

# Linear Operators in Hilbert Space

A Gemini Compilation

October 25, 2025

## 1. Densely Defined Operators

An operator is called **densely defined** if its domain is a dense subset of the Hilbert space. This means you can get arbitrarily close to *any* vector in the entire space using only vectors from the operator's domain.

### The Definition of a Dense Set

A subset  $D$  of a Hilbert space  $\mathcal{H}$  is said to be **dense** in  $\mathcal{H}$  if its closure is the entire space. We write this as:

$$\overline{D} = \mathcal{H}$$

The **closure** of a set ( $D$ ) is the set itself plus all of its limit points. This has a very practical meaning: any vector in the whole Hilbert space  $\mathcal{H}$  can be approximated with arbitrary precision by a sequence of vectors from the dense set  $D$ .

#### Analogy: Rational Numbers ( $\mathbb{Q}$ ) in Real Numbers ( $\mathbb{R}$ )

The set of rational numbers,  $\mathbb{Q}$ , is dense in the set of real numbers,  $\mathbb{R}$ . You cannot pick any real number, like  $\pi$ , without there being a rational number right next to it. For any tiny distance you choose, you can always find a rational number that is closer to  $\pi$  than that distance. A sequence like  $\{3, 3.1, 3.14, 3.141, \dots\}$  consists entirely of rational numbers, but its limit point,  $\pi$ , is not. Because you can approximate *every* real number this way, the closure of  $\mathbb{Q}$  is  $\mathbb{R}$ .

### Densely Defined Operators

Now, we connect this concept to operators. An operator  $O$  in a Hilbert space  $\mathcal{H}$  is defined on a specific set of vectors, called its **domain**, denoted  $D(O)$ . The operator is said to be **densely defined** if this domain,  $D(O)$ , is a dense subset of  $\mathcal{H}$ .

Using our definition from above, this simply means:

$$\overline{D(O)} = \mathcal{H}$$

So, even if an operator isn't defined for every single vector in the Hilbert space (which is common for unbounded operators), its domain must be "spread out" enough to approximate any vector in the space.

### Why is this important in Quantum Mechanics?

This is not just a mathematical technicality. The property of being densely defined is a crucial prerequisite for an operator to be physically meaningful.

Specifically, an operator must be densely defined in order to uniquely define its **adjoint** ( $O^\dagger$ ). The concept of the adjoint is fundamental for defining **self-adjoint operators** ( $O = O^\dagger$ ), which are the mathematical representations of physical observables (like energy, momentum, position, etc.). If an operator's domain is not dense, its adjoint is not well-defined, and the entire structure of quantum theory built upon observables would collapse.

## 2. Kernel of an Operator

The **kernel** of an operator is the set of all vectors that the operator maps to the **zero vector**. It is also commonly called the **null space**.

### Formal Definition

Let  $T$  be a linear operator that maps from a vector space  $V$  to another vector space  $W$ , written as  $T : V \rightarrow W$ .

The kernel of  $T$ , denoted as  $\text{Ker}(T)$  or  $\text{null}(T)$ , is the subset of the domain  $V$  containing all vectors that are sent to the zero vector  $0_W$  in the codomain  $W$ .

#### Definition

In set-builder notation:

$$\text{Ker}(T) = \{v \in V \mid T(v) = 0_W\}$$

### Intuitive Analogy: The Projector's Blind Spot

Imagine an operator  $T$  as a projector that takes 3D objects and casts a 2D shadow on a wall. The 3D world is your domain ( $V$ ) and the 2D wall is your codomain ( $W$ ).

The **kernel** is the set of all 3D vectors that get completely “crushed” by the projector and produce no shadow (i.e., they map to the origin, or the “zero spot,” on the wall). In this case, any vector pointing directly from the origin towards the projector's light source would have a length in 3D but would cast no shadow. The line containing all such vectors is the kernel. It's the operator's “blind spot.”

### Key Properties and Significance

The kernel of an operator is a crucial concept because it tells you about the operator's behaviour.

1. **Measures Injectivity:** The size of the kernel tells you if the operator is **one-to-one (injective)**.
  - If  $\text{Ker}(T) = \{0_V\}$  (meaning only the zero vector gets mapped to zero), then the operator is **injective**. No two different vectors are mapped to the same output.
  - If the kernel contains non-zero vectors, the operator is **not injective**. It “loses” information by mapping multiple input vectors to the same output.
2. **It's a Subspace:** The kernel is not just a random set of points; it is always a **vector subspace** of the domain  $V$ .
3. **Solving Equations:** The kernel is fundamental to understanding the solutions to linear equations of the form  $T(x) = b$ . If you find one particular solution  $x_p$ , then the set of *all* solutions is given by the set  $\{x_p + k \mid k \in \text{Ker}(T)\}$ . The kernel describes the “ambiguity” or degrees of freedom in the solution set.

## 3. Hilbert-Schmidt Operator

In functional analysis, a Hilbert-Schmidt operator is a bounded linear operator on a Hilbert space that has a finite Hilbert-Schmidt norm. These operators are an important subclass of compact operators and are, in a sense, the infinite-dimensional generalization of matrices with a finite Frobenius norm.

## Definition

Let  $\mathcal{H}$  be a separable Hilbert space and let  $\{e_i\}_{i=1}^{\infty}$  be an orthonormal basis for  $\mathcal{H}$ . A bounded linear operator  $T : \mathcal{H} \rightarrow \mathcal{H}$  is called a **Hilbert-Schmidt operator** if the following sum converges:

$$\sum_{i=1}^{\infty} \|Te_i\|^2 < \infty \quad (1)$$

A crucial property of this definition is that the value of this sum is **independent of the choice of orthonormal basis**.

## The Hilbert-Schmidt Norm and Inner Product

The set of all Hilbert-Schmidt operators on  $\mathcal{H}$ , denoted  $HS(\mathcal{H})$  or  $B_2(\mathcal{H})$ , itself forms a Hilbert space.

The condition for an operator to be Hilbert-Schmidt defines a norm, called the **Hilbert-Schmidt norm**, denoted  $\|\cdot\|_{HS}$ .

### Hilbert-Schmidt Norm

$$\|T\|_{HS} = \left( \sum_{i=1}^{\infty} \|Te_i\|^2 \right)^{1/2}$$

This can also be expressed in terms of the trace of the operator  $T^\dagger T$ :

$$\|T\|_{HS}^2 = \text{tr}(T^\dagger T)$$

The Hilbert-Schmidt norm is induced by an inner product, which gives the space of Hilbert-Schmidt operators its own Hilbert space structure. For two operators  $S, T \in HS(\mathcal{H})$ , the inner product is:

$$\langle S, T \rangle_{HS} = \text{tr}(S^\dagger T) \quad (2)$$

## Key Properties

- **Compactness:** Every Hilbert-Schmidt operator is a **compact operator**. This is one of their most important features.
- **Boundedness:** Every Hilbert-Schmidt operator is also a bounded operator, and its operator norm is bounded by its Hilbert-Schmidt norm:  $\|T\| \leq \|T\|_{HS}$ .
- **An Ideal:** The set of Hilbert-Schmidt operators forms a two-sided ideal in the algebra of bounded operators. This means if  $T$  is a Hilbert-Schmidt operator and  $A$  is any bounded operator, then  $AT$  and  $TA$  are also Hilbert-Schmidt operators.

## Analogy with Matrices

The most intuitive way to think about Hilbert-Schmidt operators is as a generalization of matrices. If we write the matrix elements of an operator  $T$  with respect to the basis  $\{e_i\}$  as  $T_{ij} = \langle e_i, Te_j \rangle$ , then the squared Hilbert-Schmidt norm is simply the sum of the squares of the absolute values of all its matrix elements:

$$\|T\|_{HS}^2 = \sum_{i,j=1}^{\infty} |T_{ij}|^2$$

This is exactly the definition of the **Frobenius norm** for finite matrices. Therefore, a Hilbert-Schmidt operator is an infinite-dimensional matrix whose entries are “square-summable.”

## 5. Unitary Operators representing Rotation

unitary operator  $U$  represents a generalized rotation in a Hilbert space  $\mathcal{H}$ . It is rigorously defined as a transformation that preserves the fundamental geometric structure of the space: the **norms** (lengths) of all vectors and the **inner products** (angles/overlaps) between them.

This is the direct generalization of a rotation in ordinary Euclidean space  $\mathbb{R}^n$ , which is defined as an orthogonal matrix  $R$  (satisfying  $R^T R = \mathbb{I}$ ) that preserves lengths and angles.

### The Geometric Structure of $\mathcal{H}$

A Hilbert space  $\mathcal{H}$  is a complete vector space endowed with an inner product  $\langle \cdot | \cdot \rangle$  that defines its geometry.

- **Norm (Length):** The norm of a state vector  $|\psi\rangle$  is  $\| |\psi\rangle \| = \sqrt{\langle \psi | \psi \rangle}$ . In quantum mechanics, the normalization condition  $\| |\psi\rangle \|^2 = 1$  (representing total probability) must be preserved by any physical evolution.
- **Inner Product (Angle/Overlap):** The inner product  $\langle \phi | \psi \rangle$  is a complex number representing the probability amplitude for a system in state  $|\psi\rangle$  to be found in state  $|\phi\rangle$ .

A "rotation" in  $\mathcal{H}$  is thus any transformation  $U$  that leaves this entire geometric structure invariant.

The derivation of the operator is as follows:-

A transformation  $U$  is a rotation if it preserves the norm (length) of *all* vectors:

$$\| U|\psi\rangle \| = \| |\psi\rangle \| \quad \text{for all } |\psi\rangle \in \mathcal{H}$$

We can prove that this single condition is equivalent to the definition of a unitary operator.

1. Squaring both sides of the requirement:

$$\| U|\psi\rangle \|^2 = \| |\psi\rangle \|^2$$

2. Writing the norms in terms of their inner products:

$$\langle U\psi | U\psi \rangle = \langle \psi | \psi \rangle$$

3. Using the definition of the Hermitian adjoint ( $U^\dagger$ ) on the left side:

$$\langle \psi | U^\dagger U | \psi \rangle = \langle \psi | \mathbb{I} | \psi \rangle$$

(where  $\mathbb{I}$  is the identity operator).

4. Since this equation must hold for *any* arbitrary vector  $|\psi\rangle$  in the space, the operators themselves must be equal.

#### Definition: Unitary Operator

An operator  $U : \mathcal{H} \rightarrow \mathcal{H}$  is **unitary** if its adjoint  $U^\dagger$  is also its inverse:

$$U^\dagger U = U U^\dagger = \mathbb{I}$$

This condition automatically guarantees the preservation of all inner products:

$$\langle U\phi | U\psi \rangle = \langle \phi | U^\dagger U | \psi \rangle = \langle \phi | \mathbb{I} | \psi \rangle = \langle \phi | \psi \rangle$$

Therefore, a unitary operator is precisely the transformation that "rotates" the state space without stretching, shrinking, or warping it, thereby preserving all physical probabilities and transition amplitudes.

## The Generator of Infinitesimal Rotations

A continuous unitary transformation (a finite "rotation" by a parameter  $\theta$ ) can be constructed by applying an infinite sequence of infinitesimal rotations. The operator that "creates" this infinitesimal rotation is called the **generator**.

We define an infinitesimal unitary transformation  $U(\delta\theta)$  as a small deviation from the identity operator  $\mathbb{I}$  by an amount proportional to the small parameter  $\delta\theta$ .

### Definition: Infinitesimal Unitary Transformation

An infinitesimal unitary transformation  $U(\delta\theta)$  is given by:

$$U(\delta\theta) = \mathbb{I} - iG\delta\theta$$

where  $G$  is the **generator** of the transformation.

The factor  $i$  is a standard convention in physics to ensure that the generator  $G$  is a Hermitian operator. We can prove this by enforcing the unitary condition  $U^\dagger U = \mathbb{I}$  on the infinitesimal transformation, keeping only terms up to the first order in  $\delta\theta$ .

$$\begin{aligned} U^\dagger(\delta\theta)U(\delta\theta) &= \mathbb{I} \\ (\mathbb{I} - iG\delta\theta)^\dagger(\mathbb{I} - iG\delta\theta) &= \mathbb{I} \\ (\mathbb{I} + iG^\dagger\delta\theta)(\mathbb{I} - iG\delta\theta) &= \mathbb{I} \\ \mathbb{I} - iG\delta\theta + iG^\dagger\delta\theta - (i)^2 G^\dagger G (\delta\theta)^2 &= \mathbb{I} \\ \mathbb{I} - i(G - G^\dagger)\delta\theta + \mathcal{O}(\delta\theta^2) &= \mathbb{I} \end{aligned}$$

For this equation to hold, the first-order term in  $\delta\theta$  must vanish:

$$i(G - G^\dagger) = 0 \implies G = G^\dagger$$

### Important Note: Generators are Hermitian

The generator  $G$  of a continuous unitary transformation  $U$  must be a **Hermitian** (or self-adjoint) **operator**. This is the fundamental link between symmetries (unitary transformations) and observables (Hermitian operators) in quantum mechanics.

A finite rotation by a parameter  $\theta$  is achieved by composing  $N$  infinitesimal rotations of size  $\delta\theta = \theta/N$  and taking the limit  $N \rightarrow \infty$ :

$$U(\theta) = \lim_{N \rightarrow \infty} \left( \mathbb{I} - iG \frac{\theta}{N} \right)^N = \exp(-iG\theta)$$

This exponential map leads to the formal definition of the generator  $G$  as the "velocity" of the transformation at the identity.

### Definition: Generator of a Transformation

Given a one-parameter family of unitary operators  $U(\theta) = \exp(-iG\theta)$ , the generator  $G$  is formally defined as the derivative of the transformation at  $\theta = 0$ :

$$G = i \left. \frac{dU(\theta)}{d\theta} \right|_{\theta=0}$$

## 6. The Spectral Theorem and Operator-Valued Measures

The Spectral Theorem is a central result in functional analysis that provides a complete description of self-adjoint operators. It generalises the finite-dimensional concept of an "eigenvalue decomposition" to all observables in a Hilbert space, including those with continuous spectra. To understand the theorem, we must first define the mathematical objects of measurement: POVMs and PVMs.

### Positive Operator-Valued Measure (POVM)

A POVM is the most general formulation of a measurement in quantum mechanics. It is a map from the measurable subsets of the outcome space (e.g., the real line  $\mathbb{R}$  for a typical observable) to positive operators on the Hilbert space.

#### Definition: Positive Operator-Valued Measure (POVM)

Let  $(\mathcal{X}, \Sigma)$  be a measurable space, where  $\mathcal{X}$  is the set of measurement outcomes and  $\Sigma$  is its Borel  $\sigma$ -algebra. A **POVM** is a function  $E : \Sigma \rightarrow \mathcal{L}(\mathcal{H})$  that assigns a bounded, positive, self-adjoint operator  $E(B)$  to each measurable set  $B \in \Sigma$ , satisfying:

1. **Positivity:**  $E(B) \geq 0$  for all  $B \in \Sigma$ . (The operator  $E(B)$  is a positive operator).
2. **Normalization (Completeness):**  $E(\mathcal{X}) = \mathbb{I}$ . (The measure of the entire outcome space is the identity operator, ensuring probabilities sum to 1).
3. **Countable Additivity:** For any sequence of pairwise disjoint sets  $\{B_i\}_{i=1}^{\infty}$  in  $\Sigma$ ,

$$E\left(\bigcup_{i=1}^{\infty} B_i\right) = \sum_{i=1}^{\infty} E(B_i)$$

where the sum converges in the strong operator topology.

For a system in a state  $\rho$ , the probability of obtaining a measurement outcome in the set  $B$  is given by  $\Pr(B) = \text{tr}(\rho E(B))$ .

### Projection-Valued Measure (PVM)

A PVM is a special, more restrictive, and "ideal" type of measurement. It is a POVM where every operator  $E(B)$  is not just positive, but is a full **projection operator**.

#### Definition: Projection-Valued Measure (PVM)

A **PVM** is a POVM  $E : \Sigma \rightarrow \mathcal{L}(\mathcal{H})$  that satisfies the additional condition of being **idempotent**:

$$E(B) = E(B)^2 \quad \text{for all } B \in \Sigma$$

This implies that  $E(B)$  is an orthogonal projector. This condition also rigorously implies that the projectors for disjoint sets are orthogonal:

$$E(B_1)E(B_2) = E(B_1 \cap B_2) = 0 \quad \text{if } B_1 \cap B_2 = \emptyset$$

PVMs represent "sharp" or "ideal" measurements that do not disturb the system if repeated. POVMs are more general and can describe "unsharp" or "noisy" measurements, as well as measurements on subsystems of an entangled state.

## The Riemann-Stieltjes Integral and Measure

The spectral integral  $\int \lambda dE(\lambda)$  is a generalisation of the **Riemann-Stieltjes integral**, which extends the concept of Riemann integration to allow integration with respect to a function of bounded variation, not just a simple variable like  $dx$ .

**Definition 0.1 (Riemann-Stieltjes Integral)** Let  $f(x)$  be a continuous function (the integrand) and  $g(x)$  be a function of bounded variation (the integrator), both defined on a finite interval  $[a, b]$ . Let  $P$  be a partition of  $[a, b]$ , given by  $a = x_0 < x_1 < \dots < x_n = b$ . Let  $\xi_k \in [x_{k-1}, x_k]$  be a sample point in each subinterval.

The **Riemann-Stieltjes sum** is defined as:

$$S(P, f, g) = \sum_{k=1}^n f(\xi_k)[g(x_k) - g(x_{k-1})]$$

Let  $\|P\| = \max_k(x_k - x_{k-1})$  be the mesh (or norm) of the partition. The Riemann-Stieltjes integral of  $f$  with respect to  $g$  exists and has the value  $L$  if, for every  $\epsilon > 0$ , there exists a  $\delta > 0$  such that for every partition  $P$  with  $\|P\| < \delta$ :

$$|S(P, f, g) - L| < \epsilon$$

If this limit  $L$  exists, we write it as:

$$L = \int_a^b f(x) dg(x)$$

### The Role of the Integrator $g(x)$

The "differential"  $dg(x)$  does not require  $g$  to be differentiable. It is a **measure** that assigns a "weight"  $\Delta g_k = [g(x_k) - g(x_{k-1})]$  to each interval  $[x_{k-1}, x_k]$ . The integral is the weighted sum of  $f(x)$  against this measure.

This is why the Stieltjes integral is essential for the Spectral Theorem:

1. **For Discrete Spectra:** The spectral family  $E(\mu)$  is a step function. The "integrator"  $g(x)$  is non-differentiable (it has "jumps"). The "weight"  $\Delta E_k = E(\lambda_k) - E(\lambda_k - \epsilon)$  is precisely the projection operator  $P_k$  at the eigenvalue  $\lambda_k$ . The Stieltjes integral  $\int \lambda dE(\lambda)$  collapses into the familiar sum:

$$\int \lambda dE(\lambda) = \sum_k f(\lambda_k) \cdot (\text{jump at } \lambda_k) = \sum_k \lambda_k P_k = A$$

2. **For Continuous Spectra:** The spectral family  $E(\mu)$  is a continuous function. Here,  $dE(\mu)$  represents an infinitesimal projection onto a continuous range of eigenvalues, and the integral  $\int \lambda dE(\lambda)$  becomes a true continuous integral.

The spectral theorem, in its full rigor, promotes this concept from a scalar-valued integrator  $g(x)$  to the **projection-valued measure (PVM)**  $E$ , which is an operator-valued measure.

## The Spectral Theorem

The Spectral Theorem establishes the fundamental link between self-adjoint operators (observables) and PVMs (ideal measurements). It states that every observable can be uniquely "decomposed" into a PVM.

This theorem generalises the finite-dimensional case,  $A = \sum_i \lambda_i P_i$ , by replacing the sum with a Stieltjes integral over the operator's spectrum,  $\sigma(A)$ .

### Definition: The Spectral Theorem

For every self-adjoint operator  $A$  (bounded or unbounded) on a Hilbert space  $\mathcal{H}$ , there exists a **unique** projection-valued measure (PVM)  $E$  on the real line  $\mathbb{R}$  such that:

$$A = \int_{-\infty}^{\infty} \lambda dE(\lambda)$$

This unique PVM,  $E$ , is called the **spectral measure** of  $A$ , and its support is the spectrum of  $A$ ,  $\sigma(A)$ .

If  $A$  is unbounded (e.g., the position or momentum operator), the integral is defined on a dense domain  $\mathcal{D}(A)$ :

$$\mathcal{D}(A) = \left\{ |f\rangle \in \mathcal{H} \mid \int_{-\infty}^{\infty} \lambda^2 d\langle f|E(\lambda)f\rangle < \infty \right\}$$

where  $d\langle f|E(\lambda)f\rangle$  is the scalar-valued measure for the vector  $|f\rangle$ .

### The Power of the Theorem: Functional Calculus

The Spectral Theorem provides a rigorous **functional calculus**. For any measurable function  $g : \mathbb{R} \rightarrow \mathbb{C}$ , we can define the (generally unbounded) operator  $g(A)$  as:

$$g(A) = \int_{-\infty}^{\infty} g(\lambda) dE(\lambda)$$

This is what gives precise mathematical meaning to crucial quantum operators, such as the time-evolution operator  $U(t) = \exp(-iHt/\hbar)$  (where  $g(\lambda) = e^{-i\lambda t/\hbar}$ ) or the square root operator  $\sqrt{A}$  (where  $g(\lambda) = \sqrt{\lambda}$ ).