

Spectral Theorem

Formal statement

Let mathcal H be a finite-dimensional complex Hilbert space, and let

$$A : \mathcal{H} \rightarrow \mathcal{H}$$

be a Hermitian operator, meaning

$$A = A^\dagger.$$

Then there exists an orthonormal basis

$$\{|e_1\rangle, \dots, |e_d\rangle\}$$

of mathcal H and real numbers

$$\lambda_1, \dots, \lambda_d \in \mathbb{R}$$

such that

$$A|e_j\rangle = \lambda_j|e_j\rangle$$

for every j. Equivalently,

$$A = \sum_{j=1}^d \lambda_j |e_j\rangle \langle e_j|.$$

If some eigenvalues are repeated, it is often better to group equal eigenvalues together. Then the same theorem becomes

$$A = \sum_{\lambda \in \text{spec}(A)} \lambda P_\lambda,$$

where P_λ is the orthogonal projector onto the eigenspace

$$E_\lambda = \{|\psi\rangle \in \mathcal{H} : A|\psi\rangle = \lambda|\psi\rangle\}.$$

The projectors satisfy

$$P_\lambda P_\mu = \delta_{\lambda\mu} P_\lambda, \quad \sum_\lambda P_\lambda = I.$$

This is the mathematical reason a finite-dimensional quantum observable can be written as a sum of possible measurement outcomes multiplied by the projectors onto the corresponding outcome subspaces.

Proof

We prove the theorem by building an orthonormal basis of eigenvectors. The proof has three main ideas. First, a Hermitian operator has real eigenvalues. Second, eigenvectors with distinct eigenvalues are orthogonal. Third, once we find one eigenvector, its orthogonal complement is preserved by the operator, so we can repeat the argument in a smaller Hilbert space.

Because mathcal H is a finite-dimensional complex vector space, the characteristic polynomial of A has at least one complex root. Therefore A has at least one eigenvalue $\lambda \in \mathbb{C}$ and one nonzero eigenvector $|v\rangle$, so

$$A|v\rangle = \lambda|v\rangle.$$

We now show that λ must be real. Since $A=A^\dagger$, the expectation-like quantity $\langle v|A|v\rangle$ satisfies

$$\langle v|A|v\rangle = \langle v|A^\dagger|v\rangle = \langle v|A|v\rangle^*.$$

Thus $\langle v|A|v\rangle$ is real. On the other hand, using the eigenvalue equation gives

$$\langle v|A|v\rangle = \lambda\langle v|v\rangle.$$

Because $|v\rangle \neq 0$, the number $\langle v|v\rangle$ is positive and real. Hence

$$\lambda = \frac{\langle v|A|v\rangle}{\langle v|v\rangle}$$

is real. So Hermiticity forces eigenvalues to be real.

Now normalize $|v\rangle$ and call the normalized vector $|e_1\rangle$. Suppose

$$A|e_1\rangle = \lambda_1|e_1\rangle.$$

Let

$$W = \{|w\rangle \in \mathcal{H} : \langle e_1|w\rangle = 0\}$$

be the orthogonal complement of $|e_1\rangle$. We claim that W is invariant under A , meaning $A|w\rangle \in W$ whenever $|w\rangle \in W$. Indeed,

$$\langle e_1|A|w\rangle = \langle Ae_1|w\rangle = \lambda_1\langle e_1|w\rangle = 0.$$

Thus $A|w\rangle$ is also orthogonal to $|e_1\rangle$. Therefore A restricts to an operator on W . This restricted operator is still Hermitian, because for any $|x\rangle, |y\rangle \in W$,

$$\langle x|Ay\rangle = \langle Ax|y\rangle.$$

If W is nonzero, we repeat the same argument inside W . We obtain a normalized eigenvector $|e_2\rangle \in W$, with a real eigenvalue λ_2 . Since $|e_2\rangle \in W$, it is orthogonal to $|e_1\rangle$. We then take the orthogonal complement of $\text{span}\{|e_1\rangle, |e_2\rangle\}$ and continue.

Because \mathcal{H} has finite dimension d , this process terminates after d steps. We obtain an orthonormal basis

$$\{|e_1\rangle, \dots, |e_d\rangle\}$$

such that

$$A|e_j\rangle = \lambda_j|e_j\rangle$$

with every $\lambda_j \in \mathbb{R}$. For an arbitrary vector

$$|\psi\rangle = \sum_{j=1}^d c_j |e_j\rangle,$$

we have

$$A|\psi\rangle = \sum_{j=1}^d c_j A|e_j\rangle = \sum_{j=1}^d c_j \lambda_j |e_j\rangle.$$

The operator

$$\sum_{j=1}^d \lambda_j |e_j\rangle\langle e_j|$$

acts on every vector in exactly the same way. Therefore

$$A = \sum_{j=1}^d \lambda_j |e_j\rangle\langle e_j|.$$

If several basis vectors have the same eigenvalue λ , the sum of their rank-one projectors is the projector P_λ onto the whole eigenspace E_λ . Grouping equal eigenvalues gives

$$A = \sum_{\lambda \in \text{spec}(A)} \lambda P_\lambda.$$

This completes the proof.

Operational meaning in quantum information

The theorem says that a Hermitian operator is not merely a matrix. It is a machine that divides Hilbert space into mutually orthogonal outcome sectors. Each sector is an eigenspace. Each sector has a real label, the corresponding eigenvalue.

This is exactly what we need from a quantum observable. A measurement outcome must be a real classical number. A Hermitian operator supplies those real numbers as eigenvalues. But a measurement also needs events: mathematical objects to which probabilities can be assigned. The spectral theorem supplies those events as orthogonal projectors.

Thus, if

$$A = \sum_{\lambda} \lambda P_{\lambda},$$

then measuring the observable A means asking which spectral subspace the state occupies. If the system is in a pure state $|\psi\rangle$, the probability of obtaining outcome λ is

$$\Pr(\lambda) = \langle \psi | P_{\lambda} | \psi \rangle.$$

If the system is in a mixed state ρ , the probability is

$$\Pr(\lambda) = \text{Tr}(P_{\lambda}\rho).$$

The expectation value is

$$\mathbb{E}[A] = \sum_{\lambda} \lambda \text{Tr}(P_{\lambda}\rho) = \text{Tr}(A\rho).$$

This is the operational bridge between linear algebra and laboratory measurement. The operator A contains two pieces of information at once: the possible numerical outcomes λ , and the projectors P_{λ} that decide the probabilities of those outcomes.

The simplest mental image is this: the spectral theorem says that every finite-dimensional observable secretly comes with its own coordinate system. In that coordinate system, the observable is diagonal. A measurement in that observable is just a measurement of which diagonal block, or eigenspace, the state belongs to.

Example: Pauli Z

The Pauli Z operator is

$$Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

It is Hermitian and already diagonal in the computational basis. Its eigenvectors are

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

with eigenvalues +1 and -1. Therefore

$$Z = |0\rangle\langle 0| - |1\rangle\langle 1|.$$

If

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \quad |\alpha|^2 + |\beta|^2 = 1,$$

then measuring Z gives outcome +1 with probability $|\alpha|^2$ and outcome -1 with probability $|\beta|^2$. The spectral theorem is doing something very simple here: it tells us that Z measures the state in the computational basis.

Example: Pauli X

The Pauli X operator is

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

It is Hermitian, but it is not diagonal in the computational basis. The spectral theorem says that there must be another orthonormal basis in which it is diagonal. That basis is

$$|+\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}, \quad |-\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}.$$

Indeed,

$$X|+\rangle = |+\rangle, \quad X|-\rangle = -|-\rangle.$$

Thus

$$X = |+\rangle\langle+| - |-\rangle\langle-|.$$

This example shows why diagonalization has physical meaning. The operator X is not a computational-basis measurement. It is a measurement in the $|+\rangle, |-\rangle$ basis. If the input state is $|0\rangle$, then

$$|0\rangle = \frac{|+\rangle + |-\rangle}{\sqrt{2}}.$$

So measuring X on $|0\rangle$ gives $+1$ with probability $1/2$ and -1 with probability $1/2$. The theorem tells us which basis the observable naturally measures.

Example: degeneracy and coarse-grained information

Consider the two-qubit observable

$$Z \otimes I.$$

It measures the first qubit in the Z -basis and ignores the second qubit. Acting on computational-basis states, it gives

$$(Z \otimes I)|00\rangle = |00\rangle, \quad (Z \otimes I)|01\rangle = |01\rangle,$$

and

$$(Z \otimes I)|10\rangle = -|10\rangle, \quad (Z \otimes I)|11\rangle = -|11\rangle.$$

The eigenvalue $+1$ has a two-dimensional eigenspace spanned by $|00\rangle$ and $|01\rangle$. The eigenvalue -1 has a two-dimensional eigenspace spanned by $|10\rangle$ and $|11\rangle$. Therefore

$$Z \otimes I = P_+ - P_-,$$

where

$$P_+ = |00\rangle\langle 00| + |01\rangle\langle 01|$$

and

$$P_- = |10\rangle\langle 10| + |11\rangle\langle 11|.$$

This example is important because it shows that a measurement outcome need not identify a single vector. It may identify a whole subspace. Measuring $Z \otimes I$ tells us the value of the first qubit in the Z-basis, but it does not tell us whether the second qubit was $|0\rangle$ or $|1\rangle$. Degeneracy is therefore a mathematical expression of coarse-grained information.

Example: density operators and entropy

A density operator ρ is Hermitian, positive semidefinite, and has trace one. By the spectral theorem,

$$\rho = \sum_j p_j |e_j\rangle\langle e_j|,$$

where

$$p_j \geq 0, \quad \sum_j p_j = 1.$$

Thus every finite-dimensional quantum state can be viewed as a classical probability distribution over an orthonormal eigenbasis, at least for the purpose of functions of ρ . For example, the von Neumann entropy becomes

$$S(\rho) = -\text{Tr}(\rho \log \rho) = -\sum_j p_j \log p_j.$$

This is one of the most common ways the spectral theorem is used in quantum information theory. It allows us to reduce many questions about a density operator to questions about its eigenvalues.

This does not mean that every mixed state is merely classical ignorance in one fixed universal basis. Different density operators generally have different eigenbases, and noncommuting states cannot all be diagonalized in the same basis. The spectral theorem diagonalizes one Hermitian operator at a time. Simultaneous diagonalization requires additional commutativity assumptions.

Example: why non-Hermitian matrices are not observables

Consider

$$N = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

This matrix is not Hermitian. It has only one eigenvalue, 1, and it does not have a full orthonormal basis of eigenvectors. Therefore it cannot be written as a spectral decomposition with orthogonal projectors.

Now consider

$$B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

This matrix is also not Hermitian. Its eigenvalues are $+i$ and $-i$, which are not real. A measurement device whose possible outcomes are $+i$ and $-i$ would not represent an ordinary real-valued physical observable.

These two examples show why Hermiticity matters. Hermiticity is not just a convenient algebraic assumption. It simultaneously guarantees real eigenvalues and orthogonal outcome subspaces.

Common mistake

A common mistake is to say that every matrix can be diagonalized. This is false. Some matrices are defective and do not have enough eigenvectors. Another common mistake is to forget the word orthonormal. In quantum mechanics, it is not enough to diagonalize an operator using some arbitrary basis. Measurement theory needs orthogonal projectors, and orthogonal projectors come from orthonormal eigenspaces.

Another important distinction is between rank-one spectral decompositions and degenerate spectral decompositions. Writing

$$A = \sum_j \lambda_j |e_j\rangle\langle e_j|$$

is mathematically correct after choosing an orthonormal eigenbasis. But when an eigenvalue is degenerate, the physically meaningful measurement outcome corresponds to the whole projector

$$P_\lambda,$$

not to a particular arbitrary basis inside the eigenspace. The basis inside a degenerate eigenspace is not uniquely determined by the observable.

Final mental image

The spectral theorem turns a Hermitian operator into a readable measurement device. The eigenvalues are the possible classical answers. The eigenspaces are the mutually exclusive quantum alternatives. The spectral projectors are the probability events. The Born rule then assigns probabilities to those events.

Thus the theorem is not only a diagonalization theorem. It is the finite-dimensional bridge from linear algebra to quantum measurement.

References

Nielsen, Michael A., and Isaac L. Chuang. Quantum Computation and Quantum Information. Cambridge University Press, 10th anniversary ed., 2010. Chapter 2 discusses the spectral decomposition of normal operators and the role of Hermitian operators in measurements. Available excerpt: <https://pages.jh.edu/rrynasi1/HealeySeminar/literature/Nielsen%2BChuang2-010QuantumComputation%2BQuantumInformation.FirstTwoChapters.pdf>

Watrous, John. The Theory of Quantum Information. Cambridge University Press, 2018. The book develops finite-dimensional quantum information theory using spectral decompositions, positive semidefinite operators, density operators, and measurements. Available draft: <https://cs.uwaterloo.ca/~watrous/TQI/TQI.pdf>

Preskill, John. Lecture Notes for Ph219/CS219: Quantum Information, Chapter 2. Preskill states that an observable in quantum mechanics is a self-adjoint operator and discusses its role in quantum measurement. Available at: <https://www.preskill.caltech.edu/ph219/chap215.pdf>

de Wolf, Ronald. Quantum Computing: Lecture Notes. 2019. The notes emphasize the correspondence between Hermitian matrices and observables through spectral decomposition. Available at:
<https://homepages.cwi.nl/~rdewolf/qcnotes.pdf>

Axler, Sheldon. Linear Algebra Done Right. Springer, 4th ed., 2024. Chapter 7 gives a clean finite-dimensional treatment of self-adjoint and normal operators and the spectral theorem. Open-access PDF:
<https://linear.axler.net/LADR4e.pdf>

Document information

Spectral Theorem

Project	[QIT 001] Mathematical Representation Theorems
Document	Primary document
Author	mujirin
Verifier	Not verified
Downloaded	July 04, 2026 17:19 KST
Status	Working
Document link	https://theorytrace.com/projects/mathematical-representation-theorems/documents/spectral-theorem/