatorname{codim}(R(T)).$$

**Theorem 1.25** (Atkinson)**.** Let $K : X \to X$ be a compact operator on a Banach space $X$. Then the operator $I - K$ is Fredholm of index 0.

### 1.11.1 Spectrum o a bounded operators

Let $X$ be a real vector space and $T : X \to X$ a real linear map. The *complexification* of $X$ is

$$X_{\mathbb{C}} = X \otimes_{\mathbb{R}} \mathbb{C} \cong \{ x + iy \mid x, y \in X\}.$$

The *complexification of* $T$, denoted $T_{\mathbb{C}}$, is the map

$$T_{\mathbb{C}} : X_{\mathbb{C}} \to X_{\mathbb{C}}, \qquad T_{\mathbb{C}}(x + iy) = Tx + iTy,$$

for all $x, y \in X$.

## Properties

1. $T_{\mathbb{C}}$ is $\mathbb{C}$-linear:

   $$T_{\mathbb{C}}\Big((a + ib)(x + iy)\Big) = T_{\mathbb{C}}\Big((ax - by) + i(ay + bx)\Big) = T(ax - by) + iT(ay + bx),$$

   which equals

   $$(a + ib)\Big(Tx + iTy\Big) = (a + ib)T_{\mathbb{C}}(x + iy).$$

2. $T_{\mathbb{C}}$ extends $T$: if we identify $X$ with $X + i0 \subset X_{\mathbb{C}}$, then

   $$T_{\mathbb{C}}(x) = T(x).$$

**Definition 1.11.** Let $X$ be a normed space and let $T \in \alpha(X, X)$. We define:

- The *resolvent set* of $T$ as

  $$\rho(T) = \{\lambda \in \mathbb{C} : \lambda I - T \text{ is bijective}\}.$$

- The *spectrum* of $T$ as

  $$\sigma(T) = \mathbb{C} \setminus \rho(T).$$

- A number $\lambda \in \mathbb{C}$ is called an *eigenvalue* of $T$ if

  $$ker(\lambda I - T) \neq \{0\}.$$

Even though $T$ is real, the spectral parameter $\lambda$ is taken in $\mathbb{C}$. This is because the characteristic polynomial of a real matrix or operator may have complex roots.

Note that if $\lambda$ is an eigenvalue of $T$ if there exists $u \in X \setminus \{0\}$ such that $Tu = \lambda u$. This agrees with the usual definition of eigenvalues for matrices.

**Theorem 1.26.** Let $T \in \alpha(X, X)$ be a bounded linear operator on a Banach space $X$. Then the resolvent set $\rho(T)$ is open.

*Proof.* Let $\lambda_0 \in \rho(T)$. Then $\lambda_0 I - T$ is invertible. Denote

$$R_0 = (\lambda_0 I - T)^{-1} \in \alpha(X, X).$$

For any $\lambda \in \mathbb{C}$, we can write

$$\lambda I - T = (\lambda_0 I - T) + (\lambda - \lambda_0)I = (\lambda_0 I - T)\Big[I + (\lambda - \lambda_0)R_0\Big].$$

Since $\lambda_0 I - T$ is invertible, the invertibility of $\lambda I - T$ is equivalent to that of

$$I - (\lambda_0 - \lambda)R_0.$$

Let $M = \|R_0\|$. if

$$|\lambda - \lambda_0| < \frac{1}{M},$$

then

$$\|(\lambda_0 - \lambda)R_0\| \leq |\lambda - \lambda_0| \, \|R_0\| < 1.$$

By the Neumann series (Theorem 1.15), $I - (\lambda_0 - \lambda)R_0$ is invertible with

$$(I - (\lambda_0 - \lambda)R_0)^{-1} = \sum_{n=0}^{\infty} ((\lambda - \lambda_0)R_0)^n,$$

which converges in operator norm. Therefore, for all $\lambda$ in the disc

$$B(\lambda_0, 1/\|R_0\|) = \{\lambda \in \mathbb{C} \mid |\lambda - \lambda_0| < 1/\|R_0\|\} \subset \rho(T).$$

Hence, $\rho(T)$ is open. □

### 1.11.2 Types of Spectrum

The spectrum can be decomposed into:

1. *Point spectrum* $\sigma_p(T)$: set of *eigenvalues*, i.e.

$$\sigma_p(T) = \{\lambda \in \mathbb{C} \mid \ker(\lambda I - T) \neq \{0\}\}.$$

2. *Continuous spectrum* $\sigma_c(T)$: $\lambda I - T$ is injective and has dense range, but not surjective.

3. *Residual spectrum* $\sigma_r(T)$: $\lambda I - T$ is injective, but its range is not dense in $X$.

**Lemma 1.6** (Spectral Radius Formula)**.** Let $X$ be a Banach space and $T \in \alpha(X, X)$ be a bounded linear operator. Then

$$\sigma(T) \subset \{\lambda \in \mathbb{C} : |\lambda| \leq \|T\|\}.$$

Hence $\sigma(T)$ is compact.

*Proof.* Take any $\lambda \in \mathbb{C}$ such that $|\lambda| > \|T\|$. We will prove that $\lambda I - T$ is invertible, which implies that $\lambda \in \rho(T)$ and therefore $\lambda \notin \sigma(T)$. We can write

$$\lambda I - T = \lambda\Big(I - \frac{1}{\lambda}T\Big).$$

Because $|\lambda| > \|T\|$, we have

$$\Big\|\frac{1}{\lambda}T\Big\| = \frac{\|T\|}{|\lambda|} < 1.$$

Now we use the Neumann series for $A = \frac{1}{\lambda}T$. Since $\|A\| < 1$, the Neumann series converges, so

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

Hence

$$(\lambda I - T)^{-1} = \frac{1}{\lambda}\Big(I - \frac{1}{\lambda}T\Big)^{-1} = \frac{1}{\lambda}\sum_{n=0}^{\infty}\Big(\frac{1}{\lambda}T\Big)^{n},$$

and this series converges in operator norm. Therefore $\lambda I - T$ is invertible, which implies that $\lambda \in \rho(T)$. Since this holds for all $|\lambda| > \|T\|$, we conclude that

$$\sigma(T) \subset \{\lambda \in \mathbb{C} : |\lambda| \leq \|T\|\}.$$

□

**Theorem 1.27** (Spectral Theorem for Compact Operators)**.** Let $X$ be an infinite-dimensional Banach space (complex or real) and $T \in \mathcal{K}(X, X)$. Then

1. $0 \in \sigma(T)$;
2. Every $\lambda \in \sigma(T) \setminus \{0\}$ is an eigenvalue of $T$;
3. If $\sigma(T)$ is infinite, then $\sigma(T) = \{\lambda_n\}_{n=1}^{\infty} \cup \{0\}$ with $\lambda_n \to 0$ as $n \to \infty$.

*Proof.* (a) If $0 \notin \sigma(T)$, then $T$ is invertible. But $I = TT^{-1}$ would then be compact, implying that the unit ball of $X$ is compact, which contradicts the infinite-dimensionality of $X$.

(b) Suppose $\lambda \neq 0$ is in $\sigma(T)$ but is not an eigenvalue. Then $\lambda I - T$ is injective, but not surjective. Since $N(\lambda I - T) = \{0\}$ and $\lambda I - T$ is Fredholm of index 0, we obtain that $R(\lambda I - T) = X$, hence $\lambda \in \rho(T)$, a contradiction.

(c) Assume that $\{\lambda_n\} \subset \sigma(T) \setminus \{0\}$ is such that $\lambda_n \to \lambda$ as $n \to \infty$. Since for each $n \geq 1$, $\lambda_n$ is an eigenvalue of $T$, choose $e_n \neq 0$ an eigenvector of $T$ corresponding to $\lambda_n$. For every $n \geq 1$, the eigenvectors $e_1, e_2, \ldots, e_n$ are linearly independent. We argue by induction. Suppose $e_1, \ldots, e_n$ are linearly independent and consider $e_{n+1}$. If $e_{n+1}$ were a linear combination of $e_1, \ldots, e_n$, relabeling if necessary, we could write

$$e_{n+1} = \sum_{i=1}^{n} \alpha_i e_i.$$

Applying $T$ to both sides,

$$Te_{n+1} = \lambda_{n+1}e_{n+1} = \sum_{i=1}^{n} \alpha_i \lambda_{n+1} e_i,$$

but also

$$Te_{n+1} = \sum_{i=1}^{n} \alpha_i \lambda_i e_i.$$

Since $e_1, \ldots, e_n$ are linearly independent, we get

$$\alpha_i(\lambda_i - \lambda_{n+1}) = 0 \quad \text{for all } i = 1, \ldots, n.$$

Because $\lambda_i \neq \lambda_{n+1}$, this forces $\alpha_i = 0$ for all $i$, contradicting $e_{n+1} \neq 0$. Thus the claim is proved.

Now, we construct a sequence violating compactness. Indeed, set $E_n = \operatorname{span}\{e_1, \ldots, e_n\}$. By the claim, the sequence $\{E_n\}$ forms a strictly increasing chain of subspaces. By the Riesz Lemma, for each $n \geq 2$ there exists $u_n \in E_n$ such that

$$\|u_n\| = 1 \quad \text{and} \quad \operatorname{dist}(u_n, E_{n-1}) \geq \frac{1}{2}.$$

For $n > m \geq 2$, note that

$$(\lambda_n I - T)(E_n) \subseteq E_{n-1}, \qquad (\lambda_m I - T)(E_n) \subseteq E_{m-1}.$$

From this it follows that

$$\frac{Tu_n}{\lambda_n} - \frac{Tu_m}{\lambda_m} \in u_n - u_m + E_{n-1} = u_n + E_{n-1}$$

and

$$\left\| \frac{Tu_n}{\lambda_n} - \frac{Tu_m}{\lambda_m} \right\| \geq \operatorname{dist}(u_n, E_{n-1}) \geq \frac{1}{2}.$$

Because $T$ is compact and $\|u_n\| = 1$, the sequence $\{Tu_n\}$ must have a convergent subsequence. However, the above inequality shows that $\{Tu_n\}$ cannot be Cauchy, a contradiction.

Finally, write

$$\sigma(T) \setminus \{0\} = \bigcup_{n=1}^{\infty} A_n, \qquad A_n = \sigma(T) \cap \{\lambda \in \mathbb{C} : |\lambda| \geq 1/n\}.$$

By Step 1, each $A_n$ is finite, hence $\sigma(T)$ is at most countable. If $\sigma(T)$ is infinite, then 0 is its only accumulation point. This completes the proof. $\square$

**Corollary 1.4.** Let $T : X \to X$ be a compact operator on a Banach space $X$. Then every $\lambda \in \sigma(T) \setminus \{0\}$ is an eigenvalue of $T$, and its eigenspace $\ker(T - \lambda I)$ is finite-dimensional.

*Proof.* Since $\ker(\lambda I - T) \neq \{0\}$ and $\lambda I - T$ is Fredholm, $\dim \ker(T - \lambda I) < \infty$. Hence the eigenvalue $\lambda$ has finite multiplicity. $\square$

**Remark 1.4.** Although 0 is always in the spectrum $\sigma(T)$ when $X$ is infinite-dimensional, it is *not always in the continuous spectrum* $\sigma_c(T)$.

- If $T$ is *not injective*, then 0 is an eigenvalue (point spectrum).

$$0 \in \sigma_p(T) \iff \ker(T) \neq \{0\}.$$

- If $T$ is *injective but not surjective*, and $\overline{\mathrm{Ran}(T)} = X$, then

$$0 \in \sigma_c(T),$$

  i.e. 0 lies in the continuous spectrum.

- If $T$ is injective but $\overline{\mathrm{Ran}(T)} \neq X$, then 0 is in the *residual spectrum*.

**Remark 1.5.** In finite-dimensional spaces, linear maps $T$ are precisely matrices. In this case, the spectrum consists only of eigenvalues:

$$\sigma(T) = \sigma_p(T),$$

and there is *no continuous spectrum* and *no residual spectrum*

# Examples

**Example 1.9.** Let $T : \ell^2 \to \ell^2$ be defined by

$$T(x_1, x_2, x_3, \dots) = \left(\frac{x_1}{2}, \frac{x_2}{3}, \frac{x_3}{4}, \dots\right).$$

Each $\lambda_n = 1/(n+1)$ is an eigenvalue with eigenvector $e_n$. The value 0 is not an eigenvalue (no nonzero vector is mapped to 0), but it is in the spectrum as an accumulation point. The *point spectrum* (set of eigenvalues) is $\{1/2, 1/3, 1/4, \dots\}$, and 0 belongs to the *continuous spectrum*.

**Example 1.10.** Consider the right shift operator $S : \ell^2 \to \ell^2$ defined by

$$S(x_1, x_2, x_3, \dots) = (0, x_1, x_2, \dots).$$

Observe that 0 is not an eigenvalue, since $Sx = 0 \implies x = 0$. The range is not dense, so 0 is not in the continuous spectrum. Hence,

$$0 \in \sigma_r(S),$$

that is, 0 lies in the *residual spectrum* of $S$.

**Example 1.11.** The operator $T : C[0,1] \to C[0,1]$, defined by

$$(Tf)(x) = \int_0^x f(y)\, dy,$$

is is compact and there are no eigenvalues of $T$. $\sigma(T) = \sigma_c(T) = \{0\}$.

### 1.11.3 Adjoint Operator in complex Banach Spaces

Let $X$ be a Banach space, and let $X^*$ denote its dual space. If $T : X \to X$ is a bounded linear operator, its *adjoint operator*

$$T^* : X^* \to X^*$$

is defined by

$$(T^* f)(x) = f(Tx) \quad \text{for all } f \in X^*,\ x \in X.$$

Equivalently, $T^* f = f \circ T$.

**Lemma 1.7** (Properties of the Adjoint Operator)**.** Let $T, S \in \alpha(X, X)$ be bounded linear operators on a Banach space $X$, and let $\alpha \in \mathbb{C}$. Then:

1. *Linearity:* $(S + T)^* = S^* + T^*, \quad (\alpha T)^* = \overline{\alpha} T^*$.
2. *Norm equality:* $\|T^*\| = \|T\|$.
3. *Composition rule:* $(S \circ T)^* = T^* \circ S^*$.
4. *Double adjoint:* If $J : X \to X^{**}$ is the canonical embedding, then

   $$J \circ T = T^{**} \circ J,$$

   where $T^{**} : X^{**} \to X^{**}$ is the double adjoint.
5. If $X$ is reflexive (i.e. $J$ is surjective), then $T^{**}$ can be naturally identified with a bounded operator on $X$ itself.

## 1.12 Hilbert Spaces, Projections and Orthogonality

A *Hilbert space* is a complete inner product space.

**Definition 1.12** (Complex Hilbert Space)**.** A *complex Hilbert space* is a vector space $H$ over the field of complex numbers $\mathbb{C}$, equipped with an inner product

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

which satisfies, for all $x, y, z \in H$ and $\alpha, \beta \in \mathbb{C}$:

1. **Conjugate symmetry:** $\langle x, y \rangle = \overline{\langle y, x \rangle}$.

2. **Linearity in the first argument:** $\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle$.

3. **Positive-definiteness:** $\langle x, x \rangle \geq 0$, and $\langle x, x \rangle = 0$ if and only if $x = 0$.

The inner product induces a norm by

$$\|x\| = \sqrt{\langle x, x \rangle}.$$

The space $H$ is called a *Hilbert space* if it is *complete* with respect to this norm, i.e., every Cauchy sequence in $H$ converges to a limit in $H$.

**Lemma 1.8** (Parallelogram Law)**.** For any $x, y \in H$,

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

*Proof.* By the definition of the norm induced by the inner product:

$$\|x + y\|^2 = \langle x + y, x + y \rangle = \|x\|^2 + 2 \operatorname{Re}\langle x, y \rangle + \|y\|^2,$$

$$\|x - y\|^2 = \langle x - y, x - y \rangle = \|x\|^2 - 2 \operatorname{Re}\langle x, y \rangle + \|y\|^2.$$

Adding these two equalities gives:

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

□

**Definition 1.13.** A normed space $X$ is called *uniformly convex* if for every $\varepsilon > 0$, there exists $\delta > 0$ such that for all $x, y \in X$ with

$$\|x\| = \|y\| = 1 \quad \text{and} \quad \|x - y\| \geq \varepsilon,$$

it holds that

$$\left\| \frac{x + y}{2} \right\| \leq 1 - \delta.$$

**Theorem 1.28.** Every Hilbert space is uniformly convex.

*Proof.* Let $H$ be a Hilbert space. Suppose $\|x\| = \|y\| = 1$. Then using the parallelogram law:

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

If $\|x - y\| \geq \varepsilon$, then:

$$\left\|\frac{x+y}{2}\right\|^2 \leq 1 - \frac{\varepsilon^2}{4} \Rightarrow \left\|\frac{x+y}{2}\right\| \leq \sqrt{1 - \frac{\varepsilon^2}{4}} < 1.$$

So for any $\varepsilon > 0$, choosing $\delta = 1 - \sqrt{1 - \frac{\varepsilon^2}{4}} > 0$ shows uniform convexity. □

**Theorem 1.29.** Every uniformly convex Banach space is reflexive.

This is a classical result in functional analysis. It follows from the Milman–Pettis theorem, which shows that the closed unit ball in a uniformly convex Banach space is weakly compact, implying reflexivity by the Eberlein–Šmulian theorem. We leave this theorem without proof.

**Corollary 1.5.** Every Hilbert space is reflexive.

**Theorem 1.30** (Cauchy–Schwarz Inequality)**.** Let $H$ be a Hilbert space. For all $x, y \in H$, we have

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

Equality holds if and only if $x$ and $y$ are linearly dependent.

*Proof.* If $x = 0$ or $y = 0$, then both sides are zero, and the inequality holds trivially. Assume $x \neq 0$ and $y \neq 0$. Define a scalar $\lambda \in \mathbb{C}$ as

$$\lambda := \frac{\langle x, y\rangle}{\|y\|^2}.$$

Now consider the norm of the vector $x - \lambda y$. Since norms are always non-negative, we have

$$0 \leq \|x - \lambda y\|^2 = \langle x - \lambda y, x - \lambda y\rangle.$$

Expanding using linearity of the inner product:

$$\|x - \lambda y\|^2 = \langle x, x\rangle - \overline{\lambda}\langle x, y\rangle - \lambda\langle y, x\rangle + |\lambda|^2\langle y, y\rangle.$$

Note that $\langle y, x\rangle = \overline{\langle x, y\rangle}$, and $\langle y, y\rangle = \|y\|^2$. Substitute:

$$= \|x\|^2 - \overline{\lambda}\langle x, y\rangle - \lambda\overline{\langle x, y\rangle} + |\lambda|^2\|y\|^2.$$

Since $\lambda = \langle x, y\rangle / \|y\|^2$, we compute:

$$|\lambda|^2\|y\|^2 = \frac{|\langle x, y\rangle|^2}{\|y\|^2}.$$

Also,

$$\overline{\lambda}\langle x, y\rangle + \lambda\overline{\langle x, y\rangle} = 2\mathrm{Re}(\overline{\lambda}\langle x, y\rangle) = 2\frac{|\langle x, y\rangle|^2}{\|y\|^2}.$$

Putting it all together:

$$\|x - \lambda y\|^2 = \|x\|^2 - 2\frac{|\langle x, y\rangle|^2}{\|y\|^2} + \frac{|\langle x, y\rangle|^2}{\|y\|^2} = \|x\|^2 - \frac{|\langle x, y\rangle|^2}{\|y\|^2}.$$

Since $\|x - \lambda y\|^2 \geq 0$, we get:

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

which implies:

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

Taking square roots gives:

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

*Equality:* Equality holds if and only if $\|x - \lambda y\|^2 = 0$, i.e., $x = \lambda y$, so $x$ and $y$ are linearly dependent. □

**Theorem 1.31** (Projection Theorem)**.** Let $H$ be a Hilbert space and $M \subseteq H$ a closed subspace. Then for every $x \in H$ there exists a unique decomposition

$$x = y + z,$$

where $y \in M$ and $z \in M^{\perp}$. Equivalently, there exists a unique $y \in M$ such that

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

*Proof. Existence.* Fix $x \in H$. Consider the set

$$d = \inf_{m \in M} \|x - m\|.$$

Because $M$ is closed and $H$ is complete, there exists a sequence $(m_n) \subseteq M$ such that

$$\|x - m_n\| \to d \text{ as } n \to \infty.$$

Since $\|x - m_n\|$ is bounded, $(m_n)$ is a bounded sequence in $H$. Note that

$$\|(x - m_n) + (x - m_k)\| = 2\left\|x - \frac{m_n + m_k}{2}\right\| \geq 2d$$

and by Lemma 1.8,

$$\|(x - m_n) - (x - m_k)\|^2 + \|(x - m_n) + (x - m_k)\|^2 = 2\|(x - m_n)\|^2 + 2\|(x - m_k)\|^2.$$

It follows that $(m_n)$ is Cauchy. Because $M$ is closed, the limit $y = \lim_{n\to\infty} m_n$ exists and lies in $M$. Thus $\|x - y\| = d$, so $y$ is a minimizer.

*Orthogonality.* Let $z = x - y$. We show that $z \in M^{\perp}$. Take any $m \in M$. For any scalar $t \in \mathbb{R}$,

$$\|x - (y + tm)\|^2 = \|z - tm\|^2 = \|z\|^2 - 2t \operatorname{Re} \langle z, m\rangle + t^2\|m\|^2.$$

Because $y$ minimizes the distance, the function $f(t) = \|x - (y + tm)\|^2$ has a minimum at $t = 0$, hence $f'(0) = 0$.

Differentiating:

$$f'(t) = -2\operatorname{Re}\langle z, m\rangle + 2t\|m\|^2.$$

Thus $f'(0) = -2\operatorname{Re}\langle z, m\rangle = 0$, giving $\langle z, m\rangle = 0$. Hence $z \in M^\perp$.

*Uniqueness.* Suppose $x = y_1 + z_1 = y_2 + z_2$ with $y_1, y_2 \in M$ and $z_1, z_2 \in M^\perp$. Then

$$(y_1 - y_2) = (z_2 - z_1).$$

The left side is in $M$, the right side in $M^\perp$. Hence both sides are in $M \cap M^\perp = \{0\}$, so $y_1 = y_2$ and $z_1 = z_2$. □

## 1.13 Riesz Representation Theorem

**Theorem 1.32** (Riesz)**.** If $H$ is a Hilbert space, then for every $f \in H^*$, there exists a unique $y \in H$ such that $f(x) = \langle x, y\rangle$. Moreover,

$$\|f\| = \|y\|.$$

*Proof.* If $f = 0$, then take $y = 0$.
Assume $f \neq 0$. Then $\ker(f)$ is a closed proper subspace of $H$. By the projection theorem, we can choose a vector $z \in H$ such that

$$z \perp \ker(f), \qquad z \neq 0.$$

For any $x \in H$, we can write $x = u + \alpha z$, where $u \in \ker(f)$ and $\alpha \in \mathbb{C}$. Then

$$f(x) = f(u + \alpha z) = f(u) + \alpha f(z) = \alpha f(z),$$

since $f(u) = 0$ for $u \in \ker(f)$. We define

$$y = \frac{\overline{f(z)}}{\|z\|^2} z.$$

For any $x \in H$,

$$\langle x, y\rangle = \left\langle u + \alpha z, \frac{\overline{f(z)}}{\|z\|^2} z \right\rangle = \alpha \frac{f(z)}{\|z\|^2}\langle z, z\rangle = \alpha f(z) = f(x).$$

Thus, such $y$ exists.
*Uniqueness:* If $y_1, y_2 \in H$ satisfy $f(x) = \langle x, y_1\rangle = \langle x, y_2\rangle$ for all $x$, then $\langle x, y_1 - y_2\rangle = 0$ for all $x$. Taking $x = y_1 - y_2$, we get $\|y_1 - y_2\|^2 = 0$, hence $y_1 = y_2$.
*Norm equality:* By the Cauchy–Schwarz inequality,

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

so $\|f\| \leq \|y\|$. Taking $x = y/\|y\|$ (if $y \neq 0$) gives $\|f\| \geq \|y\|$. Hence $\|f\| = \|y\|$.

□

## 1.14 Spectral Decomposition of Self-Adjoint Compact Operators

**Theorem 1.33.** Let $H$ be a Hilbert space. Then any compact operator $T \in \mathcal{K}(H, H)$ is the norm limit of a sequence of finite-rank operators. That is, there exists a sequence $(T_n) \subset \mathcal{K}(H, K)$, each of finite rank, such that

$$\|T - T_n\| \to 0 \quad \text{as } n \to \infty.$$

*Proof.* Since $T : H \to H$ is compact, the image of the unit ball $B_H = \{x \in H : \|x\| \leq 1\}$ under $T$, denoted $T(B_H)$, has compact closure in $H$. Thus, for any $\varepsilon > 0$, there exists a finite set $\{y_1, \ldots, y_n\} \subset T(B_H)$ such that every $Tx \in T(B_H)$ can be approximated within $\varepsilon$ by some $y_j$. Let $F_\varepsilon$ be the finite-dimensional subspace spanned by these $y_j$, and let $P_\varepsilon : H \to F_\varepsilon$ be the orthogonal projection.
Now, define the operator

$$T_\varepsilon = P_\varepsilon T.$$

Then $T_\varepsilon$ is a finite-rank operator, since $R(T_\varepsilon) \subset F_\varepsilon$. Observe that, for all $x \in B_H$,

$$\|Tx - T_\varepsilon x\| = \|Tx - P_\varepsilon Tx\| \leq \varepsilon.$$

Hence,$\|T - T_\varepsilon\| \leq \varepsilon$. Since $\varepsilon > 0$ was arbitrary, this shows that $T_\varepsilon \to T$ in norm, and each $T_\varepsilon$ is a finite-rank operator. □

### 1.14.1 Is spectrum real?

**Definition 1.14.** Let $H$ be a complex Hilbert space. A bounded linear operator $T : H \to H$ is called *self-adjoint* (or Hermitian) if

$$\langle Tx, y\rangle = \langle x, Ty\rangle \quad \text{for all } x, y \in H,$$

equivalently, $T = T^*$ where $T^*$ is the adjoint operator of $T$.

**Theorem 1.34.** Let $H$ be a complex Hilbert space and let $T : H \to H$ be a bounded linear operator such that $T = T^*$ (i.e., $T$ is self-adjoint). Then

$$\sigma(T) \subseteq \mathbb{R} \quad \text{and } \|T\| = r(T),$$

where $r(T) := \sup\{|\lambda| : \lambda \in \sigma(T)\}$ denoted the *spectral radius* of $T$.

*Proof.* Let $\lambda \in \mathbb{C} \setminus \mathbb{R}$. We will show that $\lambda I - T$ is invertible, which implies that $\lambda \notin \sigma(T)$. Write $\lambda = \alpha + i\beta$ with $\alpha \in \mathbb{R}$ and $\beta \in \mathbb{R} \setminus \{0\}$. For any $x \in H$,

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

Since $T = T^*$, we have

$$(\lambda I - T)^* = (\lambda I)^* - T^* = (\overline{\lambda} I - T).$$

Thus

$$\|(\lambda I - T)x\|^2 = \langle(\lambda I - T)x, (\lambda I - T)x\rangle = \langle(\lambda I - T)^*(\lambda I - T)x, x\rangle = \langle(\overline{\lambda} I - T)(\lambda I - T)x, x\rangle.$$

Expanding gives

$$(\overline{\lambda}I - T)(\lambda I - T) = (\alpha I - T)^2 + \beta^2 I.$$

Hence

$$\|(\lambda I - T)x\|^2 = \|(\alpha I - T)x\|^2 + |\beta|^2\|x\|^2.$$

Since $|\beta|^2 > 0$, we obtain

$$\|(\lambda I - T)x\|^2 \geq |\beta|^2\|x\|^2,$$

which shows that $\lambda I - T$ is injective and has closed range. The latter statement follows from the arguments below. Let $(y_n) \subseteq R(\lambda I - T)$ be a sequence such that $y_n \to y$ as $n \to \infty$. To show that $R(\lambda I - T)$ is closed, we must prove that $y \in R(\lambda I - T)$. Since $y_n \in R(\lambda I - T)$, there exists a sequence $(x_n) \subseteq H$ such that $(\lambda I - T)x_n = y_n$ for each $n \in \mathbb{N}$. Using the inequality $\|(\lambda I - T)x_n - (\lambda I - T)x_m\| \geq |\beta|\|x_n - x_m\|$, we observe that:

$$\|x_n - x_m\| \leq \frac{1}{|\beta|}\|y_n - y_m\|.$$

Since $(y_n)$ is a convergent sequence, it is a Cauchy sequence. The inequality above implies that $(x_n)$ must also be a Cauchy sequence in $H$, thus convergent to $x$, and by the continuity of the operator $\lambda I - T$ we get

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

Therefore, $y \in R(\lambda I - T)$, which proves that the range of the operator is closed.

Its range is also dense, indeed let $y \in R(\lambda I - T)^{\perp}$, i.e.

$$\langle(\lambda I - T)x, y\rangle = 0 \quad \text{for all } x \in H.$$

This implies

$$\langle x, (\lambda I - T)^* y\rangle = 0 \quad \text{for all } x \in H.$$

Thus $(\lambda I - T)^* y = 0$. Since $T = T^*$, we have

$$(\lambda I - T)^* = \overline{\lambda}I - T$$

so

$$(\overline{\lambda}I - T)y = 0.$$

Now compute the inner product:

$$\langle(\overline{\lambda}I - T)y, y\rangle = 0.$$

So we have:

$$\langle Ty, y\rangle = \overline{\lambda}\|y\|^2.$$

Since $\langle Ty, y\rangle \in \mathbb{R}$, we get $y = 0$. (In fact: $R(A)^{\perp} = \ker(A^*)$.)

As the range is also closed, it follows that $R(T - \lambda I) = H$. Therefore $T - \lambda I$ is bijective and has a bounded inverse. Thus every $\lambda \in \mathbb{C}\backslash\mathbb{R}$ lies in the resolvent set $\rho(T)$. Consequently,

$$\sigma(T) \subseteq \mathbb{R}.$$

□

**Lemma 1.9.** Let $H$ be a complex Hilbert space and $T \in \alpha(H, H)$ be a self-adjoint operator ($T = T^*$). Then:

$$\|T^2\| = \|T\|^2.$$

*Proof.* For any bounded linear operator $T$ on a Hilbert space, the $C^*$-property of the norm holds:

$$\|T^*T\| = \|T\|^2.$$

Since $T$ is self-adjoint, we have $T = T^*$ and obtain

$$\|T^2\| = \|T \cdot T\| = \|T^*T\| = \|T\|^2.$$

□

**Theorem 1.35** (Beurling-Gelfand Formula)**.** Let $H$ be a complex Hilbert space and $T \in \alpha(H, H)$. The spectral radius $r(T)$, defined as $r(T) = \sup\{|\lambda| : \lambda \in \sigma(T)\}$, satisfies:

$$r(T) = \lim_{n\to\infty} \|T^n\|^{1/n}.$$

The lower bound is easy to obtain. Indeed, let $\lambda \in \sigma(T)$. Then $\lambda^n \in \sigma(T^n)$ for every $n \in \mathbb{N}$. Then $|\lambda|^n \leq r(T^n) \leq \|T^n\|$, hence $|\lambda| \leq \|T^n\|^{1/n}$. Consequently:

$$r(T) \leq \liminf_{n\to\infty} \|T^n\|^{1/n}.$$

The upper bound use the holomorphicity of the resolvent operator $R(z) = (zI - T)^{-1}$, see [3].

**Theorem 1.36.** Let $H$ be a complex Hilbert space and let $T : H \to H$ be a bounded linear operator such that $T = T^*$ (i.e., $T$ is self-adjoint). Then

$$\|T\| = r(T),$$

where $r(T) := \sup\{|\lambda| : \lambda \in \sigma(T)\}$ denoted the *spectral radius* of $T$.

*Proof.* It is always true that:

$$r(T) := \sup\{|\lambda| : \lambda \in \sigma(T)\} \leq \|T\|.$$

We show the reverse inequality for self-adjoint $T$.

*Proof of* $r(T) \geq \|T\|$ *no. 1:* In view of the Beurling-Gelfand Formula and Lemma 1.9 we have $\|T^{2^k}\| = \|T\|^{2^k}$ and

$$r(T) = \lim_{k\to\infty} \left(\|T^{2^k}\|\right)^{1/2^k} = \lim_{k\to\infty} \left(\|T\|^{2^k}\right)^{1/2^k} = \|T\|.$$

*Proof of* $r(T) \geq \|T\|$ *no. 2:*

The spectral theorem provides a spectral measure $E$ (see Section 1.14.2 below) such that:

$$T = \int_{\sigma(T)} \lambda \, dE(\lambda).$$

Then for any $u \in H$, $\|u\| = 1$, we have

$$\langle Tu, u\rangle = \int_{\sigma(T)} \lambda \, d\mu_u(\lambda), \quad \text{where } \mu_u(\cdot) := \langle E(\cdot)u, u\rangle.$$

Using this representation:

$$\|Tu\|^2 = \langle T^*Tu, u\rangle = \langle T^2u, u\rangle = \int_{\sigma(T)} \lambda^2 \, d\mu_u(\lambda) \leq \sup_{\lambda\in\sigma(T)} \lambda^2 \cdot \int_{\sigma(T)} 1 \, d\mu_u(\lambda) = \sup_{\lambda\in\sigma(T)} \lambda^2,$$

since

$$\int_{\sigma(T)} 1 \, d\mu_u(\lambda) = \mu_u(\sigma(T)) = \langle E(\sigma(T))u, u\rangle = \langle Iu, u\rangle = \langle u, u\rangle = 1.$$

Taking supremum over all unit vectors

$$\|T\|^2 \leq \sup_{\lambda\in\sigma(T)} \lambda^2, \quad \Rightarrow \quad \|T\| \leq \sup_{\lambda\in\sigma(T)} |\lambda| = r(T).$$

Hence

$$\|T\| = r(T).$$

□

### 1.14.2 Comments on spectral measure

Recall that in $\mathbb{R}^n$, a symmetric matrix $T$ can be diagonalized as:

$$T = \sum_{i=1}^{n} \lambda_i P_i,$$

where $\lambda_i$ are the eigenvalues, and $P_i$ are projections onto the corresponding eigenspaces.

In the infinite-dimensional case, the sum becomes an integral:

$$T = \int_{\sigma(T)} \lambda \, dE(\lambda),$$

where

$$E : \mathcal{B}(\mathbb{R}) \longrightarrow \alpha(H, H)$$

is the spectral measure. $E(\cdot)$ plays the role of *continuous projections* indexed by subsets of $\sigma(T)$. This generalizes the diagonalization of symmetric matrices and is fundamental to quantum mechanics, PDEs, and functional analysis.

For any bounded Borel measurable function $f : \sigma(T) \to \mathbb{C}$ and any Borel set $B \in \mathcal{B}(\mathbb{R})$, the operator $\int_B f(\lambda) \, dE(\lambda)$ is defined through the following steps:

- Simple functions over $B$: For a Borel measurable partition $(B_i)_{i=1}^n$ of the set $B \cap \sigma(T)$ and a simple function $f_n(\lambda) = \sum_{i=1}^n \alpha_i \chi_{B_i}(\lambda)$, where $\alpha_i \in \mathbb{C}$ and $\chi_{B_i}$ is the characteristic function of $B_i$, the integral over $B$ is defined as:

  $$\int_B f_n(\lambda) \, dE(\lambda) = \sum_{i=1}^{n} \alpha_i E(B_i).$$

  Since $E(B_i)$ are orthogonal projections, this sum is a well-defined bounded operator in $\mathcal{B}(H, H)$.

- General bounded Borel functions: For a general bounded Borel function $f$, there exists a sequence of simple functions $(f_n)$ that converges uniformly to $f$ on $B \cap \sigma(T)$. The spectral integral over $B$ is then defined as the limit in the uniform operator topology:

$$\int_B f(\lambda)\, dE(\lambda) = \lim_{n\to\infty} \int_B f_n(\lambda)\, dE(\lambda).$$

Key properties:

- **Multiplicativity:**

$$\int_B (fg)\, dE(\lambda) = \left(\int_B f\, dE(\lambda)\right)\left(\int_B g\, dE(\lambda)\right).$$

- **Isometric Property:**

$$\left\|\int_B f(\lambda)\, dE(\lambda)\right\| = \sup_{\lambda\in B\cap\sigma(T)} |f(\lambda)|.$$

- **Adjoint:**

$$\left(\int_B f(\lambda)\, dE(\lambda)\right)^* = \int_B \overline{f(\lambda)}\, dE(\lambda).$$

- **Set Additivity:** If $B = B_1 \cup B_2$ with $B_1 \cap B_2 = \emptyset$, then:

$$\int_{B_1\cup B_2} f\, dE(\lambda) = \int_{B_1} f\, dE(\lambda) + \int_{B_2} f\, dE(\lambda).$$

The following *spectral theorem* tells us that there is a way to write $T$ as an *infinite-dimensional diagonal operator*– using an *integral over its spectrum* rather than a sum over eigenvalues.

**Theorem 1.37** (Spectral Theorem)**.** Let $H$ be a complex Hilbert space and let $T \in \alpha(H, H)$ be a bounded self-adjoint operator ($T = T^*$). Then there exists a unique *projection-valued measure*

$$E : \mathcal{B}(\mathbb{R}) \longrightarrow \alpha(H, H)$$

on the Borel subsets of $\mathbb{R}$, supported on the spectrum $\sigma(T)$, such that:

1. *Spectral representation:*

$$T = \int_{\sigma(T)} \lambda\, dE(\lambda)$$

   where the integral is understood in the sense of the strong operator topology.

2. *Functional calculus:* For every bounded Borel measurable function $f : \sigma(T) \to \mathbb{C}$, there exists a bounded operator

$$f(T) = \int_{\sigma(T)} f(\lambda)\, dE(\lambda)$$

   satisfying

$$\|f(T)\| \le \sup_{\lambda\in\sigma(T)} |f(\lambda)|.$$

3. *Scalar spectral measures:* For each $u \in H$, the map

$$\mu_u(B) := \langle E(B)u, u \rangle$$

defines a positive Borel measure on $\mathbb{R}$ such that

$$\langle f(T)u, u \rangle = \int_{\sigma(T)} f(\lambda)\, d\mu_u(\lambda)$$

for all bounded Borel measurable $f$.

Moreover, the spectral measure $E$ is uniquely determined by $T$.

In other words, the spectral theorem states that there exists a spectral measure assigning a projection operator $E(B) \in \alpha(H, H)$ to each Borel set $B \subset \mathbb{R}$, such that

$$T = \int_{\sigma(T)} \lambda\, dE(\lambda).$$

This integral is an *operator-valued integral*, and it gives a complete representation of $T$ in terms of its spectral properties.

This is a generalization of the concept of spectral decomposition in finite-dimensional spaces.

Given $u \in H$, the scalar-valued measure

$$\mu_u(B) := \langle E(B)u, u \rangle$$

is a positive Borel measure on $\mathbb{R}$. It allows us to represent quantities like:

$$\langle Tu, u \rangle = \int_{\sigma(T)} \lambda\, d\mu_u(\lambda), \quad \|Tu\|^2 = \int_{\sigma(T)} \lambda^2\, d\mu_u(\lambda).$$

This turns inner products and norms into integrals — a key tool in spectral analysis.

- It lets us treat $T$ as a "diagonal" operator, even in infinite dimensions.
- It provides a natural way to define *functions of operators*:

  $$f(T) := \int_{\sigma(T)} f(\lambda)\, dE(\lambda),$$

  for any bounded Borel measurable function $f : \mathbb{R} \to \mathbb{C}$.
- It reveals how different parts of the spectrum influence the behavior of $T$.

### 1.14.3 Spectral properties of self-adjoint operators

**Theorem 1.38.** If $T \in \alpha(H, H)$ is self-adjoint, then the residual spectrum is empty:

$$\sigma_r(T) = \emptyset.$$

*Proof.* Assume $\lambda \in \mathbb{R}$ and $T - \lambda I$ is injective, but $R(\lambda I - T)$ is not dense in $H$. Then there exists $v \in \overline{R(\lambda I - T)}^{\perp} \setminus \{0\}$ such that

$$\langle (\lambda I - T)u, v \rangle = 0 \quad \text{for all } u \in H.$$

By the definition of the adjoint and the fact that $T = T^*$, we have

$$\langle u, (\lambda I - T)v \rangle = 0 \quad \text{for all } u \in H.$$

Hence,

$$(\lambda I - T)v = 0 \quad \Rightarrow \quad v \in \ker(\lambda I - T).$$

This contradicts the assumption that $\lambda I - T$ is injective. Therefore, our assumption that $R(\lambda I - T)$ is not dense must be false. □

**Theorem 1.39** (Weyl's Criterion for the Continuous Spectrum)**.** Let $T \in \alpha(H, H)$ be a bounded self-adjoint operator on a Hilbert space $H$, and let $\lambda \in \mathbb{R}$. Then $\lambda \in \sigma_c(T)$ (the continuous spectrum of $T$) if and only if:

- $T - \lambda I$ is injective (i.e., $\ker(T - \lambda I) = \{0\}$),
- There exists a sequence $\{u_n\} \subset H$ with $\|u_n\| = 1$ such that

$$\|(T - \lambda I)u_n\| \to 0 \quad \text{as } n \to \infty.$$

*Proof.* **(⇒)** Suppose $\lambda \in \sigma_c(T)$. Then:

- $\lambda I - T$ is injective,
- $\overline{R(\lambda I - T)} = H$, but
- $R(\lambda I - T) \neq H$, so $T - \lambda I$ is not surjective.

Thus, $(\lambda I - T)$ is injective but does not have a bounded inverse. Hence, there exists a sequence $y_n \in R(\lambda I - T)$, with $\|y_n\| \to 0$, such that the preimages $x_n = (\lambda I - T)^{-1} y_n$ satisfy $\|x_n\| \geq \delta > 0$. Define $u_n := x_n / \|x_n\|$, then $\|u_n\| = 1$, and

$$\|(\lambda I - T)u_n\| = \left\| \frac{y_n}{\|x_n\|} \right\| \to 0,$$

since $\|y_n\| \to 0$. Hence, $\{u_n\}$ is a Weyl sequence for $\lambda$.

**(⇐)** Now suppose $\lambda \in \mathbb{R}$, $\lambda I - T$ is injective, and there exists $\{u_n\} \subset H$ with $\|u_n\| = 1$, and

$$\|(\lambda I - T)u_n\| \to 0.$$

We show that $\lambda \in \sigma_c(T)$. Suppose for contradiction that $\lambda \notin \sigma_c(T)$. Then either:

1. $\lambda \in \rho(T)$, i.e., $T - \lambda I$ is bijective with bounded inverse, or
2. $\lambda \in \sigma_p(T)$, i.e., $T - \lambda I$ is not injective.

Case (2) contradicts our assumption that $T - \lambda I$ is injective. In Case (1), if $(T - \lambda I)^{-1}$ exists and is bounded, then

$$u_n = (T - \lambda I)^{-1}(T - \lambda I)u_n \to 0,$$

since $\|(T - \lambda I)u_n\| \to 0$, and the inverse is bounded. But this contradicts $\|u_n\| = 1$ for all $n$. Therefore, $\lambda \notin \rho(T)$ or $\sigma_p(T)$, so it must be in $\sigma_c(T)$. □

**Theorem 1.40.** Let $H$ be a Hilbert space, and let $T \in \alpha(H, H)$ be a bounded self-adjoint operator. Then the operator norm of $T$ is given by

$$\|T\| = \sup_{\|u\|=1} |\langle Tu, u\rangle|$$

and

$$\|T\| = r(T) := \sup\{|\lambda| : \lambda \in \sigma(T)\}.$$

*Proof.* Let $\|u\| = 1$. Since $T$ is self-adjoint, we have $\langle Tu, u\rangle \in \mathbb{R}$, and by the Cauchy–Schwarz inequality:

$$|\langle Tu, u\rangle| \leq \|Tu\| \cdot \|u\| = \|Tu\| \leq \|T\|.$$

Hence

$$\sup_{\|u\|=1} |\langle Tu, u\rangle| \leq \|T\|.$$

To prove the reverse inequality, we use Theorem 1.38 and Theorem 1.39. Indeed, for $\varepsilon > 0$, there exists a unit vector $u_\varepsilon \in H$ such that

$$|\langle Tu_\varepsilon, u_\varepsilon\rangle| > \|T\| - \varepsilon.$$

Therefore:

$$\sup_{\|u\|=1} |\langle Tu, u\rangle| \geq \|T\| - \varepsilon.$$

Since this holds for all $\varepsilon > 0$, we conclude:

$$\sup_{\|u\|=1} |\langle Tu, u\rangle| = \|T\|.$$

□

This result shows that for self-adjoint operators, the norm and spectral radius coincide, and both can be computed using the maximum absolute value of $\langle Tu, u\rangle$ over unit vectors. This reflects the fact that self-adjoint operators behave much like real symmetric matrices in finite-dimensional spaces.

**Definition 1.15** (Numerical Range)**.** Let $H$ be a complex Hilbert space and $T \in \alpha(H, H)$. The *numerical range* of $T$ is

$$W(T) := \{\langle Tu, u\rangle : u \in H,\ \|u\| = 1\}.$$

**Theorem 1.41** (Toeplitz–Hausdorff)**.** For any bounded linear operator $T \in \alpha(H,H)$ on a complex Hilbert space $H$, the numerical range $W(T)$ is a convex subset of $\mathbb{C}$. Moreover $\sigma(T) \subset \overline{W(T)}$.

**Theorem 1.42.** Let $H$ be a Hilbert space, and let $T \in \alpha(H,H)$ be a bounded self-adjoint operator. Define:

$$M := \sup_{\|u\|=1} \langle Tu, u\rangle, \quad m := \inf_{\|u\|=1} \langle Tu, u\rangle.$$

Then:

(a) The norm of $T$ satisfies

$$\|T\| = \max\{|m|, |M|\}.$$

(b) The spectrum $\sigma(T) \subset \mathbb{R}$ lies in the interval $[m, M]$, and both endpoints belong to the spectrum:

$$\{m, M\} \subset \sigma(T) \subset [m, M].$$

*Proof.* *(a)* Since $T$ is self-adjoint, all values $\langle Tu, u\rangle$ with $\|u\| = 1$ are real. It is known that:

$$\|T\| = \sup_{\|u\|=1} |\langle Tu, u\rangle|.$$

Hence

$$\|T\| = \max\{|M|, |m|\}.$$

*(b)* For self-adjoint operators, $W(T) \subset \mathbb{R}$, and $W(T)$ is convex. Moreover, the spectrum of $T$ is contained in the closure of the numerical range:

$$\sigma(T) \subset \overline{W(T)} = [m, M].$$

To show that $m, M \in \sigma(T)$, suppose for contradiction that $M \notin \sigma(T)$. Then the operator $(T - MI)$ is bounded and invertible. That implies there exists $\delta > 0$ such that:

$$\|(T - MI)u\| \geq \delta\|u\| \quad \text{for all } u \in H. \tag{1}$$

Hence $a(u, v) := \langle (T - MI)u, v\rangle$ is a symmetric and positive definite bilinear form on $H$. Hence, it satisfies the Cauchy-Schwarz inequality, i.e.

$$|\langle (T - MI)u, v\rangle| \leq |\langle (T - MI)u, u\rangle|^{1/2} |\langle (T - MI)v, v\rangle|^{1/2}.$$

Observe that if $\|v\| \leq 1$, then

$$\langle (MI - T)v, v\rangle = M\|v\|^2 - \langle Tv, v\rangle \leq M - m.$$

Then we get

$$|\langle (MI - T)u, v\rangle| \leq \sqrt{M - m} \cdot \langle (MI - T)u, u\rangle^{1/2}.$$

Taking supremum for $\|v\| \leq 1$, we get $(MI - T)u$:

$$\|(MI - T)u\| \leq \sqrt{M - m} \cdot \langle (MI - T)u, u\rangle^{1/2},$$

By the definition of $M$, there is a sequence $(u_n)$ in $H$ such that $\|u_n\| = 1$ and

$$\langle Tu_n, u_n \rangle \to M \quad \text{if } n \to \infty.$$

Then

$$\langle (MI - T)u_n, u_n \rangle = M\|u_n\|^2 - \langle Tu_n, u_n \rangle = M - \langle Tu_n, u_n \rangle \to 0$$

contradicting (1). Therefore, $M \in \sigma(T)$, and a similar argument shows $m \in \sigma(T)$. Thus:

$$\sigma(T) \subset [m, M], \quad \{m, M\} \subset \sigma(T).$$

□

**Remark 1.6.** This result is fundamental in spectral theory. Part (a) gives a simple way to compute the norm of a self-adjoint operator. Part (b) says that the spectrum, which describes the *values* associated with the operator, lies within the interval of the smallest and largest expected values $\langle Tu, u \rangle$, and includes the endpoints.

## 1.15 Coercivity and Lax-Milgram theorem

This result is fundamental in the theory of weak (variational) solutions to PDEs. It ensures that, under mild conditions, variational formulations of boundary value problems have unique solutions as we shall see later.

**Theorem 1.43** (Lax-Milgram)**.** Let $H$ be a real Hilbert space, and let $a : H \times H \to \mathbb{R}$ be a bilinear form satisfying:

- *Boundedness:* There exists a constant $M > 0$ such that

$$|a(u, v)| \leq M\|u\|\|v\| \quad \text{for all } u, v \in H.$$

- *Coercivity:* There exists a constant $\alpha > 0$ such that

$$a(u, u) \geq \alpha\|u\|^2 \quad \text{for all } u \in H.$$

Then, for every functional $f \in H^*$, there exists a unique $u \in H$ such that

$$a(u, v) = f(v) \quad \text{for all } v \in H.$$

*Proof.* Since $f \in H^*$, by the Riesz Representation Theorem, there exists a unique $w \in H$ such that

$$f(v) = \langle v, w \rangle \quad \text{for all } v \in H.$$

Define an operator $A : H \to H$ by the condition:

$$\langle Au, v \rangle = a(u, v) \quad \text{for all } v \in H.$$

We first show $A$ is well-defined and bounded. From the boundedness of $a$, for all $u \in H$:

$$\|Au\| = \sup_{\|v\|=1} |a(u, v)| \leq M\|u\|.$$

Next, we show $A$ is coercive:

$$\langle Au, u\rangle = a(u,u) \geq \alpha\|u\|^2.$$

Hence, $A$ is bounded, linear, and coercive. These properties imply that $A$ is an isomorphism from $H$ onto $H$. Thus, for each $f \in H^*$, there exists a unique $u \in H$ such that

$$\langle Au, v\rangle = f(v) \quad \text{for all } v \in H,$$

i.e.,

$$a(u,v) = f(v) \quad \text{for all } v \in H.$$

□

**Remark 1.7.** This result is also true in complex Hilbert spaces provided that we replace the coercivity condition by $\operatorname{Re} a(u,u) \geq \alpha\|u\|^2$.

### 1.15.1 Orthonormal basis

**Definition 1.16.** Let $H$ be a Hilbert space. A sequence $\{e_n\}_{n=1}^{\infty} \subset H$ is called an *orthonormal basis* of $H$ (or a *Hilbert basis* of $H$) if:

(a) $\langle e_n, e_m\rangle = \delta_{mn}$ for all $m, n \in \mathbb{N}$;

(b) the linear span of $\{e_n\}_{n=1}^{\infty}$ is dense in $H$.

**Lemma 1.10.** Let $H$ be a Hilbert space and let $\{e_n\}_{n=1}^{\infty} \subset H$ be an orthonormal basis of $H$. Then for any $u \in H$, we have

$$u = \sum_{n=1}^{\infty} \langle u, e_n\rangle e_n,$$

and

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

*Proof.* Since $\{e_n\}$ is an orthonormal basis of $H$, by definition the linear span of $\{e_n\}$ is dense in $H$. Define the partial sums:

$$u_N := \sum_{n=1}^{N} \langle u, e_n\rangle e_n.$$

Each $u_N$ lies in the finite-dimensional subspace spanned by $\{e_1, \ldots, e_N\}$, so $u_N \in H$. Now we compute the norm of the difference:

$$\|u - u_N\|^2 = \left\| u - \sum_{n=1}^{N} \langle u, e_n\rangle e_n \right\|^2.$$

Using the Pythagorean theorem (since the projection onto the subspace is orthogonal), we get:

$$\|u\|^2 = \|u_N\|^2 + \|u - u_N\|^2.$$

Hence:

$$\|u - u_N\|^2 = \|u\|^2 - \sum_{n=1}^{N} |\langle u, e_n\rangle|^2.$$

This shows that:

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

But since $\{e_n\}$ is an orthonormal basis (not just orthonormal), we also know that if $\|u - u_N\| \to 0$, then $u_N \to u$ in norm. Therefore:

$$\lim_{N\to\infty} u_N = u.$$

This proves the first identity:

$$u = \sum_{n=1}^{\infty} \langle u, e_n\rangle e_n.$$

Substituting this into the norm yields:

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

Thus, both identities are proven. □

**Theorem 1.44.** Every separable Hilbert space possesses an orthonormal basis.

*Proof.* Since the Hilbert space $H$ is separable, there exists a countable dense subset, which we denote by $\{v_n\}_{n=1}^{\infty} \subset H$. For each positive integer $n$, let $F_n$ represent the finite-dimensional subspace of $H$ consisting of all linear combinations of the first $n$ vectors:

$$F_n = \text{span}\{v_1, v_2, \ldots, v_n\}.$$

Define

$$F = \bigcup_{n=1}^{\infty} F_n.$$

Since $\{v_n\}$ is dense in $H$, so is $F$. We now apply the Gram-Schmidt orthonormalization procedure to construct an orthonormal set from $\{v_n\}$ that spans a dense subspace in $H$. Begin by choosing $e_1 \in F_1$ with $\|e_1\| = 1$. Assume that we have already constructed an orthonormal set $\{e_1, \ldots, e_k\} \subset F$. If $F_{k+1}$ contains a vector not in the span of $\{e_1, \ldots, e_k\}$, we can find such a vector and orthonormalize it against the existing ones to produce $e_{k+1}$, thereby extending our orthonormal set. Continuing this process inductively, we obtain a countable orthonormal set $\{e_n\}_{n=1}^{\infty} \subset H$ whose linear combinations are dense in $H$, meaning that it forms an orthonormal basis. □

**Theorem 1.45.** Let $H$ be a Hilbert space, and let $T : H \to H$ be a compact, self-adjoint linear operator. Then there exists an orthonormal basis of $H$ consisting of eigenvectors of $T$. Moreover, all the corresponding eigenvalues are real and the only possible accumulation point of the spectrum $\sigma(T)$ is zero.

*Proof.* Use the spectral theorem for compact operators and the fact that self-adjointness guarantees real eigenvalues and orthogonality of eigenspaces. Indeed we sonsider the following steps.
*Step 1: Properties of compact self-adjoint operators.* Since $T$ is self-adjoint, the spectrum $\sigma(T) \subset \mathbb{R}$. As a compact operator on an infinite-dimensional Hilbert space, its spectrum consists of a countable set of real eigenvalues with only possible accumulation point at zero. Moreover, every nonzero element of $\sigma(T)$ is an eigenvalue with finite multiplicity.
*Step 2: Existence of a maximal eigenvalue.* Consider the quantity:

$$\lambda_1 := \sup \{ |\langle Tx, x \rangle| : x \in H, \ \|x\| = 1 \}.$$

By the Riesz representation theorem and compactness of $T$, this supremum is attained. Thus, there exists $x_1 \in H$, $\|x_1\| = 1$, such that:

$$\langle Tx_1, x_1 \rangle = \lambda_1.$$

We now show that $x_1$ is an eigenvector. Define the functional $\phi(x) := \langle Tx, x_1 \rangle$. Since $T$ is self-adjoint,

$$\langle Tx_1, x \rangle = \langle x_1, Tx \rangle = \langle Tx, x_1 \rangle = \phi(x).$$

Then $\phi$ is continuous and linear, hence by Riesz, there exists $y \in H$ such that $\phi(x) = \langle x, y \rangle$. But then $\langle Tx_1 - y, x \rangle = 0$ for all $x \in H$, so $Tx_1 = y = \lambda_1 x_1$. Hence, $x_1$ is an eigenvector with eigenvalue $\lambda_1$.
*Step 3: Orthogonal decomposition.* Let $H_1 := \operatorname{span}\{x_1\}$, and consider $H_1^\perp$, the orthogonal complement of $H_1$. Since $T$ is self-adjoint, $H_1^\perp$ is invariant under $T$: if $x \in H_1^\perp$, then for any $y \in H_1$,

$$\langle Ty, x \rangle = \langle y, Tx \rangle = 0,$$

so $Tx \in H_1^\perp$. Restrict $T$ to $H_1^\perp$, denoted $T|_{H_1^\perp}$, which is still compact and self-adjoint. Repeating the argument, there exists a unit vector $x_2 \in H_1^\perp$ such that $Tx_2 = \lambda_2 x_2$ for some $\lambda_2 \in \mathbb{R}$. Proceed inductively: at each step, define $H_n := \operatorname{span}\{x_1, \ldots, x_n\}$, and consider $H_n^\perp$, which is invariant under $T$, and on which $T$ remains compact and self-adjoint. If $T|_{H_n^\perp} \neq 0$, then there exists a unit eigenvector $x_{n+1} \in H_n^\perp$ such that $Tx_{n+1} = \lambda_{n+1} x_{n+1}$.
*Step 4: Completeness.* This process produces an orthonormal set $\{x_n\}$ of eigenvectors of $T$ with corresponding eigenvalues $\lambda_n \in \mathbb{R}$, possibly zero. Suppose the closed span of $\{x_n\}$ is a proper subspace $M \subsetneq H$. Then $M^\perp \neq \{0\}$, and $T|_{M^\perp}$ is again compact and self-adjoint. But by construction, $T|_{M^\perp} = 0$, since we exhausted all eigenvectors. Hence, for all $x \in M^\perp$, $Tx = 0$. So $x$ is an eigenvector with eigenvalue 0. Thus, including such vectors, we obtain a complete orthonormal set of eigenvectors of $T$.
*Conclusion:* The set $\{x_n\}$ (including all with eigenvalue 0) forms an orthonormal basis for $H$, and each $x_n$ is an eigenvector of $T$. Therefore, $T$ is diagonalizable in an orthonormal basis of eigenvectors, completing the proof. □

**Remark 1.8.** This theorem fails for general bounded self-adjoint operators that are not compact. For example, the identity operator on an infinite-dimensional Hilbert space is self-adjoint but has no eigenvalues.

# 2 Sobolev Spaces and Elliptic PDEs

## 2.1 Sobolev Spaces

Sobolev spaces generalize the idea of differentiability by allowing us to work with functions whose derivatives may exist only in a weak (distributional) sense. They provide a natural setting for studying partial differential equations, variational problems, and modern mathematical physics.

## 2.2 Geometric meaning

Intuitively, a Sobolev space $W^{1,p}(\Omega)$ can be thought of as the set of functions whose graphs are *not too rough*. The parameter $p$ measures *how much oscillation* or *how big the values* of the function and its derivatives can be. For example

- Large $p$ values penalize large peaks strongly, so functions are more regular.
- Smaller $p$ allows more variation, but still controls the average size of the derivatives.

In geometric terms, $W^{1,p}(\Omega)$ contains functions whose slopes are $p$-integrable.

## 2.3 Definition and Basic Properties

**Definition 2.1** ($C_c^\infty(\Omega)$ or $C_0^\infty(\Omega)$)**.** Let $\Omega \subset \mathbb{R}^N$ be an open set. We denote by $C_c^\infty(\Omega)$ or by $C_0^\infty(\Omega)$ the set of all functions

$$C_c^\infty(\Omega) = \{\varphi : \Omega \to \mathbb{R} \mid \varphi \text{ is infinitely differentiable and has compact support in } \Omega\}.$$

That is:

- $\varphi \in C^\infty(\Omega)$ (all partial derivatives of all orders exist and are continuous);
- $\operatorname{supp}(\varphi)$, the closure of the set where $\varphi \neq 0$, is compact and contained in $\Omega$.

Intuitively, $C_c^\infty(\Omega)$ consists of *smooth bump functions* that vanish outside some bounded region strictly inside $\Omega$. They are used as *test functions* in distribution theory and in the definition of weak derivatives.

**Example 2.1.** If $\Omega = (-1, 1)$, the function

$$\varphi(x) = \begin{cases} e^{-\frac{1}{1-x^2}}, & |x| < 1, \\ 0, & |x| \geq 1, \end{cases}$$

belongs to $C_c^\infty(-1, 1)$: it is smooth and its support is the closed interval $[-1, 1]$.

**Definition 2.2** (Weak derivative)**.** Let $u \in L^1_{\text{loc}}(\Omega)$ and $1 \leq i \leq N$. We say that $u$ has a *weak partial derivative* $\frac{\partial u}{\partial x_i} = g \in L^1_{\text{loc}}(\Omega)$ if

$$\int_\Omega u \frac{\partial \varphi}{\partial x_i}\, dx = -\int_\Omega g\, \varphi\, dx \quad \text{for all } \varphi \in C_c^\infty(\Omega).$$

In this case, $g$ is called the *weak derivative* of $u$ with respect to $x_i$.

**Remark 2.1.** This definition generalizes the classical derivative:

- If $u$ is differentiable in the classical sense, its classical derivative coincides with the weak derivative.
- The weak derivative can exist even if $u$ is not differentiable pointwise (e.g., functions with corners or cusps).

**Example 2.2** (Absolute value)**.** Consider $u(x) = |x|$ on $\Omega = (-1, 1)$. Classically, $u$ is not differentiable at $x = 0$. However, in the weak sense:

$$u'(x) = \begin{cases} 1, & x > 0, \\ -1, & x < 0, \end{cases}$$

and this $u'$ satisfies the weak derivative identity

$$\int_{-1}^{1} |x| \, \varphi'(x) \, dx = -\int_{-1}^{1} \operatorname{sign}(x) \, \varphi(x) \, dx \quad \forall \varphi \in C_c^\infty(-1, 1).$$

**Definition 2.3.** Let $1 \le p \le \infty$ and $\Omega \subset \mathbb{R}^N$ open. We define the *first-order Sobolev space*

$$W^{1,p}(\Omega) = \left\{ u \in L^p(\Omega) \; : \; \frac{\partial u}{\partial x_i} \text{ exists in the weak sense and } \frac{\partial u}{\partial x_i} \in L^p(\Omega), \quad \forall \, 1 \le i \le N \right\}.$$

**Remark 2.2.** Functions equal almost everywhere are identified as the same element in $W^{1,p}(\Omega)$. When $p = 2$, we write $H^1(\Omega) = W^{1,2}(\Omega)$. The letter $H$ stands for Hilbert, since $H^1(\Omega)$ is a Hilbert space.

**Remark 2.3.** If $u : \Omega \to \mathbb{R}$ is differentiable a.e., then

$$u \in W^{1,p}(\Omega) \qquad \Longleftrightarrow \qquad u \in L^p(\Omega) \text{ and } |\nabla u| \in L^p(\Omega).$$

Here $|\nabla u|$ is the Euclidean length of the gradient vector.

**Example 2.3** (Singular power function)**.** Let $\alpha > 0$, $\Omega = B_1(0) \subset \mathbb{R}^N$ and $u(x) = |x|^{-\alpha}$.
From $u \in L^p(\Omega)$ we get $0 < p\alpha < N$. Also $u \in C^1(B_1(0) \setminus \{0\})$ and

$$\frac{\partial u}{\partial x_i}(x) = -\alpha \frac{x_i}{|x|^{\alpha+2}}.$$

We require $(\alpha + 1)p < N$.

$$u \in W^{1,p}(\Omega) \iff 1 \le p < \frac{N}{1 + \alpha}.$$

**Example 2.4** (Function in $L^p$ but not in $W^{1,p}$)**.** Let $\Omega = (0, 1)$ and $u(x) = \sqrt{x}$. Then $u \in L^2(0, 1)$, but $u'(x) = \frac{1}{2\sqrt{x}} \notin L^2(0, 1)$, so $u \notin W^{1,2}(0, 1)$.

**Example 2.5** (Smooth compactly supported functions)**.** If $\varphi \in C_c^\infty(\Omega)$, then $\varphi \in W^{k,p}(\Omega)$ for all $k \ge 0$, $1 \le p \le \infty$, because all derivatives are smooth and bounded.

**Theorem 2.1** (Banach/Hilbert structure)**.** For each $1 \le p \le \infty$, $W^{1,p}(\Omega)$ is a Banach space with the norm

$$\|u\|_{W^{1,p}(\Omega)} = \begin{cases} \left( \int_\Omega |u|^p \, dx + \sum_{i=1}^N \int_\Omega \left| \frac{\partial u}{\partial x_i} \right|^p dx \right)^{1/p}, & 1 \le p < \infty, \\ \operatorname{ess\,sup}_\Omega |u| + \sum_{i=1}^N \operatorname{ess\,sup}_\Omega \left| \frac{\partial u}{\partial x_i} \right|, & p = \infty. \end{cases}$$

When $p = 2$, $H^1(\Omega)$ is a Hilbert space with the inner product

$$\langle u, v \rangle_{H^1(\Omega)} = \int_\Omega uv \, dx + \sum_{i=1}^N \int_\Omega \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_i} \, dx.$$

## 2.4 Density and mollifiers

Let $\rho \in C_c^\infty(\mathbb{R}^N)$ be a non-negative function satisfying:

- $\operatorname{supp} \rho \subset B(0,1)$,
- $\int_{\mathbb{R}^N} \rho(x) \, dx = 1$.

For $\varepsilon > 0$, define the scaled mollifier:

$$\rho_\varepsilon(x) := \frac{1}{\varepsilon^N} \rho\left(\frac{x}{\varepsilon}\right).$$

**Convolution with a mollifier:** For $u \in L^1_{\text{loc}}(\mathbb{R}^N)$, define the mollification:

$$u_\varepsilon(x) := (\rho_\varepsilon * u)(x) = \int_{\mathbb{R}^N} \rho_\varepsilon(x-y) u(y) \, dy.$$

Let $u \in L^p(\mathbb{R}^N)$ for some $1 \le p < \infty$. Then:

1. $u_\varepsilon \in C^\infty(\mathbb{R}^N)$ for all $\varepsilon > 0$.
2. $u_\varepsilon \to u$ in $L^p(\mathbb{R}^N)$ as $\varepsilon \to 0$.
3. If $u \in W^{1,p}(\mathbb{R}^N)$, then:

$$u_\varepsilon \in C^\infty(\mathbb{R}^N) \cap W^{1,p}(\mathbb{R}^N), \quad \text{and} \quad \nabla u_\varepsilon = \rho_\varepsilon * \nabla u.$$

4. $\|u_\varepsilon - u\|_{W^{1,p}(\mathbb{R}^N)} \to 0$ as $\varepsilon \to 0$.

**Theorem 2.2.** Let $1 \le p < \infty$. Then the space $C_c^\infty(\mathbb{R}^N)$ is dense in $W^{1,p}(\mathbb{R}^N)$. That is,

$$\forall u \in W^{1,p}(\mathbb{R}^N), \ \exists \{u_k\} \subset C_c^\infty(\mathbb{R}^N) \text{ such that } \|u_k - u\|_{W^{1,p}(\mathbb{R}^N)} \to 0.$$

*Proof.* Let $u \in W^{1,p}(\mathbb{R}^N)$. The proof proceeds in two steps:
Define a cut-off function $\chi_k \in C_c^\infty(\mathbb{R}^N)$ such that:

- $\chi_k(x) = 1$ for $|x| \le k$,

- $\chi_k(x) = 0$ for $|x| \geq 2k$,
- $|\nabla \chi_k(x)| \leq \frac{C}{k}$.

Set $u_k := \chi_k u \in W^{1,p}(\mathbb{R}^N)$, with compact support.

Then $u_k \to u$ in $W^{1,p}(\mathbb{R}^N)$ as $k \to \infty$.

Let $\rho_\varepsilon$ be a mollifier. Define:

$$u_{k,\varepsilon} := \rho_\varepsilon * u_k \in C_c^\infty(\mathbb{R}^N).$$

Then:

$$\|u_{k,\varepsilon} - u_k\|_{W^{1,p}(\mathbb{R}^N)} \to 0 \quad \text{as } \varepsilon \to 0.$$

Hence, the sequence $u_{k,\varepsilon} \in C_c^\infty(\mathbb{R}^N)$ approximates $u$ in $W^{1,p}(\mathbb{R}^N)$.

Combining the cut-off and mollification steps, we conclude that any function $u \in W^{1,p}(\mathbb{R}^N)$ can be approximated arbitrarily well in the Sobolev norm by smooth, compactly supported functions. Thus,

$$C_c^\infty(\mathbb{R}^N) \text{ is dense in } W^{1,p}(\mathbb{R}^N).$$

□

## 2.5 Higher-order Sobolev spaces

**Definition 2.4** (4.14)**.** Let $k \in \mathbb{N}$, $1 \leq p \leq \infty$, $\Omega \subset \mathbb{R}^N$ open. The *higher-order Sobolev space* $W^{k,p}(\Omega)$ is

$$W^{k,p}(\Omega) = \{u \in L^p(\Omega) : D^\alpha u \text{ exists in the weak sense and } D^\alpha u \in L^p(\Omega)\ \forall\, |\alpha| \leq k\},$$

where $\alpha$ is a multi-index and

$$D^\alpha u = \frac{\partial^{|\alpha|} u}{\partial x_1^{\alpha_1} \dots \partial x_N^{\alpha_N}}.$$

Norm:

$$\|u\|_{W^{k,p}(\Omega)} = \begin{cases} \left( \sum_{|\alpha| \leq k} \int_\Omega |D^\alpha u|^p \, dx \right)^{1/p}, & 1 \leq p < \infty, \\ \sum_{|\alpha| \leq k} \operatorname{ess\,sup}_\Omega |D^\alpha u|, & p = \infty. \end{cases}$$

**Remark 2.4.** For $p = 2$, $W^{k,2}(\Omega)$ is denoted $H^k(\Omega)$ and is a Hilbert space.

**Definition 2.5.** Let $\Omega \subset \mathbb{R}^n$ be an open set. The Sobolev space $W^{k,p}(\Omega)$ for $k \in \mathbb{N}$ and $1 \leq p \leq \infty$ is defined as the space of functions $u \in L^p(\Omega)$ whose weak derivatives up to order $k$ also belong to $L^p(\Omega)$:

$$W^{k,p}(\Omega) = \{u \in L^p(\Omega) : D^\alpha u \in L^p(\Omega), \text{ for all multi-indices } |\alpha| \leq k\}.$$

The space $H^k(\Omega) := W^{k,2}(\Omega)$ is a Hilbert space with the inner product:

$$\langle u, v\rangle_{H^k} = \sum_{|\alpha|\le k} \int_\Omega D^\alpha u \cdot D^\alpha v \, dx,$$

which induces the norm:

$$\|u\|_{H^k} = \left( \sum_{|\alpha|\le k} \|D^\alpha u\|^2_{L^2(\Omega)} \right)^{1/2}.$$

**Example 2.6.** Let $u(x) = |x|$ on $\Omega = (-1, 1)$. Then $u \in H^1((-1, 1))$ because $u \in L^2$ and its weak derivative $u' = \operatorname{sign}(x)$ is also in $L^2$.

## 2.6 Sobolev Embeddings and Poincaré inequality

Let $\Omega \subset \mathbb{R}^n$ be a bounded Lipschitz domain.

**Theorem 2.3** (Sobolev Embedding Theorem)**.** The following holds:

- If $kp < n$, then $W^{k,p}(\Omega) \hookrightarrow L^q(\Omega)$ for $q$ such that $p \le q \le \frac{np}{n-kp}$.
- If $kp = n$, then $W^{k,p}(\Omega) \hookrightarrow L^q(\Omega)$ for all $q < \infty$.
- If $kp > n$, then $W^{k,p}(\Omega) \hookrightarrow C^0(\bar{\Omega})$.

The embedding is compact if $q < \frac{np}{n-kp}$.

**Theorem 2.4** (Poincaré–Wirtinger)**.** Let $1 \le p < \infty$. Define the integral average

$$u_\Omega := \frac{1}{|\Omega|} \int_\Omega u(x)\, dx \qquad (u \in L^1(\Omega)).$$

Then there exists a constant $C_{PW} = C_{PW}(\Omega, p) > 0$ such that for all $u \in W^{1,p}(\Omega)$

$$\|u - u_\Omega\|_{L^p(\Omega)} \;\le\; C_{PW}\, \|\nabla u\|_{L^p(\Omega)}.$$

## 2.7 Weak Solutions and Critical Points

**Definition 2.6.** A function $u \in H^1_0(\Omega)$ is a *weak solution* of the elliptic equation $-\Delta u = f$ in $\Omega$ if:

$$\int_\Omega \nabla u \cdot \nabla v \, dx = \int_\Omega f v \, dx \quad \text{for all } v \in H^1_0(\Omega).$$

This corresponds to finding critical points of the energy functional:

$$E(u) = \frac{1}{2} \int_\Omega |\nabla u|^2 dx - \int_\Omega f u \, dx.$$

## 2.8 Spectrum of the Laplace Operator

Let $\Omega \subset \mathbb{R}^n$ be a bounded Lipschitz domain. Consider the Dirichlet Laplacian eigenvalue problem:

$$-\Delta u = \lambda u \quad \text{in } \Omega, \qquad u = 0 \quad \text{on } \partial\Omega.$$

We aim to prove that the spectrum consists of a discrete set of positive eigenvalues:

$$0 < \lambda_1 < \lambda_2 \leq \lambda_3 \leq \cdots \to \infty,$$

with corresponding eigenfunctions forming an orthonormal basis in $L^2(\Omega)$.

Define the bilinear form:

$$a(u, v) = \int_\Omega \nabla u \cdot \nabla v \, dx,$$

on the Hilbert space $H_0^1(\Omega)$. The eigenvalue problem becomes: find $u \in H_0^1(\Omega)$, $u \neq 0$, and $\lambda \in \mathbb{R}$ such that

$$a(u, v) = \lambda \int_\Omega uv \, dx \quad \forall v \in H_0^1(\Omega).$$

Define the operator $T : L^2(\Omega) \to L^2(\Omega)$ as follows: for $f \in L^2(\Omega)$, let $u \in H_0^1(\Omega)$ be the unique weak solution of

$$-\Delta u = f \quad \text{in } \Omega, \qquad u = 0 \quad \text{on } \partial\Omega,$$

and set $T(f) = u$.

We verify the following:

- **Boundedness:** $T$ is a bounded operator from $L^2(\Omega)$ into $H_0^1(\Omega)$, hence into $L^2(\Omega)$.
- **Self-adjointness:** For $f, g \in L^2(\Omega)$, with $u = T(f)$ and $v = T(g)$, we have:

  $$\langle T(f), g \rangle = \int_\Omega ug \, dx = \int_\Omega \nabla u \cdot \nabla v \, dx = \int_\Omega fv \, dx = \langle f, T(g) \rangle.$$

  So $T$ is self-adjoint.
- **Compactness:** The inclusion $H_0^1(\Omega) \hookrightarrow L^2(\Omega)$ is compact (Rellich–Kondrachov theorem). Since $T$ maps bounded subsets of $L^2$ into bounded subsets of $H_0^1(\Omega)$, and then into $L^2$, $T$ is a compact operator on $L^2(\Omega)$.

By the spectral theorem for compact self-adjoint operators on Hilbert spaces, the operator $T$ has a discrete spectrum $\{\mu_k\} \subset \mathbb{R}$ with $\mu_k \to 0$, and the corresponding eigenfunctions $\{\phi_k\}$ form an orthonormal basis in $L^2(\Omega)$. Note that $m := \inf_{\|u\|=1} \langle T(u), u \rangle > 0$ and $\sigma(T) \subset [m, \infty)$.

From $T(\phi_k) = \mu_k \phi_k$, we have

$$-\Delta \phi_k = \frac{1}{\mu_k} \phi_k = \lambda_k \phi_k.$$

Therefore, the eigenvalues of the Laplacian are $\lambda_k = \frac{1}{\mu_k} \to \infty$ and they are positive.

**Remark 2.5.** One can show that $\lambda_1$ is simple, and by the strong maximum principle for elliptic equations, either $u > 0$ in $\Omega$ or $u < 0$ in $\Omega$.

The Dirichlet Laplacian has a countable set of positive eigenvalues $\lambda_k > 0$ with finite multiplicities, satisfying $\lambda_k \to \infty$, and the corresponding eigenfunctions form an orthonormal basis of $L^2(\Omega)$. Hence, the spectrum is discrete.

### 2.8.1 Dirichlet eigenvalue problem on $(0, \pi)$

Consider the Dirichlet eigenvalue problem on $\Omega = (0, \pi)$:

$$\begin{cases} -u''(x) = \lambda u(x), & x \in (0, \pi), \\ u(0) = u(\pi) = 0. \end{cases}$$

We seek nontrivial solutions $u(x)$ such that $-u''(x) = \lambda u(x)$ and $u(0) = u(\pi) = 0$. Assume $u(x) = \sin(kx)$, then:

$$u''(x) = -k^2 \sin(kx) \Rightarrow -u'' = k^2 \sin(kx) = \lambda u(x) \Rightarrow \lambda = k^2.$$

Boundary condition $u(\pi) = \sin(k\pi) = 0$ implies $k = n \in \mathbb{N}$. The eigenvalues are

$$\lambda_n = n^2, \quad n = 1, 2, 3, \ldots$$

The corresponding eigenfunctions are:

$$\phi_n(x) = \sin(nx).$$

These form an orthogonal basis of $H_0^1(0, \pi)$. The spectrum is discrete, real, and unbounded above.

## 2.9 Variational Characterization of the First Eigenvalue

Let $\Omega$ be a bounded Lipschitz domain.

**Theorem 2.5.** The first eigenvalue $\lambda_1$ of the Dirichlet Laplacian satisfies:

$$\lambda_1 = \inf_{u \in H_0^1(\Omega) \setminus \{0\}} \frac{\int_\Omega |\nabla u|^2 dx}{\int_\Omega u^2 dx}.$$

*Proof.* Let $\mathcal{R}(u) = \frac{\int_\Omega |\nabla u|^2 dx}{\int_\Omega u^2 dx}$ be the Rayleigh quotient. Minimizing $\mathcal{R}(u)$ over $H_0^1(\Omega) \setminus \{0\}$ yields $\lambda_1$, and the minimizer solves $-\Delta u = \lambda_1 u$. □

**Remark 2.6.** The minimizer can be chosen positive and is unique up to a constant multiple.

In particular any norm

$$\|u\|_\lambda := \left( \int_\Omega |\nabla u|^2 + \lambda |u|^2 \, dx \right)^{1/2}, \quad \text{for } \lambda > -\lambda_1$$

is equivalent in $H_0^1(\Omega)$.

## 2.10 From the Nonlinear Schrödinger Equation to an Elliptic PDE

The nonlinear Schrödinger equation (NLS) arises as an effective model in various physical contexts:

- **Nonlinear optics:** In a Kerr medium, the electric field envelope $u(x,t)$ satisfies a paraxial wave equation that, under suitable scaling, reduces to the cubic NLS. The cubic term $\mp b|u|^2 u$ models the intensity-dependent refractive index; the "$-$" sign corresponds to self-focusing (formation of spatial or temporal solitons), while the "$+$" sign describes self-defocusing.

- **Bose–Einstein condensates (BECs):** At ultra-low temperatures, the macroscopic wavefunction of a dilute Bose gas is governed by the Gross–Pitaevskii equation

  $$i\hbar\,\partial_t\psi = -\frac{\hbar^2}{2m}\Delta\psi + V_{\mathrm{trap}}(x)\psi + g|\psi|^2\psi,$$

  where $g$ is proportional to the $s$-wave scattering length. In nondimensional form, this is the cubic NLS (2). The sign of $g$ determines whether the atomic interactions are repulsive (defocusing) or attractive (focusing). *Nobel prize:* The first creation of a BEC in dilute gases (Eric A. Cornell, Wolfgang Ketterle, and Carl E. Wieman) was awarded the 2001 Nobel Prize in Physics, with the Gross–Pitaevskii model playing a central theoretical role.

*Nonlinear Schrödinger equation* We consider the cubic NLS with real parameter $b > 0$:

$$i\,\partial_t u = -\Delta u + V(x)u \;\mp\; b\,|u|^2 u, \qquad x \in \mathbb{R}^n,\ t > 0. \tag{2}$$

The *focusing* case corresponds to the "$-$" sign and the *defocusing* case to "$+$".

*Standing wave ansatz* Let

$$u(x,t) = e^{-i\omega t}\,\phi(x), \quad \omega > 0.$$

Substituting into (2) yields the semilinear elliptic PDE

$$-\Delta\phi + \Big(V(x) + \omega\Big)\,\phi \;\mp\; b\,|\phi|^2\phi = 0.$$

$$\operatorname{sech}(x) = \frac{1}{\cosh(x)} = \frac{2}{e^x + e^{-x}}, \quad \text{where} \quad \cosh(x) = \frac{e^x + e^{-x}}{2}.$$

*Typical boundary conditions*

- Whole space: $x \in \mathbb{R}^n$, $\phi \in H^1(\mathbb{R}^n)$, $\phi(x) \to 0$ as $|x| \to \infty$.

- Bounded domain $\Omega$: Dirichlet $\phi|_{\partial\Omega} = 0$, or Neumann/periodic.

*1D explicit solution (focusing case)* In one spatial dimension, the solution to

$$-\phi'' + \omega\phi - b\,\phi^3 = 0$$

is given by the solitary wave

$$\phi(x) = \sqrt{\frac{2\omega}{b}}\;\operatorname{sech}\Big(\sqrt{\omega}\,x\Big).$$

## 2.11 Direct Method in the Calculus of Variations

**Theorem 2.6.** Let $X$ be a reflexive Banach space and $K \subset X$ be nonempty, closed, and weakly sequentially closed. Suppose $J : K \to \mathbb{R} \cup \{+\infty\}$ is

- *coercive*: $\|u\| \to \infty$ implies $J(u) \to +\infty$,
- *weakly lower semicontinuous (w.l.s.c.)*: if $u_n \rightharpoonup u$ in $X$, then $J(u) \le \liminf_{n\to\infty} J(u_n)$.

Then $J$ attains its minimum on $K$, i.e., there exists $u^* \in K$ with $J(u^*) = \inf_{u\in K} J(u)$.

*Proof.* Let $m := \inf_{u\in K} J(u)$ and pick a minimizing sequence $(u_n) \subset K$ such that $J(u_n) \downarrow m$. Coercivity implies $(u_n)$ is bounded in $X$. Since $X$ is reflexive, there exists a subsequence (not relabeled) and $u^* \in X$ such that $u_n \rightharpoonup u^*$ in $X$. Because $K$ is weakly sequentially closed, $u^* \in K$. By weak lower semicontinuity,

$$J(u^*) \le \liminf_{n\to\infty} J(u_n) = m.$$

Hence $J(u^*) = m$ and $u^*$ is a minimizer. □

Common sufficient conditions for w.l.s.c. are convexity of $J$ and continuity.

**Theorem 2.7.** Let $X$ be a Banach space and let $J : X \to (-\infty, +\infty]$ be convex and continuous. Then $J$ is *weakly lower semicontinuous* (w.l.s.c.), i.e.,

$$u_n \rightharpoonup u \text{ in } X \quad \Longrightarrow \quad J(u) \le \liminf_{n\to\infty} J(u_n).$$

*Proof.* For $\alpha \in \mathbb{R}$ consider the sublevel set

$$A_\alpha := \{\, u \in X : \; J(u) \le \alpha \,\}.$$

Since $J$ is convex, each $A_\alpha$ is convex; since $J$ is (strongly) continuous, each $A_\alpha$ is also closed in the norm topology. We now recall a classical geometric fact (Mazur's theorem): for convex subsets of a Banach space, the weak closure coincides with the norm closure; in particular, if a convex set is strongly closed, it is also weakly closed. Consequently, each $A_\alpha$ is *weakly closed*. By the definition of weak lower semicontinuity, a functional $J$ is w.l.s.c. if and only if all its sublevel sets $A_\alpha$ are weakly closed. Therefore $J$ is w.l.s.c. □

**Example 2.7.** Let $\Omega \subset \mathbb{R}^d$ be bounded with Lipschitz boundary and $f \in H^{-1}(\Omega)$. Consider

$$J(u) = \int_\Omega \left(|\nabla u|^2 + |u|^2\right) dx \; - \; 2\langle f, u\rangle_{H^{-1},H_0^1}, \qquad u \in H_0^1(\Omega).$$

Then $J$ has a unique minimizer $u^* \in H_0^1(\Omega)$. Moreover $u^*$ is a unique weak solution to

$$-\Delta u + u = f, \quad \text{in } H_0^1(\Omega).$$

*Proof.* Work in the reflexive space $X = H^1_0(\Omega)$ with the norm $\|u\|^2_{H^1_0} = \int_\Omega |\nabla u|^2\, dx$ (equivalent to the standard norm by Poincaré). For coercivity, by Cauchy–Schwarz and the Riesz representation of $H^{-1}$,

$$J(u) \geq \|u\|^2_{H^1_0} - 2\|f\|_{H^{-1}}\|u\|_{H^1_0} \;\geq\; \tfrac{1}{2}\|u\|^2_{H^1_0} - C\|f\|^2_{H^{-1}},$$

so $J(u) \to +\infty$ as $\|u\|_{H^1_0} \to \infty$. For w.l.s.c., observe that the quadratic form $u \mapsto \int_\Omega (|\nabla u|^2 + |u|^2)\, dx$ is convex and w.l.s.c. on $H^1_0(\Omega)$, while the linear term $u \mapsto -2\langle f, u\rangle$ is weakly continuous. Hence $J$ is w.l.s.c. By the Direct Method Theorem of Calculus of Variations, a minimizer $u^*$ exists. Moreover, $J$ is strictly convex (sum of a strictly convex quadratic form and a linear functional), so the minimizer is unique. □

### 2.11.1 Dirichlet problem with a defocuscing nonlinearity

Let $\Omega \subset \mathbb{R}^N$ be bounded with Lipschitz boundary, and let

$$2 < p < 2^* = \frac{2N}{N-2} \quad \text{(with the usual convention } 2^* = \infty \text{ if } N \leq 2).$$

Work in $H^1_0(\Omega)$ with the equivalent norm

$$\|u\|^2_{H^1_0} = \int_\Omega |\nabla u|^2\, dx.$$

Consider the energy functional

$$J(u) = \int_\Omega \left( \frac{1}{2}|\nabla u|^2 + \frac{\lambda}{2}|u|^2 + \frac{1}{p}|u|^p \right) dx, \quad u \in H^1_0(\Omega).$$

Critical points of $J$ solve

$$-\Delta u + \lambda u + |u|^{p-2}u = 0 \quad \text{in } \Omega, \qquad u = 0 \text{ on } \partial\Omega,$$

i.e. the are *weak solutions* to $-\Delta u + \lambda u = -|u|^{p-2}u$.

**Theorem 2.8.** Assume $2 < p < 2^*$. Then $J$ attains its minimum on $H^1_0(\Omega)$. Let $\lambda_1 = \lambda_1(\Omega)$ be the first Dirichlet eigenvalue of $-\Delta$.

- If $\lambda \geq -\lambda_1$, then the unique minimizer is $u^* \equiv 0$, hence the unique weak solution is $u = 0$.
- If $\lambda < -\lambda_1$, then there exists a nontrivial minimizer $u^* \not\equiv 0$. Any minimizer is a weak solution of the PDE.

*Proof. Coercivity.* By Sobolev embedding $H^1_0(\Omega) \hookrightarrow L^p(\Omega)$, we have $\|u\|_{L^p} \leq C\|u\|_{H^1_0}$. By Poincaré's inequality,

$$\int_\Omega |u|^2\, dx \leq \lambda_1^{-1} \int_\Omega |\nabla u|^2\, dx.$$

Hence

$$J(u) \;\geq\; \frac{1}{2}\|\nabla u\|^2_{L^2} + \frac{\lambda}{2\lambda_1}\|\nabla u\|^2_{L^2} + \frac{1}{p}\|u\|^p_{L^p} \;\geq\; c_1\|\nabla u\|^2_{L^2} - c_2,$$

for some $c_1 > 0$ (use Young's inequality if $\lambda < 0$). Thus $J(u) \to +\infty$ when $\|u\|_{H_0^1} \to \infty$.

*Weak lower semicontinuity.* The maps $u \mapsto \int |\nabla u|^2$ and $u \mapsto \int |u|^p$ are convex and continuous, hence weakly lower semicontinuous. With the compact embedding $H_0^1 \hookrightarrow L^2$, the quadratic term is weakly continuous. Therefore $J$ is weakly l.s.c.

*Existence.* Take a minimizing sequence, use reflexivity to extract $u_n \rightharpoonup u^*$ in $H_0^1$, then weak l.s.c. yields $J(u^*) = \inf J$.

*Characterization by $\lambda_1$.* We can write

$$J(u) = \frac{1}{2}\int_\Omega \Big(|\nabla u|^2 + \lambda |u|^2\Big) dx + \frac{1}{p}\int_\Omega |u|^p \, dx.$$

By Poincaré,

$$\int_\Omega (|\nabla u|^2 + \lambda |u|^2) \, dx \geq (\lambda_1 + \lambda) \int_\Omega |u|^2 \, dx.$$

If $\lambda \geq -\lambda_1$, then both the quadratic and the $p$-power terms are nonnegative, so $J(u) \geq 0$ with equality only at $u = 0$. Hence $u^* = 0$ is the unique minimizer/solution.

If $\lambda < -\lambda_1$, pick the first eigenfunction $\phi_1 > 0$ with $\|\phi_1\|_{L^2} = 1$. For small $t > 0$,

$$J(t\phi_1) = \frac{t^2}{2}(\lambda_1 + \lambda) + \frac{t^p}{p}\|\phi_1\|_{L^p}^p < 0,$$

since $\lambda_1 + \lambda < 0$ and $p > 2$. Thus $\inf J < 0 = J(0)$, so any minimizer satisfies $u^* \not\equiv 0$. □

## 2.12 Minimax Methods and the Mountain Pass Theorem

**Theorem 2.9** (Mountain Pass Theorem)**.** Let $X$ be a Banach space and $J \in C^1(X, \mathbb{R})$ satisfy:

- $J(0) = 0$,
- There exist $\rho, \alpha > 0$ such that $J(u) \geq \alpha$ for $\|u\| = \rho$,
- There exists $v$ with $\|v\| > \rho$ such that $J(v) < 0$.

If $J$ satisfies the *Palais–Smale condition* (defined below), then $J$ has a critical point $u \neq 0$ at level

$$c = \inf_{\gamma \in \Gamma} \sup_{t \in [0,1]} E(\gamma(t)),$$

where $\Gamma = \{\gamma \in C([0,1], X) : \gamma(0) = 0, \ \gamma(1) = v\}$.

**Definition 2.7** (Palais–Smale condition)**.** A functional $J \in C^1(X, \mathbb{R})$ is said to satisfy the *Palais–Smale (PS) condition* if every sequence $(u_n) \subset X$ such that $(E(u_n))$ is bounded and $J'(u_n) \to 0$ in $X^*$ contains a convergent subsequence in $X$.

**Illustration: Mountain Pass Geometry.**

### 2.12.1 Dirichlet problem with a focuscing nonlinearity

Let us consider

$$J(u) = \int_\Omega \left( \frac{1}{2}|\nabla u|^2 + \frac{\lambda}{2}|u|^2 - \frac{1}{p}|u|^p \right) dx, \quad u \in H_0^1(\Omega).$$

**Lemma 2.1** (Positivity of the mountain pass level)**.** Let $2 < p < 2^*$ and $\lambda > -\lambda_1$. Define

$$\Gamma = \{\gamma \in C([0,1], H_0^1(\Omega)) : \ \gamma(0) = 0, \ J(\gamma(1)) < 0\}, \qquad c = \inf_{\gamma \in \Gamma} \sup_{t \in [0,1]} J(\gamma(t)).$$

Then $c > 0$.

*Proof.* By Sobolev embedding and Poincaré, for small $\|u\|_{H_0^1}$ we have

$$J(u) \ \geq \ \frac{1}{2}(\lambda_1 + \lambda)\|u\|_{L^2}^2 - \frac{C}{p}\|u\|_{H_0^1}^p \ \geq \ \alpha > 0 \quad \text{whenever} \quad \|u\|_{H_0^1} = \rho$$

for some $\rho, \alpha > 0$. Any $\gamma \in \Gamma$ must cross the sphere $\|u\|_{H_0^1} = \rho$; hence $\sup_t J(\gamma(t)) \geq \alpha$ and taking the infimum over $\Gamma$ yields $c \geq \alpha > 0$. □

**Lemma 2.2** (Boundedness of $(\mathrm{PS})_c$ sequences)**.** Let $(u_n) \subset H_0^1(\Omega)$ satisfy $J(u_n) \to c$ and $J'(u_n) \to 0$ in $H^{-1}(\Omega)$. Then $(u_n)$ is bounded in $H_0^1(\Omega)$.

*Proof.* Compute, for each $u \in H_0^1(\Omega)$,

$$\langle J'(u), u\rangle = \int_\Omega \Big(|\nabla u|^2 + \lambda|u|^2 - |u|^p\Big)\,dx,$$

and

$$J(u) - \frac{1}{p}\langle J'(u), u\rangle = \frac{p-2}{2p}\int_\Omega \Big(|\nabla u|^2 + \lambda|u|^2\Big)\,dx. \qquad (*)$$

Apply $(*)$ with $u = u_n$:

$$\frac{p-2}{2p}\int_\Omega \Big(|\nabla u_n|^2 + \lambda|u_n|^2\Big)\,dx = J(u_n) - \frac{1}{p}\langle J'(u_n), u_n\rangle = c + o(1) + o(1)\,\|u_n\|_{H_0^1}.$$

Using Poincaré,

$$\int_\Omega \Big(|\nabla u_n|^2 + \lambda|u_n|^2\Big)\,dx \geq (\lambda_1+\lambda)\|u_n\|_{L^2}^2,$$

with $\lambda_1 + \lambda > 0$. Hence, for some constants $C_1, C_2 > 0$

$$\|u_n\|_{H_0^1}^2 \;\leq C_1 \;\int_\Omega \Big(|\nabla u_n|^2 + \lambda|u_n|^2\Big)\,dx \;\leq\; C_2\Big(1 + \|u_n\|_{H_0^1}\Big),$$

which implies $\sup_n \|u_n\|_{H_0^1} < \infty$. □

**Remark 2.7.** The identity *(∗)* is the standard trick that yields boundedness for any (PS) sequence of $J$ and works thanks to $p \in (2, 2^*)$ and $\lambda > -\lambda_1$. Combined with Rellich–Kondrachov, this also gives (after extracting a subsequence) $u_n \rightharpoonup u$ in $H_0^1$ and $u_n \to u$ in $L^p$, which is the key step in verifying the Palais-Smale condition for $J$.

**Theorem 2.10** (Mountain pass solution for the focusing case)**.** Let $\Omega \subset \mathbb{R}^N$ be bounded with Lipschitz boundary and $2 < p < 2^*$. Assume $\lambda > -\lambda_1$. Then $J$ has a critical point $u \not\equiv 0$ obtained by the Mountain Pass Theorem.

*Proof. Mountain pass geometry.* By Poincaré, $\int(|\nabla u|^2 + \lambda|u|^2)\,dx \geq (\lambda_1+\lambda)\|u\|_{L^2}^2$. Hence for small $\|u\|_{H_0^1}$,

$$J(u) \geq \tfrac{1}{2}(\lambda_1+\lambda)\|u\|_{L^2}^2 - \tfrac{1}{p}\|u\|_{L^p}^p > 0,$$

so there exist $\rho, \alpha > 0$ with $J(u) \geq \alpha$ whenever $\|u\|_{H_0^1} = \rho$. On the other hand, for any fixed $w \in H_0^1(\Omega) \setminus \{0\}$, $J(tw) = \frac{t^2}{2}\int(|\nabla w|^2 + \lambda|w|^2) - \frac{t^p}{p}\|w\|_{L^p}^p \to -\infty$ as $t \to \infty$, so there exists $v$ with $J(v) < 0$. Thus the Mountain Pass geometry holds.

*Palais–Smale condition.* Let $(u_n)$ be a sequence with $J(u_n)$ bounded and $J'(u_n) \to 0$ in $H^{-1}$. In view of Lemma 2.2, we deduce that $(u_n)$ is bounded in $H_0^1(\Omega)$. By Rellich–Kondrachov, up to a subsequence $u_n \rightharpoonup u$ in $H_0^1$ and $u_n \to u$ in $L^p(\Omega)$ (since $p < 2^*$). Then the standard variational argument gives $u_n \to u$ in $H_0^1(\Omega)$, i.e. the *(PS)* condition holds.

□

## 2.13 Introduction to Spectral Theory of Schrödinger Operators

Consider the operator

$$H = -\Delta + V(x), \quad x \in \mathbb{R}^n,$$

where the potential $V : \mathbb{R}^n \to \mathbb{R}$ is periodic with respect to a lattice $\Gamma \subset \mathbb{R}^n$, i.e.

$$V(x+\gamma) = V(x) \quad \forall \gamma \in \Gamma.$$

For simplicity, take $\Gamma = \mathbb{Z}^n$.

By the Floquet–Bloch theory, the spectrum of $H$ is the union of disjoint intervals:

$$\sigma(H) = \bigcup_{n \geq 1} [a_k,\, b_k].$$

Each interval $[a_k,\, b_k]$ is called an *energy band*. The spectrum is purely absolutely continuous, and *spectral gaps* may occur between consecutive bands.

**Physical interpretation.** In solid-state physics, $V$ models the periodic potential of a crystal lattice. The band–gap structure explains the difference between conductors, semiconductors, and insulators.

## Spectrum of $-\Delta$ on $\mathbb{R}^N$

Let $-\Delta$ act on $L^2(\mathbb{R}^N)$ with domain $H^2(\mathbb{R}^N)$. Then

$$\sigma(-\Delta) = \sigma_{\mathrm{c}}(-\Delta) = [0, \infty).$$

In words: the spectrum is purely absolutely continuous, equal to the half–line $[0, \infty)$; there are no eigenvalues and no singular continuous spectrum.

## 2.14 Radial solutions in $\mathbb{R}^N$ via variational methods

Let $N \geq 2$, $2 < p < 2^* := \frac{2N}{N-2}$ (with the usual convention $2^* = \infty$ if $N = 2$). Assume $V : \mathbb{R}^N \to \mathbb{R}$ is measurable, radial, and

$$V(x) \geq V_0 > 0 \quad \text{for a.e. } x \in \mathbb{R}^N.$$

Work in the radial subspace

$$H^1_{\mathrm{rad}}(\mathbb{R}^N) := \{u \in H^1(\mathbb{R}^N) : u(x) = u(|x|)\}.$$

Consider the functional

$$J(u) = \frac{1}{2}\int_{\mathbb{R}^N} |\nabla u|^2 + V(x)|u|^2\, dx - \frac{1}{p}\int_{\mathbb{R}^N} |u|^p\, dx, \qquad u \in H^1_{\mathrm{rad}}(\mathbb{R}^N).$$

Critical points of $J$ are weak radial solutions of

$$-\Delta u + V(x)u = |u|^{p-2}u \quad \text{in } \mathbb{R}^N.$$

**Theorem 2.11.** Under the assumptions above, $J$ satisfies the Mountain Pass geometry and the Palais–Smale condition on $H^1_{\text{rad}}(\mathbb{R}^N)$. Consequently, there exists $u \in H^1_{\text{rad}}(\mathbb{R}^N) \setminus \{0\}$ with $J'(u) = 0$.

*Proof. Mountain Pass geometry.* For small $\|u\|_{H^1}$,

$$J(u) \geq \tfrac{1}{2}\int(|\nabla u|^2 + V_0|u|^2) - \tfrac{C}{p}\|u\|^p_{H^1} > 0,$$

so $\exists\, \rho, \alpha > 0$ with $J(u) \geq \alpha$ if $\|u\|_{H^1} = \rho$. Fix $w \in H^1_{\text{rad}} \setminus \{0\}$; then $J(tw) \to -\infty$ as $t \to \infty$, so $\exists\, v$ with $J(v) < 0$.

*Palais–Smale on the radial subspace.* Let $(u_n) \subset H^1_{\text{rad}}$ be a (PS) sequence: $J(u_n)$ bounded, $J'(u_n) \to 0$. The identity

$$J(u_n) - \frac{1}{p}\langle J'(u_n), u_n\rangle = \frac{p-2}{2p}\int \left(|\nabla u_n|^2 + V(x)|u_n|^2\right)$$

gives boundedness of $(u_n)$ in $H^1$. By the *radial compact embedding* (Strauss lemma), $H^1_{\text{rad}}(\mathbb{R}^N) \hookrightarrow L^q(\mathbb{R}^N)$ is compact for every $2 < q < 2^*$. Hence, up to a subsequence, $u_n \rightharpoonup u$ in $H^1$ and $u_n \to u$ in $L^p$, which yields $J'(u) = 0$ and then $u_n \to u$ in $H^1$ by standard arguments. □

**Remark 2.8.**

- If $V$ is constant, all arguments above hold.
- The restriction to $H^1_{\text{rad}}$ restores compactness in $\mathbb{R}^N$; without radial symmetry one typically uses concentration–compactness (Lions).
- At the critical exponent $p = 2^*$ the compactness fails; existence requires additional structure (e.g. potentials $V$ with traps) or refined tools.

**Lemma 2.3** (Lions)**.** Let $(u_n)$ be a bounded sequence in $H^1(\mathbb{R}^N)$, $N \geq 2$. Assume that there exists $R > 0$ such that

$$\lim_{n\to\infty} \sup_{y\in\mathbb{R}^N} \int_{B_R(y)} |u_n(x)|^2\, dx = 0.$$

Then $u_n \to 0$ strongly in $L^q(\mathbb{R}^N)$ for every $q \in (2, 2^*)$.

**Theorem 2.12** (Existence via Mountain Pass for periodic $V$)**.** Let $V : \mathbb{R}^N \to \mathbb{R}$ be $\mathbb{Z}^N$–periodic and continuous. Assume that the Schrödinger operator

$$L := -\Delta + V(x)$$

on $L^2(\mathbb{R}^N)$ satisfies

$$0 < \inf \sigma(L).$$

Let $2 < p < 2^*$. Then the equation

$$-\Delta u + V(x)u = |u|^{p-2}u, \quad u \in H^1(\mathbb{R}^N),$$

admits a nontrivial weak solution obtained by the Mountain Pass Theorem.

For the interested reader, we refer to the relevant literature [1, 5, 11, 12] and welcome further inquiries, especially regarding the fascinating world of nonlinear phenomena associated with Maxwell and Schrödinger equations [4].

# 3 Problems

**Exercise 3.1.** The sequence space $\ell^p$ for $1 \le p \le \infty$ consists of all sequences $x = (x_n)_{n=1}^{\infty}$ of scalars such that:

$$\|x\|_p = \begin{cases} \left(\sum_{n=1}^{\infty} |x_n|^p\right)^{1/p} & \text{if } 1 \le p < \infty, \\ \sup_{n\in\mathbb{N}} |x_n| & \text{if } p = \infty, \end{cases}$$

is finite. Each $\ell^p$ space is a Banach space.

**Exercise 3.2.** The space $L^p([a,b])$ for $1 \le p \le \infty$ consists of (equivalence classes of) measurable functions $f : [a,b] \to \mathbb{R}$ (or $\mathbb{C}$) such that the $p$-th power of the absolute value is integrable:

$$\|f\|_p = \begin{cases} \left(\int_a^b |f(x)|^p \, dx\right)^{1/p} & \text{if } 1 \le p < \infty, \\ \text{ess sup}_{x\in[a,b]}|f(x)| & \text{if } p = \infty. \end{cases}$$

These spaces are Banach spaces.

**Exercise 3.3.** The space $C([a,b])$ of continuous real (or complex) functions on $[a,b]$ equipped with the *supremum norm*

$$\|f\|_\infty = \sup_{x\in[a,b]} |f(x)|$$

is also a Banach space.

**Exercise 3.4** (Non-complete normed space)**.** Let $c_{00}$ denote the space of sequences with only finitely many nonzero terms, equipped with the $\ell^p$ norm for some $1 \le p < \infty$. Then $(c_{00}, \|\cdot\|_p)$ is a normed space, but it is not complete — its completion is $\ell^p$.

**Exercise 3.5.** Consider the space of polynomials $P([0,1])$ with the sup norm. Is this a Banach space?

**Exercise 3.6.** The space of polynomials is dense in $C([0,1])$, but it is not complete, hence not a Banach space.

**Exercise 3.7.** Let $X = c$, the space of convergent sequences with the sup norm. Show that $X$ is a Banach space.

**Exercise 3.8.** Define $T : \ell^2 \to \ell^2$ by $T(x_1, x_2, x_3, ...) = (x_1, x_2/2, x_3/3, ...)$. Prove that $T$ is bounded and compute its norm.

**Exercise 3.9.** Let $T$ be defined on $L^2([0,1])$ by $(Tf)(x) = \int_0^x f(t)dt$. Show that $T$ is a bounded linear operator.

**Exercise 3.10.** Let $(X, \|\cdot\|)$ be a normed space. For every $x_0 \in X$ there exists a continuous functional $f_0 : X \to \mathbb{R}$ such that $\|f_0\| = \|x_0\|$ and $f_0(x_0) = \|x_0\|^2$.

**Exercise 3.11.** Let $c_0$ be the space of all sequences converging to zero, equipped with the supremum norm. Is $c_0$ a closed subspace of $\ell^\infty$? Is it a Banach space?

**Exercise 3.12.** Let $X = \mathbb{R}^2$ and $M = \text{span}((1,1))$. Define the quotient space $X/M$. Describe the equivalence classes and the geometry of this space.

**Exercise 3.13.** Let $X = \ell^2$ and let $M \subset X$ be the subspace consisting of all sequences with only the first coordinate possibly nonzero. Describe the quotient space $X/M$. Is it a Banach space?

**Exercise 3.14.** Find the dual space $(\ell^1)^*$. Prove that

$$(\ell^1)^* \cong \ell^\infty.$$

**Exercise 3.15.** Define the operator $T : \ell^2 \to \ell^2$ by

$$T((x_1, x_2, x_3, \dots)) = (0, x_1, x_2, \dots).$$

Is $T$ linear? Is it continuous? Is it invertible? Is it an open map?

**Exercise 3.16.** Define the operator $T : \ell^2 \to \ell^2$ by

$$T((x_1, x_2, x_3, \dots)) = (x_2, x_3, \dots).$$

Is $T$ linear? Is it continuous? Is it invertible? Is it an open map?

**Exercise 3.17.** Define a Banach space $D(T) \subset C[0,1]$ such that $T : D(T) \to C[0,1]$ defined by $T(f) = f'$ is well-defined. Is $T$ bounded?

**Exercise 3.18.** Define the operator $T : L^2[0,1] \to L^2[0,1]$ by

$$T(f)(x) = \int_0^x f(t)\,dt.$$

Is $T$ linear? Is it bounded? Is it invertible?

**Exercise 3.19.** Let $T : C[0,1] \to C[0,1]$ be defined by

$$T(f)(x) = \int_0^x f(t)\,dt.$$

Show that $T$ is not an open map.

**Exercise 3.20.** Let $X = Y = L^p[0,1]$, where $1 < p < \infty$, and let $T : X \to Y$ be a bounded surjective linear operator. Prove that the image $T(B_X(0,1))$ contains a ball around 0 in $Y$.

**Exercise 3.21.** Let

$$c_{00} = \{x = \{x_n\}_{n\in\mathbb{N}} : \#\{n : x_n \neq 0\} < \infty\}$$

be the space of finitely supported sequences, equipped with the supremum norm

$$\|x\|_\infty = \sup_{n\in\mathbb{N}} |x_n|.$$

For each $n \in \mathbb{N}$, define a linear operator $T_n : c_{00} \to \mathbb{R}$ by

$$T_n(x) = nx_n.$$

(a) Show that the family $\{T_n\}_{n\in\mathbb{N}}$ is pointwise bounded.

(b) Compute the operator norm $\|T_n\|$ for each $n \in \mathbb{N}$. Conclude that the family $\{T_n\}$ is not uniformly bounded.

(c) Explain why this example does not contradict the Banach-Steinhaus theorem.

**Exercise 3.22.** Let $T : \ell^1 \to \ell^\infty$ be defined by $T(x) = x$. Is that $T$ a bounded operator, does it has a closed graph?

**Exercise 3.23.** Let $T : D(T) \subset L^2(0,1) \to L^2(0,1)$, with $D(T) = C_c^\infty(0,1)$, and $T(f) = f'$. Determine whether $T$ has a closed graph and whether $T$ is continuous.

**Example 3.1.** In $\ell^2$, the sequence $x_n = (0, 0, ..., 0, 1, 0, ...)$ with 1 in the $n$-th place converges weakly to 0 but not strongly.

**Exercise 3.24.** Let $x_n = (1/n, 1/n, 1/n, ..., 1/n, 0, \cdots)$ in $\ell^2$, $n$ times $\frac{1}{n}$. Determine whether $x_n$ converges strongly, weakly, or both.

**Exercise 3.25.** Let $X = \mathbb{R}^N$ for $N \geq 1$. Prove that a sequence $\{x_n\} \subset \mathbb{R}^N$ converges weakly to $x \in \mathbb{R}^N$ if and only if it converges strongly.

**Exercise 3.26.** Let $e_n = (0, 0, \dots, 1, 0, \dots) \in \ell^\infty = (\ell^1)^*$ be the $n$-th standard basis vector. Then $e_n \overset{*}{\rightharpoonup} 0$.

**Exercise 3.27.** Consider the unilateral shift operator $S : \ell^2 \to \ell^2$ defined by

$$S(x_1, x_2, x_3, \dots) = (0, x_1, x_2, x_3, \dots).$$

1. Show that $S$ is a bounded linear operator.
2. Show that $S$ is Fredholm and find its index.

**Exercise 3.28.** We consider the standard norms in $C[0,1]$ and $C^1[0,1]$:

$$\|g\|_{C[0,1]} = \sup_{x\in[0,1]} |g(x)| \quad \text{for } g \in C[0,1],$$

$$\|f\|_{C^1[0,1]} = \sup_{x\in[0,1]} |f(x)| + \sup_{x\in[0,1]} |f'(x)| \quad \text{for } f \in C^1[0,1].$$

Define $T : C^1[0,1] \to C[0,1]$ by $Tf = f'$.

1. Show that $T$ is bounded and determine $\ker T$.
2. Describe the range $R(T)$.
3. Show that $T$ is Fredholm and compute its index.

**Exercise 3.29.** Let $T : \ell^2 \to \ell^2$ be defined by

$$T(x_1, x_2, x_3, \dots) = \left(\frac{x_1}{2}, \frac{x_2}{3}, \frac{x_3}{4}, \dots\right).$$

Determine the spectrum $\sigma(T)$ and all eigenvalues of $T$.

We recall the following theorem.

**Theorem 3.1** (Arzelà–Ascoli)**.** Let $(X, d)$ be a compact metric space and let $\mathcal{F} \subset C(X)$, where $C(X)$ is the space of continuous real-valued functions on $X$ equipped with the sup norm

$$\|f\|_\infty = \sup_{x \in X} |f(x)|.$$

Then $\mathcal{F}$ is relatively compact in $C(X)$ (i.e., its closure is compact) if and only if:

1. **Equicontinuity:** For every $\varepsilon > 0$ there exists $\delta > 0$ such that for all $f \in \mathcal{F}$ and all $x, y \in X$ with $d(x, y) < \delta$,

$$|f(x) - f(y)| < \varepsilon.$$

2. **Pointwise boundedness:** For every $x \in X$, the set

$$\{f(x) : f \in \mathcal{F}\}$$

is bounded in $\mathbb{R}$.

**Exercise 3.30.** For the operator $T : C[0, 1] \to C[0, 1]$, defined by

$$(Tf)(x) = \int_0^x f(y)\, dy,$$

prove that $T$ is compact and find all eigenvalues of $T$.

**Exercise 3.31.** Let $T : C[0, 1] \to C[0, 1]$ be defined by

$$(Tf)(x) = \int_0^1 K(x, y) f(y)\, dy,$$

where $K \in C([0, 1] \times [0, 1])$. Show that $T$ is compact and, if $K(x, y)$ can be written as a finite sum $\sum_{i=1}^m g_i(x) h_i(y)$, then $T$ has finite rank.

**Exercise 3.32.** Let $T : \ell^2 \to \ell^2$ be given by

$$T(x_1, x_2, x_3, \dots) = (0, x_1, x_2, \dots),$$

i.e. the right shift operator. Prove that $T$ is bounded but not compact, and determine its spectrum $\sigma(T)$.

**Exercise 3.33.** Let $T : \ell^2 \to \ell^2$ be a diagonal operator defined by

$$T(x_1, x_2, x_3, \dots) = (\lambda_1 x_1, \lambda_2 x_2, \lambda_3 x_3, \dots),$$

where $(\lambda_n)$ is a bounded sequence of scalars. Determine when $T$ is compact and describe $\sigma(T)$ in that case.

**Exercise 3.34.** Let $H = \ell^2$, and let $M \subset H$ be the subspace defined by

$$M = \Big\{ x = (x_1, x_2, x_3, \dots) \in \ell^2 : x_n = 0 \text{ for all } n \geq 2 \Big\}.$$

That is, $M = \operatorname{span}\{e_1\}$, where $e_1 = (1, 0, 0, \dots)$.
Given the vector $x = (3, 4, 0, 0, \dots) \in \ell^2$, find the orthogonal projection $P_M x$ of $x$ onto $M$, and compute the distance $\|x - P_M x\|$.

**Exercise 3.35.** Find $\sigma(A)$, where

$$A : \ell^2 \to \ell^2$$
$$(Ax)_1 = x_2,$$
$$(Ax)_n = x_{n-1} + x_{n+1}, \qquad n \geq 2.$$

**Exercise 3.36.** Find $\sigma(A)$, where

$$A : \ell^2 \to \ell^2$$
$$(Ax)_1 = x_1,$$
$$(Ax)_n = x_{n-1} + x_{n+1}, \qquad n \geq 2.$$

**Exercise 3.37.** Let $H = L^2([0,1])$. Define the linear functional

$$\phi(f) = \int_0^1 f(x) \cdot x^2 \, dx.$$

Show that $\phi$ is a bounded linear functional on $H$, and find the unique function $g \in H$ such that

$$\phi(f) = \langle f, g \rangle_{L^2} \quad \text{for all } f \in H.$$

**Exercise 3.38.** Let $H = \ell^2$. Define the linear functional $\phi : H \to \mathbb{C}$ by

$$\phi(x) = 2x_1 - x_3 + ix_4, \quad \text{for } x = (x_1, x_2, x_3, x_4, \dots).$$

Show that $\phi$ is bounded and find the vector $y \in \ell^2$ such that

$$\phi(x) = \langle x, y \rangle_{\ell^2} \quad \text{for all } x \in \ell^2.$$

**Exercise 3.39.** Let $T : \ell^2 \to \ell^2$ be the diagonal operator defined by

$$T(x_1, x_2, x_3, \dots) = \left( \frac{1}{1}x_1, \frac{1}{2}x_2, \frac{1}{3}x_3, \dots \right).$$

(a) Show that $T$ is compact and self-adjoint.

(b) Find the spectrum $\sigma(T)$.

(c) Use the spectral theorem to describe an orthonormal basis of eigenvectors for $T$.

**Exercise 3.40.** Let $T : L^2([0,1]) \to L^2([0,1])$ be the integral operator defined by

$$(Tf)(x) = \int_0^1 \min(x,y) f(y) \, dy.$$

(a) Show that $T$ is compact and self-adjoint.

(b) State what the spectral theorem says about the structure of $T$.

**Exercise 3.41.** Let $H = L^2([0,1])$, and $M = \{f \in H : \int_0^1 f(x)dx = 0\}$. Find the orthogonal complement $M^\perp$.

**Exercise 3.42.** Let $H = L^2[0,1]$. Then the sequence

$$\{1\} \cup \left\{\sqrt{2}\cos(2\pi nt), \sqrt{2}\sin(2\pi nt)\right\}_{n=1}^{\infty}$$

is an orthonormal basis of $H$.

**Exercise 3.43.** Let $H = \ell^2$ and define the bilinear form:

$$a(u,v) = \sum_{n=1}^{\infty} \lambda_n u_n v_n,$$

where $\lambda_n \geq \lambda > 0$ for all $n \in \mathbb{N}$, and $(\lambda_n)$ is a bounded sequence. Let $f = (f_1, f_2, \ldots) \in \ell^2$, and define the linear functional:

$$f(v) = \sum_{n=1}^{\infty} f_n v_n.$$

(a) Show that $a(u,v)$ is a bounded bilinear form on $\ell^2$.

(b) Show that $a$ is coercive, i.e., $a(u,u) \geq \alpha \|u\|^2$ for some $\alpha > 0$.

(c) Show that $f(v)$ is a bounded linear functional on $\ell^2$.

(d) Use the Lax-Milgram theorem to prove that there exists a unique $u \in \ell^2$ such that

$$a(u,v) = f(v) \quad \text{for all } v \in \ell^2.$$

Moreover, find an explicit formula for $u$.

**Exercise 3.44.** Prove that the operator $T : L^2(\Omega) \to L^2(\Omega)$ from Section 2.8 is bounded and compact.

# References

[1] M. Badiale and E. Serra, *Semilinear Elliptic Equations for Beginners: Existence Results via the Variational Approach*, Universitext, Springer, London, 2010.

[2] H. Brezis, *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Universitext, Springer, New York, 2010.

[3] J. B. Conway, *A Course in Functional Analysis*, 2nd ed., Graduate Texts in Mathematics, Vol. 96, Springer, New York, 1990.

[4] J. Mederski, *Variational Methods for Nonlinear Maxwell Equations* (lecture notes), Karlsruhe Institute of Technology, Summer Semester 2018. Available at: https://www.math.kit.edu/iana2/lehre/varimaxwell2018s/media/maxwell_lectures_nov_15.pdf

[5] A. A. Pankov, *Introduction to Spectral Theory of Schrödinger Operators*, Mathematics Preprint Archive, Report No. MP-ARC-2002-449 (original lecture notes from 1999–2000; published December 2001).

[6] M. Reed and B. Simon, *Methods of Modern Mathematical Physics, Vol. I: Functional Analysis*, 2nd ed., Academic Press, San Diego, 1980.

[7] M. Reed and B. Simon, *Methods of Modern Mathematical Physics, Vol. II: Fourier Analysis, Self-Adjointness*, Academic Press, New York, 1975.

[8] M. Reed and B. Simon, *Methods of Modern Mathematical Physics, Vol. III: Scattering Theory*, Academic Press, New York, 1979.

[9] M. Reed and B. Simon, *Methods of Modern Mathematical Physics, Vol. IV: Analysis of Operators*, Academic Press, New York, 1978.

[10] W. Rudin, *Functional Analysis*, 2nd ed., International Series in Pure and Applied Mathematics, McGraw-Hill, New York, 1991.

[11] M. Struwe, *Variational Methods: Applications to Nonlinear Partial Differential Equations and Hamiltonian Systems*, 4th ed., Ergebnisse der Mathematik und ihrer Grenzgebiete, Springer, Berlin, 2008.

[12] M. Willem, *Minimax Theorems*, Progress in Nonlinear Differential Equations and Their Applications, Vol. 24, Birkhäuser, Boston, 1996.