University of Massachusetts, Amherst - Spring 2025

PHYSICS 614: Intermediate Quantum Mechanics I

Agustin Garcia Flores

These are lecture notes taken during *PHYSICS 614: Intermediate Quantum Mechanics I*, a graduate course taught by Ben Heidenreich in the Spring 2025 semester.

Contents

2025-01-30: Paradoxical Behavior

When we think quantum, we instantly also think the Schrödinger equation:

$$i\hbar\partial_t\Psi = -\frac{\hbar^2}{2m}\nabla^2\Psi + V\Psi. \tag{1}$$

However, this is not all quantum mechanics is about! Quantum mechanics is about *behavior*.

Double Slit Experiment

Light comes with quantized energy,

$$E = hf. \tag{2}$$

This is particularly important when considering the double slit experiment, shown in Figure 1. In this experiment, light is said to behave light both a particle and a wave. When several photons are allowed to pass through the double slit with no external interference, they form an intensity like

$$I_{\text{ntensity}} \propto \text{sinc}^2\left(\frac{\pi w x}{\lambda L}\right)\cos^2\left(\frac{\pi d x}{\lambda L}\right) \tag{3}$$

where w is the width of each slit, x is the displacement from the central axis, λ is the wavelength of light, L is the distance from the slits to the observation screen, d is the distance between the two centers of the slits $(d > w)$, and

$$\text{sinc}\, x := \frac{\sin x}{x}. \tag{4}$$

The sinc term comes from the single-slit diffraction and modulates the overall intensity envelope while the cos term represents the interference pattern from the two slits. Equation 3 is shown in Figure 2 with arbitrary values for the parameters.

Stern-Gerlach Experiment

The Stern-Gerlach experiment is an experiment which runs silver atoms through an inhomogeneous magnetic field (created by the magnet orientation shown in Figure 3) and observes their deflection. It essentially filter atoms by their angular momentum. The essence of it is captured beautifully by David Z. Albert[1] so we ommit the details here.

$h = 6.626\,070\,15 \times 10^{-34}$ J.s, $\hbar = h.(2\pi)^{-1}$
$\implies \hbar = 1.054\,571\,817 \times 10^{-34}$ J.s.

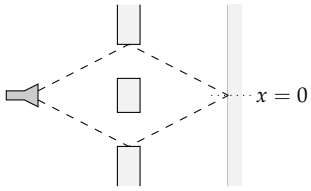

Figure 1: A rough illustration of the double slit experiment, pointing out the position of $x = 0$.

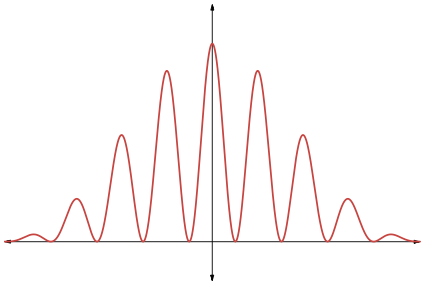

Figure 2: Light intensity as a function of position from the origin.

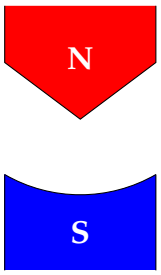

Figure 3: A front view of the magnetic setup in the Stern-Gerlach experiment.

[1] David Z. Albert. *Quantum Mechanics and Experience*. Harvard University Press, 2021

2025-02-04: Finite-dimensional Hilbert Spaces

In short, a Hilbert Space, henceforward denoted by $\mathcal{H}$, is a vector space with an inner product. For now, we focus on the finite-dimensional case.

Vector Spaces

Following notational conventions invented by Paul Dirac in 1939, we write vectors in $\mathcal{H}$ as **kets**:

$$|\psi\rangle, |\chi\rangle, |\phi\rangle \in \mathcal{H}. \tag{5}$$

Addition and commutative[2] scalar[3] multiplication hold: for $|\psi\rangle, |\chi\rangle \in \mathcal{H}$, $\lambda \in \mathbb{C}$,

$$|\psi\rangle + \lambda |\chi\rangle \in \mathcal{H}. \tag{6}$$

We write covectors[4] as **bras**:

$$\langle\omega|, \langle\xi|, \langle\zeta| \in \mathcal{H}^*, \tag{7}$$

where $\mathcal{H}^*$ is the space of covectors, also known as the vector space "dual" to $\mathcal{H}$. **Contracting** a bra with a ket produces a complex number:

$$\langle\omega|\chi\rangle := \langle\omega|(|\chi\rangle) \in \mathbb{C}. \tag{10}$$

Operators[5] are linear maps from $\mathcal{H}$ to itself. The result of applying an operator $\boldsymbol{O}$ to a state $|\psi\rangle$[6] is written as:

$$\boldsymbol{O}|\psi\rangle := \boldsymbol{O}(|\psi\rangle). \tag{11}$$

Operators likewise act *on the right* on bras:

$$\langle\omega|\boldsymbol{O} \text{ is defined by } (\langle\omega|\boldsymbol{O})(|\psi\rangle) := \langle\omega|\boldsymbol{O}|\psi\rangle, \tag{12}$$

where here the notation is chosen in such a way that is naturally *associative*.

Definition 1 (Equal operators)**.** Two operators $\boldsymbol{X}, \boldsymbol{Y}$ are **equal** if

$$\forall\ |\psi\rangle \in \mathcal{H},\ \boldsymbol{X}|\psi\rangle = \boldsymbol{Y}|\psi\rangle. \tag{13}$$

Operators can be added together and multiplied by scalars as usual:

$$(\boldsymbol{X} + \lambda\boldsymbol{Y})(|\psi\rangle) := \boldsymbol{X}|\psi\rangle + \lambda\boldsymbol{Y}|\psi\rangle = \boldsymbol{X}(|\psi\rangle) + \lambda\boldsymbol{Y}(|\psi\rangle). \tag{14}$$

They can also be multiplied by other operators:

$$(\boldsymbol{XY})|\psi\rangle = \boldsymbol{XY}|\psi\rangle := \boldsymbol{X}(\boldsymbol{Y}(|\psi\rangle)). \tag{15}$$

In general, operator multiplication does not commute, $\boldsymbol{XY} \neq \boldsymbol{YX}$. The failure of two operators to commute is measured by their **commutator**:

$$[\boldsymbol{X}, \boldsymbol{Y}] := \boldsymbol{XY} - \boldsymbol{YX} = -[\boldsymbol{Y}, \boldsymbol{X}]. \tag{16}$$

Definition 2 (Outer product)**.** The **outer product**, denoted by $|\psi\rangle\langle\omega|$ of a $|\psi\rangle$ with a $\langle\omega|$, is an operator defined by

$$(|\psi\rangle\langle\omega|)(|\chi\rangle) = |\psi\rangle\ \langle\omega|\chi\rangle = \langle\omega|\chi\rangle\ |\psi\rangle. \tag{18}$$

[2] $\lambda|\psi\rangle = |\psi\rangle\lambda$.

[3] Complex numbers!

[4] Linear maps from $\mathcal{H}$ to $\mathbb{C}$.

It's useful to use parenthesis when contracting a bra with a ket. Doing so emphasizes the fact that bras are linear maps, allowing for distribution,

$$\langle\omega|(|\psi\rangle + \lambda|\chi\rangle) = \langle\omega|\psi\rangle + \lambda\ \langle\omega|\chi\rangle, \tag{8}$$

and addition and scalar multiplication of bras,

$$(\langle\omega| + \lambda\langle\xi|)|\psi\rangle = \langle\omega|\psi\rangle + \lambda\ \langle\xi|\psi\rangle. \tag{9}$$

[5] Operators are written in **boldface**.

[6] Notice how the operator acts *on the left* on kets.

Examples of operators include the identity operator, $\boldsymbol{I}$, which does nothing, $\boldsymbol{I}|\psi\rangle = |\psi\rangle$, and the zero operator $\boldsymbol{0}$, which annihilates all states, $\boldsymbol{0}|\psi\rangle = 0$.

A non-zero operator that can be written as an outer product is said to be **rank one**. More generally, an operator that can be written as the sum of N rank one operators but *not* as the sum of fewer than N rank one operators is rank N:

$$\boldsymbol{O} = \sum_{i=1}^{N} |\psi_i\rangle\langle\omega_i|. \tag{17}$$

One can show that if $\mathcal{H}$ is N dimensional then *every* operator on $\mathcal{H}$ has rank $\leq N$, and so can be written (non-uniquely) in the form Equation 17.

Hilbert Spaces

So far, we have only invoked the properties of a (complex) vector space. We now get into a necessary requirement for a Hilbert space.

Definition 3 (Inner product)**.** Now consider the **inner product** on $\mathcal{H}$, which takes a pair of kets, $|\psi\rangle$ and $|\chi\rangle$, as arguments and produces a complex number $(|\psi\rangle, |\chi\rangle) \in \mathbb{C}$, s.t. the following properties hold:

1. $(|\psi\rangle, |\chi\rangle)^* = (|\chi\rangle, |\psi\rangle)$,
2. $(|\psi\rangle, |\chi\rangle + \lambda |\phi\rangle) = (|\psi\rangle, |\chi\rangle) + \lambda (|\psi\rangle, |\phi\rangle)$,
3. $(|\psi\rangle, |\psi\rangle) \geq 0$ and $(|\psi\rangle, |\psi\rangle) = 0 \implies |\psi\rangle = 0$.

The inner product being sequilinear ensures that the **norm**,

$$\| |\psi\rangle \| := \sqrt{(|\psi\rangle, |\psi\rangle)}, \tag{22}$$

is real and positive for $|\psi\rangle \neq 0$.[7] Using the inner product, we define the **conjugate** of a $|\psi\rangle$ to be the bra:

$$|\psi\rangle^\dagger := \langle\psi| \text{ defined by } |\psi\rangle^\dagger (|\chi\rangle) := (|\psi\rangle, |\chi\rangle). \tag{23}$$

This is a linear map $\mathcal{H} \to \mathbb{C}$ as required, since the inner product is linear in its second argument. Likewise, since the inner product is *antilinear* in its first argument, we find:

$$(|\psi\rangle + \lambda |\chi\rangle)^\dagger = \langle\psi| + \lambda^* \langle\chi|, \tag{24}$$

thus $\dagger : \mathcal{H} \to \mathcal{H}^*$ is an *antilinear* map from kets to bras.

Theorem 5 (Riesz representation theorem)**.** *For every bra, $\langle\omega|$, there is a ket, $|\omega\rangle$, s.t.*

$$\langle\omega| = |\omega\rangle^\dagger. \tag{29}$$

In other words, the map $\dagger : \mathcal{H} \to \mathcal{H}^*$ is surjective (onto).

Since $\dagger$ is both surjective and injective, it is bijective by definition and therefore also an invertible map. Abusing notation, we write the inverse map $\dagger^{-1}$ from bras to kets using the same symbol, s.t.

$$\langle\omega|^\dagger \text{ is the unique ket } |\omega\rangle \text{ that satisfies } |\omega\rangle^\dagger = \langle\omega|. \tag{30}$$

Thus, $\left(|\psi\rangle^\dagger\right)^\dagger = |\psi\rangle$, and there is a direct, antilinear correspondence between bras and kets.[8]

Since $(|\psi\rangle, |\chi\rangle) = |\psi\rangle^\dagger (|\chi\rangle) = \langle\psi|\chi\rangle$ by definition of $\dagger : \mathcal{H} \to \mathcal{H}^*$, we can write the inner product $(|\psi\rangle, |\chi\rangle)$ equivalently as:

$$\langle\psi|\chi\rangle = \langle\chi|\psi\rangle^*. \tag{31}$$

The is the notation we will typically use henceforward, where now the antilinearity of the inner product in its first argument manifests itself as the antilinearity of the map $\dagger : |\psi\rangle \to \langle\psi|$.

Note that properties 1 and 2 from **Definition 3** together imply that

$$(|\chi\rangle + \lambda |\phi\rangle, |\psi\rangle) = (|\chi\rangle, |\psi\rangle) + \lambda^* (|\phi\rangle, |\psi\rangle), \tag{19}$$

i.e., the inner product is *antilinear* (*linear*) in its first (second) argument. This is called **sequilinear**.

Definition 4 (Antilinear)**.** A function $f : V \to W$ between two complex vector spaces is said to be **antilinear**, or **conjugate-linear**, if, $\forall\, x, y \in V$ and $\forall\, s \in \mathbb{C}$,

$$f(x+y) = f(x) + f(y) \qquad \text{(additivity)} \tag{20}$$

$$f(sx) = \bar{s} f(x) \qquad \text{(conjugate homogeneity)} \tag{21}$$

where $\bar{s}$ denotes the complex conjugate of s.

[7] $\| |\psi\rangle \| = 0 \iff |\psi\rangle = 0$.

Question 6. *Is $\dagger : \mathcal{H} \to \mathcal{H}^*$ injective (one-to-one)?*

Proof. Suppose that $|\psi\rangle^\dagger = |\chi\rangle^\dagger$. We find, $\forall\, |\xi\rangle \in \mathcal{H}$,

$$\left(|\psi\rangle^\dagger - |\chi\rangle^\dagger\right)(|\xi\rangle) = 0. \tag{25}$$

Taking the complex conjugate and applying the definition of $\dagger : \mathcal{H} \to \mathcal{H}^*$ as well as property 1 of **Definition 3**, we conclude that $\forall\, |\xi\rangle \in \mathcal{H}$

$$|\xi\rangle^\dagger (|\psi\rangle - |\chi\rangle) = 0. \tag{26}$$

By the Riesz representation theorem, we rewrite this as, $\forall\, \langle\zeta| \in \mathcal{H}^*$,

$$\langle\zeta| (|\psi\rangle - |\chi\rangle) = 0. \tag{27}$$

However, the only vector that is mapped to zero by every linear map $\mathcal{H} \to \mathbb{C}$ is the zero vector, so

$$|\psi\rangle - |\chi\rangle = 0 \implies |\psi\rangle = |\chi\rangle. \tag{28}$$

$\therefore$, $\dagger$ is injective. □

[8] Equivalently, $\dagger$ defines an *antilinear* isomorphism $\mathcal{H}^* \cong \mathcal{H}$.

Definition 7. The **adjoint** of an operator $\boldsymbol{X}$ is the operator $\boldsymbol{X}^\dagger$ defined by

$$\boldsymbol{X}^\dagger |\psi\rangle := (\langle\psi| \boldsymbol{X})^\dagger . \tag{32}$$

Definition 8. An operator $\boldsymbol{X}$ is **self-adjoint** if $\boldsymbol{X} = \boldsymbol{X}^\dagger$, **unitary** if $\boldsymbol{X}^\dagger \boldsymbol{X} = \boldsymbol{X}\boldsymbol{X}^\dagger = \boldsymbol{I}$, and more generally **normal** if $\boldsymbol{X}^\dagger \boldsymbol{X} = \boldsymbol{X}\boldsymbol{X}^\dagger$.[9]

Equivalently, an operator $\boldsymbol{U}$ is unitary iff it preserves inner products: $\boldsymbol{U}$ is unitary iff, $\forall\, |\psi\rangle, |\chi\rangle \in \mathcal{H}$,[10]

$$(\boldsymbol{U}|\psi\rangle, \boldsymbol{U}|\chi\rangle) = (|\psi\rangle, |\chi\rangle). \tag{33}$$

Definition 9. A ket $|\psi\rangle$ is **normalized** if $\langle\psi|\psi\rangle = 1$. Any non-zero ket can be normalized:

$$|\hat{\boldsymbol{\psi}}\rangle = \frac{1}{\sqrt{\langle\psi|\psi\rangle}} |\psi\rangle . \tag{34}$$

Definition 11. Kets $|\psi\rangle, |\chi\rangle$ are **orthogonal** if $\langle\psi|\chi\rangle = 0$.

Definition 12. An **orthonormal set** is a set of kets, $\{|1\rangle, \ldots, |N\rangle\}$, that are normalized and mutually orthogonal, i.e., s.t.

$$\langle i|j\rangle = \delta_{ij}. \tag{35}$$

This set is an **orthonormal basis** if any ket $|\psi\rangle$ can be written as a linear combination of its elements:

$$|\psi\rangle = \psi_1 |1\rangle + \ldots + \psi_N |N\rangle , \tag{36}$$

where ψ_i are the **components** of $|\psi\rangle$ in the basis $\{|1\rangle, \ldots, |N\rangle\}$.

When $\mathcal{H}$ is finite dimensional, then the size of any orthonormal set is at most the dimension of $\mathcal{H}$, $N \leq \dim \mathcal{H}$, and the set is an orthonormal *basis* iff $N = \dim \mathcal{H}$. Such an orthonormal set is said to be **complete** (i.e., orthonormal basis $\iff$ complete).

Given an orthonormal basis $\{|1\rangle, \ldots, |N\rangle\}$, we can extract the components of $|\psi\rangle$ by its inner product with each basis element and using $\langle i|h\rangle = \delta_{ij}$:

$$\langle i|\psi\rangle = \sum_{j=1}^{N} \psi_j \langle i|j\rangle = \psi_i. \tag{37}$$

Substituting what we found in Equation 37 into Equation 36, we obtain

$$\begin{aligned} |\psi\rangle &= \langle 1|\psi\rangle \cdot |1\rangle + \ldots + \langle N|\psi\rangle \cdot |N\rangle \\ &= |1\rangle \langle 1|\psi\rangle + \ldots + |N\rangle \langle N|\psi\rangle \\ &= (|1\rangle\langle 1| + \ldots + |N\rangle\langle N|)\, |\psi\rangle , \end{aligned} \tag{38}$$

where we first move the scalar factors to the right and then reinterpret the result as an outer product acting on $|\psi\rangle$. Since this equality holds for any $|\psi\rangle$, we read off

$$\boldsymbol{I} = |1\rangle\langle 1| + \ldots + |N\rangle\langle N| = \sum_{i=1}^{N} |i\rangle\langle i| , \tag{39}$$

which is known as the **completeness relation** for the orthonormal basis $\{|1\rangle, \ldots, |N\rangle\}$. An orthonormal set is complete iff it satisfies such a completeness relation.

[9] Both self-adjoint and unitary operators are normal.

[10] Equation 33 can equivalently be shown using our new notation since $(\boldsymbol{U}|\psi\rangle, \boldsymbol{U}|\chi\rangle) = \langle\psi|U^\dagger U|\chi\rangle = \langle\psi|\chi\rangle = (|\psi\rangle, |\chi\rangle)$, where $U^\dagger U = I$ by **Definition 8**.

Question 10. *Show the following:*

1. $(|\psi\rangle, \boldsymbol{X}|\chi\rangle) = (\boldsymbol{X}^\dagger|\psi\rangle, |\chi\rangle)$.

   Proof. First, note that **Definition 7** $\implies \langle\psi|\boldsymbol{X} = (\boldsymbol{X}^\dagger|\psi\rangle)^\dagger$. $(|\psi\rangle, \boldsymbol{X}|\chi\rangle) = \langle\psi|\boldsymbol{X}|\chi\rangle = (\boldsymbol{X}^\dagger|\psi\rangle)^\dagger|\chi\rangle = \boldsymbol{X}\langle\psi|\chi\rangle = (\boldsymbol{X}^\dagger|\psi\rangle, |\chi\rangle)$. □

2. $(\langle\psi|\boldsymbol{X}|\chi\rangle)^* = \langle\chi|\boldsymbol{X}^\dagger|\psi\rangle$.

   Proof. $(\langle\psi|\boldsymbol{X}|\chi\rangle)^* = (|\psi\rangle, \boldsymbol{X}|\chi\rangle)^* = (\boldsymbol{X}^\dagger|\psi\rangle, |\chi\rangle)^*$ (by part (1)) $= (\boldsymbol{X}\langle\psi|\chi\rangle)^* = \boldsymbol{X}^*\langle\chi|\psi\rangle = \langle\chi|\boldsymbol{X}^\dagger|\psi\rangle$. □

3. $(\boldsymbol{X}\boldsymbol{Y})^\dagger = \boldsymbol{Y}^\dagger\boldsymbol{X}^\dagger$.

   Proof. $\left\langle\psi\middle|(\boldsymbol{X}\boldsymbol{Y})^\dagger\middle|\chi\right\rangle = \left(|\psi\rangle, (\boldsymbol{X}\boldsymbol{Y})^\dagger|\chi\rangle\right) = (\boldsymbol{X}\boldsymbol{Y}|\psi\rangle, |\chi\rangle)$ (by part (1)) $= (\boldsymbol{Y}|\psi\rangle, \boldsymbol{X}^\dagger|\chi\rangle)$ (by part (1) again) $= \left(|\psi\rangle, \boldsymbol{Y}^\dagger\boldsymbol{X}^\dagger|\chi\rangle\right)$ (by part (1) again) $= \left\langle\psi\middle|\boldsymbol{Y}^\dagger\boldsymbol{X}^\dagger\middle|\chi\right\rangle$. $\therefore$, $(\boldsymbol{X}\boldsymbol{Y})^\dagger = \boldsymbol{Y}^\dagger\boldsymbol{X}^\dagger$. □

4. $(\boldsymbol{X} + \lambda\boldsymbol{Y})^\dagger = \boldsymbol{X}^\dagger + \lambda^*\boldsymbol{Y}^\dagger$.

   Proof. $\left(|\psi\rangle, (\boldsymbol{X} + \lambda\boldsymbol{Y})^\dagger|\chi\rangle\right) = ((\boldsymbol{X} + \lambda\boldsymbol{Y})|\psi\rangle, |\chi\rangle) = (\boldsymbol{X}|\psi\rangle, |\chi\rangle) + (\lambda\boldsymbol{Y}|\psi\rangle, |\chi\rangle)$ (by (2) of **Definition 3**) $= (|\psi\rangle, \boldsymbol{X}^\dagger|\chi\rangle) + \left(|\psi\rangle, \lambda^*\boldsymbol{Y}^\dagger|\chi\rangle\right) = \left(|\psi\rangle, \left(\boldsymbol{X}^\dagger + \lambda^*\boldsymbol{Y}^\dagger\right)|\chi\rangle\right)$. $\therefore$, $(\boldsymbol{X} + \lambda\boldsymbol{Y})^\dagger = \boldsymbol{X}^\dagger + \lambda^*\boldsymbol{Y}^\dagger$. □

5. $(\langle\psi|\omega\rangle)^\dagger = \langle\omega|\psi\rangle$.

   Proof. Recall $\langle\psi|\omega\rangle \in \mathbb{C}$. $\forall\, c \in \mathbb{C}$, $c^\dagger = c^*$. $\therefore$, $(\langle\psi|\omega\rangle)^\dagger = (\langle\psi|\omega\rangle)^* = \langle\omega|\psi\rangle$ by Equation 31. □

2025-02-06: Class Canceled (Snow Day)

2025-02-11: Finite Dimensional Hilbert Spaces (cont.)

Projection Operators

Definition 13. A **projection operator $\boldsymbol{\Pi}$** is one that satisfies $\boldsymbol{\Pi}^2 = \boldsymbol{\Pi}$.[11]

[11] I.e., projection operators are idempotent!

Definition 15. An **orthogonal projector** is a self-adjoint operator, $\boldsymbol{\Pi}^2 = \boldsymbol{\Pi} = \boldsymbol{\Pi}^\dagger$.

One can then show (via the spectral theorem, see **Theorem 18**) that *any* orthogonal projector takes the form shown in **Question 14**.

Question 14. *If $\{|1\rangle, \ldots, |k\rangle\}$ is an orthonormal set, show that $\boldsymbol{\Pi} = \sum_{i=1}^{k} |i\rangle\langle i|$ is an orthogonal projector.*

Proof. □

Definition 16. An **eigenket** of an operator $\boldsymbol{X}$ is a non-zero ket $|\psi\rangle$ s.t.

$$\boldsymbol{X}|\psi\rangle = \lambda|\psi\rangle \tag{40}$$

for some **eigenvalue** $\lambda \in \mathbb{C}$.

Spectral Theorem

Proposition 17. *If $\boldsymbol{X}$ is **self-adjoint** $\left(\boldsymbol{X} = \boldsymbol{X}^\dagger\right)$ then*

1. *its eigenvalues are real, and*
2. *its eigenvectors with distinct eigenvalues are orthogonal.*

Proof. 1. $\boldsymbol{X}|\psi\rangle = \lambda|\psi\rangle \implies (\boldsymbol{X}|\psi\rangle)^\dagger = \lambda^*\langle\psi|$. Note that $(\boldsymbol{X}|\psi\rangle)^\dagger = \left(\boldsymbol{X}^\dagger|\psi\rangle\right)^\dagger$ (since $\boldsymbol{X}$ is **self-adjoint**) $= \langle\psi|\boldsymbol{X}$ (by **Definition 7**), so $\langle\psi|\boldsymbol{X} = \lambda^*\langle\psi| \implies \langle\psi|\boldsymbol{X}|\psi\rangle = \lambda^*\langle\psi|\psi\rangle$. Also, $\langle\psi|\boldsymbol{X}|\psi\rangle = \lambda\langle\psi|\psi\rangle$ (by **Definition 16**), $\implies \lambda\langle\psi|\psi\rangle = \lambda^*\langle\psi|\psi\rangle \implies$ either $\lambda \in \mathbb{R}$ or $\langle\psi|\psi\rangle = 0$. $|\psi\rangle$ is non-zero by **Definition 16**, so $\lambda \in \mathbb{R}$ must be true.

2. Let $\boldsymbol{X}|\psi_1\rangle = \lambda_1|\psi_1\rangle$ and $\boldsymbol{X}|\psi_2\rangle = \lambda_2|\psi_2\rangle$. $\boldsymbol{X}|\psi_2\rangle = \lambda_2|\psi_2\rangle \implies \langle\psi_1|\boldsymbol{X}|\psi_2\rangle = \lambda_2\langle\psi_1|\psi_2\rangle$, call this relation $(*)$. Also, $\boldsymbol{X}|\psi_1\rangle = \lambda_1|\psi_1\rangle \implies \langle\psi_2|\boldsymbol{X}|\psi_1\rangle = \lambda_1\langle\psi_2|\psi_1\rangle$. Hence, $(\langle\psi_2|\boldsymbol{X}|\psi_1\rangle)^* = (\lambda_1\langle\psi_2|\psi_1\rangle)^* \implies \langle\psi_1|\boldsymbol{X}^\dagger|\psi_2\rangle = \lambda_1^*(\langle\psi_2|\psi_1\rangle)^*$ (by (2) in **Question 10**) $\implies \langle\psi_1|\boldsymbol{X}|\psi_2\rangle = \lambda_1\langle\psi_1|\psi_2\rangle$ since $\boldsymbol{X}$ is self-adjoint, $\lambda_1 \in \mathbb{R}$ by (1), and by Equation 31. Relating this most recent relation back to relation $(*)$, we see that either $\lambda_1 = \lambda_2$ or $\langle\psi_1|\psi_2\rangle = 0$. $\lambda_1 = \lambda_2$ is not true since these are distinct by definition. $\therefore$, $\langle\psi_1|\psi_2\rangle = 0$ must be true.

□

More generally,

Theorem 18 (Spectral Theorem)**.** *For any self-adjoint operator $\boldsymbol{X} = \boldsymbol{X}^\dagger$, $\exists$ an orthonormal basis $\{|1\rangle, \ldots, |N\rangle\}$ and $\lambda_1, \ldots, \lambda_N \in \mathbb{R}$ s.t.*

$$\boldsymbol{X} = \sum_{i=1}^{N} \lambda_i |i\rangle\langle i| . \tag{41}$$

The proof of this theorem is non-trivial, especially in the infinite-dimensional case (where the statement has to be modified somewhat from that given above).

Note that the orthonormal basis produced by the spectral theorem consists of $\boldsymbol{X}$ eigenvectors, $\boldsymbol{X}|i\rangle = \lambda_i |i\rangle$, with eigenvalues λ_i. Thus, a self-adjoint operator has an *orthonormal basis of eigenvectors*.

The spectral theorem has a couple important generalizations:

1. It applies equally well to *normal* operators (satisfying $\boldsymbol{X}\boldsymbol{X}^\dagger = \boldsymbol{X}^\dagger\boldsymbol{X}$), except that the eigenvalues become complex, $\lambda_i \in \mathbb{C}$. (A normal operator is self-adjoint iff its eigenvalues are all real.)

2. If $\boldsymbol{X}$, $\boldsymbol{Y}$ are self-adjoint operators that commute, $[\boldsymbol{X}, \boldsymbol{Y}] = 0$, then $\exists$ a **simultaneous eigenbasis**:

$$\boldsymbol{X} = \sum_{i=1}^{N} \lambda_i^X |i\rangle\langle i|\,,\ \boldsymbol{Y} = \sum_{i=1}^{N} \lambda_i^Y |i\rangle\langle i|\,,\ \left(\text{s.t. } \langle i|\,j\rangle = \delta_{ij},\ \sum_{i=1}^{N} |i\rangle\langle i| = \boldsymbol{I}\right). \tag{47}$$

   (The statements in parentheses express the fact that $\{|1\rangle, \ldots, |N\rangle\}$ is an orthonormal basis in equations.)

   More generally, for any set of self-adjoint (or normal) operators $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_k$, all of which commute, $[\boldsymbol{X}_i, \boldsymbol{X}_j] = 0$, there is a simultaneous eigenbasis.

Degeneracies

Note that the basis in Equation 47 is not necessarily the same as that in Equation 41 because $\boldsymbol{X}$ may have multiple eigenvectors with the same eigenvalue. Such eigenvalues are said to be **degenerate**. If, for instance, $\lambda_1 = \lambda_2$, then:

$$\lambda_1 |1\rangle\langle 1| + \lambda_2 |2\rangle\langle 2| = \lambda_1 (|1\rangle\langle 1| + |2\rangle\langle 2|) = \lambda_1 \left(|1'\rangle\langle 1'| + |2'\rangle\langle 2'|\right), \tag{48}$$

where $|1'\rangle = a|1\rangle + b|2\rangle$, $|2'\rangle = c|1\rangle + d|2\rangle$ for some $a, b, c, d \in \mathbb{C}$ s.t.

$$\mathcal{U} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{49}$$

is a unitary matrix (satisfying $\mathcal{U}^\dagger\mathcal{U} = \mathcal{I}$), ensuring that $\{|1'\rangle, |2'\rangle\}$ is an orthonormal set. Thus, if $\boldsymbol{X}$ has degenerate eigenvalues then its eigenbasis is not unique. The vector space generated by all $\boldsymbol{X}$ eigenvectors with the same eigenvalue is known as a **degenerate eigenspace**, and any unitary rotation on the eigenvectors generating this eigenspace produces another eigenbasis.

However, even if both $\boldsymbol{X}$ and $\boldsymbol{Y}$ have degenerate eigenvalues, their simultaneous eigenbasis may be unique, as in the following example:

$$\boldsymbol{X} = |1\rangle\langle 1| + |2\rangle\langle 2| - |3\rangle\langle 3|\,,\quad \boldsymbol{Y} = |1\rangle\langle 1| - |2\rangle\langle 2| - |3\rangle\langle 3|\,. \tag{50}$$

A linear combination $|1'\rangle = a|1\rangle + b|2\rangle$ of $|1\rangle$, $|2\rangle$ is still an $\boldsymbol{X} = +1$ eigenket, but for $a, b \neq 0$ it is *not* a $\boldsymbol{Y}$ eigenket. Likewise, a linear combination of $|2\rangle$, $|3\rangle$ remains a $\boldsymbol{Y} = -1$ eigenket, but is no longer an $\boldsymbol{X}$ eigenket.

Question 19.

1. *Show that every eigenvalue of $\boldsymbol{X}$ appears in the spectral decomposition (Equation 41).*

   Proof. Equation 41 tells us that $\boldsymbol{X}$ acting on some arbitrary eigenvector $|j\rangle \in \{|1\rangle, \ldots, |N\rangle\}$ is equivalent to

$$\begin{aligned} \boldsymbol{X}|j\rangle &= \left(\sum_{i=1}^{N} \lambda_i |i\rangle\langle i|\right)|j\rangle \\ &= \sum_{i=1}^{N} \lambda_i |i\rangle\,(\langle i|\,j\rangle) \\ &= \sum_{i=1}^{N} \lambda_i |i\rangle\,\delta_{i,j} \\ &= \lambda_j |j\rangle\,. \end{aligned} \tag{42}$$

   Since $|j\rangle$ can be *any* eigenvector $\in \{|1\rangle, \ldots, |N\rangle\}$, this means that every eigenvalue of $\boldsymbol{X}$ appears in this decomposition. □

2. *Must every eigenvector of $\boldsymbol{X}$ appear in the basis $\{|1\rangle, \ldots, |N\rangle\}$? If so, prove it. If not, provide a counterexample.*

   Proof. Let $\boldsymbol{X}|\psi\rangle = \mu|\psi\rangle$ where $|\psi\rangle \neq 0$ is any arbitrary eigenvector of $\boldsymbol{X}$ and μ is its associated eigenvalue. Since $\{|1\rangle, \ldots, |N\rangle\}$ forms a complete set of the Hilbert space, any vector can be written as a linear combination of this basis. In particular,

$$|\psi\rangle = \sum_{i=1}^{N} \psi_i |i\rangle\,, \tag{43}$$

   where $\psi_i \in \mathbb{C}$ are the components of $|\psi\rangle$ in the basis $\{|1\rangle, \ldots, |N\rangle\}$. This means that

$$\boldsymbol{X}|\psi\rangle = \sum_{i=1}^{N} \psi_i \lambda_i |i\rangle \tag{44}$$

   by Equation 42. Also recall that

$$\boldsymbol{X}|\psi\rangle = \mu|\psi\rangle = \mu \sum_{i=1}^{N} \psi_i |i\rangle\,. \tag{45}$$

   Equations 44 and 45 together tell us that

$$\psi_i (\mu - \lambda_i) = 0 \tag{46}$$

   since $\{|1\rangle, \ldots, |N\rangle\}$ is an orthogonal eigenbasis so all eigenvectors must have a coefficient of 0 (e.g., there can be no interference between $|1\rangle$ and $|2\rangle$). By definition, $|\psi\rangle \neq 0$ so $\psi_i \neq 0$. $\implies \mu = \lambda_i$ which means that $|\psi\rangle \in \{|1\rangle, \ldots, |N\rangle\}$. □

In general, a simultaneous eigenbasis will be unique (up to multiplying each basis vector by a phase factor) when no two eigenvectors have the same *eigenvalue list*, e.g.,

$$\left(\lambda^X, \lambda^Y\right) = (1,1)\,,\ (1,-1)\,,\ (-1,-1)\,, \tag{51}$$

for the three eigenvectors in Equation 50, so no two of them can mix while remaining eigenkets of both $\boldsymbol{X}$ and $\boldsymbol{Y}$.

When there is degeneracy, we can write the spectral decomposition (Equation 41) in a unique way by recognizing, e.g., $|1\rangle\langle 1| + |2\rangle\langle 2|$ in the spectral decomposition of $\boldsymbol{X}$ in Equation 50 as an orthogonal projector. More generally, a **complete set of orthogonal projectors**[12] is a set of operators $\{\boldsymbol{\Pi}_1, \ldots, \boldsymbol{\Pi}_k\}$ satisfying

$$\boldsymbol{\Pi}_i^\dagger = \boldsymbol{\Pi}_i, \quad \boldsymbol{\Pi}_i\boldsymbol{\Pi}_j = \delta_{ij}\boldsymbol{\Pi}_i, \quad \sum_{i=1}^{k} \boldsymbol{\Pi}_i = \boldsymbol{I}. \tag{52}$$

Thus, the operators in this set are orthogonal projectors, each orthogonal to the others, s.t. their sum is the identity operator. Note that some of these conditions follow from the others (see **Question 20**).

For each degenerate eigenspace $|1\rangle$, ..., $|k\rangle$ in the spectral decomposition (Equation 41), there is an orthogonal projector

$$\boldsymbol{\Pi} = |1\rangle\langle 1| + \ldots + |k\rangle\langle k|\,. \tag{53}$$

The spectral decomposition (Equation 41) can thus be written *uniquely* as

$$\boldsymbol{X} = \sum_{i=1}^{n} \lambda_i \boldsymbol{\Pi}_i, \tag{54}$$

where $\lambda_i \neq \lambda_j$ are the *distinct* $\boldsymbol{X}$ eigenvalues and $\boldsymbol{\Pi}_i$ are a complete set of orthogonal projectors, each one projecting onto one of the degenerate eigenspaces of $\boldsymbol{X}$, with $\mathrm{rank}(\boldsymbol{\Pi}_i)$ equal to the dimension of the eigenspace in question. The orthogonal projectors $\boldsymbol{\Pi}_i$ thus defined are known as **eigenspace projectors** (or "Riesz projectors").

Question 20. *Let $\{\boldsymbol{\Pi}_1, \ldots, \boldsymbol{\Pi}_k\}$ be a set of orthogonal projectors summing to the identity, $\sum_{i=1}^{k} \boldsymbol{\Pi}_i = \boldsymbol{I}$. Show that $\boldsymbol{\Pi}_i\boldsymbol{\Pi}_j = 0$ for $i \neq j$, hence $\{\boldsymbol{\Pi}_1, \ldots, \boldsymbol{\Pi}_k\}$ is a complete set of orthogonal projectors.*

(HINT: Write out $\left(\boldsymbol{\Pi}_i\boldsymbol{\Pi}_j\right)^\dagger \boldsymbol{\Pi}_i\boldsymbol{\Pi}_j$ and use the properties of non-negative operators.)

Proof. □

[12] This can also be viewed as an example of a more general concept known as **projection-valued measure** (PVM).

Question 21. *Let $\boldsymbol{X}$, $\boldsymbol{Y}$ be normal operators with eigenspace projectors $\boldsymbol{\Pi}_1^X$, ..., $\boldsymbol{\Pi}_m^X$ and $\boldsymbol{\Pi}_1^Y$, ..., $\boldsymbol{\Pi}_n^Y$, respectively.*

1. *Show that $\boldsymbol{X}$, $\boldsymbol{Y}$ commute iff their eigenspace projectors commute.*

   Proof. □

2. *When $\boldsymbol{X}$, $\boldsymbol{Y}$ commute, show that $\boldsymbol{\Pi}_{i\alpha}^{XY} = \boldsymbol{\Pi}_i^Y\boldsymbol{\Pi}_\alpha^Y$ is a complete set of orthogonal projectors projecting onto the simultaneous eigenspaces.*

   Proof. □

Unitary Transformations

Recall that an operator $\boldsymbol{U}$ is unitary if $\boldsymbol{U}\boldsymbol{U}^\dagger = \boldsymbol{U}^\dagger\boldsymbol{U} = \boldsymbol{I}$. Equivalently, $\boldsymbol{U}$ is unitary iff it preserves inner products.

Theorem 22. *$\boldsymbol{U}$ is unitary iff, $\forall\, |\psi\rangle, |\chi\rangle \in \mathcal{H}$,*[13]

$$(\boldsymbol{U}|\psi\rangle, \boldsymbol{U}|\chi\rangle) = (|\psi\rangle, |\chi\rangle). \tag{55}$$

Suppose that $\{|1\rangle, \ldots, |N\rangle\}$ is an orthonormal basis. Then: [14]

$$|i'\rangle := \boldsymbol{U}|i\rangle \text{ satisfies } \langle i'|j'\rangle = \left\langle i \middle| \boldsymbol{U}^\dagger\boldsymbol{U} \middle| j \right\rangle = \delta_{ij}. \tag{56}$$

By inserting the completeness relation for the original basis, we obtain a decomposition of $\boldsymbol{U}$ in terms of these two bases:

$$\boldsymbol{U} = \boldsymbol{U}\sum_{i=1}^{N} |i\rangle\langle i| = \sum_{i=1}^{N} |i'\rangle\langle i|. \tag{57}$$

In fact, for *any* two orthonormal bases $\{|1\rangle, \ldots, |N\rangle\}$, $\{|1'\rangle, \ldots, |N'\rangle\}$ there is an operator $\boldsymbol{U} := \sum_{i=1}^{n} |i'\rangle\langle i|$ that relates them.[15]

Definition 23. The **unitary transform** of an operator $\boldsymbol{X}$ by a unitary operator $\boldsymbol{U}$ is defined as

$$\boldsymbol{X}_U := \boldsymbol{U}\boldsymbol{X}\boldsymbol{U}^\dagger. \tag{59}$$

If $[\boldsymbol{X}, \boldsymbol{U}] = 0$ then $\boldsymbol{X}_U = \boldsymbol{X}$, i.e., $\boldsymbol{X}$ is **$\boldsymbol{U}$-invariant**. More generally, $\boldsymbol{X}_U \neq \boldsymbol{X}$, but $\boldsymbol{X}_U$ has very similar properties to $\boldsymbol{X}$ (see **Question 24**).

Question 24. *Show the following:*

1. *$\boldsymbol{X}_U$ is (self-adjoint | unitary | normal) iff $\boldsymbol{X}$ is (self-adjoint | unitary | normal).*

   Proof. □

2. *$|\psi\rangle$ is an $\boldsymbol{X}$ eigenvector iff $\boldsymbol{U}|\psi\rangle$ is an $\boldsymbol{X}_U$ eigenvectors, where the eigenvalues are the same.*[16]

   Proof. □

3. $\boldsymbol{X}_U\boldsymbol{Y}_U = (\boldsymbol{X}\boldsymbol{Y})_U$.

   Proof. □

Operators related by a unitary transformation are **unitarily equivalent**. Self-adjoint (or normal) operators are unitarily equivalent iff they have the same **spectrum** of eigenvalues and degeneracies.

Note that a unitary transformation is what a mathematician would call an *automorphism* on the Hilbert space $\mathcal{H}$. They close to what we might call a **symmetry** in physics, but actual physical symmetries will satisfy a further dynamical requirement, as we will see.

[13] This was already stated in Equation 33, but it's worth re-stating in this section.

[14] Equation 56 tells us that $\{|1'\rangle, \ldots, |N'\rangle\}$ is another orthonormal basis!

[15] This operator is unitary since

$$\begin{aligned}\boldsymbol{U}^\dagger\boldsymbol{U} &= \sum_{i=1}^{N} |i\rangle\langle i'| \sum_{j=1}^{N} |j'\rangle\langle j| \\ &= \sum_{i,j=1}^{N} |i\rangle\, \delta_{ij}\, \langle j| \\ &= \sum_{i=1}^{N} |i\rangle\langle i| \\ &= \boldsymbol{I},\end{aligned} \tag{58}$$

and likewise $\boldsymbol{U}\boldsymbol{U}^\dagger = \boldsymbol{I}$ by a similar calculation.

[16] In the special case where $\boldsymbol{X}$ is self-adjoint and $\boldsymbol{U}$-invariant then $\boldsymbol{U}$ only acts within each degenerate eigenspace of $\boldsymbol{X}$.

Matrices

Given an orthonormal basis, we can use the associated completeness relation to decompose bras and kets into their components:

$$|\psi\rangle = \sum_{i=1}^{N} |i\rangle \langle i|\psi\rangle = \sum_{i=1}^{N} \langle i|\psi\rangle |i\rangle\,, \quad \langle\chi| = \sum_{i=1}^{N} \langle i|\chi\rangle \langle i|\,, \tag{60}$$

where $\langle i|\psi\rangle$ and $\langle i|\chi\rangle$ are the components of $|\psi\rangle$ and $|\chi\rangle$, respectively. Likewise, for a linear operator $\boldsymbol{X}$:

$$\boldsymbol{X} = \sum_{i,j=1}^{N} |i\rangle\langle i|\boldsymbol{X}|j\rangle\langle j| = \sum_{i,j=1}^{N} \langle i|\boldsymbol{X}|j\rangle |i\rangle\langle j|\,, \tag{62}$$

where $\langle i|\boldsymbol{X}|j\rangle$ are the matrix elements of $\boldsymbol{X}$. Similarly, extracting the components of $\boldsymbol{X}|\psi\rangle$, we find

$$\langle i|\boldsymbol{X}|\psi\rangle = \sum_{j=1}^{N} \langle i|\boldsymbol{X}|j\rangle \langle j|\psi\rangle\,, \tag{63}$$

which the matrix product of $\langle i|\boldsymbol{X}|j\rangle$ with the column vector $\langle j|\psi\rangle$. Dealing with bras analogously, we see that we can realize the entire Hilbert space concretely as:

$$|\psi\rangle \to \begin{pmatrix} \psi_1 \\ \vdots \\ \psi_N \end{pmatrix}, \quad \langle\chi| \to \begin{pmatrix} \chi_1 & \cdots & \chi_N \end{pmatrix}, \quad \boldsymbol{X} \to \begin{pmatrix} X_{11} & \cdots & X_{1N} \\ \vdots & \ddots & \vdots \\ X_{N1} & \cdots & X_{NN} \end{pmatrix}. \tag{64}$$

Moreover, both the conjugate map and the adjoint correspond to taking the conjugate transpose of the matrix in question.[17]

Despite the pleasing simplicity of this, it is important not to conflate the *components* of each object with bras, kets, and operators themselves. For instance, suppose that ψ_i, X_{ij}, etc., change by a unitary transformation. This could mean different things. Either

1. $|\psi\rangle \to \boldsymbol{U}|\psi\rangle$, $\boldsymbol{X} \to \boldsymbol{U}\boldsymbol{X}\boldsymbol{U}^\dagger$, etc., undergo a unitary transformation with the basis $\{|1\rangle, \ldots, |N\rangle\}$ held fixed[18] or
2. $|\psi\rangle$, $\boldsymbol{X}$, etc., are *fixed* while the basis $|i\rangle \to \boldsymbol{U}^\dagger|i\rangle$ undergoes an (inverse) unitary transformation.[19]

Physically, these are very different: either (1) the objects (kets, operators, etc.) actually change or (2) the objects do not change at all, but our *perspective* of them changes. Combining 1 and 2, it is even possible for both the objects and the basis to change while leaving the components invariant.[20]

Working with matrices makes it easier to define certain operations. In particular, the **trace** and the **determinant** of a linear operator $\boldsymbol{X}$ are defined as the trace and determinant of the corresponding matrix $\langle i|\boldsymbol{X}|j\rangle$.

Question 25. *Show that*

When we multiply operators $\boldsymbol{X}$, $\boldsymbol{Y}$ together, the matrix elements of the product $\boldsymbol{X}\boldsymbol{Y}$ are:

$$\langle i|\boldsymbol{X}\boldsymbol{Y}|j\rangle = \sum_{k=1}^{N} \langle i|\boldsymbol{X}|k\rangle \langle k|\boldsymbol{Y}|j\rangle\,, \tag{61}$$

which is the usual matrix multiplication rule.

[17] Hence the common notation $(\ldots)^\dagger$ for all of these operations!

[18] This is called an **active transformation**.

[19] This is called a **passive transformation**.

[20] This is still an "active" transformation, as the actual objects change. It differs from the active transformation described in 1 by a passive transformation.

1. $\boldsymbol{X}$ *and* $\operatorname{Det}\boldsymbol{X}$ *so defined are independent of the choice of orthonormal basis.*

   Proof. □

2. $(|\psi\rangle\langle\omega|) = \langle\omega|\,\psi\rangle$.[21]

   Proof. □

[21] One could also take this property to define the trace.

2025-02-13: The Postulates of Quantum Mechanics

We start by seeing how the postulates of quantum mechanics emerge from the *Stern-Gerlach* thought experiments, which are described by a quantum two-state system (or **qubit**). We will then generalize the postulates to any finite-dimensional quantum system.

Analysis of the Stern-Gerlach Thought Experiments

To analyze the thought experiments, we assign a quantum state vector $|\psi\rangle \in \mathcal{H}$ to each atomic beam in our device, where the properties of the Hilbert space $\mathcal{H}$ remain to be determined. We normalize these vectors so that the squared norm of $|\psi\rangle$ equals the *number* of atoms that would be intercepted by a plate blocking the beamline in question:

$$N_{\textbf{beam}} = \langle\psi|\,\psi\rangle\,. \tag{65}$$

The original Stern-Gerlach experiment takes an *unpolarized*[22] beam of N atoms and splits it into two polarized components: $\sqrt{\frac{N}{2}}\,|+\hat{z}\rangle$ and $\sqrt{\frac{N}{2}}\,|-\hat{z}\rangle$, each with 50% of the beam, see Figure 4.

Here we take the quantum states of the two output beams (up to normalization) to *define* the state vectors $|+\hat{z}\rangle$, $|-\hat{z}\rangle$, which we take to be normalized:

$$\langle+\hat{z}|+\hat{z}\rangle = \langle-\hat{z}|-\hat{z}\rangle = 1. \tag{66}$$

Thus, the prefactor $\sqrt{\frac{N}{2}}$ in each beam corresponds to the fact that the beam contains $\left(\sqrt{\frac{N}{2}}\right)^2 = \frac{N}{2}$ atoms.

Now consider a *polarized* beam into the apparatus which takes in a state $|\psi_{\text{in}}\rangle$ and gives the outputs of $|\psi_+\rangle$ and $|\psi_-\rangle$, see Figure 5. The results of the first though experiment tell us that:[23]

1. If $|\psi_{\text{in}}\rangle = |+\hat{z}\rangle$ then $|\psi_+\rangle = |+\hat{z}\rangle$, $|\psi_-\rangle = 0$.
2. If $|\psi_{\text{in}}\rangle = |-\hat{z}\rangle$ then $|\psi_+\rangle = 0$, $|\psi_-\rangle = |-\hat{z}\rangle$.

Since the number of atome coming *out* of the device should scale linearly with the number of atoms coming *in*, the final state should be *linear* in the inital state. Therefore, for a *quantum superposition* of $|+\hat{z}\rangle$, $|-\hat{z}\rangle$ (i.e., $a\,|+\hat{z}\rangle + b\,|-\hat{z}\rangle$) coming in, the output should be the same, as seen in Figure 6.

Thus, the number of atoms coming out up and down are:

$$N_+ = |a|^2\,,\quad N_- = |b|^2\,. \tag{67}$$

Let us compare this with the number of atoms entering the device:

$$\begin{aligned} N &= \left(a^*\,\langle+\hat{z}| + b^*\,\langle-\hat{z}|\right)\left(a\,|+\hat{z}\rangle + b\,|-\hat{z}\rangle\right) \\ &= |a|^2 + |b|^2 + a^*b\,\langle+\hat{z}|-\hat{z}\rangle + b^*a\,\langle-\hat{z}|+\hat{z}\rangle \\ &= |a|^2 + |b|^2 + a^*b\,\langle+\hat{z}|-\hat{z}\rangle + \left(a^*b\,\langle+\hat{z}|-\hat{z}\rangle\right)^* \\ &= |a|^2 + |b|^2 + 2\,\text{Re}\,\left[a^*b\,\langle+\hat{z}|-\hat{z}\rangle\right]. \end{aligned} \tag{68}$$

[22] The quantum state of the unpolarized beam can be described by a density matrix, but we defer this for later.

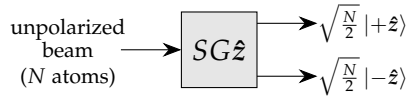

Figure 4: Original Stern-Gerlach experiment.

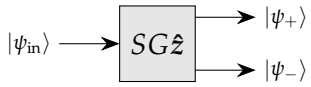

Figure 5: Polarized Stern-Gerlach experiment.

[23] To be precise, the Stern-Gerlach device might introduce phase factors, but this will not matter until we consider quantum interference, so we ignore it for now.

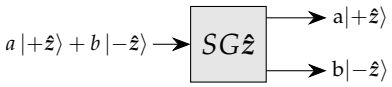

Figure 6: Superposition through the Stern-Gerlach device.

The number of atoms coming out had better be the same as the number going in! This implies that $N = N_+ + N_-$, so that $\mathrm{Re}\left[a^* b \left\langle +\hat{z}\right| - \hat{z}\right\rangle] = 0$. Moreover, this should be true for *any* quantum superposition, i.e., for any $a, b \in \mathbb{C}$. This is only possible if

$$\left\langle +\hat{z}\right| - \hat{z}\right\rangle = 0. \tag{69}$$

This illustrates a generic principle: *distinct* measurement outcomes correspond to *orthogonal* quantum states.[24] This is required to avoid violating the laws of probability.

Therefore $\left|+\hat{z}\right\rangle, \left|-\hat{z}\right\rangle \in \mathcal{H}$ form an orthonormal set, implying that $\dim \mathcal{H} \geqslant 2$. We now consider the *minimal* scenario where $\dim \mathcal{H} = 2$ and see whether we can explain all the experiments this was.[25]

Thus, by assumption $\left\{\left|+\hat{z}\right\rangle, \left|-\hat{z}\right\rangle\right\}$ is an orthonormal basis, and the effect of the SG device is to split the beam into orthogonal components: $\mathbf{\Pi}_{+\hat{z}} \left|\psi\right\rangle$ and $\mathbf{\Pi}_{-\hat{z}} \left|\psi\right\rangle$ (see Figure 7), where $\mathbf{\Pi}_{\pm\hat{z}} = \left|\pm\hat{z}\right\rangle\!\left\langle\pm\hat{z}\right|$ is a complete set of orthogonal projectors. Recall that the two beams correspond to the atomic angular momenta $\pm\frac{\hbar}{2}$. We represent this fact in Hilbert space language by defining the self-adjoint operator:

$$S_z := \frac{\hbar}{2}\left(\left|+\hat{z}\right\rangle\!\left\langle+\hat{z}\right| - \left|-\hat{z}\right\rangle\!\left\langle-\hat{z}\right|\right), \tag{70}$$

where it is convention to use the letter $\underline{S}$ for this type of angular momentum (due to the intrinsic spin of one of the electrons).

This is an **observable**, a self-adjoint operator representing a measurable quantity, whose *eigenvalues* are possible outcomes and whose *eigenvectors* are the quantum states that can result after the measurement is made. Measuring many identically prepared atoms in this way, the *average* of the measure S_z values is

$$\begin{aligned}
\left\langle S_z\right\rangle &= \frac{1}{N}\left[\frac{\hbar}{2}N_+ - \frac{\hbar}{2}N_-\right] \\
&= \frac{1}{\left\langle\psi\right|\left.\psi\right\rangle}\left[\frac{\hbar}{2}\left\langle\psi\right|\left.+\hat{z}\right\rangle\left\langle+\hat{z}\right|\left.\psi\right\rangle - \frac{\hbar}{2}\left\langle\psi\right|\left.-\hat{z}\right\rangle\left\langle-\hat{z}\right|\left.\psi\right\rangle\right] \\
&= \frac{1}{\left\langle\psi\right|\left.\psi\right\rangle}\cdot\left\langle\psi\right|\left[\frac{\hbar}{2}\left|+\hat{z}\right\rangle\!\left\langle+\hat{z}\right| - \frac{\hbar}{2}\left|-\hat{z}\right\rangle\!\left\langle-\hat{z}\right|\right]\left|\psi\right\rangle \\
&= \frac{\left\langle\psi\right|S_z\left|\psi\right\rangle}{\left\langle\psi\right|\left.\psi\right\rangle} \\
&:= \left\langle S_z\right\rangle,
\end{aligned} \tag{71}$$

known as the **expectation value** of S_z.

Now consider the second thought experiment, the essential part of which is we send in a state in the $\left\{\left|+\hat{z}\right\rangle, \left|-\hat{z}\right\rangle\right\}$ basis and out comes a state in the $\left\{\left|+\hat{x}\right\rangle, \left|-\hat{x}\right\rangle\right\}$ basis.[26] As before, we take the quantum states of the two output beams of this experiment to *define* the normalized kets $\left|\pm\hat{x}\right\rangle$. For the same reason as before, these must be orthogonal, so $\left\{\left|+\hat{x}\right\rangle, \left|-\hat{x}\right\rangle\right\}$ is an orthonormal basis, and the SG$\hat{x}$ likewise splits the beam into orthogonal states, see Figure 9.[27] Applying this to the particular

[24] This is exactly the property that self-adjoint operators carry, from 2 in **Proposition 17**!

[25] In reality, the atomic Hilbert space is much larger than this, but here we are only interested in the parts of it that can be measured by Stern-Gerlach devices.

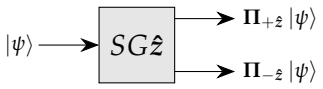

Figure 7: Orthogonal components from Stern-Gerlach device.

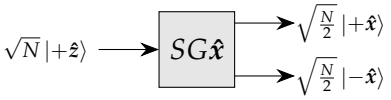

Figure 8: Second Stern-Gerlach device, this time along the $\hat{x}$ axis.

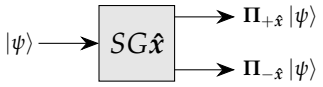

Figure 9: Orthogonal components from Stern-Gerlach device, along the $\hat{x}$ axis.

[26] For example, if we send in the state $\sqrt{N}\left|+\hat{z}\right\rangle$ then the outputs are $\sqrt{\frac{N}{2}}\left|+\hat{x}\right\rangle$ and $\sqrt{\frac{N}{2}}\left|-\hat{x}\right\rangle$, each with 50% of the beam, as seen in Figure 8

[27] Like before, $\mathbf{\Pi}_{+\hat{x}}\left|\psi\right\rangle$ and $\mathbf{\Pi}_{-\hat{x}}\left|\psi\right\rangle$, where $\mathbf{\Pi}_{\pm\hat{x}} = \left|\pm\hat{x}\right\rangle\!\left\langle\pm\hat{x}\right|$ is a complete set of orthogonal projectors.

situation shown in **Footnote 26**, we conclude that

$$\frac{1}{\sqrt{2}}|+\hat{\boldsymbol{x}}\rangle = \boldsymbol{\Pi}_{+\hat{\boldsymbol{x}}}|+\hat{\boldsymbol{z}}\rangle = |+\hat{\boldsymbol{x}}\rangle\langle+\hat{\boldsymbol{x}}|+\hat{\boldsymbol{z}}\rangle \implies \langle+\hat{\boldsymbol{x}}|+\hat{\boldsymbol{z}}\rangle = \frac{1}{\sqrt{2}}, \tag{72}$$

as well as $\langle-\hat{\boldsymbol{x}}|+\hat{\boldsymbol{z}}\rangle = \frac{1}{\sqrt{2}}$ by the same reasoning. Therefore,

$$|+\hat{\boldsymbol{z}}\rangle = \frac{1}{\sqrt{2}}\left(|+\hat{\boldsymbol{x}}\rangle + |-\hat{\boldsymbol{x}}\rangle\right). \tag{73}$$

Since $|-\hat{\boldsymbol{z}}\rangle$ is orthogonal to $|+\hat{\boldsymbol{z}}\rangle$, it follows that

$$|-\hat{\boldsymbol{z}}\rangle = \frac{1}{\sqrt{2}}e^{i\phi}\left(|+\hat{\boldsymbol{x}}\rangle + |-\hat{\boldsymbol{x}}\rangle\right). \tag{74}$$

up to an *a priori*[28] unknown phase factor ψ. So far, we have not fixed the phase of $|-\hat{\boldsymbol{z}}\rangle$ relative to $|+\hat{\boldsymbol{z}}\rangle$, so we can *redefine* the ket $|-\hat{\boldsymbol{z}}\rangle$ to absorb the phase factor, leaving:

$$|\pm\hat{\boldsymbol{z}}\rangle = \frac{1}{\sqrt{2}}\left(|+\hat{\boldsymbol{x}}\rangle \pm |-\hat{\boldsymbol{x}}\rangle\right), \quad \text{or equivalently} \quad |\pm\hat{\boldsymbol{x}}\rangle = \frac{1}{\sqrt{2}}\left(|+\hat{\boldsymbol{z}}\rangle \pm |-\hat{\boldsymbol{z}}\rangle\right). \tag{75}$$

As before, we define an observable associated to this S_x measurement:

$$\boldsymbol{S}_x = \frac{\hbar}{2}|+\hat{\boldsymbol{x}}\rangle\langle+\hat{\boldsymbol{x}}| - \frac{\hbar}{2}|-\hat{\boldsymbol{x}}\rangle\langle-\hat{\boldsymbol{x}}| = \frac{\hbar}{2}\left(|+\hat{\boldsymbol{z}}\rangle\langle-\hat{\boldsymbol{z}}| + |-\hat{\boldsymbol{z}}\rangle\langle+\hat{\boldsymbol{z}}|\right). \tag{77}$$

Now consider what happens when we instead measure S_y. With the entering atoms in the $|+\hat{\boldsymbol{z}}\rangle$ state, this is a repetition of our second thought experiment in Figure 8, but with the axis relabeled. I.e., recalling **Footnote 26**, the outputs are $\sqrt{\frac{N}{2}}|+\hat{\boldsymbol{y}}\rangle$ and $\sqrt{\frac{N}{2}}|-\hat{\boldsymbol{y}}\rangle$, each with 50% of the beam. An illustration of this is shown in Figure 10. Now we proceed just as before, defining the states $|\pm\hat{\boldsymbol{y}}\rangle$ using the two output beams, with the final result:

$$|+\hat{\boldsymbol{z}}\rangle = \frac{1}{\sqrt{2}}\left(|+\hat{\boldsymbol{y}}\rangle + |-\hat{\boldsymbol{y}}\rangle\right), \; |-\hat{\boldsymbol{z}}\rangle = \frac{1}{\sqrt{2}}e^{i\delta}\left(|+\hat{\boldsymbol{y}}\rangle + |-\hat{\boldsymbol{y}}\rangle\right). \tag{78}$$

Here δ is another *a priori* undetermined phase. This time, however, we no longer have the freedom to redefine the phase of $|-\hat{\boldsymbol{z}}\rangle$ relative to $|+\hat{\boldsymbol{z}}\rangle$, since we already used it above. Thus, we obtain

$$|\pm\hat{\boldsymbol{y}}\rangle = \frac{1}{\sqrt{2}}\left(|+\hat{\boldsymbol{z}}\rangle \pm e^{i\delta}|-\hat{\boldsymbol{z}}\rangle\right). \tag{79}$$

To fix δ, we note that by analogy with Equation 72 we must have $|\langle+\hat{\boldsymbol{x}}|+\hat{\boldsymbol{y}}\rangle| = \frac{1}{\sqrt{2}}$.[29] We find:

$$\langle+\hat{\boldsymbol{x}}|+\hat{\boldsymbol{y}}\rangle = \frac{1}{2}\left(1+e^{-i\delta}\right) = e^{-i\delta}\cos\left(\frac{\delta}{2}\right). \tag{80}$$

Thus, we require $\cos^2(\delta/2) = 1/2$, which has two solutions, $\delta = \pm\pi/2$, corresponding to $e^{i\delta} = \pm i$.

[28] Relating to or denoting reasoning or knowledge which proceeds from theoretical deduction rather than from observation or experience.

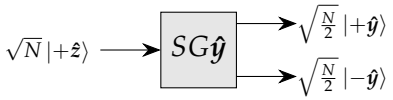

Figure 10: Third Stern-Gerlach device, this time along the $\hat{\boldsymbol{y}}$ axis.

Note that S_z and S_x do not commute:

$$[\boldsymbol{S}_z, \boldsymbol{S}_x] = \frac{\hbar^2}{2}\left(|+\hat{\boldsymbol{z}}\rangle\langle-\hat{\boldsymbol{z}}| - |-\hat{\boldsymbol{z}}\rangle\langle+\hat{\boldsymbol{z}}|\right) \neq 0. \tag{76}$$

Mathematically, this reflects the fact that their eigenspace projectors do not commute, i.e., when measuring both S_z and S_x, the outcome depends on the *order* in which the two measurements are made. For this reason, we say that S_z, S_x are **simultaneously observable.** Equivalently, such observables are said to be **incompatible**.

[29] Equivalently, this can be deduced by analyzing yet another rotated version of the second thought experiment.

Which choice we make for δ does not matter at this stage, since the two choices are related to each other by the antilinear map on $\mathcal{H}$:

$$a\,|+\hat{z}\rangle + b\,|-\hat{z}\rangle \to a^*\,|+\hat{z}\rangle + b^*\,|-\hat{z}\rangle\,, \tag{81}$$

which has not effect on measurement of S_z and S_x. Thus, we just have to make a choice. Conventionally, we take $\delta = -\pi/2$ so that

$$|\pm\hat{y}\rangle = \frac{1}{\sqrt{2}}\left(|+\hat{z}\rangle \pm i\,|-\hat{z}\rangle\right). \tag{82}$$

Thus, the observable associated to a measurement of S_y is:

$$S_y = \frac{\hbar}{2}\,|+\hat{y}\rangle\langle+\hat{y}| - \frac{\hbar}{2}\,|-\hat{y}\rangle\langle-\hat{y}| = \frac{\hbar}{2i}\left(|+\hat{z}\rangle\langle-\hat{z}| - |-\hat{z}\rangle\langle+\hat{z}|\right). \tag{83}$$

Notice that this is the same as $\frac{1}{i\hbar}[S_z, S_x]$, in Equation 76. In fact, one can show the commutator of ant two of the S_x, S_y, S_z is $\pm i\hbar$ times the third. Thus, no two components of the spin can be simultaneously measured.

Question 26.

1. *Verify the second equality in both Equation 77 and Equation 83.*

   Proof. Equation 75 tells us that

   $$\begin{aligned} |+\hat{x}\rangle\langle+\hat{x}| &= \frac{1}{2}\left(|+\hat{z}\rangle + |-\hat{z}\rangle\right)\left(\langle+\hat{z}| + \langle-\hat{z}|\right) \\ &= \frac{1}{2}\left(|+\hat{z}\rangle\langle+\hat{z}| + |+\hat{z}\rangle\langle-\hat{z}| + |-\hat{z}\rangle\langle+\hat{z}| + |-\hat{z}\rangle\langle-\hat{z}|\right) \end{aligned} \tag{84}$$

   and

   $$\begin{aligned} |-\hat{x}\rangle\langle-\hat{x}| &= \frac{1}{2}\left(|+\hat{z}\rangle - |-\hat{z}\rangle\right)\left(\langle+\hat{z}| - \langle-\hat{z}|\right) \\ &= \frac{1}{2}\left(|+\hat{z}\rangle\langle+\hat{z}| - |+\hat{z}\rangle\langle-\hat{z}| - |-\hat{z}\rangle\langle+\hat{z}| + |-\hat{z}\rangle\langle-\hat{z}|\right), \end{aligned} \tag{85}$$

   so

   $$|+\hat{x}\rangle\langle+\hat{x}| - |-\hat{x}\rangle\langle-\hat{x}| = |+\hat{z}\rangle\langle-\hat{z}| + |-\hat{z}\rangle\langle+\hat{z}|\,, \tag{86}$$

   $\implies$ Equation 77.

   Additionally, Equation 82 tells us that

   $$\begin{aligned} |+\hat{y}\rangle\langle+\hat{y}| &= \frac{1}{2}\left(|+\hat{z}\rangle + i\,|-\hat{z}\rangle\right)\left(\langle+\hat{z}| - i\,\langle-\hat{z}|\right) \\ &= \frac{1}{2}\left(|+\hat{z}\rangle\langle+\hat{z}| - i\,|+\hat{z}\rangle\langle-\hat{z}| + i\,|-\hat{z}\rangle\langle+\hat{z}| + |-\hat{z}\rangle\langle-\hat{z}|\right) \end{aligned} \tag{87}$$

   and

   $$\begin{aligned} |-\hat{y}\rangle\langle-\hat{y}| &= \frac{1}{2}\left(|+\hat{z}\rangle - i\,|-\hat{z}\rangle\right)\left(\langle+\hat{z}| + i\,\langle-\hat{z}|\right) \\ &= \frac{1}{2}\left(|+\hat{z}\rangle\langle+\hat{z}| + i\,|+\hat{z}\rangle\langle-\hat{z}| - i\,|-\hat{z}\rangle\langle+\hat{z}| + |-\hat{z}\rangle\langle-\hat{z}|\right), \end{aligned} \tag{88}$$

so

$$\begin{aligned} |+\hat{y}\rangle\langle+\hat{y}| - |-\hat{y}\rangle\langle-\hat{y}| &= -i\,|+\hat{z}\rangle\langle-\hat{z}| + i\,|-\hat{z}\rangle\langle+\hat{z}| \\ &= -i\,(|+\hat{z}\rangle\langle-\hat{z}| - |-\hat{z}\rangle\langle+\hat{z}|) \\ &= \frac{1}{i}\,(|+\hat{z}\rangle\langle-\hat{z}| - |-\hat{z}\rangle\langle+\hat{z}|)\,, \end{aligned} \tag{89}$$

$\implies$ Equation 83. □

2. *Using Equations 70, 77, 83, derive the* ***spin algebra****:*

$$[S_x, S_y] = i\hbar S_z,\ [S_y, S_z] = i\hbar S_x,\ [S_z, S_x] = i\hbar S_y. \tag{90}$$

Proof.

$$\begin{aligned} [S_x, S_y] &= S_x S_y - S_y S_x \\ &= \frac{\hbar^2}{2i}\,(-|+\hat{z}\rangle\langle+\hat{z}| + |-\hat{z}\rangle\langle-\hat{z}|) - \frac{\hbar^2}{2i}\,(|+\hat{z}\rangle\langle+\hat{z}| - |-\hat{z}\rangle\langle-\hat{z}|) \\ &= \frac{\hbar^2}{2i}\,(-|+\hat{z}\rangle\langle+\hat{z}| + |-\hat{z}\rangle\langle-\hat{z}|) \\ &= \frac{i\hbar^2}{2}\,(|+\hat{z}\rangle\langle+\hat{z}| - |-\hat{z}\rangle\langle-\hat{z}|) \\ &= i\hbar S_z, \end{aligned} \tag{91}$$

$$\begin{aligned} [S_y, S_z] &= S_y S_z - S_z S_y \\ &= \frac{\hbar^2}{4i}\,(-|+\hat{z}\rangle\langle-\hat{z}| - |-\hat{z}\rangle\langle+\hat{z}|) - \frac{\hbar^2}{4i}\,(|+\hat{z}\rangle\langle-\hat{z}| + |-\hat{z}\rangle\langle-\hat{z}|) \\ &= -\frac{\hbar^2}{2i}\,(|+\hat{z}\rangle\langle-\hat{z}| + |-\hat{z}\rangle\langle+\hat{z}|) \\ &= \frac{i\hbar^2}{2}\,(|+\hat{z}\rangle\langle-\hat{z}| + |-\hat{z}\rangle\langle+\hat{z}|) \\ &= i\hbar S_x, \end{aligned} \tag{92}$$

and

$$\begin{aligned} [S_z, S_x] &= S_z S_x - S_x S_z \\ &= \frac{\hbar^2}{4}\,(|+\hat{z}\rangle\langle-\hat{z}| - |-\hat{z}\rangle\langle+\hat{z}|) - \frac{\hbar^2}{4}\,(-|+\hat{z}\rangle\langle-\hat{z}| + |-\hat{z}\rangle\langle+\hat{z}|) \\ &= \frac{\hbar^2}{2}\,(|+\hat{z}\rangle\langle-\hat{z}| - |-\hat{z}\rangle\langle+\hat{z}|) \\ &= \frac{i\hbar^2}{2i}\,(|+\hat{z}\rangle\langle-\hat{z}| - |-\hat{z}\rangle\langle+\hat{z}|) \\ &= i\hbar S_y. \end{aligned} \tag{93}$$

□

Working, e.g., in the $\{|+\hat{z}\rangle\,, |-\hat{z}\rangle\}$ basis, we can equivalently specify the operators S_x, S_y, S_z by their matrix elements. From Equations 70, 77, 83, we read off:

$$S_x \to \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad S_y \to \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad S_z \to \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{94}$$

Stripping off the factor of $\frac{\hbar}{2}$ yields the **Pauli matrices**:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{95}$$

Any real 2×2 Hermitian matrix is a real linear combination of $\mathcal{I}$, σ_x, σ_y, σ_z. Furthermore, one can show that

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = \mathcal{I}, \tag{96}$$

$$\sigma_x \sigma_y = -\sigma_y \sigma_x = i\sigma_z, \tag{97}$$

$$\sigma_y \sigma_z = -\sigma_z \sigma_y = i\sigma_x \tag{98}$$

$$\sigma_z \sigma_x = -\sigma_x \sigma_z = i\sigma_y, \tag{99}$$

which can be summarized as $\sigma_i \sigma_j = \delta_{ij} + i\varepsilon_{ijk}\sigma_k$.[30]

Note, however, that S_x, S_y, and S_z will have different matrix elements if we work in a different basis.

[30] Here, ε_{ijk} is the totally antisymmetric **Levi-Civita symbol**, with $\varepsilon_{123} = \varepsilon_{231} = \varepsilon_{312} = +1$ and $\varepsilon_{321} = \varepsilon_{213} = \varepsilon_{132} = -1$, and repeated indices are summed over.

Question 27. *Find the matrix representations of S_x, S_y, and S_z in the $\{|+\hat{x}\rangle, |-\hat{x}\rangle\}$ basis.*

Proof. By Equation 77,

$$S_x = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{100}$$

Also, recall Equations 73 and 74, which tell us what $|\pm\hat{z}\rangle$ are in the $\{|+\hat{x}\rangle, |-\hat{x}\rangle\}$ basis. We can than substitute these expressions into the current definition of S_z in Equation 70 to find its representation in the $\{|+\hat{x}\rangle, |-\hat{x}\rangle\}$ basis. First, note that

$$\begin{aligned} |+\hat{z}\rangle\langle+\hat{z}| &= \frac{1}{2}\left(|+\hat{x}\rangle + |-\hat{x}\rangle\right)\left(\langle+\hat{x}| + \langle-\hat{x}|\right) \\ &= \frac{1}{2}\left(|+\hat{x}\rangle\langle+\hat{x}| + |+\hat{x}\rangle\langle-\hat{x}| + |-\hat{x}\rangle\langle+\hat{x}| + |-\hat{x}\rangle\langle-\hat{x}|\right) \end{aligned} \tag{101}$$

and

$$\begin{aligned} |-\hat{z}\rangle\langle-\hat{z}| &= \frac{1}{2}\left(|+\hat{x}\rangle - |-\hat{x}\rangle\right)\left(\langle+\hat{x}| - \langle-\hat{x}|\right) \\ &= \frac{1}{2}\left(|+\hat{x}\rangle\langle+\hat{x}| - |+\hat{x}\rangle\langle-\hat{x}| - |-\hat{x}\rangle\langle+\hat{x}| + |-\hat{x}\rangle\langle-\hat{x}|\right). \end{aligned} \tag{102}$$

Then,

$$|+\hat{z}\rangle = \frac{1}{\sqrt{2}}\left(|+\hat{x}\rangle + |-\hat{x}\rangle\right). \tag{103}$$

$$|-\hat{z}\rangle = \frac{1}{\sqrt{2}} e^{i\phi}\left(|+\hat{x}\rangle + |-\hat{x}\rangle\right). \tag{104}$$

$$\begin{aligned} S_z &= \frac{\hbar}{2}\left(|+\hat{z}\rangle\langle+\hat{z}| - |-\hat{z}\rangle\langle-\hat{z}|\right) \\ &= \frac{\hbar}{2}\left(|+\hat{x}\rangle\langle-\hat{x}| + |-\hat{x}\rangle\langle+\hat{x}|\right) \\ &= \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \end{aligned} \tag{105}$$

Finally, to compute S_y, we recall that we already have it defined in terms of $|\pm\hat{z}\rangle$ in Equation 83, so we can simply reuse Equations 73 and 74 to express it in the $\{|+\hat{x}\rangle, |-\hat{x}\rangle\}$ basis. Again, we first note that

$$\begin{aligned} |+\hat{z}\rangle\langle-\hat{z}| &= \frac{1}{2}\left(|+\hat{x}\rangle + |-\hat{x}\rangle\right)\left(\langle+\hat{x}| - \langle-\hat{x}|\right) \\ &= \frac{1}{2}\left(|+\hat{x}\rangle\langle+\hat{x}| - |+\hat{x}\rangle\langle-\hat{x}| + |-\hat{x}\rangle\langle+\hat{x}| - |-\hat{x}\rangle\langle-\hat{x}|\right) \end{aligned} \tag{106}$$

and

$$\begin{aligned} |-\hat{z}\rangle\langle+\hat{z}| &= \frac{1}{2}\left(|+\hat{x}\rangle - |-\hat{x}\rangle\right)\left(\langle+\hat{x}| + \langle-\hat{x}|\right) \\ &= \frac{1}{2}\left(|+\hat{x}\rangle\langle+\hat{x}| + |+\hat{x}\rangle\langle-\hat{x}| - |-\hat{x}\rangle\langle+\hat{x}| - |-\hat{x}\rangle\langle-\hat{x}|\right). \end{aligned} \tag{107}$$

Thus,

$$\begin{aligned} S_y &= \frac{\hbar}{2i}\left(|+\hat{z}\rangle\langle-\hat{z}| - |-\hat{z}\rangle\langle+\hat{z}|\right) \\ &= \frac{\hbar}{2i}\left(|-\hat{x}\rangle\langle+\hat{x}| - |+\hat{x}\rangle\langle-\hat{x}|\right) \\ &= \frac{\hbar}{2i}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \\ &= \frac{\hbar}{2}\begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}. \end{aligned} \tag{108}$$

□

Question 28. *Consider the modified Stern-Gerlach apparatus used in the fourth thought experiment, seen in Figure 11. In this idealized device, the output state equals the input state with both beams unblocked:*

$$|\psi_{out}\rangle = \mathbf{\Pi}_{+\hat{x}}|\psi_{in}\rangle + \mathbf{\Pi}_{-\hat{x}}|\psi_{in}\rangle = \left(\mathbf{\Pi}_{+\hat{x}} + \mathbf{\Pi}_{-\hat{x}}\right)|\psi_{in}\rangle = |\psi_{in}\rangle. \tag{109}$$

1. *In a real device, the two beams may pick up different phase factors $e^{i\theta_+}$, $e^{i\theta_-}$ as they travel through it. Show that in this case*

$$|\psi_{\text{out}}\rangle = \boldsymbol{U}|\psi_{\text{in}}\rangle \tag{110}$$

and determine the operator $\boldsymbol{U}$. What kind of operator is it (e.g., self-adjoint, normal, unitary)?

Proof. □

2. *An S_z measurement is performed on the outgoing beam. Determine the probability of each outcome as a function of θ_+, θ_-.*

Proof. □

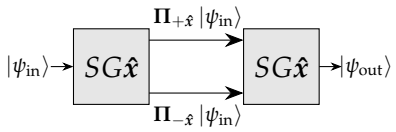

Figure 11: Fourth Stern-Gerlach experiment.

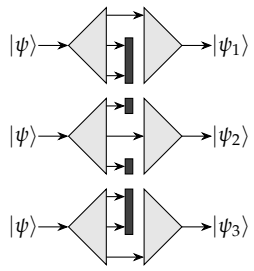

Figure 12: Fourth Stern-Gerlach experiment: showing all possible paths.

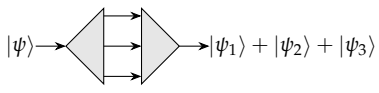

Figure 13: Fourth Stern-Gerlach experiment: combining all paths from Figure 12.

Postulate 1: Quantum States

Quantum states are vectors in some abstract Hilbert space. When a given outcome can be reached in several different ways, the final state is the *sum* of the state vectors arising from restricting to each individual scenario, allowing for *quantum interference*. Schematically, this is seen in Figure 13, where Figure 12 shows all of the outcomes independently.

The overall phase of a vector is not meaningful; only the *relative phase* between interfering components can be measured. Likewise, the overall normalization is not meaningful (though we may sometimes fix it by convention), but of course the relative normalizations of interfering components *will* matter.

The essential nature of quantum states is encapsulated by our first postulate.[31]

[31] This postulate can be generalized to allow for *real* Hilbert spaces. However, all known quantum systems are described by complex Hilbert spaces.

Postulate 1.

Associated to a quantum system, there is a complex, projective Hilbert space $\mathcal{H}$ of possible **quantum states** $|\psi\rangle$.

Here "projective" means that the state $\lambda|\psi\rangle$ is physically equivalent to $|\psi\rangle$ for any non-zero complex number λ (i.e., $\lambda \in \mathbb{C}^* = \mathbb{C} \setminus \{0\}$) and that $|\psi\rangle = 0$ is not a valid quantum state.

Note that **Postulate 1** does not tell us how to *identify* the appropriate Hilbert space $\mathcal{H}$ to use in a given situation. If we know the classical limit, the process of finding $\mathcal{H}$ (and other quantum data) is known as **quantization**. This somewhat imprecise process often involves a degree of intuition, so we will not consider it on the same level as the postulates. However, it is useful to formalize it as much as possible. Summarizing our prior observations about quantum interference, we observe that:

Quantization Rule 1.

If a quantum system is allowed to take multiple class paths to the same final configuration then then resulting state vector is the *sum* of the state vectors resulting from each individual path.

This rule plays a central role in an alternative formulation of quantum mechanics invented by Richard Feynman, known as **path integral quantization**. Suitable interpreted, it can be used to construct the Hilbert space $\mathcal{H}$ of a quantum system from the classical limit of the system. Note, however, that while **Postulate 1** is a purely mathematical statement, **Quantization Rule 1** requires significant interpretation to make it precise in any given scenario.

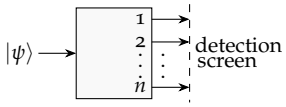

Figure 14: A generalized Stern-Gerlach device producing n possible outcomes.

Postulate 2: Measurements

A **measurement** is something like performing a generalized Stern-Gerlach experiment. Such an experiment has n possible outcomes, each with probability p_i that

depends on the input state, $|\psi\rangle$. See Figure 14. To analyze it, suppose we perform in $N \gg 1$ with the same incoming state $|\psi\rangle$, which we choose to normalize as

$$\langle\psi|\,\psi\rangle = N. \tag{111}$$

Let $|\psi_i\rangle$ denote the final state when the measurement outcome is i, normalized so that $\langle\psi_i|\,\psi_i\rangle = N_i$ is the **expected** (average) number of times that the measurement results in this outcome. We notate thus diagrammatically in Figure 15.

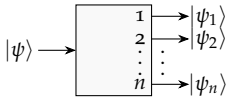

Figure 15: The quantum states resulting from each measurement outcome.

With these normalizations, the final states for each measurement outcome must be a *linear* function of the incoming state, i.e.,

$$|\psi_i\rangle = \boldsymbol{O}_i\,|\psi\rangle \tag{112}$$

for some linear operator $\boldsymbol{O}_i$ that is independent of $|\psi\rangle$.

Why must measurements be linear?

To see why this must be the case, consider the more general ansatz[32] $|\psi_i\rangle = f_i\left(|\psi\rangle\right)$ for some unknown (potentially non-linear) function $f_i : \mathcal{H} \to \mathcal{H}$. We now make the following observations:

[32] An educated guess, or an additional assumption, made to help solve the problem. It may later be verified to be part of the solution by its results.

1. Rescaling $|\psi\rangle \to \lambda\,|\psi\rangle$ for real $\lambda > 0$ corresponds to changing the number of iterations of the experiment, $N \to \lambda^2 N$, which changes the expected number of occurrence of each output proportionately, $N_i \to \lambda^2 N_i$, thus:[33]

$$f_i\left(\lambda\,|\psi\rangle\right) = \lambda f_i\left(|\psi\rangle\right), \forall\,\lambda \in \mathbb{R}_{>0}. \tag{113}$$

[33] Note that the different outcomes *cannot* pick up relative phases under this rescaling, since that would cause a measurable change in out probabilities (via quantum interference) if we subsequently recombined the beams, in contradiction with the fact that we merely performed the *same* experiment a different total number of times.

2. Now let the output of a quantum interference experiment be directed into our measuring device, as shown in Figure 16. With only one shutter open, we get certain states (depending on the setup) going into and coming out of the device:

$$\begin{aligned} &\text{A open}: \quad |\psi\rangle = |\psi_A\rangle\,, \quad |\psi_i\rangle = \left|\psi_i^{(A)}\right\rangle, \\ &\text{B open}: \quad |\psi\rangle = |\psi_B\rangle\,, \quad |\psi_i\rangle = \left|\psi_i^{(B)}\right\rangle. \end{aligned} \tag{114}$$

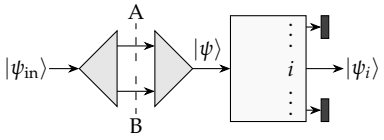

Figure 16: Directing interfering beams into a measuring device.

Therefore, with both shutters open the **Quantization Rule 1** dictates that:

$$|\psi\rangle = |\psi_A\rangle + |\psi_B\rangle\,, \text{ but also } |\psi_i\rangle = \left|\psi_i^{(A)}\right\rangle + \left|\psi_i^{(B)}\right\rangle. \tag{115}$$

Thus,

$$f_i\left(|\phi\rangle + |\chi\rangle\right) = f_i\left(|\phi\rangle\right) + f_i\left(|\chi\rangle\right). \tag{116}$$

3. All that remains is to understand how f_i is affected by phase factors. Because the overall phase is unobservable, we must have:

$$f_i\left(\mathrm{e}^{i\phi}\,|\psi\rangle\right) = \mathrm{e}^{i\phi'}\,f_i\left(|\psi\rangle\right), \tag{117}$$

where the relationship between ϕ' and ϕ is still to be determined. Using Equations 113 and 116, one can work out this relationship almost completely.

Question 29. *Show that Equations 113, 116, and 117 imply that*

$$f_i\left(e^{i\phi}\left|\psi\right\rangle\right) = e^{\pm i\phi} f_i\left(\left|\psi\right\rangle\right), \tag{118}$$

and that this in turn implies that, depending on the sign in $e^{\pm i\phi}$*:*

$$\begin{aligned} +: \; & f_i\left(\left|\psi\right\rangle + \lambda\left|\chi\right\rangle\right) = f_i\left(\left|\psi\right\rangle\right) + \lambda f_i\left(\left|\chi\right\rangle\right), \\ -: \; & f_i\left(\left|\psi\right\rangle + \lambda\left|\chi\right\rangle\right) = f_i\left(\left|\psi\right\rangle\right) + \lambda^* f_i\left(\left|\chi\right\rangle\right). \end{aligned} \tag{119}$$

In other words, f_i *is either a linear operator or an antilinear operator.*

Proof. □

Thus, the output states $\left|\psi_i\right\rangle$ must either be a *linear* or *antilinear* of the input state $\left|\psi\right\rangle$. This is almost what we wanted to show, but it raises an interesting question: could be build an *anti*linear measuring device? To see why not, one must consider the interference experiment shown in Figure 17. There, we interfere paths that do and do not pass through the measuring device. If the device were antilinear then

$$\left|\psi_{\text{in}}\right\rangle \to e^{i\phi}\left|\psi_{\text{in}}\right\rangle \text{ would lead to } \left|\psi_i\right\rangle \to e^{-i\phi}\left|\psi_i\right\rangle. \tag{120}$$

However, since the other path does not pass through the antilinear device, we would have $\left|\psi_0\right\rangle \to e^{i\phi}\left|\psi_0\right\rangle$, which would produce *measurable* interference between the two paths. This contradicts the assertion, **Postulate 1**, that the overall phase of the state vector is unobservable. Thus, measurements must be *linear*, exactly as seen in Figure 18.

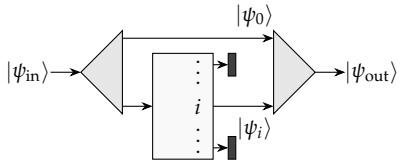

Figure 17: Paths that do and do not pass through the measuring device can interfere.

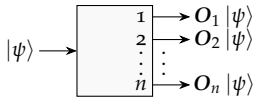

Figure 18: The states produced by each measurement outcome must be linear functions of the input state.

2025-02-18: The Postulates of Quantum Mechanics (cont.) & Time Evolution and Symmetries

Today we finished the section on the postulates of quantum mechanics and started a new section on time evolution and symmetries.

The Born Rule

Per Equation 111 (and Equation 112), the expected number of occurences of each measurement outcome is

$$N_i = \langle \psi_i | \, \psi_i \rangle = \langle \psi | \, \boldsymbol{O}_i^\dagger \boldsymbol{O} \, | \psi \rangle \, . \tag{121}$$

Since each measurement results in one outcome, these must add up to N. To say more, we impose requirements on our measuring device. First, we insist that the measurement be *repeatable*, i.e., chanining two of the measuring devices together produces the same outcome with 100% certainty, seen in Figure 19.

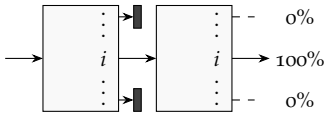

Figure 19: Repeating the measurement produces the same result with certainty.

Question 30.

1. *Show that the requirements of (1) each measurement producing one outcome and (2) repeatability are equivalent to:*

$$\sum_{i=1}^{n} \boldsymbol{O}_i^\dagger \boldsymbol{O}_i = \boldsymbol{I} \;\text{ and }\; \boldsymbol{O}_i \boldsymbol{O}_j = 0 \;\; (i \neq j) \, . \tag{122}$$

   Proof. □

2. *Show that $\boldsymbol{O}_i^\dagger \boldsymbol{O}_i$ for $j \neq i$, and hence that $\boldsymbol{U} = \sum_{i=1}^{n} \boldsymbol{O}_i$ is unitary.*

   Proof. □

3. *Defining $\boldsymbol{\Pi}_i := \boldsymbol{U}^\dagger \boldsymbol{O}_i$, show that (i) $\boldsymbol{\Pi}_i = \boldsymbol{\Pi}_i^\dagger$, (ii) $\sum_{i=1}^{n} \boldsymbol{\Pi}_i = \boldsymbol{I}$, and (iii) $\boldsymbol{\Pi}_i \boldsymbol{\Pi}_j = \delta_{ij} \boldsymbol{\Pi}_i$.*

   Proof. □

4. *Show that $[\boldsymbol{U}, \boldsymbol{\Pi}_i] = 0$.*

   Proof. □

Thus, $\boldsymbol{O}_i = \boldsymbol{U} \boldsymbol{\Pi}_i$ where $\{\boldsymbol{\Pi}_1, \ldots, \boldsymbol{\Pi}_n\}$ is a complete set of orthogonal projectors and $\boldsymbol{U}$ is a unitary operator that commutes with them.

We can now compute:

$$\boldsymbol{O}_i^\dagger \boldsymbol{O}_i = \boldsymbol{\Pi}_i^\dagger \boldsymbol{U}^\dagger \boldsymbol{U} \boldsymbol{\Pi}_i = \boldsymbol{\Pi}_i^2 = \boldsymbol{\Pi}_i \implies N_i = \langle \psi_i | \, \psi_i \rangle = \langle \psi_i | \, \boldsymbol{O}_i^\dagger \boldsymbol{O}_i \, | \psi_i \rangle = \langle \psi | \, \boldsymbol{\Pi}_i \, | \psi \rangle \, . \tag{123}$$

Thus, the probability of measuring the outcome i is:

$$p_i = \frac{N_i}{N} = \frac{\langle\psi|\,\boldsymbol{\Pi}_i\,|\psi\rangle}{\langle\psi|\,\psi\rangle}. \tag{124}$$

This is known as the **Born rule**. After a measurement resulting in this outcome, the state vector becomes

$$|\psi_i\rangle = \boldsymbol{U}\boldsymbol{\Pi}_i\,|\psi\rangle\,. \tag{125}$$

What is the effect of the unitary operator $\boldsymbol{U}$? Performing the measurement *again* as in Figure 19, we get

$$|\psi_i'\rangle = \boldsymbol{U}\boldsymbol{\Pi}_i\,|\psi_i\rangle = \boldsymbol{U}\boldsymbol{\Pi}_i\boldsymbol{U}\boldsymbol{\Pi}_i\,|\psi\rangle = \boldsymbol{U}^2\boldsymbol{\Pi}_i^2\,|\psi\rangle = \boldsymbol{U}^2\boldsymbol{\Pi}_i\,|\psi\rangle\,, \tag{126}$$

where we use $[\boldsymbol{U}, \boldsymbol{\Pi}_i] = 0$ in the penultimate step. Thus, each subsequent measurement after the first has the effect of applying the unitary operator $\boldsymbol{U}$ to the state.

If $\boldsymbol{\Pi}_i$ has rank one, $\boldsymbol{\Pi}_i = |\chi\rangle\langle\chi|$, then because $\boldsymbol{U}$ commutes with it, $|\chi\rangle$ must be a $\boldsymbol{U}$ eigenvector, i.e.,

$$\boldsymbol{U}\,|\chi\rangle = \mathrm{e}^{i\phi}\,|\chi\rangle \implies \boldsymbol{U}\boldsymbol{\Pi}_i\,|\chi\rangle = \mathrm{e}^{i\phi}\,\boldsymbol{\Pi}_i\,|\chi\rangle\,. \tag{127}$$

Thus, the unitary operator only contributes an overall phase to the state, having no physical effect. However, if $\boldsymbol{\Pi}_i$ has rank greater than 1 then $\boldsymbol{U}$ can act non-trivially on the $\boldsymbol{\Pi}_i = +1$ eigenspace corresponding to the outcome i, chaining the state with each repetition of the measurement.

For simplicity let us consider only "ideal" measuring devices with $\boldsymbol{U} = \boldsymbol{I}$. This can be thought of as a strong version of the repeatability requirement: we require that repeatedly remeasuring the same quantity does not change the quantum state (up to a phase).

For an ideal measuring device of this kind, each output state $|\psi_i\rangle = \boldsymbol{\Pi}_i\,|\psi\rangle$ is a projection of the input state. Often, we assign distinct real numbers λ_i to each measurement outcome so that

$$\boldsymbol{L} = \sum_{i=1}^{n} \lambda_i \boldsymbol{\Pi}_i, \tag{128}$$

is the spectral decomposition of the self-adjoint operator $\boldsymbol{L}$, λ_i are its eigenvalues, and $\boldsymbol{\Pi}_i$ are its eigenspace projectors. Every such **observable** defines a possible idealized measurement.

Postulate 2: Measurements (cont.)

We can summarize what we have learned with another postulate:[34]

[34] The ideas expressed in **Postulate 2**, especially that of "wavefunction collapse," go by the name of the "Copenhagen interpretation of quantum mechanics."

Postulate 2.

Possible measurements are described by self-adjoint operators (observables) $\boldsymbol{L} = \boldsymbol{L}^\dagger$, where:

1. The outcome of each $\boldsymbol{L}$ measurement is one of its eigenvalues.
2. The probability of measuring eigenvalue λ_i when in the state $|\psi\rangle$ is

$$p_i = \frac{\langle\psi|\,\boldsymbol{\Pi}_i\,|\psi\rangle}{\langle\psi|\,\psi\rangle}, \quad \text{(the Born rule)}\,, \tag{129}$$

   where $\boldsymbol{\Pi}_i$ is the $\boldsymbol{L} = \lambda_i$ eigenspace projector.
3. After measuring $\boldsymbol{L} = \lambda_i$ the state vector changes to:

$$|\psi'\rangle = \boldsymbol{\Pi}_i\,|\psi\rangle\,, \quad \text{(“wavefunction collapse”)}\,. \tag{130}$$

Here we've chosen to describe measurement in the conventional way, using observables. Alternately, **Postulate 2** could be rephrased using only a complete set of orthogonal projectors to describe possible measurements. However, observables come closer to classical intuition.[35]

[35] Note that self-adjoint operators are *real* observables (i.e., they have real eigenvalues), One can also define *complex* observables via normal operators, which is sometimes convenient depending on the context.

Quantization Rule 2.

Classically measurable quantities correspond to quantum observables.

Like **Quantization Rule 1**, this is a less precise statement than any of the postulates, but it helps specify how quantum mechanics relates to classical physics.

Postulate 3: Unitary Evolution

There are many ways to manipulate a quantum system *without* performing a measurement on it. For instance, in the case of the silver atoms in the Stern-Gerlach experiment, we can induce Larmor precession[36] by applying a magnetic field without detecting either their initial or their final spin states. Let us represent such a manipulation with the diagram in Figure 20. Such a process is said to be **closed**.

[36] The precession of the magnetic moment of an object about an external magnetic field.

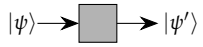

Figure 20: A device that manipulates a quantum system without measuring it.

By exactly the same reasoning as before, we can argue that the output state must be linearly-related to the input state:

$$|\psi'\rangle = \boldsymbol{U}\,|\psi\rangle\,, \tag{131}$$

for some linear operator $\boldsymbol{U}$. Moreover, since every input produes an output with 100% certainty, we must have:

$$\langle\psi'|\,\psi'\rangle = \langle\psi|\,\boldsymbol{U}^\dagger\boldsymbol{U}\,|\psi\rangle = \langle\psi|\,\psi\rangle\,, \tag{132}$$

for any $|\psi\rangle \in \mathcal{H}$. Thus, $\boldsymbol{U}$ must be a unitary operator.[37] This is our third postulate:

[37] To be precise, this implies only that $\boldsymbol{U}^\dagger\boldsymbol{U} = \boldsymbol{I}$, which need not imply that $\boldsymbol{U}\boldsymbol{U}^\dagger = \boldsymbol{I}$ in the case of an infinite-dimensional Hilbert space. However, if we assume that the process is *reversible*, i.e., if it is possible to build a second device which undoes the first one $|\psi''\rangle = \boldsymbol{U}^\dagger\,|\psi'\rangle = |\psi\rangle$, then the consistency of this second device requires $\boldsymbol{U}\boldsymbol{U}^\dagger = \boldsymbol{I}$, hence $\boldsymbol{U}$ must be unitary.

Postulate 3.

In any closed process the state vector of a quantum system changes by a unitary transformation:

$$|\psi'\rangle = \boldsymbol{U}\,|\psi\rangle\,, \tag{133}$$

where $\boldsymbol{U}$ is a unitary operator that depends on the process.

To interpret this postulate, we need to know what a "closed" process is. So far, we defined it to be a process in which no measurement takes place. More generally, a closed process is one in which the Hilbert space, $\mathcal{H}$, of our quantum system does not get mixed up with the Hilbert space, $\mathcal{H}'$, of some other quantum system. The result of such a mixing would be similar to a measurement, as the other system would thereby gain information about the initial state $|psi\rangle$ of our system, a phenomenon known as **entanglement**.

To account for processes that are *not* closed, we can either (1) enlarge $\mathcal{H}$ to include the states we are mixing with, or (2) introduce so-called "mixed-states," which we will discuss later.

A simple closed process is merely to *leave the system alone* for some period of time and let it evolve according to its own internal dynamics. We can also manipulate it more directly, e.g., by applying a magnetic field as discussed above, so long as we do not measure it. (It does not matter *who* or *what* is doing the measurement. Any outside access to $|\psi\rangle$ will create entanglement between $\mathcal{H}$ and the rest of the universe, in which case the process is not closed.)

Remarkably, the three postulates discussed above (with suitable technical modifications to accommodate continuous observables in infinite-dimensional Hilbert spaces) apply to *any* (closed) quantum system, and in some sense *define* what quantum mechanics is, much like Newton's laws define Newtonian mechanics. We will see how these postulates lead to more familiar aspects of quantum mechanics in subsequent sections.

Time Evolution and Symmetries

You may have been surprised to find that the none of the postulates enumerated involved Schrödinger's equation equation or anything like it.[38]

[38] In fact, **Postulate 3** secretely implies a generalized form of Schrödinger's equation.

The Time Evolution Operator and the Hamiltonian

Consider some quantum system, and let it be in the state $|\psi;t_0\rangle$ at some initial time $t = t_0$. If we do not measure it or otherwise mix it with other quantum systems, then **Postulate 3** specifies that the state at any other time t is

$$|\psi;t\rangle = \boldsymbol{U}\,(t,t_0)\,|\psi;t_0\rangle\,, \tag{134}$$

for some unitary operator $\boldsymbol{U}(t,t_0)$, called the *time evolution operator*. here we assume that the system undergoes only closed processes, making it a **closed quantum sys-**

tem. We can still control these processes externally, e.g., by applying electromagnetic fields, which will affect what $\boldsymbol{U}(t, t_0)$ is.

We now define the operator:

$$\boldsymbol{H}(t) := i\hbar \left[\frac{\partial}{\partial t}\boldsymbol{U}(t, t_0)\right] \boldsymbol{U}(t, t_0)^\dagger, \tag{135}$$

called the **quantum Hamiltonian**.

From Equation 134 and the unitarity of $\boldsymbol{U}(t, t_0)$, we find that the time evolution operator has the following properties:

1. $\boldsymbol{U}(t, t) = \boldsymbol{I}$,
2. $\boldsymbol{U}(t_3, t_2)\,\boldsymbol{U}(t_2, t_1) = \boldsymbol{U}(t_3, t_1)$,
3. $\boldsymbol{U}(t_1, t_2)^\dagger = \boldsymbol{U}(t_1, t_2)^{-1} = \boldsymbol{U}(t_2, t_1)$.

Question 31. *From Equation 135 and the properties of the time evolution operator, show that*

1. $\boldsymbol{H}(t) = \boldsymbol{H}(t)^\dagger$.

   Proof. □

2. $\boldsymbol{H}(t)$ *does not depend on* t_0.

   Proof. □

Equation 135 immediately implies that

$$i\hbar \frac{\partial}{\partial t}\boldsymbol{U}(t, t_0) = \boldsymbol{H}(t)\,\boldsymbol{U}(t, t_0). \tag{136}$$

Applying both sides to the ket $|\psi; t_0\rangle$, we obtain

$$i\hbar \frac{\partial}{\partial t}\boldsymbol{U}(t, t_0)\,|\psi; t_0\rangle = \boldsymbol{H}(t)\,\boldsymbol{U}(t, t_0)\,|\psi; t_0\rangle. \tag{137}$$

Thus, by Equation 135,[39]

$$i\hbar \frac{\partial}{\partial t}|\psi; t\rangle = \boldsymbol{H}(t)\,|\psi; t\rangle \quad (\text{the } \textbf{Schrödinger equation}). \tag{138}$$

Thus, if we know that quantum Hamiltonian $\boldsymbol{H}(t)$, then Equation 138 specifies how the quantum state vector changes in time.

[39] Note that Equation 138 is really a *generalization* of the wave equation published by Erwin Schrödinger in 1926. We will come to the Schrödinger wave equation later on.

So far we have shown that *some* operator $\boldsymbol{H}(t)$ exists s.t. Equation 138 is satisfied, but (despite the name) the physical significance of the quantum Hamiltonian remains to be determined. Likewise, the constant $\hbar$ in Equation 135 could so far be any convenient number; despite the choice of symbol, it might be unrelated to the angular momentum quantum discussed previously. We will fix these ambiguities presently.

In the special case where $\boldsymbol{H}(t)$ is *time indpendent*, its spectral decomposition is likewise time independent:

$$\boldsymbol{H} = \sum_n E_n\,|n\rangle\langle n|, \tag{139}$$

where the eigenvalues E_n and eigenvectors $|n\rangle$ are time-independent. Then, decomposing $|\psi; t\rangle$ in this eigenvasis:

$$|\psi; t\rangle = \sum_n \psi_n(t)\,|n\rangle \implies \sum_n i\hbar \frac{\partial \psi_n}{\partial t}|n\rangle = \sum_n E_n \psi_n\,|n\rangle, \tag{140}$$

since $|n\rangle$ is time-independent. Extracting each component of this equation, we find

$$i\hbar\frac{\partial\psi_n}{\partial t} = E_n\psi_n \implies \psi_n(t) = e^{-\frac{iE_nt}{\hbar}}\psi_n(0), \tag{141}$$

where we use the fact that E_n is time-independent. Thus, we can explicitly solve Schrödinger's equation:

$$|\psi;t\rangle = \sum_n e^{-\frac{iE_nt}{\hbar}}\psi_n(0)|n\rangle = \sum_n e^{-\frac{iE_nt}{\hbar}}|n\rangle\langle n|\psi;0\rangle. \tag{142}$$

Equivalently, the time evolution operator in this case is:[40]

$$\boldsymbol{U}(t,0) = \sum_n e^{-\frac{iE_nt}{\hbar}}|n\rangle\langle n| = \exp\left(-\frac{i\boldsymbol{H}t}{\hbar}\right), \tag{143}$$

which is the unique solution to the differential equation in Equation 136 with the boundary condition $\boldsymbol{U}(0,0) = 1$.

[40] Here for any analytic function $f(x) = \sum_{n=0}^{\infty}\frac{1}{n!}x^n f^{(n)}(0)$, we define $f(\boldsymbol{X}) := \sum_{n=0}^{\infty}\frac{1}{n!}\boldsymbol{X}^n f^{(n)}(0)$, where one can show that when $\boldsymbol{X}$ is normal with spectral decomposition $\boldsymbol{X} = \sum_i \lambda_i|i\rangle\langle i|$ then $f(\boldsymbol{X}) = \sum_i f(\lambda_i)|i\rangle\langle i|$.

Thus, when $\boldsymbol{H}(t)$ is time-independent, we can fully describe the quantum dynamics so long as we can solve the eigenvector equation:

$$\boldsymbol{H}|\psi\rangle = E|\psi\rangle, \tag{144}$$

which is also known as the **time-independent Schrödinger equation**. Given such a solution $|\psi\rangle$, the time-evolved state is $|\psi;t\rangle = e^{-\frac{iEt}{\hbar}}|\psi\rangle$. Since only the overall phase changes, this a **stationary state**, with properties that do not depend on time. By taking linear combinations of stationary states with different eigenvalues E, the relative phases become important, and the state changes with time. In this way, we can describe the quantum dynamics of any initial state we are interested in.

Example: Spin Precession

Consider the spin system from the SG experiment, and the the quantum Hamiltonian be

$$\boldsymbol{H} = \frac{E_0}{2}|+\hat{z}\rangle\langle+\hat{z}| - \frac{E_0}{2}|-\hat{z}\rangle\langle-\hat{z}|. \tag{145}$$

Consider, for instance, the initial state

$$|\psi;0\rangle = |+\hat{x}\rangle = \frac{1}{\sqrt{2}}(|+\hat{z}\rangle + |i\hat{z}\rangle) \implies |\psi;t\rangle = \frac{1}{\sqrt{2}}\left(e^{-\frac{iE_0t}{2\hbar}}|+\hat{z}\rangle + e^{\frac{iE_0t}{2\hbar}}|-\hat{z}\rangle\right). \tag{146}$$

To understand what happens to the spin, we compute the expectation value of $\boldsymbol{S}_x$:

$$\begin{aligned}
\langle\boldsymbol{S}_x\rangle &= \langle\psi;t|\boldsymbol{S}_x|\psi;t\rangle \\
&= \frac{1}{2}\left(e^{\frac{iE_0t}{2\hbar}}\langle+\hat{z}| + e^{-\frac{iE_0t}{2\hbar}}\langle-\hat{z}|\right)\boldsymbol{S}_x\left(e^{-\frac{iE_0t}{2\hbar}}|+\hat{z}\rangle + e^{\frac{iE_0t}{2\hbar}}|-\hat{z}\rangle\right) \\
&= \frac{1}{2}\left(e^{\frac{iE_0t}{2\hbar}}\langle+\hat{z}| + e^{-\frac{iE_0t}{2\hbar}}\langle-\hat{z}|\right)\left(e^{-\frac{iE_0t}{2\hbar}}\frac{\hbar}{2}|-\hat{z}\rangle + e^{\frac{iE_0t}{2\hbar}}\frac{\hbar}{2}|+\hat{z}\rangle\right) \\
&= \frac{\hbar}{4}\left(e^{\frac{iE_0t}{\hbar}} + e^{-\frac{iE_0t}{\hbar}}\right) \\
&= \frac{\hbar}{2}\cos\left(\frac{E_0t}{\hbar}\right).
\end{aligned} \tag{147}$$

Likewise, one can easily show that

$$\langle \boldsymbol{S}_y \rangle = \frac{\hbar}{2} \sin\left(\frac{E_0 t}{\hbar}\right), \quad \langle \boldsymbol{S}_z \rangle = 0. \tag{148}$$

Thus, the spin precesses about the $\hat{\boldsymbol{z}}$ axis with the angular frequency $\underline{\boldsymbol{\omega}} = \frac{E_0}{\hbar}\hat{\boldsymbol{z}}$, as seen in Figure 21.

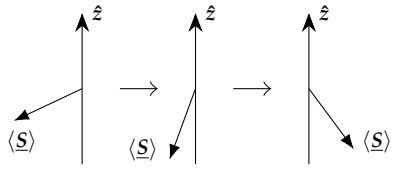

Figure 21: The precession of the spin expectation value due to the quantum Hamiltonian in Equation 145.

Let us compare this result with the classical Larmor precession $\underline{\boldsymbol{\omega}} = -\gamma \underline{\boldsymbol{B}}$ of the atom in an external magnetic field. To obtain a match, the magnetic field $\underline{\boldsymbol{B}} = B_z \hat{\boldsymbol{z}}$ should point in the $\hat{\boldsymbol{z}}$ direction, with $E_0 = -\hbar\gamma B_z$. In this case we find

$$\boldsymbol{H} = -\gamma B_z \left(\frac{\hbar}{2}|+\hat{\boldsymbol{z}}\rangle\langle+\hat{\boldsymbol{z}}| - \frac{\hbar}{2}|-\hat{\boldsymbol{z}}\rangle\langle-\hat{\boldsymbol{z}}|\right) = -\gamma B_z \boldsymbol{S}_z = -\underline{\boldsymbol{\mu}} \cdot \underline{\boldsymbol{B}}, \tag{149}$$

where $\underline{\boldsymbol{\mu}} = \gamma \underline{\boldsymbol{S}}$ is the magnetic moment operator for the atom. This exactly matches the classical Hamiltonian (i.e., the interaction energy between the magnetic moment and the external field), where it was important that the constant $\hbar$ in Equations 135 and 138 is precisely the same as the angular momentum quantum $\hbar$.

2025-02-20: Time Evolution and Symmetries (cont.)

Today was a make-up day for the class session we missed on 2025-02-06.

The Heisenberg picture and canonical quantization

To understand the relationship between $\boldsymbol{H}(t)$ and the classical Hamiltonian more systematically, we now re-express the quantum dynamics in a different way. Consider a time-dependent unitary transformation applied to all states and observables simultaneously:

$$|\psi\rangle \to \boldsymbol{U}(t)|\psi\rangle, \quad \boldsymbol{X} \to \boldsymbol{U}(t)\boldsymbol{X}\boldsymbol{U}(t)^\dagger. \tag{150}$$

Observe that this has *no effect* on any observable quantity, e.g., consider an expectation value

$$\langle \boldsymbol{X}\rangle = \langle\psi|\boldsymbol{X}|\psi\rangle \to \langle\psi|\boldsymbol{U}^\dagger\left(\boldsymbol{U}\boldsymbol{X}\boldsymbol{U}^\dagger\right)\boldsymbol{U}|\psi\rangle = \langle\psi|\boldsymbol{X}|\psi\rangle, \tag{151}$$

where we assume that $|\psi\rangle$ is normalized for simplicity (where the norm $\langle\psi|\psi\rangle$ is likewise preserved by the unitary transformation).

The above observation is somewhat counterintuitive, because we can pick the time-dependent unitary transformation to make the time-dependence of the state vector $|\psi;t\rangle$ to be anything we choose, somehow without affecting the physics! This is possible because the transformed observables $\boldsymbol{U}(t)\boldsymbol{X}\boldsymbol{U}(t)^\dagger$ are generally *time-dependent* as well.

In particular, it is natural to choose $\boldsymbol{U}(t)$ s.t. $|\psi\rangle$ becomes *time-independent*, resulting in the **Heisenberg picture**. To see how this works, consider a (generally time-dependent) expectation value:

$$\langle \boldsymbol{X}\rangle(t) = \langle\psi;t|\boldsymbol{X}|\psi;t\rangle = \langle\psi|\boldsymbol{U}(t,0)^\dagger\boldsymbol{X}\boldsymbol{U}(t,0)|\psi\rangle = \langle\psi|\boldsymbol{X}_H(t)|\psi\rangle \tag{152}$$

where $|\psi\rangle := |\psi;0\rangle$ is the Heisenberg-picture state vector and
$\boldsymbol{X}_H(t) := \boldsymbol{U}(t,0)^\dagger\boldsymbol{X}\boldsymbol{U}(t,0)$ is the Heisenberg-picture operator corresponding to $\boldsymbol{X}$. To disambiguate the different pictures, the operator $\boldsymbol{X}$ is known as the **Schrödinger pictures**, sometimes written as $\boldsymbol{X}_S$.

Thus, in the Heisenberg picture the dynamics lies in the operators, not in the states. There is no Schrödinger equation, since $|\psi\rangle$ is time-independent. Instead, the observables themselves change with time:

$$\boldsymbol{X}_H(t) = \boldsymbol{U}(t)^\dagger\boldsymbol{X}_S\boldsymbol{U}(t) \tag{153}$$

$$\begin{aligned}
\implies \dot{\boldsymbol{X}}_H(t) &= \dot{\boldsymbol{U}}^\dagger\boldsymbol{X}_S\boldsymbol{U} + \boldsymbol{U}^\dagger\boldsymbol{X}_S\dot{\boldsymbol{U}} \\
&= \left(-\frac{i}{\hbar}\boldsymbol{H}_S\boldsymbol{U}\right)^\dagger\boldsymbol{X}_S\boldsymbol{U} + \boldsymbol{U}^\dagger\boldsymbol{X}_S\left(-\frac{i}{\hbar}\boldsymbol{H}_S\boldsymbol{U}\right) \\
&= \frac{i}{\hbar}\left(\boldsymbol{U}^\dagger\boldsymbol{H}_S\boldsymbol{X}_S\boldsymbol{U} - \boldsymbol{U}^\dagger\boldsymbol{X}_S\boldsymbol{H}_S\boldsymbol{U}\right) \\
&= \frac{i}{\hbar}\boldsymbol{U}^\dagger\left(\boldsymbol{H}_S\boldsymbol{X}_S - \boldsymbol{X}_S\boldsymbol{H}_S\right)\boldsymbol{U} \\
&= \frac{i}{\hbar}\left(\boldsymbol{H}_H\boldsymbol{X}_H - \boldsymbol{X}_H\boldsymbol{H}_H\right)
\end{aligned} \tag{154}$$

Where $\dot{} := \frac{d}{dt}$, $\boldsymbol{U}(t) = \boldsymbol{U}(t,0)$, we assume that $\dot{\boldsymbol{X}}_S = 0$, and we applied the relation from Equation 136. Thus, we conclude that

$$\frac{d\boldsymbol{X}_H}{dt} = \frac{i}{\hbar}\left[\boldsymbol{H}_H(t), \boldsymbol{X}_H(t)\right], \quad (\text{the } \textbf{Heisenberg equation}). \tag{155}$$

This is first order differential equation for $\boldsymbol{X}_H(t)$ with the initial condition $\boldsymbol{X}_H(0) = \boldsymbol{X}_S$. To specify the quantum dynamics in the Heisenberg picture, we typically solve it to obtain $\boldsymbol{X}_H(t)$ in terms of some fixed operator basis, e.g., a basis consisting of Schrödinger-picture observables.

Above, we assumed that the observable $\boldsymbol{X}_S$ in question was time-independent in the Schrödinger picture. More generally, one finds

$$\frac{d\boldsymbol{X}_H}{dt} = \frac{i}{\hbar}\left[\boldsymbol{H}_H(t), \boldsymbol{X}_H(t)\right] + \frac{\partial \boldsymbol{X}_H}{\partial t}, \quad (\textbf{Heisenberg equation} \text{ - general case}) \tag{157}$$

where

$$\frac{\partial \boldsymbol{X}_H}{\partial t} := \boldsymbol{U}(t)^\dagger \frac{d\boldsymbol{X}_S}{dt}\boldsymbol{U}(t), \tag{158}$$

measures the **explicit time-dependence** of the observable $\boldsymbol{X}_H(t)$, i.e., the time-dependence that was present in the definition of the operator, rather than being dynamically induced.

Note that while in general the Heisenberg-picture Hamiltonian $\boldsymbol{H}_H(t)$ differs from the Schrödinger picture Hamiltonian $\boldsymbol{H}_S(t)$, if the latter is time-independent then

$$\begin{aligned} &\boldsymbol{U}(t) = \exp -\frac{-i\boldsymbol{H}_S t}{\hbar} \\ \implies &[\boldsymbol{U}(t), \boldsymbol{H}_s] = 0 \\ \implies &\boldsymbol{H}_H = \boldsymbol{U}^\dagger \boldsymbol{H}_S \boldsymbol{U} = \boldsymbol{H}_S, \end{aligned} \tag{156}$$

so in the time-independent case the Hamiltonian is the same in both pictures.

Example: Spin Precession Again

Let us revisit the example of spin precession, with the Schrödinger picture Hamiltonian:

$$\boldsymbol{H}_S = \frac{E_0}{\hbar}\boldsymbol{S}_z, \tag{159}$$

as in Equation 145, where $\boldsymbol{S}_z$ denotes the fixed, Schrödinger-picture operator that we constructed previously. Note that, since $\boldsymbol{H}_S$ is time-independent, we have $\boldsymbol{H}_H = \boldsymbol{H}_S$, and we can omit the subscript henceforward without any chance of confusion.

Since $[\boldsymbol{H}, \boldsymbol{S}_z] = 0$, solving the Heisenberg equation with the initial condition $\boldsymbol{S}_z^H(0) = \boldsymbol{S}_z$, we conclude that

$$\boldsymbol{S}_z^H(t) = \boldsymbol{S}_z. \tag{160}$$

Now consider the commutators:

$$[\boldsymbol{H}, \boldsymbol{S}_z] = iE_0\boldsymbol{S}_y, \quad [\boldsymbol{H}, \boldsymbol{S}_y] = -iE_0\boldsymbol{S}_x. \tag{161}$$

Because these are non-zero, the corresponding Heisenberg picture operators $\boldsymbol{S}_x^H(t)$ and $\boldsymbol{S}_y^H(t)$ will depend on time, as specified by the Heisenberg equation:

$$\frac{d\boldsymbol{S}_x^H}{dt} = \frac{i}{\hbar}\left[\boldsymbol{H}, \boldsymbol{S}_x^H\right], \tag{162}$$

with the initial condition $\boldsymbol{S}_x^H(0) = \boldsymbol{S}_x$, in the case of $\boldsymbol{S}_x^H(t)$, and likewise for $\boldsymbol{S}_y^H(t)$. To solve these equations, note that at $t = 0$ they reduce to

$$\frac{d\boldsymbol{S}_x^H}{dt} = -\frac{E_0}{\hbar}\boldsymbol{S}_y,. \tag{163}$$

Thus $S_x^H(t)$, which is initially equal to S_x, will begin to pick up a component in the S_y direction. By the same token, once this component is present, the S_x component will begin to change with time, etc., just as in a classical oscillatory system. This suggests the ansatz:

$$S_x^H(t) = A(t) S_x + B(t) S_y. \tag{165}$$

Substituting this ansatz into the Heisenberg equation, we find:

$$\dot{A} S_x + \dot{B} S_y = -\frac{E_0}{\hbar} A S_y + \frac{E_0}{\hbar} B S_x. \tag{166}$$

Since the operators S_x, S_y are linearly independent, this implies two equations:

$$\dot{A} = \frac{E_0}{\hbar} B, \; \dot{B} = -\frac{E_0}{\hbar} A, \implies \ddot{A} = -\frac{E_0^2}{\hbar^2} A, \tag{167}$$

upon eliminating B by taking the time derivative of the first equation and substituting it into the second. Since $S_x^H(0) = S_x$, the initial conditions are $A(0) = 1$ and $B(0) = 0$, where the latter is equivalent to $\dot{A}(0) = 0$ by Equation 167. Thus, we obtain the solution

$$A(t) = \cos\left(\frac{E_0 t}{\hbar}\right), \implies B(t) = -\sin\left(\frac{E_0 t}{\hbar}\right) \tag{168}$$

by Equation 167. We can solve for $S_y^H(t)$ by exactly the same process, with the final result:

$$\begin{aligned} S_y^H(t) &= \cos\left(\frac{E_0 t}{\hbar}\right) S_x^S - \sin\left(\frac{E_0 t}{\hbar}\right) S_y^S, \\ S_y^H(t) &= \sin\left(\frac{E_0 t}{\hbar}\right) S_x^S + \cos\left(\frac{E_0 t}{\hbar}\right) S_y^S, \\ S_z^H(t) &= S_z^S, \end{aligned} \tag{169}$$

where we now include the explicit "S" superscripts on the fixed operators S_x, S_y, S_z as a reminder of the difference between the two pictures.

We can now describe the Larmor precession in very general terms. Consider any quantum state $|\psi\rangle$ in the Heisenberg picture. Taking the expectation value of Equation 168 in this state, we conclude that

$$\begin{pmatrix} \langle S_x \rangle(t) \\ \langle S_y \rangle(t) \\ \langle S_z \rangle(t) \end{pmatrix} = \begin{pmatrix} \cos\frac{E_0 t}{\hbar} & -\sin\frac{E_0 t}{\hbar} & 0 \\ \sin\frac{E_0 t}{\hbar} & \cos\frac{E_0 t}{\hbar} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \langle S_x \rangle(t=0) \\ \langle S_y \rangle(t=0) \\ \langle S_z \rangle(t=0) \end{pmatrix}, \tag{170}$$

where we use the fact that $|\psi; 0\rangle = |\psi\rangle$ in the Schrödinger picture, so that $\langle\psi| S_i^S |\psi\rangle = \langle S_i \rangle(t=0)$. Thus, *regardless* of the initial state, the expectation values of the spin precesses about the applied magnetic field exactly as predicted by the classical equations of motion.

Question 32. *Find a normalized ket $|+\hat{\mathbf{x}}\rangle_H$ solving the eigenvalue equation*

$$S_x^H(t) |+\hat{\mathbf{x}}\rangle_H = +\frac{\hbar}{2} |+\hat{\mathbf{x}}\rangle_H. \tag{164}$$

Show that $|+\hat{\mathbf{x}}\rangle_H$ is necessarily time-dependent. How is this consistent with the claim that quantum states are time-independent in the Heisenberg picture?

Proof. □

Canonical Quantization

To interpret the Heisenberg equation, Equation 157, we recall a few facts from classical Hamiltonian mechanics. Let $H = H(p, q, t)$ be the classical Hamiltonian, where p denotes the generalized momenta and q denotes the generalized coordinates. A "classical observable" (such as the angular momentum of the system) is a specified function $X = X(p, q, t)$ of p, q, t. The chain rule and Hamilton's equations[41] then imply that:

$$\frac{dX}{dt} = \frac{\partial X}{\partial p}\frac{dp}{dt} + \frac{\partial X}{\partial q}\frac{dq}{dt} + \frac{\partial X}{\partial t} = -\frac{\partial X}{\partial p}\frac{\partial H}{\partial q} + \frac{\partial X}{\partial q}\frac{\partial H}{\partial p} + \frac{\partial X}{\partial t} = [X, H]_{\text{PB}} + \frac{\partial X}{\partial t}, \quad (172)$$

where we define the **Poisson bracket**:

$$[X, Y]_{\text{PB}} := \frac{\partial X}{\partial q}\frac{\partial Y}{\partial p} - \frac{\partial X}{\partial p}\frac{\partial Y}{\partial q}. \quad (173)$$

Above we assumed a single generalized coordinate q and associated canonical momentum p, but these equations carry over to any number of generalized coordinates $q_1, \ldots, q_n$, where:

$$[X, Y]_{\text{PB}} := \sum_i \left(\frac{\partial X}{\partial q_i}\frac{\partial Y}{\partial p_i} - \frac{\partial X}{\partial p_i}\frac{\partial Y}{\partial q_i} \right). \quad (174)$$

One can show that the Poisson bracket satisfies[42]

$$[X, Y]_{\text{PB}} = -[Y, X]_{\text{PB}} \qquad (\textbf{antisymmetry}), \quad (175)$$

$$0 = [[X, Y]_{\text{PB}}, Z]_{\text{PB}} + [[Y, Z]_{\text{PB}}, X]_{\text{PB}} + [[Z, X]_{\text{PB}}, Y]_{\text{PB}} \qquad (\textbf{Jacobi identity}). \quad (176)$$

With this in mind, we observe that the classical equation, Equation 172, – equivalent to Hamilton's equations – are remarkably similar to the general Heisenberg equation, Equation 157. In particular, the former maps the latter with the replacements:

$$X_{\text{cl}} \to \boldsymbol{X}_H, \quad H_{\text{cl}} \to \boldsymbol{H}_H, \quad [\cdot, \cdot]_{\text{PB}} \to \frac{1}{i\hbar}[\cdot, \cdot], \quad \frac{\partial X_{\text{cl}}}{\partial t} \to \frac{\partial \boldsymbol{X}_H}{\partial t}. \quad (177)$$

In other words, if we take the classical system and promote each classical observable (including the Hamiltonian itself) to a quantum operator in the Heisenberg picture satisfying commutation relations $[\boldsymbol{X}_H, \boldsymbol{Y}_H] = i\hbar\,[X_{\text{cl}}, Y_{\text{cl}}]_{\text{PB}}$ then the Heisenberg equation will reproduce the classical equations of motion. Any explicit time dependence in the classical observable likewise maps to explicit time dependence in the Heisenberg picture operators, i.e., to a time-dependent Schrödinger picture operator.

This procedure, known as **canonical quantization**, can be summarized as follows:

[41] In $2n$ dimensions, **Hamilton's equations** are

$$\frac{d\boldsymbol{q}}{dt} = \frac{\partial H}{\partial \boldsymbol{p}}, \quad \frac{d\boldsymbol{p}}{dt} = -\frac{\partial H}{\partial \boldsymbol{q}}, \quad (171)$$

for the n-dimensional phase space coordinates $(\boldsymbol{p}, \boldsymbol{q})$.

[42] These same properties are satisfied by commutators, inspiring the notation "$[X, Y]_{\text{PB}}$."

Question 33. *Consider a classical particle in 3d, with angular momentum $\underline{\boldsymbol{L}} = \boldsymbol{q} \times \boldsymbol{p}$. Compute the Poisson brackets between L_x, L_y, and L_z and compare your answer with the spin algebra (Equation 90 in **Question 26**).*

Proof. □

Quantization Rule 3. (Canonical quantization procedure).

To quantize a classical system with Hamiltonian H and observables X, we promote H, X to operators in the Heisenberg picture satisfying commutation relations given by the rule:

$$[\cdot,\cdot]_{\text{PB}} \to \frac{1}{i\hbar}[\cdot,\cdot], \qquad \text{(the } \textbf{canonical commutation relations}\text{)}, \quad (178)$$

where the classical Hamiltonian is promoted to the quantum Hamiltonian operator.

Note that canonical quantization can be *ambiguous*, because, e.g., the classical observable pq might plausibly get promoted to either $\boldsymbol{pq}$ or $\boldsymbol{qp}$, but of course $\boldsymbol{pq} \neq \boldsymbol{qp}$, since $[\boldsymbol{q}, \boldsymbol{p}] = i\hbar$ by the canonical commutation relations. Such **ordering ambiguities** can be resolved by various prescriptions (such as "Weyl ordering"), but it is perhaps better to think of the procedure in reverse: given a suitable quantum system, the canonical quantization procedure allows us to recognize the classical physics to which it reduces in the appropriate limit.[43]. There can even be *more than one* quantum system with the same classical limit, so we cannot hope to do better than this.

[43] What it means to take such a **classical limit** will be discussed more fully later on

Taking the expectation value of the Heisenberg equation, Equation 157, we find that:[44]

$$\frac{d}{dt}\langle \boldsymbol{X}\rangle = \frac{i}{\hbar}\langle[\boldsymbol{H}, \boldsymbol{X}]\rangle + \left\langle \frac{\partial \boldsymbol{X}}{\partial t}\right\rangle, \qquad \text{(the } \textbf{Ehrenfest theorem}\text{)}. \quad (179)$$

[44] Here we drop the picture labels, since the expectation values is the same regardless of the picture, provided that the states and operators are all evaluated in the same picture.

Thus, for a canonically-quantized quantum system, the expectation values of quantum observables will follow trajectories dictated by the *classical* equations of motion. We already showed this to be true for Larmor precession, but now we see that the result is completely general.

Quantum Symmetries

Just as in classical physics, quantum dynamics are constrained by symmetries. Taking this into account often leads to import insights and simplifications, as we see in the following example.

Motivating Example: Ammonia Maser

The ammonia molecule, NH_3, consists of three hydrogen atoms bounds to a single nitrogen atom. Due to the nature of the bond, the nitrogen atom preferentially sits out of the plane of the three hydrogen atoms. Consequently, fixing the hydrogen atoms in place, the nitrogen atom has to stable positions at which it likes to sit. These two positions are related by a reflection through the plan of the three hydrogen atoms, as seen in Figure 22.

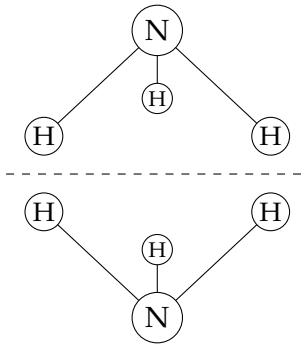

Figure 22: The two stable configurations of the nitrogen atom in ammonia (NH_3), which are related to each other by a reflection symmetry.

Of course, the hydrogen atoms can also move, and once we allow this we see that the two stable configurations discussed above are *also* related by a 180° rotation about an axis connecting one of the hydrogen atoms in the midpoint between the other two. However, such motions really represent *separate* degrees of freedom of the molecule, not related to the bi-stability of the nitrogen atom position. To illustrate this fact, imagine giving the atom some angular momentum perpendicular to the hydrogen plane. Angular momentum conservation then prevents the hydrogen plane from flipping over in the manner described above, but it does not present the nitrogen atom from occupying either of its two stable positions.

Thus, for the purpose of understanding the nitrogen position, we can crudely model NH_3 as a two-state system, with basis states

$$\begin{aligned} |1\rangle &:= \text{ Nitrogen atom above the plane,} \\ |2\rangle &:= \text{ Nitrogen atom below the plane.} \end{aligned} \tag{180}$$

The true Hilbert space of the ammonia molecule $\mathcal{H}_{\text{actual}}$ is of course *far* more complicated than this, but this crude two-state Hilbert space $\mathcal{H}_{\text{model}}$ captures the essential physics related to the nitrogen position.

The reflection is represented by an operator $\boldsymbol{U}$ acting on $\mathcal{H}_{\text{model}}$ as follows:

$$\boldsymbol{U}|1\rangle = e^{i\phi_{21}}|2\rangle\,, \;\; \boldsymbol{U}|2\rangle = e^{i\phi_{12}}|1\rangle\,, \;\; \Longrightarrow \boldsymbol{U} = e^{i\phi_{12}}|1\rangle\langle 2| + e^{i\phi_{21}}|2\rangle\langle 1| \tag{181}$$

where ϕ_{12} and ϕ_{21} are *a priori* unknown phases. Redefining the relative phase of $|1\rangle$ vs $|2\rangle$, we can set $\phi_{12} = \phi_{12} = \phi$, so that $\boldsymbol{U} = e^{i\phi}\left(|1\rangle\langle 2| + |2\rangle\langle 1|\right)$. However, the overall phase can be dropped, since as usual it has no physical significance. Thus, in the end we obtain[45]

$$\boldsymbol{U} = |1\rangle\langle 2| + |2\rangle\langle 1|\,. \tag{182}$$

[45] Notice that $\boldsymbol{U}$ is *unitary*, as required for the reflection to conserve probabilities.

The Hamiltonian must be reflection invariant, i.e.,

$$\boldsymbol{U}\boldsymbol{H}\boldsymbol{U}^\dagger = \boldsymbol{H}, \tag{183}$$

which is the same as $[\boldsymbol{U}, \boldsymbol{H}] = 0$. The most general solution to this constraint is[46]

$$\boldsymbol{H} = E_0\left(|1\rangle\langle 1| + |2\rangle\langle 2|\right) + A\left(|1\rangle\langle 2| + |2\rangle\langle 1|\right), \tag{185}$$

[46] In matrix form,

$$\boldsymbol{H} = \begin{pmatrix} E_0 & A \\ A & E_0 \end{pmatrix}. \tag{184}$$

where E_0, A are constants, which must be *real* to ensure $\boldsymbol{H} = \boldsymbol{H}^\dagger$. We can interpret these constants physically as follows:

$$\begin{aligned} E_0 &:= \text{ energy of the up and down configurations,} \\ A &:= \text{ effect of } \textbf{quantum tunneling} \text{ between the two configurations.} \end{aligned} \tag{186}$$

Here, A represents a *wholly quantum effect*. Classically, the nitrogen atom will stay in one of its stable positions unless we give it a big enough push to carry it over the potential barrier between the two. Quantum mechanically, however, such a term mixing the two configurations is possible, and indeed it turns out that $A \neq 0$. Even though A is much smaller than E_0 (in fact, *exponentially* smaller), the quantum dynamics are crucially affected by it.

To understand these dynamics, we diagonalize $\boldsymbol{H}$.[47] The eigenvalues are solutions to

$$\det(\boldsymbol{H}-\lambda\boldsymbol{I})=0 \implies \begin{pmatrix} E_0-\lambda & A \\ A & E_0-\lambda \end{pmatrix} = (E_0-\lambda)^2+A^2=0 \implies \lambda_\pm = E_0 \pm A. \tag{188}$$

The eigenvectors solve

$$\begin{pmatrix} E_0-\lambda_\pm & A \\ A & E_0-\lambda_\pm \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \mp A & A \\ A & \mp A \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \implies b=\pm a. \tag{189}$$

Thus, the normalized energy eigenkets are

$$|\pm\rangle = \frac{1}{\sqrt{2}}\left(|1\rangle + |2\rangle\right), \tag{190}$$

with energies $E_\pm = E_0 \pm A$.

Note that these are *also* $\boldsymbol{U}$ eigenkets:

$$\boldsymbol{U}|\pm\rangle = \pm|\pm\rangle, \tag{191}$$

where the $|+\rangle$ ($|-\rangle$) ket is said to be **even** (**odd**) under the reflection. Of course, this is hardly an accident, since $[\boldsymbol{U},\boldsymbol{H}]=0$ by construction. Indeed, we could have used this fact at the outset to diagonalize $\boldsymbol{H}$ more quickly!

By convention, we set $A<0$.[48] The lowest energy state – i.e., the **ground state** – is then the even ket $|+\rangle$. Notice that this state is $\boldsymbol{U}$ invariant. This turns out to be quite a general statement (though not quite a theorem): the ground state of a quantum system with finitely many "degrees of freedom" is invariant[49] under all symmetries of the Hamiltonian. Roughly speaking, this is because the states related by the symmetry are always connected by quantum tunneling terms (however weakly), resulting in a true ground state that is a symmetry-invariant mixture of these states. In QFT, by contrast, **spontaneous symmetry breaking** – in which the ground state (i.e., the vacuum state) is *not* invariant under some symmetry – can occur. This is *not* because QFT violates the laws of quantum mechanics (it does not), but simply because the quantum system in question is "big" enough that tunneling between sufficiently disparate configurations becomes infinitely suppressed.

In summary, the quantum picture of this molecule is *radiccally* different from the classical one: there is no degeneracy, and the lowest energy state is a *quantum superposition* of both nitrogen positions:

$$|+\rangle = \frac{1}{\sqrt{2}}\left(|1\rangle + |2\rangle\right). \tag{192}$$

If we prepare the system in one of its "classical" states, such as $|\psi;0\rangle = |1\rangle$ (which can be done by applying an inhomogeneous electric field), then

$$|\psi;0\rangle = |1\rangle = \frac{1}{\sqrt{2}}\left(|+\rangle + |-\rangle\right) \implies |\psi;t\rangle = \frac{1}{\sqrt{2}}\left(\mathrm{e}^{-0\frac{iE_+t}{\hbar}}|+\rangle + \mathrm{e}^{-0\frac{iE_-t}{\hbar}}|-\rangle\right). \tag{193}$$

[47] Recall that for a 2×2 matrix, the eigenvalues are given by

$$\lambda_\pm = \frac{\mathrm{tr}\pm\sqrt{\mathrm{tr}^2-4\det}}{2}, \tag{187}$$

where tr is the trace of said matrix and det is the determinant.

[48] We can switch the sign of A by redefining $|2\rangle \to -|2\rangle$, which exchanges the even and odd kets. This has the effect of redefining $\boldsymbol{U}\to -\boldsymbol{U}$, but this makes no difference as the overall phase of $\boldsymbol{U}$ is nonphysical. We take $A<0$ by convention because the ground state has an *even* wave function, $\psi(-z)=\psi(z)$.

[49] Here we mean "invariant" in the projective sense, i.e., the symmetry operators act by a (nonphysical) phase on the ground state. This is necessarily true if the ground state is non-degenerate.

Thus, the probability to fins the nitrogen atom above the plane at any later time is

$$|\langle 1|\,\psi;t\rangle|^2 = \left|\frac{e^{-0\frac{iE_+t}{\hbar}} + e^{-0\frac{iE_-t}{\hbar}}}{2}\right|^2 = \cos^2\left(\frac{E_- - E_+}{2}\frac{t}{\hbar}\right) = \frac{1}{2}\left[1+\cos\left(\frac{2\,|A|\,t}{\hbar}\right)\right]. \tag{194}$$

By a similar calculation, the probability to find it below the plane is:

$$|\langle 2|\,\psi;t\rangle|^2 = \frac{1}{2}\left[1-\cos\left(\frac{2\,|A|\,t}{\hbar}\right)\right]. \tag{195}$$

Thus, the nitrogen atome oscillates back and forth with angular frequency $\omega = \frac{2|A|}{\hbar}$. The *ammonia maser*[50] works by coherently stimulating this oscillation using electromagnetic radiation.

[50] A "maser" is a microwave-frequency laser.

2025-02-25: Quantum Symmetries (cont.)

Symmetries and Selection Rules

More generally, a **symmetry** of a quantum system is a *unitary operator* $\boldsymbol{U}$ that preserves the quantum dynamics. Explicitly, $\boldsymbol{U}$ should map a solution to Schrödinger's equation, Equation 138, onto another solution, so that:

$$i\hbar\frac{\partial}{\partial t}\left|\psi;t\right\rangle = \boldsymbol{H}\left(t\right)\left|\psi;t\right\rangle \implies i\hbar\frac{\partial}{\partial t}\left(\boldsymbol{U}\left|\psi;t\right\rangle\right) = \boldsymbol{H}\left(t\right)\left(\boldsymbol{U}\left|\psi;t\right\rangle\right). \tag{196}$$

Applying the product rule to the left-hand sideo of the second relation and simplifying using the first, we see that this is true iff

$$\boldsymbol{U}\boldsymbol{H}\left(t\right)\left|\psi;t\right\rangle = \boldsymbol{H}\left(t\right)\boldsymbol{U}\left|\psi;t\right\rangle \implies \forall\, t,\ \left[\boldsymbol{U},\boldsymbol{H}\left(t\right)\right] = 0, \tag{197}$$

where we assume that $\boldsymbol{U}$ is time-independent, and we can cancel off the ket since for an fixed t, $\left|\psi;t\right\rangle$ can take any value we choose.

Thus, a (time-independent) quantum symmetry is a unitary operator that commutes with the Hamiltonian $\boldsymbol{H}(t)$ for all times t. Such a symmetry might correspond to a spatial reflection or rotation, but could also involve some "internal" transformation on the Hilbert space unrelated to these.

The existence of a quantum symmetry $\boldsymbol{U}$ has immediate dynamical consequences: since $\left[\boldsymbol{U},\boldsymbol{H}\left(t\right)\right] = 0$, $\boldsymbol{H}\left(t\right)$ maps each $\boldsymbol{U}$ eigenspace to itself. Thus, for instance, if $\left|\psi;0\right\rangle$ lies in one of these eigenspaces then $\left|\psi;t\right\rangle$ remains within the same eigenspace for all times t. If $\boldsymbol{U}$ happens to have a non-degenerate spectrum then the dynamics reduce to a time-dependent phase of $\boldsymbol{U}$ eigenket. This occurred in the ammonia example, where the $\boldsymbol{U}$ eigenbasis $\left|\pm\right\rangle$ was the same as the $\boldsymbol{H}$ eigenbasis.

Just like a classical symmetry, a quantum symmetry can be interpreted in either an active or a passive way. In fact, there are *three* natural viewpoints:

1. In the **state viewpoint**, the symmetry acts on the quantum state while holding the observables fixed:

   $$\left|\psi\right\rangle \to \boldsymbol{U}\left|\psi\right\rangle, \quad \boldsymbol{X}\to\boldsymbol{X}. \tag{198}$$

   This is an *active* viewpoint, changing physical quantities such as expectation values.

2. In the **operator viewpoint**, the symmetry acts on the *observables*, while holding the quantum state fixed:

   $$\left|\psi\right\rangle \to \left|\psi\right\rangle, \quad \boldsymbol{X}\to\boldsymbol{U}^{\dagger}\boldsymbol{X}\boldsymbol{U}. \tag{199}$$

   This is again an active viewpoint, and results in the *same* changes in physical quantities as Equation 198. The difference between the state and operator viewpoints is analogous to (bu *independent of*) the difference between the Schrödinger and the Heisenberg pictures.

3. In the **passive viewpoint**, the symmetry acts on the quantum state and also acts *inversely* on the observables:

$$|\psi\rangle \to \boldsymbol{U}|\psi\rangle, \quad \boldsymbol{X} \to \boldsymbol{U}^\dagger \boldsymbol{X} \boldsymbol{U}. \tag{200}$$

As the name suggests, this is a passive transformation, leaving all physical quantities (such as the matrix elements $\langle\psi|\boldsymbol{X}|\chi\rangle$) invariant.

Either active viewpoint correctly specifies how the symmetry acts on physical quantities. For example, define the "position operator" for the nitrogen atom in NH_3:

$$\boldsymbol{Z} := |1\rangle\langle 1| - |2\rangle\langle 2|. \tag{201}$$

This is $+1$ when the nitrogen atom is above the hydrogen plane and -1 when it is below it. One finds:

$$\boldsymbol{U}^\dagger \boldsymbol{Z} \boldsymbol{U} = -\boldsymbol{Z}, \tag{202}$$

so $\boldsymbol{Z}$ is odd under the reflection symmetry, i.e., using the operator viewpoint we conclude that $\langle \boldsymbol{Z}\rangle \to -\langle \boldsymbol{Z}\rangle$ under the reflection, swapping the up/down positions of the nitrogen atom just as expected. The same result follows from the state viewpoint, of course.

By contrast, in the passive viewpoint all physical quantities are invariant. This is actually very useful because it allows one to derive so-called **selection rules**, such as:

$$\langle +|\boldsymbol{H}|-\rangle = 0 \quad \text{or} \quad \langle +|\boldsymbol{Z}|+\rangle = \langle -|\boldsymbol{Z}|-\rangle = 0. \tag{203}$$

Here we use the fact that $|+\rangle$ and $\boldsymbol{H}$ are even under the reflection, whereas $|-\rangle$ and $\boldsymbol{Z}$ are odd. Thus, e.g., in the passive viewpoint $\langle +|\boldsymbol{H}|-\rangle \to -\langle +|\boldsymbol{H}|-\rangle$. However, as all matrix elements are necessarily invariant in this viewpoint, this requires $\langle +|\boldsymbol{H}|-\rangle = -\langle +|\boldsymbol{H}|-\rangle$, i.e., $\langle +|\boldsymbol{H}|-\rangle$ must vanish! The same argument applies to $\langle +|\boldsymbol{Z}|+\rangle$ and $\langle -|\boldsymbol{Z}|-\rangle$.[51]

[51] Of course, the relations in Equation 203 are easy to verify in this example, but the technique they illustrate is bother general and powerful.

Symmetry Groups and Representations

Quatum symmetries have more profound consequences when there are several of them. To better understand these, consequences, we make some observations. First, note that if $\boldsymbol{U}_1$, $\boldsymbol{U}_2$ are both symmetries, then $\boldsymbol{U}_1\boldsymbol{U}_2$ is also a symmetry, since the product of two unitary operators is unitary and

$$[\boldsymbol{U}_1\boldsymbol{U}_2, \boldsymbol{H}(t)] = \boldsymbol{U}_1[\boldsymbol{U}_2, \boldsymbol{H}(t)] + [\boldsymbol{U}_1, \boldsymbol{H}(t)]\boldsymbol{U}_2 = 0. \tag{204}$$

Moreover, if $\boldsymbol{U}$ is a symmetry, then so is $\boldsymbol{U}^\dagger = \boldsymbol{U}^{-1}$:

$$\boldsymbol{U}\boldsymbol{H} - \boldsymbol{H}\boldsymbol{U} = 0 \implies \boldsymbol{U}^\dagger(\boldsymbol{U}\boldsymbol{H} - \boldsymbol{H}\boldsymbol{U})\boldsymbol{U} = \boldsymbol{H}\boldsymbol{U}^\dagger - \boldsymbol{U}^\dagger\boldsymbol{H} = 0. \tag{205}$$

Thus, the symmetries of a quantum system form a **group** G, i.e., a set equipped with a multiplication law $G \times G \to G$ s.t., $\forall\, a, b, c \in G$,

1. $a(bc) = (ab)c$, (associative multiplication)

2. $\exists 1 \in G$ s.t. $a1 = 1a = a$, (multiplicative identity)

3. $\exists a^{-1} \in G$ s.t. $aa^{-1} = a^{-1}a = 1$. (multiplicative inverse)

In the case of a symmetry group, the multiplication law is just that of operator multiplication. Note that this multiplication need not be commutative, i.e., $ab \neq ba$ in general. A group with a commutative (non-commutative) multiplication law is said to be **abelian** (**non-abelian**).

The symmetry group appearing in the ammonia maser example is $\mathbb{Z}_2$, consisting of two elements $\boldsymbol{I}$ and $\boldsymbol{U}$ with $\boldsymbol{U}^2 = \boldsymbol{I}$. To generalize this, consider an analogous problem where an atom can sit at *three* different sites arranges in an equilateral triangle. Now there are three different reflection symmetries, as well as a 120° rotational symmetry. These symmetries can be seen in Figure 23 where the reflection symmetries are labeled R_i and one can see the 120° rotational symmetry. We refer to this 120° rotational symmetry as $\boldsymbol{C}$, though not labeled in Figure 23.

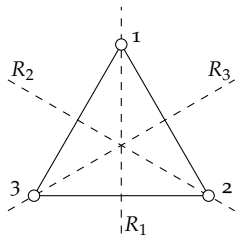

Figure 23: Reflections and rotations interchanging the three sites

Writing the states corresponding to the atom at each site as $|1\rangle$, $|2\rangle$, $|3\rangle$, these symmetries take the form:[52]

$$\begin{aligned} \boldsymbol{R}_1 &= |1\rangle\langle 1| + |2\rangle\langle 3| + |3\rangle\langle 2|\,, \\ \boldsymbol{R}_2 &= |1\rangle\langle 3| + |2\rangle\langle 2| + |3\rangle\langle 1|\,, \\ \boldsymbol{R}_3 &= |1\rangle\langle 2| + |2\rangle\langle 1| + |3\rangle\langle 3|\,, \\ \boldsymbol{C} &= |1\rangle\langle 3| + |2\rangle\langle 1| + |3\rangle\langle 2|\,. \end{aligned} \tag{207}$$

The symmetry group generated by these rotations and reflections, consisting of $\left\{\boldsymbol{I}, \boldsymbol{C}, \boldsymbol{C}^{-1}, \boldsymbol{R}_1, \boldsymbol{R}_2, \boldsymbol{R}_3\right\}$, is the **dihedral group** D_3. Note every group element can be written in terms of a pair of suitable chosen **generators**, e.g., in terms of $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$, we have $\boldsymbol{C} = \boldsymbol{R}_1\boldsymbol{R}_2$, $\boldsymbol{C}^{-1} = \boldsymbol{R}_2\boldsymbol{R}_1$, and $\boldsymbol{R}_3 = \boldsymbol{R}_2\boldsymbol{R}_1\boldsymbol{R}_2 = \boldsymbol{R}_1\boldsymbol{R}_2\boldsymbol{R}_1$. A **presentation** of a group is a list of elements generating it, together with the relations they satisfy.[53]

Imposing D_3 symmetry, the most general Hamiltonian would take the form:

$$\boldsymbol{H} = E_0\left(|1\rangle\langle 1| + |2\rangle\langle 2| + |3\rangle\langle 3|\right) + \frac{A}{2}\left(|1\rangle\langle 2| + |2\rangle\langle 1| + |1\rangle\langle 3| + |3\rangle\langle 1| + |2\rangle\langle 3| + |3\rangle\langle 2|\right) \tag{209}$$

for real E_0, A. To simplify the task of diagonalizing $\boldsymbol{H}$, we might start by diagonalizing, e.g., $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$. However, since $\boldsymbol{R}_1\boldsymbol{R}_2 \neq \boldsymbol{R}_2\boldsymbol{R}_1$, these are not simultaneously diagonalizable. Instead, let us start with the simple observation that the ket

$$|0\rangle := \frac{1}{\sqrt{3}}\left(|1\rangle + |2\rangle + |3\rangle\right) \tag{210}$$

is an eigenket of both $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$ with eigenvalue $+1$, i.e., $|0\rangle$ is *invariant* under the entire D_3 symmetry group. One finds:

$$\boldsymbol{H}\,|0\rangle = (E_0 + A)\,|0\rangle\,, \tag{211}$$

so this indeed produces an $\boldsymbol{H}$ eigenvector.

Now consider the **orthogonal complement** to the 1d subspace $\mathcal{H}_0$ generated by $|0\rangle$, i.e., the subspace $\mathcal{H}_{\overline{0}}$ selected by the projector $\boldsymbol{\Pi}_{\overline{0}} = \boldsymbol{I} - |0\rangle\langle 0|$. This 2d subspace

[52] These expressions for $\boldsymbol{R}_i$ and $\boldsymbol{C}$ give rise to the following relation:

$$\boldsymbol{C} = \boldsymbol{R}_1\boldsymbol{R}_2 = \boldsymbol{R}_2\boldsymbol{R}_3 = \boldsymbol{R}_3\boldsymbol{R}_1. \tag{206}$$

[53] For D_3, one presentation is:

$$D_3 : \left\{r_1, r_2 \;:\; r_1^2 = r_2^2 = (r_1 r_2)^3 = 1\right\}. \tag{208}$$

This readily generalizes to the dihedral group $D_n : \{r_1, r_2 \;:\; r_1^2 = r_2^2 = (r_1 r_2)^n = 1\}$, which is the symmetry group of a regular n-gon.

has an orthonormal basis

$$|1'\rangle = \frac{1}{\sqrt{2}}\left(|1\rangle - |3\rangle\right), \quad |2'\rangle = \frac{1}{\sqrt{6}}\left(2\,|2\rangle - |1\rangle - |3\rangle\right). \tag{212}$$

Now observe that

$$\begin{aligned} \boldsymbol{R}_1\,|1'\rangle &= \frac{1}{2}\,|1'\rangle - \frac{\sqrt{3}}{2}\,|2'\rangle, \qquad \boldsymbol{R}_2\,|1'\rangle = -\,|1'\rangle, \\ \boldsymbol{R}_1\,|2'\rangle &= -\frac{\sqrt{3}}{2}\,|1'\rangle - \frac{1}{2}\,|2'\rangle, \quad \boldsymbol{R}_2\,|2'\rangle = |2'\rangle. \end{aligned} \tag{213}$$

In other words, D_3 maps $\mathcal{H}_{\bar{0}}$ onto itself. We find:

$$\boldsymbol{H}\,|1'\rangle = \left(E_0 - \frac{A}{2}\right)|1'\rangle, \quad \boldsymbol{H}\,|2'\rangle = \left(E_0 - \frac{A}{2}\right)|2'\rangle. \tag{214}$$

Thus, $\mathcal{H}_{\bar{0}}$ is an $\boldsymbol{H}$ eigenspace, and we have succeeded in diagonlizing the Hamiltonian.[54]

[54] As in the ammonia example, one finds $A < 0$ typically, so that the invariant state $|0\rangle$ is the ground state.

To understand what is going on here, note that the quantum symmetries form a **unitary representation** of the abstract symmetry group acting on $\mathcal{H}$, i.e., for each $g \in G$ there is a unitary operator $\boldsymbol{U}_g$ acting on $\mathcal{H}$ s.t.

$$\boldsymbol{U}_g \boldsymbol{U}_h = \boldsymbol{U}_{gh}. \tag{215}$$

In the above example, the representation of D_3 on $\mathcal{H}$ was **reducible**, i.e., $\mathcal{H} = \mathcal{H}_0 \oplus \mathcal{H}_{\bar{0}}$, where $\mathcal{H}_0$ and $\mathcal{H}_{\bar{0}}$ are D_3 invariant subspaces and thus themselves furnish representations of D_3. More generally, any unitary G representation is the direct sum of **irreducible** G representations (G irreps). If G is abelian then the irrep components of $\mathcal{H}$ are 1d and are simply the simultaneous eigenvectors of the symmetry operators. Thus, the G irrep components of $\mathcal{H}$ (some of which will be multidimensional when G is not abelian) are the appropriate *non-abelian generalization* of these simultaneous eigenvectors. Likewise, if the *same* G irrep appears more than once in the decomposition of $\mathcal{H}$ then the subspace generated by identical G irreps is the appropriate non-abelian generalization of a degenerate eigenspace.

With this in mind, we can now concisely state the dynamical consequences of a symmetry group G. By definition of a symmetry, $\left[\boldsymbol{U}_g, \boldsymbol{H}(t)\right] = 0$ for all $g \in G$. As a consequence of "Schur's lemma"[55], this implies that $\boldsymbol{H}(t)$ cannot mix together inequivalent G irreps. Moreover, when acting on a single G irrep, it can only produce an overall factor, i.e., the states in the irrep are $\boldsymbol{H}(t)$ eigenvectors. This si what we observed above. Thus, the dynamics preserve G irreps, just as they preserve the simultaneous eigenvalues of the symmetry operators in the abelian case.

[55] A technical result that we will not state.

Continuous Symmetries, Conserved Charges, and Spins

A **continuous** symmetry is a family of symmetry operators parameterized by one or more real numbers, $\boldsymbol{U} = \boldsymbol{U}(\theta_1, \ldots, \theta_k)$ for $\theta_1, \ldots, \theta_k \in \mathbb{R}$, s.t. the identity operator lies within the family.

One Parameter Groups

A **one-parameter group** of symmetries is a family parameterized by a single real variable, $\theta \in \mathbb{R}$, s.t. the group multiplication law corresponds to addition of the parameters:

$$\boldsymbol{U}_{\theta_1}\boldsymbol{U}_{\theta_2} = \boldsymbol{U}_{\theta_1+\theta_2}. \tag{216}$$

Such a subgroup is always abelian. It can be shown that *any* continuous symmetry lies in some one-parameter subgroup.

For example, the Hamiltonian in Equation 145 has a one-parameter continuous symmetry:

$$\boldsymbol{U}_\theta = |+\hat{\boldsymbol{z}}\rangle\langle+\hat{\boldsymbol{z}}| + e^{i\theta}\,|-\hat{\boldsymbol{z}}\rangle\langle-\hat{\boldsymbol{z}}|\,. \tag{217}$$

This symmetry alters the relative phase between $|+\hat{\boldsymbol{z}}\rangle$ and $|-\hat{\boldsymbol{z}}\rangle$. Abstractly, this can be thought of as the group of unit-magnitude complex numbers under multiplication, $e^{i\theta}\,e^{i\theta'} = e^{i(\theta+\theta')}$, known as $U(1)$.

In the operator viewpoint, we find

$$\boldsymbol{S}_x \to \boldsymbol{U}_\theta^\dagger \boldsymbol{S}_x \boldsymbol{U}_\theta = \frac{\hbar}{2}\left(e^{i\theta}\,|+\hat{\boldsymbol{z}}\rangle\langle-\hat{\boldsymbol{z}}| + e^{-i\theta}\,|-\hat{\boldsymbol{z}}\rangle\langle+\hat{\boldsymbol{z}}|\right) = \cos\theta\boldsymbol{S}_x - \sin\theta\boldsymbol{S}_y. \tag{218}$$

Likewise,

$$\boldsymbol{S}_y \to \boldsymbol{U}_\theta^\dagger \boldsymbol{S}_y \boldsymbol{U}_\theta = \frac{\hbar}{2i}\left(e^{i\theta}\,|+\hat{\boldsymbol{z}}\rangle\langle-\hat{\boldsymbol{z}}| - e^{-i\theta}\,|-\hat{\boldsymbol{z}}\rangle\langle+\hat{\boldsymbol{z}}|\right) = \sin\theta\boldsymbol{S}_x + \cos\theta\boldsymbol{S}_y, \tag{219}$$

whereas $\boldsymbol{S}_z$ is invariant. Written in terms of expectation values,

$$\begin{pmatrix}\langle S_x\rangle \\ \langle S_y\rangle\end{pmatrix} \to \begin{pmatrix}\langle S_x'\rangle \\ \langle S_y'\rangle\end{pmatrix} = \begin{pmatrix}\cos\theta & -\sin\theta \\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}\langle S_x\rangle \\ \langle S_y\rangle\end{pmatrix}, \tag{220}$$

where the result is now valid in either active viewpoint. We recognize this as an *active spatial rotation* in the counterclockwise direction about the $\hat{\boldsymbol{z}}$ axis by an angle θ, shown in Figure 24.

The **generator** of 1-parameter symmetry $\boldsymbol{U}_\theta$ is defined as:

$$\boldsymbol{Q} := i\hbar\frac{d\boldsymbol{U}_\theta}{d\theta}\boldsymbol{U}_\theta^\dagger. \tag{221}$$

Notice the similarity with how the quantum Hamiltonian was defined in Equation 135. Similar to before, this definition immediately implies that $\boldsymbol{Q}$ has certain properties, seen in the following question.

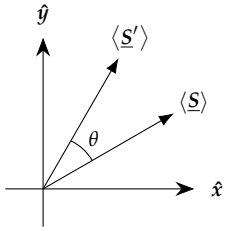

Figure 24: $\boldsymbol{U}_\theta$ rotates the expectation value of the spin $\langle \underline{\boldsymbol{S}} \rangle$ counterclockwise by an angle θ about the $\hat{\boldsymbol{z}}$ axis.

Question 34. *Consider Equation 221.*

1. *Show that $\boldsymbol{Q}$ is independent of θ, and therefore equivalently:*

$$\boldsymbol{Q} := i\hbar\frac{d\boldsymbol{U}_\theta}{d\theta}\bigg|_{\theta=0}. \tag{222}$$

Proof. □

2. *Show that* $\boldsymbol{Q}^\dagger = \boldsymbol{Q}$.

Proof. □

3. *Show that* $[\boldsymbol{Q}, \boldsymbol{H}] = 0$.

Proof. □

Since $\boldsymbol{Q}$ has no explicit time-dependence by assumption, this implies that it is **conserved**:

$$\frac{d\boldsymbol{Q}_H}{dt} = 0. \tag{223}$$

A self-adjoint operator that satisfies this equation (and therefore has a time-independent expectation value) is known as a **conserved charge**. Thus, the generator of a 1-parameter continuous symmetry is a conserved charge, which is a quantum version of **Noether's theorem**.[56]

[56] The discussion here excludes certain "time-independent" continuous symmetries.

Let us find the conserved charge associated to the rotational symmetry discussed above. To make things more symmetric, we first rewrite

$$\boldsymbol{U}_\theta = \mathrm{e}^{-i\theta/2} \left|+\hat{\boldsymbol{z}}\right\rangle\!\left\langle+\hat{\boldsymbol{z}}\right| + \mathrm{e}^{i\theta/2} \left|-\hat{\boldsymbol{z}}\right\rangle\!\left\langle-\hat{\boldsymbol{z}}\right|, \tag{224}$$

after redefining the symmetry operator by a nonphysical overall phase. We have

$$\frac{d\boldsymbol{U}_\theta}{d\theta} = -\frac{i}{2}\mathrm{e}^{-i\theta/2} \left|+\hat{\boldsymbol{z}}\right\rangle\!\left\langle+\hat{\boldsymbol{z}}\right| + \frac{i}{2}\mathrm{e}^{i\theta/2} \left|-\hat{\boldsymbol{z}}\right\rangle\!\left\langle-\hat{\boldsymbol{z}}\right| \xrightarrow{\theta\to 0} -\frac{i}{2}\left|+\hat{\boldsymbol{z}}\right\rangle\!\left\langle+\hat{\boldsymbol{z}}\right| + \frac{i}{2}\left|-\hat{\boldsymbol{z}}\right\rangle\!\left\langle-\hat{\boldsymbol{z}}\right|, \tag{225}$$

so, applying Equation 222,

$$\boldsymbol{Q} = \frac{\hbar}{2}\left(\left|+\hat{\boldsymbol{z}}\right\rangle\!\left\langle+\hat{\boldsymbol{z}}\right| - \left|-\hat{\boldsymbol{z}}\right\rangle\!\left\langle-\hat{\boldsymbol{z}}\right|\right) = \boldsymbol{S}_z. \tag{226}$$

Note that, since the overall phase of a symmetry operator is nonphysical, we can always shift our generator by some multiple of the identity operator. To eliminate this ambiguity, it is common to enforce $\operatorname{tr}\boldsymbol{Q} = 0$, although this is not mandatory. Moreover, note that the constraint $\operatorname{tr}\boldsymbol{Q} = 0$ is only well-defined in the case where $\mathcal{H}$ is finite-dimensional.

Therefore the generator of rotations about the $\hat{\boldsymbol{z}}$ axis is nothing but the $\hat{\boldsymbol{z}}$ component of the (spin) angular momentum.

We can follow the same logic to describe rotations about the $\hat{\boldsymbol{x}}$ or $\hat{\boldsymbol{y}}$ axis. For instance, by analogy with the above, a clockwise rotation by an angle θ about the $\hat{\boldsymbol{x}}$ axis should be described by 1-parameter symmetry:

$$\boldsymbol{U}_\theta^{(\hat{\boldsymbol{x}})} = \mathrm{e}^{-i\theta/2} \left|+\hat{\boldsymbol{x}}\right\rangle\!\left\langle+\hat{\boldsymbol{x}}\right| + \mathrm{e}^{i\theta/2} \left|-\hat{\boldsymbol{x}}\right\rangle\!\left\langle-\hat{\boldsymbol{x}}\right|. \tag{227}$$

It is easy to check that the one-parameter subgroups associated to different rotation axes do not commute, just as rotations about different axes in 3d do not commute.

Question 35.

1. *Verify that Equation 227 acts on* $\boldsymbol{S}_x$, $\boldsymbol{S}_y$, $\boldsymbol{S}_z$ *as expected in the operator viewpoint.*

Proof. □

2. *Show that Equation 227 can be written in the* $\left|+\hat{\boldsymbol{z}}\right\rangle$, $\left|-\hat{\boldsymbol{z}}\right\rangle$ *basis as*

$$\boldsymbol{U}_\theta^{(\hat{\boldsymbol{x}})} = \cos\left(\frac{\theta}{2}\right)\left(\left|+\hat{\boldsymbol{z}}\right\rangle\!\left\langle+\hat{\boldsymbol{z}}\right| + \left|-\hat{\boldsymbol{z}}\right\rangle\!\left\langle-\hat{\boldsymbol{z}}\right|\right) - i\sin\left(\frac{\theta}{2}\right)\left(\left|+\hat{\boldsymbol{z}}\right\rangle\!\left\langle-\hat{\boldsymbol{z}}\right| + \left|-\hat{\boldsymbol{z}}\right\rangle\!\left\langle+\hat{\boldsymbol{z}}\right|\right). \tag{228}$$

Proof. □

3. *Likewise, argue that*

$$\boldsymbol{U}_{\theta}^{(\hat{\boldsymbol{y}})} = \cos\left(\frac{\theta}{2}\right)\left(|+\hat{\boldsymbol{z}}\rangle\langle+\hat{\boldsymbol{z}}| + |-\hat{\boldsymbol{z}}\rangle\langle-\hat{\boldsymbol{z}}|\right) - \sin\left(\frac{\theta}{2}\right)\left(|+\hat{\boldsymbol{z}}\rangle\langle-\hat{\boldsymbol{z}}| - |-\hat{\boldsymbol{z}}\rangle\langle+\hat{\boldsymbol{z}}|\right) \tag{229}$$

and use this form to show that $\boldsymbol{Q}^{(\hat{\boldsymbol{y}})} = \boldsymbol{S}_y$.

Proof. □

2025-02-27: Quantum Symmetries (cont.)

One Parameter Groups (cont.)

If we know the generator $\boldsymbol{Q}$ of a symmetry then the symmetry operator $\boldsymbol{U}_\theta$ itself can be constructed by solving the differential equation

$$i\hbar \frac{d\boldsymbol{U}_\theta}{d\theta} = \boldsymbol{Q}\boldsymbol{U}_\theta. \tag{230}$$

Because $\boldsymbol{Q}$ is independent of θ, the solution is just an exponential:

$$\boldsymbol{U}_\theta = \exp\left[-\frac{i\theta}{\hbar}\boldsymbol{Q}\right]. \tag{231}$$

This is analogous to how the time-evolution operator is related to the Hamiltonian when $\boldsymbol{H}$ is time-independent, see Equation 143. Indeed, when $\frac{\partial \boldsymbol{H}}{\partial t} = 0$, the Heisenberg equation implies that $\frac{d\boldsymbol{H}_H}{dt} = 0$, and therefore $\boldsymbol{H}$ is a conserved charges in this case; a.k.a the conserved energy. The associated symmetry operator is

$$\boldsymbol{U}_{\Delta t} = \underbrace{\exp\left[-\frac{i\Delta t}{\hbar}\boldsymbol{H}\right]}_{=\boldsymbol{U}(t+\Delta t,t)} \implies \boldsymbol{U}_{\Delta t}\left|\psi;t\right\rangle = \left|\psi;t+\Delta t\right\rangle. \tag{232}$$

Thus, the associated symmetry is *time translation invariance*. This is the quantum manifestation of the fact that energy conservation is a consequence of time-translation invariance, as implied by Noether's theorem.

Question 36. *Show that for infinitesimal θ, the quantum states and observables transform under a one-parameter symmetry as follows:*

$$\begin{aligned}
\left|\psi\right\rangle \to \left|\psi\right\rangle - \frac{i\theta}{\hbar}\boldsymbol{Q}\left|\psi\right\rangle, \qquad & \boldsymbol{X}\to\boldsymbol{X}, && (\textit{state viewpoint}),\\
\left|\psi\right\rangle \to \left|\psi\right\rangle, \qquad & \boldsymbol{X}\to\boldsymbol{X} - \frac{i\theta}{\hbar}\left[\boldsymbol{X},\boldsymbol{Q}\right], && (\textit{operator viewpoint}),\\
\left|\psi\right\rangle \to \left|\psi\right\rangle - \frac{i\theta}{\hbar}\boldsymbol{Q}\left|\psi\right\rangle, \qquad & \boldsymbol{X}\to\boldsymbol{X} + \frac{i\theta}{\hbar}\left[\boldsymbol{X},\boldsymbol{Q}\right], && (\textit{passive viewpoint}).
\end{aligned}$$

Proof. □

Lie Groups and Lie Algebras

Now suppose there are several conserved charges $\boldsymbol{Q}_1, \ldots, \boldsymbol{Q}_n$. Clearly, any linear combination of these is also a conserved charge.[57] Thus, we have a multi-parameter family of symmetry operators:

$$\boldsymbol{U}_{\theta_1,\ldots,\theta_n} := \exp\left[-\frac{i}{\hbar}\left(\theta_1\boldsymbol{Q}_1 + \ldots + \theta_n\boldsymbol{Q}_n\right)\right]. \tag{233}$$

What happens when we multiply these operators together? If the charges commute then the parameters just add, e.g.,

$$\exp\left[-\frac{i}{\hbar}\theta_1\boldsymbol{Q}_1\right]\exp\left[-\frac{i}{\hbar}\theta_2\boldsymbol{Q}_2\right] = \exp\left[-\frac{i}{\hbar}\left(\theta_1\boldsymbol{Q}_1 + \theta_2\boldsymbol{Q}_2\right)\right] \tag{234}$$

[57] I.e., a self-adjoint operator satisfying Equation 223.

if $[Q_1, Q_2] = 0$. If the charges do not commute then the product of exponentials can be computed using the **Baker-Campbell-Hausdorff (BCH) formula**:

$$\exp[X]\exp[Y] = \exp\left[X + Y + \frac{1}{2}[X,Y] + \frac{1}{12}[X,[X,Y]] - \frac{1}{12}[Y,[X,Y]] + \cdots\right] \tag{235}$$

where the omitted terms – specified in the full BCH formula first obtained by Eugen Dynkin (1947) – involve higher-order commutators.

Applying the BCH formula,

$$\exp\left[-\frac{i}{\hbar}\theta_1 Q_1\right]\exp\left[-\frac{i}{\hbar}\theta_2 Q_2\right] = \exp\left[-\frac{i}{\hbar}\left(\theta_1 Q_1 + \theta_2 Q_2 + \frac{\theta_1\theta_2}{i\hbar}[Q_1, Q_2] + \cdots\right)\right] \tag{236}$$

where the omitted terms are cubic or higher order in the parameters $\theta_{1,2}$. We can neglect these terms when $\theta_{1,2}$ are small, so this formula is meaningful sufficiently near the identity operator in the symmetry group. There are two important things to notice about it:

1. The two symmetry operators do not commute, since by the same reasoning

$$\exp\left[-\frac{i}{\hbar}\theta_2 Q_2\right]\exp\left[-\frac{i}{\hbar}\theta_1 Q_1\right] = \exp\left[-\frac{i}{\hbar}\left(\theta_1 Q_1 + \theta_2 Q_2 - \frac{\theta_1\theta_2}{i\hbar}[Q_1, Q_2] + \cdots\right)\right]. \tag{237}$$

   The failure of the two symmetry operators to commute near the identity (for small $\theta_{1,2}$) is a direct consequence of (and equivalent to) the failure of the charges to commute.

2. In order for the result in Equation 236 to be of the form described in Equation 233, $\frac{1}{i\hbar}[Q_1, Q_2]$ must be a linear combination of the conserved charges $Q_1, \ldots, Q_n$. While the commutator of two conserved charges is indeed conserved (see **Question 37**), we see that to generate a symmetry group closed under multiplication, we must select a set of conserved charges that is closed under commutation:[58]

$$\left[Q_i, Q_j\right] = i\hbar\sum_k f_{ij}^k Q_k. \tag{238}$$

   This is the **Lie algebra** associated to this continuous symmetry group, which depends on the **structure constants** $f_{ij}^k = -f_{ji}^k \in \mathbb{R}$. Further properties of these structure constants are explored in **Question 38** and **Question 39**.

Question 37. *Use the Jacobi identity (Equation 176) to show that if $Q_{1,2}$ are conserved charges then $\frac{1}{i\hbar}[Q_1, Q_2]$ is a conserved charge.*

Proof. □

Question 38. *Using the Jacobi identity, show that*

$$\sum_l \left(f_{ij}^l f_{lk}^m + f_{jk}^l f_{li}^m + f_{ki}^l f_{lj}^m\right) = 0. \tag{239}$$

Proof. □

[58] Note that the identity operator I may have to be included in the set of conserved charges to make the Lie algebra closed – even though it generates a trivial symmetry under which all kets pick up the same phase – as in the Heisenberg algebra, $[x, p] = i\hbar I$, $[x, I] = [p, I] = 0$. However, when $\dim\mathcal{H} < \infty$, $\operatorname{tr}\left(\left[Q_i, Q_j\right]\right) = \operatorname{tr}\left(Q_i Q_j\right) - \operatorname{tr}\left(Q_j Q_i\right) = 0$ whereas $\operatorname{tr}(I) = \dim\mathcal{H} \neq 0$, so this does not occur. (In the infinite-dimensional case, the trace is not well-defined.)

Question 39. *Assuming the Hilbert space is finite-dimensional, we can define* $h_{ij} := \mathrm{tr}\,[Q_i, Q)j]$.

1. *For linearly-independent* $Q_1, \ldots, Q_n$*, show that* h_{ij} *is a positive-definite real symmetric matrix.*

   Proof. □

2. *Show that* $\mathrm{tr}\left([\boldsymbol{A}, \boldsymbol{B}]\, \boldsymbol{C} + [\boldsymbol{A}, \boldsymbol{C}]\, \boldsymbol{B}\right) = 0$ *for any operators* $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$*. Use this to show that*

$$\sum_k \left(f^l_{ij} h_{lk} + f^l_{ik} h_{ij} \right) = 0. \tag{240}$$

   Proof. □

3. *Argue that one can pick a basis for the conserved charges s.t.* $h_{ij} = \delta_{ij}$*. Show that tin this basis* $f_{ijk} = -f_{ikj}$*, hence the structure constants are completely antisymmetric.*[59]

   Proof. □

[59] A Lie algebra of this type is said to be **reductive**.

Note that the parameterization in Equation 233 provides a continuous, invertible map between the region near the origin in $\mathbb{R}^n$ and the region "near" the identity in the group G. Using the group multiplication law, the region "near" any element of the group can similarly be mapped onto a small patch of $\mathbb{R}^n$. This makes the symmetry group of G into a **manifold**.

Definition 40. A group that is also a manifold s.t. the group multiplication law is continuous is a **Lie group**.

Abstractly, a Lie algebra[60] is a vector space over $\mathbb{R}$ equipped with a bilinear **Lie bracket** $[X, Y]_{\mathrm{LB}}$ satisfying the same properties, Equation 174 and Equation 175, as the Poisson bracket in Hamiltonian mechanics. Since it determines the group multiplication law infinitesimally near the identity, the abstract Lie algebra $\mathfrak{g}$ defined by the commutation relations of the conserved charges depends only on the abstract Lie group G, and not on its particular representation in terms of unitary operators acting on $\mathcal{H}$.

[60] What we describe here is a *real* Lie algebra, which is what arises in quantum mechanics. For a complex Lie algebra, we replace $\mathbb{R}$ with $\mathbb{C}$. (Other fields may be used, but this is not common in physics.)

Definition 41. A **Unitary representation** of $\mathfrak{g}$ on a Hilbert space $\mathcal{H}$ is a linear map $x \mapsto Q_x$ from $\mathfrak{g}$ to self-adjoint opertors on $\mathcal{H}$, s.t.

$$[x, y]_{\mathrm{LB}} \mapsto i\hbar \left[Q_x, Q_y \right]. \tag{241}$$

This is much like the canonical quantization procedure, bit the Lie algebra $\mathfrak{g}$ need not (but *could*) arise from the Poisson bracket of any classical system.

We now revisit rotational symmetries. We saw previously that the generators for right-handed rotations about the $\hat{x}$, $\hat{y}$, and $\hat{z}$ aces acting on the qubit are the

spin operators S_x, S_y, and S_z. From their commutator algebra we can read off the associated Lie bracket:

$$\left[S_i, S_j\right] = i\hbar\varepsilon_{ijk}S_k \quad \Longrightarrow \quad \left[L_i, L_j\right]_{\text{LB}} = \varepsilon_{ijk}L_k, \tag{242}$$

where we denote the corresponding elements of the abstract Lie algebra as L_x, L_y, and L_z. This is the Lie algebra $\mathfrak{so}(3)$ of the abstract symmetry group $\mathrm{SO}(3)$ of rotations in three spatial dimensions. Thus, the qubit is a particular *representation* of this Lie algebra.

Indeed, it makes perfect sense for the Hilbert space of each atom in the Stern-Gerlach experiment to furnish a representation of $\mathfrak{so}(3)$; since in the absence of an external magnetic field the atom sees no preferred direction(s) in space, that space of quantum states should map to itself under rotations. It is very interesting to turn this observation around, and ask what are *all* the possible Hilbert spaces that can be mapped to themselves under spatial rotations. As before, such a Hilbert space $\mathcal{H}$ will split into *irreducible* unitary representations of the Lie algebra $\mathfrak{so}(3)$, so we need only classify these irreps.

Question 42. *Consider the spin algebra in Equation 243.*

1. *Using this algebra and* ***Definition 43****, show that* $\left[\boldsymbol{L}^2, \boldsymbol{L}_i\right] = 0$.

   Proof. □

2. *Show that this algebra is equivalent to the ladder operator algebra in* ***Definition 44****.*

   Proof. □

Spin Representations

To represent the Lie algebra in Equation 242, we seek self-adjoint operators $\boldsymbol{L}_x$, $\boldsymbol{L}_y$, and $\boldsymbol{L}_z$ acting on some *a priori* unknown Hilbert space $\mathcal{H}$, s.t.

$$\left[\boldsymbol{L}_i, \boldsymbol{L}_j\right] = i\hbar\varepsilon_{ijk}\boldsymbol{L}_k. \tag{243}$$

Note that these are no longer the spin operators $\boldsymbol{S}_x$, $\boldsymbol{S}_y$, and $\boldsymbol{S}_z$ acting on the qubit! We will only impose these commutation relations and try to reconstruct what $\mathcal{H}$ is from that.

We can extend the algebra in Equation 243 via the following definition:

Definition 43. The **total angular momentum** operator is defined in terms of the angular momentum "components" $\boldsymbol{L}_x$, $\boldsymbol{L}_y$, and $\boldsymbol{L}_z$ as

$$\boldsymbol{L}^2 L = \boldsymbol{L}_x^2 + \boldsymbol{L}_y^2 + \boldsymbol{L}_z^2 \quad \Longrightarrow \quad \left[\boldsymbol{L}^2, \boldsymbol{L}_i\right] = 0, \tag{244}$$

where the implication follows from **Question 42**.

Thus, the operator $\boldsymbol{L}^2$ is *central*, i.e., it commutes with all the other operators in the algebra. Because of this, we can simultaneously diagonalize $\boldsymbol{L}^2$ and $\boldsymbol{L}_z$. Consider a normalized simultaneous eigenvector:

$$\boldsymbol{L}^2 \left|\lambda; m\right\rangle = \hbar^2\lambda \left|\lambda; m\right\rangle, \quad \boldsymbol{L}_z \left|\lambda; m\right\rangle = \hbar m \left|\lambda; m\right\rangle, \tag{245}$$

where λ and m fix the $\boldsymbol{L}^2$ and $\boldsymbol{L}_z$ eigenvalues, up to convenient factors of $\hbar$ that we have chose to include.

Definition 44. We define the **ladder operators**

$$\boldsymbol{L}_\pm := \boldsymbol{L}_x \pm i\boldsymbol{L}_y \quad \Longrightarrow \quad \left[\boldsymbol{L}_z, \boldsymbol{L}_\pm\right] = \pm\hbar\boldsymbol{L}_\pm, \quad \left[\boldsymbol{L}_+, \boldsymbol{L}_-\right] = 2\hbar\boldsymbol{L}_z, \tag{246}$$

where $\boldsymbol{L}_- := \boldsymbol{L}_+^\dagger$. The implication here also follows from **Question 42**.

Using the ladder operator algebra from **Definition 44**, we find

$$\boldsymbol{L}_z \boldsymbol{L}_\pm = \boldsymbol{L}_\pm \left(\boldsymbol{L}_z \pm \hbar\right) \implies \boldsymbol{L}_z \left(\boldsymbol{L}_\pm \left|\lambda; m\right\rangle\right) = \boldsymbol{L}_\pm \left(\boldsymbol{L}_z \pm \hbar\right) \left|\lambda; m\right\rangle = \hbar \left(m \pm 1\right) \boldsymbol{L}_\pm \left|\lambda; m\right\rangle . \tag{247}$$

Thus, provided it does not vanish, we conclude that $\boldsymbol{L}_\pm \left|\lambda; m\right\rangle$ is an $\boldsymbol{L}_z$ eigenvector with eigenvalue $\hbar\left(m \pm 1\right)$.[61] Let us label this eigenvector $\left|\lambda; m \pm 1\right\rangle$ up to a normalization factor so that

$$\boldsymbol{L}_\pm \left|\lambda; m\right\rangle = \mathcal{N} \left|\lambda; m \pm 1\right\rangle , \tag{248}$$

for $\mathcal{N} > 0$ to be determined. Since $\boldsymbol{L}_+$ *increases* the $\boldsymbol{L}_z$ eigenvalue, it is known as the **raising operator**, whereas $\boldsymbol{L}_-$ (which *decreases* the $\boldsymbol{L}_z$ eigenvalue) is the **lowering operator**.

To find the value of $\mathcal{N}$, we compute

$$\left|\mathcal{N}\right|^2 = \left\|\boldsymbol{L}_\pm \left|\lambda; m\right\rangle\right\|^2 = \left\langle\lambda; m\right| \boldsymbol{L}_\pm^\dagger \boldsymbol{L}_\pm \left|\lambda; m\right\rangle = \left\langle\lambda; m\right| \boldsymbol{L}_\mp \boldsymbol{L}_\pm \left|\lambda; m\right\rangle . \tag{249}$$

This can be simplified as follows:

$$\begin{aligned} \boldsymbol{L}_\mp \boldsymbol{L}_\pm &= \left(\boldsymbol{L}_x \mp i\boldsymbol{L}_y\right) \left(\boldsymbol{L}_x \pm i\boldsymbol{L}_y\right) \\ &= \boldsymbol{L}_x^2 + \boldsymbol{L}_y^2 \pm i\left[\boldsymbol{L}_x, \boldsymbol{L}_y\right] \\ &= \boldsymbol{L}^2 - \boldsymbol{L}_z^2 \mp \hbar \boldsymbol{L}_z \\ &= \boldsymbol{L}^2 - \boldsymbol{L}_z \left(\boldsymbol{L}_z \pm \hbar\right) . \end{aligned} \tag{250}$$

Thus,

$$\left|\mathcal{N}\right|^2 = \hbar^2 \left|\lambda - m\left(m \pm 1\right)\right| . \tag{251}$$

Now there is an interesting conundrum: clearly $\left|\mathcal{N}\right|^2 \geqslant 0$, but as we repeatedly apply the raising $\boldsymbol{L}_+$, we obtain kets with larger and larger values of m until $m\left(m+1\right)$ becomes larger than λ, implying that $\left|\mathcal{N}\right|^2 < 0$. The only way to avoid this contradiction is if there is a **highest weight state** that is annihilated by the raising operator:

$$\boldsymbol{L}_+ \left|\lambda; m_{\text{high}}\right\rangle = 0. \tag{252}$$

Comparing with Equations 248 and 251, we see that

$$\lambda = m_{\text{high}} \left(m_{\text{high}} + 1\right) . \tag{253}$$

Since λ is the same for *all* the eigenkets in the ladder, it is convenient to relabel them using the weight $\ell := m_{\text{high}}$ of the highest weight state instead of the $\boldsymbol{L}^2/\hbar^2$ eigenvalue λ:

$$\boldsymbol{L}^2 \left|\ell, m\right\rangle = \hbar^2 \ell \left(\ell + 1\right) \left|\ell, m\right\rangle , \quad \boldsymbol{L}_z \left|\lambda; m\right\rangle = \hbar m \left|\ell, m\right\rangle , \tag{254}$$

where $\ell \geqslant 0$. Since the highest weight state $\left|\ell, \ell\right\rangle \propto \boldsymbol{L}_+^k \left|\ell, m\right\rangle$ can be reached by applying $\boldsymbol{L}_+$ an integer number of times, we must also have $m \leqslant \ell$ with $\ell - m$ an integer.

[61] Similarly, since $\boldsymbol{L}^2 \boldsymbol{L}_\pm = \boldsymbol{L}_\pm \boldsymbol{L}^2$ (which follows from $\boldsymbol{L}^2$ being central), $\boldsymbol{L}_\pm \left|\lambda; m\right\rangle$ is also an $\boldsymbol{L}^2$ eigenvector with the same $\hbar^2\lambda$ eigenvalue as before.

Since $m(m \pm 1)$ also increases without bound as m becomes increasingly negative, we can apply the same logic to the lowering operator. Thus, there must be a lowest-weight state:

$$\boldsymbol{L}_- |\ell, m_{\text{low}}\rangle = 0 \implies \ell(\ell+1) = m_{\text{low}}(m_{\text{low}} - 1), \tag{255}$$

where the implication follows from Equations 248 and 251 as before. The two possible solutions are then $m_{\text{low}} = -\ell$ and $m_{\text{low}} = \ell + 1$, but clearly $m_{\text{low}} \leqslant m_{\text{high}}$, so we must have $m_{\text{low}} = -\ell$. Thus, $\ell - m_{\text{low}} = 2\ell$ is an integer, so the list of all *possible* highest weights are

$$\ell = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots. \tag{256}$$

For each value of ℓ, we obtain an orthonormal set

$$|\ell, -\ell\rangle, \quad |\ell, -\ell+1\rangle, \quad \ldots, \quad |\ell, \ell\rangle \tag{257}$$

of size $2\ell + 1$, where

$$\boldsymbol{L}_z |\ell, m\rangle = \hbar m |\ell, m\rangle, \quad \boldsymbol{L}_\pm |\ell, m\rangle = \sqrt{\ell(\ell+1) - m(m \pm 1)}\, |\ell, m \pm 1\rangle. \tag{258}$$

Since the subspace spanned by Equation 257 is mapped into itself by $\boldsymbol{L}_{x,y,z}$, this defines a representation fo the spin algebra, known as the **spin-ℓ** representation.

Question 45. *Let $\boldsymbol{O}$ be an arbitrary **spin-ℓ** representation that commutes with $\boldsymbol{L}_z$, $\boldsymbol{L}_\pm$.*

1. *Show that $\boldsymbol{O} \propto \boldsymbol{I}$.*

   Proof. □

2. *Argue that any reducible representation has non-trivial operators $\boldsymbol{O} \not\propto \boldsymbol{I}$ that commute with all the charges / symmetry operators. (HINT: consider subspace projectors.) Thus, spin-ℓ rep is irreducible.*

   Proof. □

Question 46. *"Ordinary" vectors. [HW4, prob 2.]*

Question 47. *Let $|\psi\rangle \in \mathcal{H}_\ell$ be a non-zero ket in the spin-ℓ representation $\mathcal{H}_\ell$. Consider the overlap of $|\psi\rangle$ with the state $|\psi'\rangle$ obtained by rotating about a fixed axis $\hat{\boldsymbol{n}}$ by an angle θ:*

$$\langle\psi|\,\psi'\rangle = \left\langle \psi \left| e^{-\frac{i\theta}{\hbar}\hat{\boldsymbol{n}}\cdot\underline{\boldsymbol{L}}} \right| \psi \right\rangle. \tag{259}$$

1. *What is the minimum rotation angle $|\theta| \geqslant \theta_{\text{min}}$ for which this overlap can vanish? For which states $|\psi\rangle$ foes it vanish when $|\theta| = \theta_{\text{min}}$. (HINT: Write* Re $\langle\psi|\,\psi'\rangle$ *in a convenient basis where it is manifestly positive – i.e., a sum of positive terms – for small $|\theta|$. Find the smallest value of $|\theta|$ for which this ceases to be true.)*

   Proof. □

2. *Is it possible to find an orthonormal basis of $\mathcal{H}_\ell$ of the form*:

$$|\psi\rangle\,,\ \mathrm{e}^{-\frac{i\theta_1}{\hbar}\hat{\boldsymbol{n}}\cdot\underline{\boldsymbol{L}}}\,|\psi\rangle\,,\ \ldots,\ \mathrm{e}^{-\frac{i\theta_{2\ell}}{\hbar}\hat{\boldsymbol{n}}\cdot\underline{\boldsymbol{L}}}\,|\psi\rangle\,,\quad 0<\theta_1<\cdots<\theta_{2\ell}<2\pi? \tag{260}$$

What are the necessary and sufficient conditions on $|\psi\rangle$ and $\theta_1, \ldots, \theta_{2\ell}$ for this to occur? Find an example or prove none exists.[62]

Proof. □

[62] The "discrete Fourier series" identity,

$$\frac{1}{n}\sum_{k=0}^{n-1}\mathrm{e}^{\frac{2\pi i p k}{n}}=\begin{cases}0 & p \text{ not a multiple of } n\\ 1 & p \text{ a multiple of } n\end{cases} \tag{261}$$

may be useful.

2025-03-04: Quantum Symmetries (cont.)

Half-integer Spins and Projective Representations

We now discuss a peculiar feature of some of the spin representations we have constructed during the previous class. Performing a 2π rotation about the $\hat{z}$ axis, we obtain:

$$|\ell, m\rangle \to \mathrm{e}^{-\frac{2\pi i}{\hbar} L_z} |\ell, m\rangle = \mathrm{e}^{-2\pi i m} |\ell, m\rangle . \tag{262}$$

In the case of **integer spin**, $\ell = 0, 1, 2, \ldots$, then m is also an integer, and so the ket is unchanged, just as we would naively expect. However, for **half-integer spin**, $\ell = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \ldots$, then m is also a half-integer, so

$$|\ell, m\rangle \to \mathrm{e}^{-\frac{2\pi i}{\hbar} L_z} |\ell, m\rangle = \mathrm{e}^{-2\pi i m} |\ell, m\rangle = - |\ell, m\rangle . \tag{263}$$

Thus, every ket in the representation picks up a minus sign under a 2π rotation. The same peculiarity occurs for a 2π rotation about any axis.

Backtracking, we see that this minus sign entered between Equation 217 and Equation 224, where we redefined the overall phase of the symmetry operators $\boldsymbol{U}_\theta$ for convenience. One might think that the solution is to undo this change, e.g., by redefining the symmetry operators as

$$\boldsymbol{U}'_{\underline{\theta}} = \mathrm{e}^{if(\underline{\theta})} \boldsymbol{U}_{\underline{\theta}} = \mathrm{e}^{if(\underline{\theta})} \exp\left[-\frac{i}{\hbar} \underline{\theta} \cdot \underline{\boldsymbol{L}}\right], \tag{264}$$

where $f(\underline{\theta})$ is some function that equals 0 when $|\underline{\theta}| = 0$ and π when $|\underline{\theta}| = 2\pi$. However, one can show that there is no such function that satisfies the group law $\boldsymbol{U}_g \boldsymbol{U}_h = \boldsymbol{U}_{gh}$ for arbitrary rotations $g, h \in \mathrm{SO}(3)$.

This implies that the half-integer spins are *not* representations of the rotation group $\mathrm{SO}(3)$. For instance, in the $\ell = \frac{1}{2}$ case, the symmetry operators are:

$$\exp\left[-\frac{i}{\hbar} \underline{\theta} \cdot \underline{S}\right] \to \exp\left[-\frac{i}{2} \underline{\theta} \cdot \underline{\sigma}\right] = \exp\left[-\frac{i}{2} \begin{pmatrix} \theta_z & \theta_x - i\theta_y \\ \theta_x + i\theta_y & -\theta_z \end{pmatrix}\right] \tag{265}$$

in matrix form. Since the matrix in the exponent is a *traceless* Hermitian matrix, the result is a 2×2 unitary matrix with determinant 1. The group of such matrices is called $\mathrm{SU}(2)$. Thus, spin-$\frac{1}{2}$ is a representation of $\mathrm{SU}(2)$. In fact, one can show that *all* the half-integer spins are representations of this group.

What have already seen that $\mathrm{SU}(2)$ and $\mathrm{SO}(3)$ have the same Lie algebra. To understand how the two groups are related, it is useful to think of them as manifolds. First, note that an arbitrary 2×2 special unitary matrix can be written as

$$\begin{pmatrix} \alpha & \beta \\ -\beta^* & \alpha^* \end{pmatrix}, \tag{266}$$

for any $\alpha, \beta \in \mathbb{C}$ s.t. $|\alpha|^2 + |\beta|^2 = 1$. The condition $|\alpha|^2 + |\beta|^2 = 1$ can be written as

$$(\mathrm{Re}\,\alpha)^2 + (\mathrm{Im}\,\alpha)^2 + (\mathrm{Re}\,\beta)^2 + (\mathrm{Im}\,\beta)^2 = 1, \tag{267}$$

but this is just the equation that defines the unit 3-sphere S^3 embedded in $\mathbb{R}^4$. Thus, the group SU (2) is a 3-sphere S^3![63]

What about SO (3)? To picture the manifold of 3d rotations, we reason as follows

1. We assign a rotation about the axis $\hat{\boldsymbol{n}}$ by an angle $\theta \geqslant 0$ (in the right-handed sense) to the point $\underline{\boldsymbol{\theta}} = \theta\hat{\boldsymbol{n}}$ in $\mathbb{R}^3$.

2. Since $\theta > 2\pi$ redundantly describes a rotation with $\theta < 2\pi$, we restrict to the ball of radius $|\theta| = 2\pi$ with $\mathbb{R}^3$. Moreover, a rotation by 2π about any axis $\hat{\boldsymbol{n}}$ is the same, so we identify the surface of the ball with a single point.

   Not coincidentally, the space we have described so far, notated mathematically as B^3/S^2, is topologically a 3-sphere, S^3. This can be seen in Figure 25. To understand *why* this space is a 3-sphere, one can compare with the simpler case of $B^2/S^1 \cong S^2$, seen in Figure 26.

3. Because a rotation by about $\hat{\boldsymbol{n}}$ is the same as a rotation by $2\pi - \theta$ about $-\hat{\boldsymbol{n}}$ (in the right-handed sense in both cases), there is still a two-fold redundancy:

$$(\theta, \hat{\boldsymbol{n}}) \cong (2\pi - \theta, -\hat{\boldsymbol{n}}) \,. \tag{268}$$

   Note that this includes, as a special case, the fact that the surface S^2 of the ball of radius 2π describes a trivial rotation, just like the point at the origin.

Thus, SU (2) is a **double cover** of SO (3), i.e., for every rotation matrix in SO (3) there are *two* distinct elements of SU (2). Note that SU (2) $\cong S^3$ is "simply connected," i.e., any closed loop within the can be contracted to a point. By contrast, $\mathbb{RP}^3 \cong S^3/\mathbb{Z}_2$ is not simply connected, because the loop shown in Figure 27 is non-contractable.

Thus, SU (2) is the "universal cover" of SO (3). Precisely because it distinguishes rotations that differ by 2π, SU (2) has both integer and half-integer spin representations, where SO (3) has only integer spin representations. For this reason, SU (2) is also known as the **spin group** Spin (3). Similarly, SO (n) has a double cover Spin (n), which is the universal cover for $n > 2$.

This all sounds rather strange until we remember that the overall phase of the quantum state vector is unphysical. Thus, we can relax the group law $\boldsymbol{U}_g\boldsymbol{U}_h = \boldsymbol{U}_{gh}$ to

$$\boldsymbol{U}_g\boldsymbol{U}_h = \mathrm{e}^{i\phi_{g,h}}\,\boldsymbol{U}_{gh}, \tag{269}$$

for some set of arbitrary phases $\phi_{g,h}$. This defines a **projective unitary representation**. Quantum symmetries are represented on the physical Hilbert space $\mathcal{H}$ projectively. Moreover, SO (3) *does* have projective representations of both integer and half-integer spins.

The reason we have not discussed projective representations previously is because they do not always straighforwardly divide into irreps in the same way the non-projective (unitary) representations do. For example, the effective Hilbert space of nitrogen atom positions in the ammonia molecule divided into two 1d irreps of

[63] Recall that the group $U(1)$ is also a sphere, S^1. These are the only two Lie groups that are spheres, however.

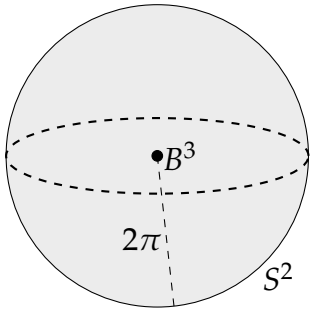

Figure 25: The 3-sphere S^3 described topologically as a 3-ball with its outer surface (a 2-sphere) identified as a single point.

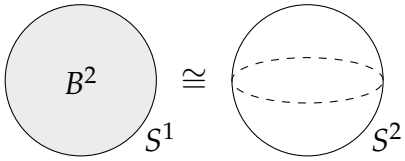

Figure 26: The 2-sphere S^2 can be described topologically as a disk (2-ball) with its outer surface (a circle, or 1-sphere S^1) identified as a single point.

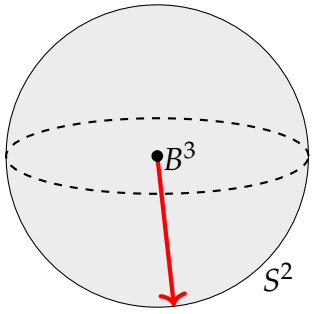

Figure 27: The red line is a closed loop in $\mathbb{RP}^3$ (pictured as a two-fold indentification of $B^3/S^2 \cong S^3$) since the origin is identified with the outer surface of the ball. This loop cannot be contracted to a point within $\mathbb{RP}^3 \cong$ SO (3).

the $\mathbb{Z}_2$ reflection symmetry, $\boldsymbol{U}|\pm\rangle = \pm|\pm\rangle$. However, both of these irreps are projectively trivial, since $\boldsymbol{U}$ acts as an overall phase within each! On the full Hilbert space, by contrast, $\boldsymbol{U}$ is *not* projectively trivial, since it gives different phases to different basis kets.

Thus, the best way to understand projective representations is to *first* think about non-projective representations, which can safely be broken into irreps, and *then* account for the fact that the overall phase of the state vector is unphysical (which is a statement about the *entire* Hilbert space, not any single irrep within it). On price we must pay for this approach is replacing the group of rotations $\mathrm{SO}(3)$ with the spin group $\mathrm{Spin}(3) \cong \mathrm{SU}(2)$ when we work non-projectively.[64,65]

Finally, note that classical symmetries do *not* act projectively. Thus, half-integer spin representations are in some sense *intrinsically quantum*. These representations do no occur in classical physics.[66]

[64] More generally and technically, the projective representations of any group G are representations of its universal cover $\tilde{G} = G/K$. Conversely, any *irreducible* representation of $\tilde{G}$ is a representation of G, or more generally any representation of $\tilde{G}$ on which the elements of K act by overall phases is a projective representation of G.

[65] However, in QFT with fermions, the *actual* symmetry group is the spin group, even accounting for the projective nature of quantum symmetries. This is because the Hilbert space includes states with both even and off numbers of fermions, and the latter pick up a sign under 2π rotations whereas the former do not. Thus, a 2π rotation defines a non-trivial operator, typically denoted as $(-1)^F$, acting differently on different states.

[66] To be precise, spin representations occur in so-called "classical" field theories with fermions, but these field theories – which involve anti-commuting Grassman variables – are really formal objects that are studies to better understand their quantum versions. They are not truly classical in the original sense.

More General Notions of Symmetry

Before moving on, let us briefly mention two generalizations that we will not consider in detail:

1. Symmetries can have explicit time-dependence, i.e., $\boldsymbol{U} = \boldsymbol{U}(t)$ can depend on time in the Schrödinger picture. Then Equation 196 becomes

$$i\hbar\frac{\partial}{\partial t}|\psi;t\rangle = \boldsymbol{H}(t)|\psi;t\rangle \implies i\hbar\frac{\partial}{\partial t}\left(\boldsymbol{U}(t)|\psi;t\rangle\right) = \boldsymbol{H}(t)\left(\boldsymbol{U}(t)|\psi;t\rangle\right). \tag{270}$$

   Following the same logic as before, we arrive at the condition

$$i\hbar\frac{\partial \boldsymbol{U}}{\partial t} + [\boldsymbol{U}(t), \boldsymbol{H}(t)] = 0. \tag{271}$$

   Translating to the Heisenberg picture, this is simply

$$\frac{d\boldsymbol{U}_H}{dt} = \frac{i}{\hbar}[\boldsymbol{H}_H(t), \boldsymbol{U}_H(t)] + \frac{\partial \boldsymbol{U}_H}{\partial t} = 0. \tag{272}$$

   Thus, in general a symmetry operator is a unitary operator that is time-independent in the Heisenberg picture.

   If the symmetry is continuous, then

$$\boldsymbol{U}_H^{(\theta)} = \exp\left[-\frac{i}{\hbar}\theta\boldsymbol{Q}_H\right] \implies \frac{d\boldsymbol{Q}_H}{dt} = 0. \tag{273}$$

   This is the quantum Noether's theorem once again, this time in full generality!

   Note that in principle *any* unitary transformation specified at a fixed time $t = t_0$ can be extended to a time-dependent symmetry by solving Equation 271. Most such "symmetries" will not be very useful for understanding quantum dynamics, in part because solving Equation 271 is typically just as hard as solving for the dynamics directly.[67] However, in some cases we can identify a naturally-defined time-dependent symmetry without full solving the dynamics. An example of such a symmetry is a *boost* symmetry in QFT.

[67] Even in the time-independent case, typically only a subset of the symmetries can be understood *before* solving the dynamics. These "identifiable" symmetries are much more important and useful than the other, and we typically refer to them as "the" symmetries of the quantum system without further qualification.

2. Systems can be *antilinear*. We saw that antilinear transformations obey all the rules of quantum mechanics *except* when we try to interfere the results of linear and antilinear transformations with each other. However, symmetries are applied to the *entire* Hilbert space at once, so no such interference can occur.

   To be precise, an **antiunitary** operator is an invertible antilinear operator $\boldsymbol{T}$ s.t.

$$(\boldsymbol{T}|\psi\rangle, \boldsymbol{T}|\chi\rangle) = (|\psi\rangle, |\chi\rangle)^* = (|\chi\rangle, |\psi\rangle). \tag{274}$$

   In other words, $\boldsymbol{T}$ preserves the inner product *up to complex conjugation*. This is sufficient to preserve all probabilities (since they are real numbers).

   An antilinear symmetry is an antiunitary operator $\boldsymbol{T}$ that preserve solutions to Schrödinger's equation *up to time reversal*, i.e.,[68]

$$i\hbar\frac{\partial}{\partial t}|\psi;t\rangle = \boldsymbol{H}(t)|\psi;t\rangle \implies i\hbar\frac{\partial}{\partial t}(\boldsymbol{T}|\psi;-t\rangle) = \boldsymbol{H}(t)(\boldsymbol{T}|\psi;-t\rangle). \tag{275}$$

   This implies that

$$[\boldsymbol{T}, \boldsymbol{H}(t)] = 0, \tag{276}$$

   in the time-independent case, just as before.

   To fully understand such **time-reversal** symmetries, one has to think carefully about antilinear operators. See, e.g., Sakurai section 4.4 for more information about them.

[68] If we tried to make $\boldsymbol{T}$ preserves solutions to Schrödinger's equation *without* reversing time, we would obtain $\boldsymbol{T}\boldsymbol{H}(t) = -\boldsymbol{H}(t)\boldsymbol{T}$. This would imply that for every state of energy E_0 there is a corresponding state of energy $-E_0$. Since the energy spectrum is typically bounded from below but not from above, this cannot be satisfied in most situations.

Quantum Information and Entanglement

Quantum Versus Classical States

Now that we understand what a quantum system is, let us examine more carefully how it differs from a classical system. In particular, quantum states differ from classical states for two related reasons:

1. Quantum states are "fuzzy," since observables that do not commute (incompatible observables) cannot be sharply specified at the same time.

We have already observed this, but this observation can be made more precise as follows:

Theorem 48 (Generalized uncertainty principle). *For any two self-adjoint operators $\boldsymbol{A}$ and $\boldsymbol{B}$ and any quantum state $|\psi\rangle$:*

$$\left\langle(\Delta\boldsymbol{A})^2\right\rangle\left\langle(\Delta\boldsymbol{B})^2\right\rangle \geqslant \frac{1}{4}|\langle[\boldsymbol{A},\boldsymbol{B}]\rangle|^2, \tag{277}$$

where the expectation values are all evaluated in the state $|\psi\rangle$.

Proof. We start with an obvious statement:

$$\forall\ |\alpha\rangle, |\beta\rangle \in \mathcal{H}, \lambda \in \mathbb{C},\ \||\alpha\rangle + \lambda|\beta\rangle\|^2 \geqslant 0. \tag{278}$$

Choosing $\lambda = -\frac{\langle\beta|\alpha\rangle}{\langle\beta|\beta\rangle}$ and writing this out, we obtain:

$$\langle a|\, a\rangle - \frac{\langle\beta|\, \alpha\rangle}{\langle\beta|\, \beta\rangle}\, \langle\alpha|\, \beta\rangle - \frac{\langle\alpha|\, \beta\rangle}{\langle\beta|\, \beta\rangle}\, \langle\beta|\, \alpha\rangle + \left|\frac{\langle\beta|\, \alpha\rangle}{\langle\beta|\, \beta\rangle}\right|^2 \langle\beta|\, \beta\rangle \geqslant 0. \tag{279}$$

Multiplying by $\langle\beta|\, \beta\rangle$, this simplifies to

$$\langle\alpha|\, \alpha\rangle\, \langle\beta|\, \beta\rangle \geqslant |\langle\alpha|\, \beta\rangle|^2, \quad (\text{the } \textbf{Cauchy-Schwarz inequality}). \tag{280}$$

This inequality is **saturated** (i.e., the two sides are equal) iff $|\alpha\rangle$ and $|\beta\rangle$ are "parallel," i.e., scalar multiples of each other.

Substituting $|\alpha\rangle = \boldsymbol{\Delta A}\,|\psi\rangle$ and $|\beta\rangle = \boldsymbol{\Delta B}\,|\psi\rangle$ into Equation 280, we obtain:

$$\left\langle (\boldsymbol{\Delta A})^2 \right\rangle \left\langle (\boldsymbol{\Delta B})^2 \right\rangle \geqslant |\langle \boldsymbol{\Delta A \Delta B}\rangle|^2, \tag{281}$$

where we use the fact that $\boldsymbol{\Delta A}$ and $\boldsymbol{\Delta B}$ are self-adjoint. We decompose the right-hand side as follows:

$$\boldsymbol{\Delta A \Delta B} = \frac{1}{2}\,[\boldsymbol{\Delta A}, \boldsymbol{\Delta B}] + \frac{1}{2}\,\{\boldsymbol{\Delta A}, \boldsymbol{\Delta B}\} = \frac{1}{2}\,[\boldsymbol{A}, \boldsymbol{B}] + \frac{1}{2}\,\{\boldsymbol{\Delta A}, \boldsymbol{\Delta B}\}, \tag{282}$$

where $\{\boldsymbol{\Delta A}, \boldsymbol{\Delta B}\} := \boldsymbol{XY} + \boldsymbol{YX}$ is the **anticommutator**, and we can replace $\boldsymbol{\Delta A} \to \boldsymbol{A}$ and $\boldsymbol{\Delta B} \to \boldsymbol{B}$ inside the commutator because the identity operator $\boldsymbol{I}$ commutes with everything. Note for for self-adjoint $\boldsymbol{X}$, $\boldsymbol{Y}$,

$$[\boldsymbol{X}, \boldsymbol{Y}]^\dagger = (\boldsymbol{XY} - \boldsymbol{YX})^\dagger = \boldsymbol{YX} - \boldsymbol{XY} = -\,[\boldsymbol{X}, \boldsymbol{Y}], \tag{283}$$

since the adjoint reverses the order of products. Thus, the commutator of two self-adjoint operators is *anti* self-adjoint. Likewise,

$$\{\boldsymbol{X}, \boldsymbol{Y}\}^\dagger = (\boldsymbol{XY} + \boldsymbol{YX})^\dagger = \boldsymbol{YX} + \boldsymbol{XY} = \{\boldsymbol{X}, \boldsymbol{Y}\}, \tag{284}$$

so the anticommutator of two self-adjoint operators is self-adjoint. Thus, since the expectation value of a self-adjoint (anti self-adjoint) operator is real (imaginary),

$$\langle \boldsymbol{\Delta A \Delta B}\rangle = \frac{1}{2}\underbrace{\langle [\boldsymbol{A}, \boldsymbol{B}]\rangle}_{\text{imaginary}} + \frac{1}{2}\underbrace{\langle \{\boldsymbol{A}, \boldsymbol{B}\}\rangle}_{\text{real}}. \tag{285}$$

Thus, taking the modulus squared:

$$|\langle \boldsymbol{\Delta A \Delta B}\rangle|^2 = \frac{1}{4}\,|\langle [\boldsymbol{A}, \boldsymbol{B}]\rangle|^2 + \frac{1}{4}\,|\langle \{\boldsymbol{A}, \boldsymbol{B}\}\rangle|^2 \geqslant \frac{1}{4}\,|\langle [\boldsymbol{A}, \boldsymbol{B}]\rangle|^2. \tag{286}$$

Substituting this back into Equation 281, we obtain the generalized uncertainty principle, Equation 277. □

Here for any operator $\boldsymbol{X}$, the state-dependent operator $\boldsymbol{\Delta X}$ is defined as $\boldsymbol{X}$ minus its expected value:

$$\boldsymbol{\Delta X} := \boldsymbol{X} - \langle \boldsymbol{X}\rangle \boldsymbol{I} = \boldsymbol{X} - \frac{\langle\psi|\, \boldsymbol{X}\, |\psi\rangle}{\langle\psi|\, \psi\rangle}\boldsymbol{I}. \tag{287}$$

This implies that $\langle \Delta X \rangle = \langle \boldsymbol{X} \rangle - \langle \boldsymbol{X} \rangle \langle \boldsymbol{I} \rangle = \langle X \rangle - \langle X \rangle = 0$ since $\langle \boldsymbol{I} \rangle = \frac{\langle \psi | \boldsymbol{I} | \psi \rangle}{\langle \psi | \psi \rangle} = \frac{\langle \psi | \psi \rangle}{\langle \psi | \psi \rangle} = 1$. Note that $\boldsymbol{\Delta X}$ is self-adjoint if $\boldsymbol{X}$ is. Moreover,

$$\left\langle (\boldsymbol{\Delta X})^2 \right\rangle = \left\langle \boldsymbol{X}^2 - 2\langle \boldsymbol{X} \rangle \boldsymbol{X} + \langle \boldsymbol{X} \rangle^2 \boldsymbol{I} \right\rangle = \left\langle \boldsymbol{X}^2 \right\rangle - 2\langle \boldsymbol{X} \rangle^2 + \langle \boldsymbol{X} \rangle^2 \langle I \rangle = \left\langle \boldsymbol{X}^2 \right\rangle - \langle \boldsymbol{X} \rangle^2, \tag{288}$$

which is the statistical **variance** of X measurements on the state $|\psi\rangle$.

Thus, the generalized uncertainty principle states that the product of the variances of two incompatible observables $\boldsymbol{A}$ and $\boldsymbol{B}$ satisfies a sharp, non-zero lower bound in states where $[\boldsymbol{A}, \boldsymbol{B}]$ has a non-zero expectation value. Equivalently,

$$\sqrt{\left\langle (\boldsymbol{\Delta A})^2 \right\rangle} \sqrt{\left\langle (\boldsymbol{\Delta B})^2 \right\rangle} \geqslant \frac{1}{2} \left| \langle [\boldsymbol{A}, \boldsymbol{B}] \rangle \right|, \tag{289}$$

where $\sqrt{\left\langle (\boldsymbol{\Delta A})^2 \right\rangle}$ is the **standard deviation** of X measurements on the state $|\psi\rangle$. Therefore, when $|\langle [\boldsymbol{A}, \boldsymbol{B}] \rangle| \neq 0$, measurements of A and B cannot both be sharp at once. As we see explicitly later, Equation 289 is a generalized and sharpened version of the Heisenberg uncertainty principle.

In this sense quantum states have *less* information in them than classical states, because the observables cannot all be specified simultaneously. However,

2. Quantum states can be *superpositions* of different classical configurations.

For instance, the states $|1\rangle$ and $|2\rangle$ of the ammonia atom – corresponding to the up and down positions of the nitrogen atom – are both classically and quantum mechanically sensible, but the ammonia ground state $|+\rangle = \frac{1}{\sqrt{2}} (|1\rangle + |2\rangle)$ is deeply quantum, with *no classical analog*.

In general, distinct classical states correspond to the elements of some orthonormal basis, whereas a general quantum state vector is an arbitrary linear combination of these. In this way, quantum states have *more* information in them than classical states.

There is an interesting interplay between quantum fuzziness and quantum superposition in the qubit Hilbert space. We started with a classical model system (a spin of fixed magnitude pointing in an arbitrary direction) with a *continuum* of states. However, as shown by the Stern-Gerlach experiment, there are actually only *two* classical distinct states, spin up $|+\hat{\boldsymbol{z}}\rangle$ and spin down $|-\hat{\boldsymbol{z}}\rangle$, for any given measurement. This was initially a puzzle, because it seemed to break rotational invariance. However, by taking quantum superpositions of these two "classical" states, we can describe the up and down spin components in *any* direction, e.g.,

$$|+\hat{\boldsymbol{x}}\rangle = \frac{1}{\sqrt{2}} (|+\hat{\boldsymbol{z}}\rangle + |-\hat{\boldsymbol{z}}\rangle). \tag{290}$$

The end result is that the projective space of quantum states is a sphere $\left(\mathbb{CP}^1\right)$, which, coincidentally, looks the same as the naive classical state space we started with.[69]

[69] It is perhaps more natural to think of a qubit as the quantum analog of a bit, in which case the quantum state vector again contains much more information than the classical state.

Of course, the quantum state of a single qubit can never be fully measured, since any attempt to measure it erases the input state while only providing one bit of (probabilistic) information about what it was. However, with access to *many* identical copies of the same states, we could determine which state it is with good accuracy by performing many different measurements.

In more complicated quantum systems, that space of quantum states is typically *much larger* than the classical state space, so there is a lot of **quantum information** in the quantum state vector, but again we can access only a very small fraction of this information[70] unless we are given many identical copies of the same state to play with.

[70] Which information we can access depends on which measurement(s) we choose to perform.

2025-03-06: Quantum Information and Entanglement (cont.)

Mixed States

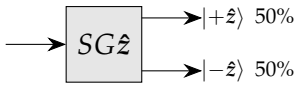

Figure 28: The outcome of the original Stern-Gerlach experiment.

To further illustrate the difference between classical and quantum notions, let us compare *classical uncertainty* with *quantum superposition*.

In the original Stern-Gerlach experiment, a beam of silver atoms split equally into "up" $|+\hat{z}\rangle$ and "down" $|-\hat{z}\rangle$ components. In our previous discussion, we never worked out what state the incoming beam was in. Evidently it must contain "an equal mixture of up and down," but there are (at least) two distinct things that this could mean:

1. Classical uncertainty.

   50% of the atoms in the incoming beam are in the $|+\hat{z}\rangle$ state, and 50% of the atoms in the incoming beam are in the $|-\hat{z}\rangle$ state.

   In this scenario, each atom in the beam has a *definite* $S_z = \pm\frac{\hbar}{2}$ value, but the spin of an atom picked randomly from the beam is random.

2. Quantum superposition.

   All the atoms in the incoming beam are in the same quantum state, which is an equal superposition of $|+\hat{z}\rangle$ and $|-\hat{z}\rangle$:

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|+\hat{z}\rangle + |-\hat{z}\rangle\right). \tag{291}$$

   Now the randomness of the outcome is not because we pick an atom to measure at random, but because measurements on superposition states produce uncertain outcomes.

Since both scenarios produce the same outcome in Figure 28, we cannot yet tell a difference.

However, we are free to perform a different measurement on some of the atoms in the beam, e.g., we can measure the $\hat{x}$ component of their spin, S_x. Let us analyze the outcome of this measurement in each scenario:

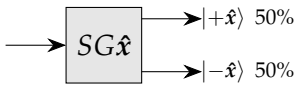

Figure 29: The result of an S_x measurement in the case of classical uncertainty.

1. Classical uncertainty.

   50% of the time, the atoms is in the state $|\psi\rangle = |+\hat{z}\rangle$. The probability of each S_x measurement outcome is then:

$$p^+_{\text{up}} = |\langle +\hat{x}|\,\psi\rangle|^2 = |\langle +\hat{x}| + \hat{z}\rangle|^2 = \frac{1}{2}, \quad p^+_{\text{down}} = |\langle -\hat{x}|\,\psi\rangle|^2 = |\langle -\hat{x}| + \hat{z}\rangle|^2 = \frac{1}{2}. \tag{292}$$

   The other 50% of the time, the atom is the state $|\psi\rangle = |-\hat{z}\rangle$, in which case the probability of each measurement outcome is

$$p^+_{\text{up}} = |\langle +\hat{x}|\,\psi\rangle|^2 = |\langle +\hat{x}| - \hat{z}\rangle|^2 = \frac{1}{2}, \quad p^+_{\text{down}} = |\langle -\hat{x}|\,\psi\rangle|^2 = |\langle -\hat{x}| - \hat{z}\rangle|^2 = \frac{1}{2}. \tag{293}$$

Thus, in net, the probability of each measurement outcome is

$$p_{\text{up}} = \frac{1}{2}p^{+}_{\text{up}} + \frac{1}{2}p^{-}_{\text{up}} = \frac{1}{2}, \quad p_{\text{down}} = \frac{1}{2}p^{+}_{\text{down}} + \frac{1}{2}p^{-}_{\text{down}} = \frac{1}{2}. \tag{294}$$

Therefore, there is an equal probability of each measurement outcome, just like the S_z measurement.

2. Quantum superposition.

   Now all the atoms are in the state $|\psi\rangle = \frac{1}{\sqrt{2}}\left(|+\hat{\boldsymbol{z}}\rangle + |-\hat{\boldsymbol{z}}\rangle\right)$, so

$$p_{\text{up}} = |\langle +\hat{\boldsymbol{x}}|\,\psi\rangle|^2 = 1, \quad p_{\text{down}} = |\langle -\hat{\boldsymbol{x}}|\,\psi\rangle|^2 = 0. \tag{295}$$

   This is a dramatically different outcome: all the atoms come out in the up beam.

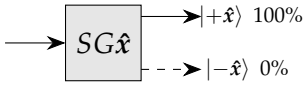

Figure 30: The result of an S_x measurement in the case of quantum superposition.

In the real Stern-Gerlach experiment, the outcome is the same regardless of how the apparatus is oriented. Thus, the first scenario (or something like it) correctly describes this experiment. However, the second scenario is precisely what we engineered in the fourth Stern-Gerlach-inspired thought experiment. Thus, classical uncertainty and quantum superposition are distinct phenomena, both of which arise in various situations. Later, we will see that they are related through the process of "entanglement."

More generally, suppose that each atom in the beam is in one of the normalized states $|\psi_1\rangle, \ldots, |\psi_k\rangle$, with probabilities $p_1, \ldots, p_k$, respectively, where $p_1 + \cdots + p_k = 1$ since each atom is in exactly one state. In this case, the probability of measuring $S_z = +\frac{\hbar}{2}$ is:

$$\begin{aligned} p_{\text{up}} &= \sum_{i=1}^{k} p_i\,|\langle +\hat{\boldsymbol{z}}|\,\psi_i\rangle|^2 = \sum_{i=1}^{k} p_i\,\langle +\hat{\boldsymbol{z}}|\,\psi_i\rangle\,\langle\psi_i|\,+\hat{\boldsymbol{z}}\rangle \\ &= \langle +\hat{\boldsymbol{z}}|\left(\sum_{i=1}^{k} p_i\,|\psi_i\rangle\langle\psi_i|\right)|+\hat{\boldsymbol{z}}\rangle = \langle +\hat{\boldsymbol{z}}|\,\boldsymbol{\rho}\,|+\hat{\boldsymbol{z}}\rangle\,, \end{aligned} \tag{296}$$

where we define the **density operator**

$$\boldsymbol{\rho} := \sum_{i=1}^{k} p_i\,|\psi_i\rangle\langle\psi_i|\,. \tag{297}$$

While we motivated it using a qubit, this definition can be applied to *any* quantum system whose state vector is randomly drawn from some statistical ensemble. One can show that:

Theorem 50 (Density operator axioms)**.** *An operator $\boldsymbol{\rho}$ can be written in the form Equation 297 for some ensemble of normalized state-vectors $|\psi_1\rangle, \ldots, |\psi_k\rangle$ with probabilities $p_1, \ldots, p_k$ iff:*

1. $\boldsymbol{\rho} = \boldsymbol{\rho}^\dagger$,
2. $\boldsymbol{\rho}$ *is non-negative* ($\boldsymbol{\rho} \geqslant \mathbf{0}$),[71]

Question 49. *Consider a spin measurement $S_{\hat{\boldsymbol{n}}}$ along the axis $\hat{\boldsymbol{n}} = \sin\theta\cos\phi\hat{\boldsymbol{x}} + \sin\theta\sin\phi\hat{\boldsymbol{y}} + \cos\theta\hat{\boldsymbol{z}}$.*

1. *If a fraction n_{up} of the atoms are in the state $|+\hat{\boldsymbol{z}}\rangle$, and a fraction $n_{\text{down}} = 1 - n_{\text{up}}$ are in the state $|-\hat{\boldsymbol{z}}\rangle$, what are the probabilities of measuring $S_{\hat{\boldsymbol{n}}} = +\frac{\hbar}{2}$ or $S_{\hat{\boldsymbol{n}}} = -\frac{\hbar}{2}$.*

   Proof. □

2. *For what value of n_{up} are the two outcomes equally likely?*

   Proof. □

[71] $\rho_{ij} \geqslant 0\ \forall\ i, j \in \dim\boldsymbol{\rho}$.

3. $\mathrm{tr}(\boldsymbol{\rho}) = 1$.

Question 51. *Prove **Theorem 50**.*

For example, in the "classical uncertainty" scenario from above, the density operator is:

$$\boldsymbol{\rho}_I = \frac{1}{2}\left(|+\hat{\boldsymbol{z}}\rangle\langle+\hat{\boldsymbol{z}}| + |-\hat{\boldsymbol{z}}\rangle\langle-\hat{\boldsymbol{z}}|\right) = \frac{1}{2}\boldsymbol{I}, \tag{298}$$

whereas in the "quantum superposition" scenario from above, it is:

$$\begin{aligned}\boldsymbol{\rho}_{II} = |\psi\rangle\langle\psi| &= \frac{1}{2}\left(|+\hat{\boldsymbol{z}}\rangle + |-\hat{\boldsymbol{z}}\rangle\right)\left(\langle+\hat{\boldsymbol{z}}| + \langle-\hat{\boldsymbol{z}}|\right) \\ &= \frac{1}{2}\left(|+\hat{\boldsymbol{z}}\rangle\langle+\hat{\boldsymbol{z}}| + |+\hat{\boldsymbol{z}}\rangle\langle-\hat{\boldsymbol{z}}| + |-\hat{\boldsymbol{z}}\rangle\langle+\hat{\boldsymbol{z}}| + |-\hat{\boldsymbol{z}}\rangle\langle-\hat{\boldsymbol{z}}|\right).\end{aligned} \tag{299}$$

Clearly, $\boldsymbol{\rho}_I \neq \boldsymbol{\rho}_{II}$. The probabilities for measuring $S_x = +\frac{\hbar}{2}$ and $S_x = -\frac{\hbar}{2}$ in each case are:

$$\begin{aligned}p^{\hat{\boldsymbol{x}}}_{\text{up}} &= \langle+\hat{\boldsymbol{x}}|\,\boldsymbol{\rho}_I\,|+\hat{\boldsymbol{x}}\rangle = \frac{1}{2}\left(\frac{1}{2} + \frac{1}{2}\right) = \frac{1}{2}, \\ p^{\hat{\boldsymbol{x}}}_{\text{down}} &= \langle-\hat{\boldsymbol{x}}|\,\boldsymbol{\rho}_I\,|-\hat{\boldsymbol{x}}\rangle = \frac{1}{2}\left(\frac{1}{2} + \frac{1}{2}\right) = \frac{1}{2},\end{aligned} \tag{300}$$

in the first scenario, and

$$\begin{aligned}p^{\hat{\boldsymbol{x}}}_{\text{up}} &= \langle+\hat{\boldsymbol{x}}|\,\boldsymbol{\rho}_{II}\,|+\hat{\boldsymbol{x}}\rangle = \frac{1}{2}\left(\frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \frac{1}{2}\right) = 1, \\ p^{\hat{\boldsymbol{x}}}_{\text{down}} &= \langle-\hat{\boldsymbol{x}}|\,\boldsymbol{\rho}_{II}\,|-\hat{\boldsymbol{x}}\rangle = \frac{1}{2}\left(\frac{1}{2} - \frac{1}{2} - \frac{1}{2} + \frac{1}{2}\right) = 0,\end{aligned} \tag{301}$$

in the second, just as we found before.

Note that the density operator contains less information than the underlying statistical ensemble of state vectors. For instance, consider a third scenario where 50% of the atoms are in the $|+\hat{\boldsymbol{x}}\rangle$ state and 50% are in the $|-\hat{\boldsymbol{x}}\rangle$ state. This is a *different* ensemble than the one from which we constructed $\boldsymbol{\rho}_I$, yet:

$$\boldsymbol{\rho}_{III} = \frac{1}{2}\left(|+\hat{\boldsymbol{x}}\rangle\langle+\hat{\boldsymbol{x}}| + |-\hat{\boldsymbol{x}}\rangle\langle-\hat{\boldsymbol{x}}|\right) = \frac{1}{2}\boldsymbol{I} = \boldsymbol{\rho}_I, \tag{302}$$

In general, there are many different statistical ensembles of state vectors that lead to the same density operator. Nonetheless, one can show that $\boldsymbol{\rho}$ contains enough information about the ensemble of state vectors to predict the outcome of *any* measurement:

Theorem 52 (Measurements on density operators)**.** *For any observable $\boldsymbol{L}$ measured on a quantum state randomly drawn from an ensemble with density operator $\boldsymbol{\rho}$,*

1. *The probability of measuring eigenvalue λ of $\boldsymbol{L}$ is*

$$p_\lambda = \mathrm{tr}\left(\boldsymbol{\rho}\boldsymbol{\Pi}_\lambda\right), \tag{303}$$

where $\boldsymbol{\Pi}_\lambda$ is the orthogonal projector onto the eigenspace of λ.

2. *The ensemble of states resulting from the measurement outcome λ has density operator*

$$\boldsymbol{\rho}' = \frac{\boldsymbol{\Pi}_\lambda \boldsymbol{\rho} \boldsymbol{\Pi}_\lambda}{\operatorname{tr}(\boldsymbol{\rho} \boldsymbol{\Pi}_\lambda)}. \tag{304}$$

Question 53. *Prove **Theorem 52**.*

As a result of **Theorem 52**, we see that the statistical ensemble fo the state vectors that generates $\boldsymbol{\rho}$ *cannot be measured*. Only the density operator can be measured. Likewise, the time evolution of the density operator can be predicted without knowing the underlying ensemble of state vectors:

Theorem 54 (Unitary evolution of density operators)**.** *In an closed process–in which by **Postulate 3** state vectors change by a unitary transformation $|\psi'\rangle = \boldsymbol{U}|\psi\rangle$–the density operator also changes by a unitary transformation:*

$$\boldsymbol{\rho}' = \boldsymbol{U} \boldsymbol{\rho} \boldsymbol{U}^\dagger. \tag{305}$$

Question 55. *Prove **Theorem 54**.*

Notice that **Theorems 50**, **52** and **54** above strongly resemble the three postulates of quantum mechanics. Indeed, some people choose to define quantum systems as density operators acting on the Hilbert space $\mathcal{H}$, rather than as (projective) vectors in $\mathcal{H}$. To see how these two concepts of states are related, note that in the special case where the normalized state vector $|\psi\rangle$ is certain, the density operator has rank 1:

$$\boldsymbol{\rho}_{\text{pure}} = |\psi\rangle\langle\psi|. \tag{306}$$

Moreover $|\psi\rangle$ is the unique (up to rescaling) eigenvector of $\boldsymbol{\rho}_{\text{pure}}$ with eigenvalue $+1$. Thus, the projective Hilbert space can be naturally identified with the space of rank-one density operators. These are called **pure states**, whereas density operators $\boldsymbol{\rho}$ with rank larger than 1 are called **mixed states** and the special case $\boldsymbol{\rho} \propto \boldsymbol{I}$ is said to be **maximally mixed**.[72]

[72] For example, the "unpolarized" input beam in our original discussion of the Stern-Gerlach experiment can be more precisely described as being in the maximally mixed spin state.

Everything we have done so far using pure states can be rephrased in terms of mixed states. For instance, the expectation value of an observable $\boldsymbol{A}$ becomes

$$\langle \boldsymbol{A} \rangle = \sum_I p_i \langle\psi_i| \boldsymbol{A} |\psi_i\rangle = \sum_i p_i \operatorname{tr}(|\psi_i\rangle\langle\psi_i| \boldsymbol{A}) = \sum_i \operatorname{tr}(p_i |\psi_i\rangle\langle\psi_i| \boldsymbol{A}) = \operatorname{tr}(\boldsymbol{\rho} \boldsymbol{A}). \tag{307}$$

In the Schrödinger picture, $\boldsymbol{\rho}$ evolves with time:

$$\begin{aligned} \frac{\partial \boldsymbol{\rho}}{\partial t} &= \sum_i p_i \left(\frac{\partial |\psi_i\rangle}{\partial t} \langle\psi_i| + |\psi_i\rangle \frac{\partial \langle\psi_i|}{\partial t} \right) = \frac{1}{i\hbar} \sum_i p_i \left(\boldsymbol{H} |\psi_i\rangle\langle\psi_i| - |\psi_i\rangle\langle\psi_i| \boldsymbol{H} \right) \\ &= \frac{1}{i\hbar} \left[\boldsymbol{H}, \sum_i p_i |\psi_i\rangle\langle\psi_i| \right] = \frac{1}{i\hbar} [\boldsymbol{H}, \boldsymbol{\rho}], \end{aligned} \tag{308}$$

using the Schrödinger equation.[73] Thus,

[73] Note that $\frac{dp_i}{dt} = 0$ since the time evolution only affects the state vector itself, not the initial population of states.

$$\boxed{i\hbar \frac{\partial \boldsymbol{\rho}}{\partial t} = -[\boldsymbol{\rho}, \boldsymbol{H}]}, \quad (\text{the } \textbf{von Neumann equation}). \tag{309}$$

This is the analog of the Schrödinger equation for mixed states.

Note that the von Neumann equation has a deceptive resemblance to the Heisenberg equation, only with the "wrong" sign for the commutator on the right-hand side. In fact, the von Neumann equation applies in the *Schrödinger*. In the Heisenberg picture,

$$\frac{d\boldsymbol{\rho}_H}{dt} = 0 \quad \text{since} \quad \frac{d\left|\psi_H\right\rangle}{dt} = 0. \tag{310}$$

This is consistent with the Heisenberg equation:

$$\frac{d\boldsymbol{\rho}_H}{dt} = \frac{i}{\hbar}\left[\boldsymbol{H}_H, \boldsymbol{\rho}_H\right] + \frac{\partial \boldsymbol{\rho}_H}{\partial t} = \frac{i}{\hbar}\left[\boldsymbol{H}_H, \boldsymbol{\rho}_H\right] + \left(\frac{i}{\hbar}\left[\boldsymbol{\rho}_S, \boldsymbol{H}_S\right]\right)_H = 0, \tag{311}$$

where the "explicit" time-dependence $\frac{\partial \boldsymbol{\rho}_H}{\partial t}$ is fixed by the von Neumann equation, Equation 309.

Moreover, notice that $\boldsymbol{\rho}_S$ is time-dependent whereas $\boldsymbol{\rho}_H$ is time-independent. This is the *opposite* fo what we would expect for a typical observable, which is time-independent in the Schrödinger picture and time-dependent in the Heisenberg picture. Indeed, even though it is self-adjoint, $\boldsymbol{\rho}$ is *not* an observable (i.e., we cannot "measure" $\boldsymbol{\rho}$ according to the rules of **Postulate 2**) since it is a state-dependent operator.

The von Neumann equation, Equation 309, is the quantum form of **Liouville's theorem**:

$$\frac{d\rho_{\text{cl}}}{dt} = \frac{\partial \rho_{\text{cl}}}{\partial t} + \left[\rho_{\text{cl}}, H_{\text{cl}}\right]_{\text{PB}} = 0. \tag{312}$$

Here $\rho_{\text{cl}} = \rho_{\text{cl}}\left(p, q, t\right)$ is the phase-space density of a statistical ensemble of *classical* particles evolving with time according to the dynamics dictated by the classical Hamiltonian H_{cl} where $\frac{\partial \rho_{\text{cl}}}{\partial t}$ is the time derivative of ρ_{cl} at a fixed location in phase space whereas $\frac{d\rho_{\text{cl}}}{dt}$ is its time derivative following the trajectory of one of the particles.

This helps to clarify the quantization procedure for classical statistical ensembles:

Quantization Rule 3 (addendum): (Canonical quantization of ensembles).

Upon canonical quantization, the classical phase space density is promoted to a density operator in the Heisenberg picture:

$$\rho_{\text{cl}} \to \boldsymbol{\rho}_H. \tag{313}$$

Mixed states representing ensembles are especially useful in quantum statistical mechanics. However they have another important interpretation, as we will see in the future.

References

[1] David Z. Albert. *Quantum Mechanics and Experience*. Harvard University Press, 2021.