

130A Quantum Physics

Part 1. Matrix Mechanics

Introduction

■ Everything Is a Vector

■ What Is Quantum Mechanics?

Quantum mechanics is a *physics theory* that describes the behavior of *quantum systems* (microscopic particles, strings, qubits ...).

What does **physics theory** do in general?

- Describe the **state** of the system: a set of *variables* encoding the relevant *information* of the system.
- Predict (i) the **observables** (measurement outcomes) and (ii) their **dynamics** (time evolution).

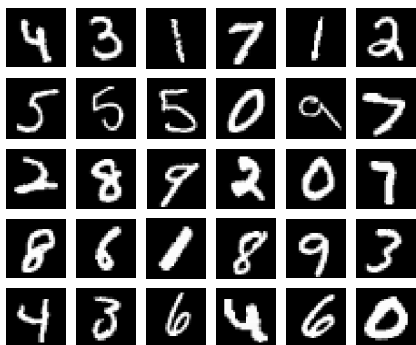
Physics theory is about encoding the physical reality in the form of **information** and making predictions about the reality based on such information.

■ How to Encode Information

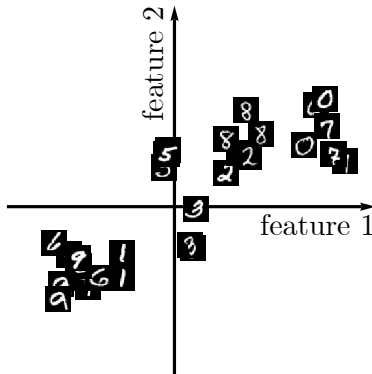
State variables encode the information about the system. They are *inferred* from observations.

- State variables may not have “physical meaning”.
- Choice of state variables may not be unique. (There can be more than one way to describe a system.)

Example: how to describe the following images?



- **Image file:** brightness of each pixel. - describe a state by all possible observables.
- **Human:** digits 0, 1, 2, ..., 9. - describe a state by a *name*.
- **Machine learning:** feature vectors in the latent space. - describe a state by a *vector* in a vector space. [This is the closest to what we do in quantum mechanics.]



In quantum mechanics, every **state** of a quantum system is *described* by a **complex vector** (an array of *complex* numbers), called the **state vector**.

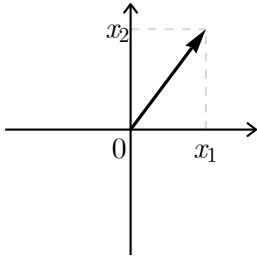
- The *vector components* are the **state variables**, and they may not need to have physical meanings. [They are also called **probability amplitudes** or **wave amplitudes**, but I don't explain what is "waving" here.]
- While the *state vector* (vector-based) approach is widely used, quantum mechanics can also be formulated using other frameworks.
 - **Examples:** **density matrix** (matrix-based), **classical shadow** (probability-based) [1], **quantum bootstrap** (observable-based) [2].
 - However, the state vector description remains a *precise* and *efficient* method for representing **pure states** of quantum systems. Therefore, we will begin our discussion with state vectors.
- The information encoded in the state vector, known as **quantum information**, forms the foundation for quantum computation and communication.

[1] Hsin-Yuan Huang, Richard Kueng, John Preskill. arXiv:2002.08953.

[2] Xizhi Han, Sean A. Hartnoll, Jorrit Kruthoff. arXiv:2004.10212.

■ What Is a Vector?

- *Geometrical* interpretation: a vector (in high-school physics) is an *arrow*, used to represent a physical quantity that has both **magnitude** and **direction**.
 - Example: the displacement vector $\mathbf{x}$ in a two-dimensional coordinate space



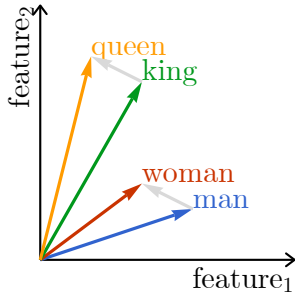
$$\mathbf{x} = (x_1, x_2) = (0.6, 0.8)$$

- *Computational* interpretation: a vector (in computer science) is an *array of numbers*, serves as a **data structure** to store and represent **information**.
 - Example 1: Color vector (the red/green/blue values form a vector)

```
■ .rgb = [1., 0.5, 0.25] # [r, g, b]
```

```
■ .rgb = [0.25, 0.5, 1.]
```

- Example 2: Word vector (in natural language processing), vector representations of words that encode the meaning and semantics of the words.



Use semantic relationship by vector arithmetic [3]:

$$|\text{king}\rangle - |\text{man}\rangle + |\text{woman}\rangle = |\text{queen}\rangle. \quad (1)$$

Vector in Quantum Mechanics:

- The notion of **state vector** in **quantum mechanics** is closer to the *computational* interpretation --- it is used to encode the *state* of a quantum system, or to store the *data* of quantum information. There is no direct physical meaning associated with its amplitude and direction.
- Real and complex vectors:

- **Real vector:** an array of *real* numbers --- the space of n -component real vectors is denoted as $\mathbb{R}^n$

$$\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n. \quad (2)$$

- **Complex vector:** an array of *complex* numbers --- the space of n -component complex vectors is denoted as $\mathbb{C}^n$

$$\mathbf{z} = (z_1, z_2, \dots, z_n) \in \mathbb{C}^n. \quad (3)$$

They are just arrays of different *data types*. It has been shown [4] that the formulation of quantum mechanics involving *complex* rather than real numbers is *necessary* to reproduce certain behavior of *quantum entanglement* observed in experiments.

[3] Ekaterina Vylomova, Laura Rimell, Trevor Cohn, Timothy Baldwin. arXiv:1509.01692

[4] Alessio Avella. Quantum Mechanics Must Be Complex. (2022)

■ Complex Algebra

■ Complex Number

A **complex number** z is made of two real numbers (x, y) that combine with a **real unit** 1 and an **imaginary unit** $i = \sqrt{-1}$ respectively,

$$x \in \mathbb{R}, y \in \mathbb{R} \rightarrow z = x + i y \in \mathbb{C}. \quad (4)$$

The real and imaginary units obey the following multiplication rules

$$1 \times 1 = 1, 1 \times i = i \times 1 = i, i \times i = -1. \quad (5)$$

• Addition:

$$\begin{cases} z = x + i y \\ w = u + i v \end{cases} \rightarrow z + w = (x + u) + i (y + v) \quad (6)$$

• Multiplication:

$$\begin{cases} z = x + i y \\ w = u + i v \end{cases} \rightarrow z w = (x u - y v) + i (x v + y u) \quad (7)$$

• Complex conjugation:

$$z = x + i y \rightarrow z^* = x - i y. \quad (8)$$

Real and imaginary parts can be extracted from

$$\begin{aligned} \operatorname{Re} z &= \frac{1}{2} (z + z^*) = x, \\ \operatorname{Im} z &= \frac{1}{2i} (z - z^*) = y. \end{aligned} \quad (9)$$

■ Complex Number in *Mathematica*

The imaginary unit i can be typeset in *Mathematica* by `EsciiEsc`. For example, here is a complex number

3 + 1 i

3 + i

Multiplying two complex numbers together (*Mathematica* treats the space between two numbers as a multiplication operator just as $a\ b = a \times b$ in algebra)

```
(3 + I) (4 + 2 I)
10 + 10 I
```

Complex conjugation is given by

```
Conjugate[3 + I]
3 - I
```

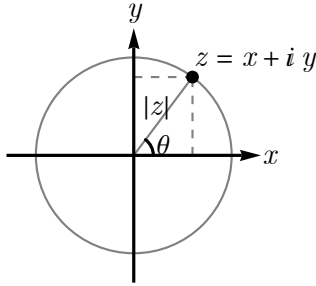
Extract real and imaginary part by

```
Re[3 + I]
Im[3 + I]
3
1
```

■ Polar Complex Form

Through Euler's formula, a complex number $z = x + i y$ may be written in a polar-coordinate form

$$z = |z| \cos \theta + i |z| \sin \theta = |z| e^{i \theta}. \quad (10)$$



- $|z|$ - **complex modulus** (or **magnitude**)

$$|z|^2 = z^* z = x^2 + y^2. \quad (11)$$

Note that $|z|^2 \geq 0$, such that its (positive) square root $|z|$ is defined.

- θ - **complex argument** (or **phase**)

$$\arg z = \theta = \text{Im} \ln z = \arctan \frac{y}{x}. \quad (12)$$

Multiplying a complex number z by $e^{i \phi}$ simply *rotates* the number by an angle ϕ in the complex plane [5],

$$\begin{aligned}
z &= |z| e^{i\theta}, \\
\Rightarrow e^{i\phi} z &= |z| e^{i(\theta+\phi)}.
\end{aligned} \tag{13}$$

- Two rotations can be *composed* into a new rotation with the rotation angles add up,

$$e^{i\phi_1} e^{i\phi_2} = e^{i(\phi_1+\phi_2)}. \tag{14}$$

- Complex conjugation simply flips the phase angle ($\phi \rightarrow -\phi$), representing an *inverse* rotation,

$$(e^{i\phi})^* = e^{-i\phi}. \tag{15}$$

[5] 3Blue1Brown. $e^{i\pi}$ in 3.14 minutes, using dynamics. (watch on YouTube)

■ Linear Algebra

■ Matrix and Vector

- A **matrix** is a two-dimensional array of numbers,

$$M = \begin{pmatrix} M_{11} & M_{12} & \cdots \\ M_{21} & M_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \in \mathbb{C}^{m \times n}. \tag{16}$$

- Matrix elements (components) M_{ij} are labeled by a *row index* $i = 1, \dots, m$ and a *column index* $j = 1, \dots, n$. Each component itself is a number. Let us consider $M_{ij} \in \mathbb{C}$ to be general, such that the space of m -row n -column matrices will be denoted as $\mathbb{C}^{m \times n}$.
- If $m = n$, the matrix is said to be a **square matrix**. In quantum mechanics, we will be mostly dealing with square matrices.
- A **vector** can be viewed as a special case of a *matrix*.
 - *Column* vectors (multi-row single-column)

$$|v\rangle = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \end{pmatrix} \in \mathbb{C}^{n \times 1} \cong \mathbb{C}^n. \tag{17}$$

- *Row* vectors (single-row multi-column)

$$\langle v| = (v_1^* \ v_2^* \ \cdots) \in \mathbb{C}^{1 \times n} \cong \mathbb{C}^n. \tag{18}$$

- *Column v.s. row*: In terms of encoding information in n numbers, it doesn't matter whether they are arranged in a column or a row. But when it comes to matrix-vector multiplication (to be discussed soon), there is a difference. So we use the $|v\rangle$ and $\langle v|$ notation to distinguish them, instead of writing both as v .

■ Linear Superposition

Matrix (or vector) space. All $m \times n$ matrices forms a matrix space $\mathbb{C}^{m \times n}$. Its defining property is that any linear combination of matrices in the space is still a matrix in the same space (same applies to vectors)

$$\begin{aligned} \forall A, B \in \mathbb{C}^{m \times n}; \alpha, \beta \in \mathbb{C}: \\ \alpha A + \beta B \in \mathbb{C}^{m \times n} \end{aligned} \quad (19)$$

A **linear combination** can be broken down into two types of basic operations:

- **Scalar multiplication:**

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots \\ A_{21} & A_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \rightarrow \alpha A = \begin{pmatrix} \alpha A_{11} & \alpha A_{12} & \cdots \\ \alpha A_{21} & \alpha A_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}, \quad (20)$$

meaning that

$$(\alpha A)_{ij} = \alpha A_{ij}. \quad (21)$$

- **Addition:**

$$\begin{aligned} A = \begin{pmatrix} A_{11} & A_{12} & \cdots \\ A_{21} & A_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}, B = \begin{pmatrix} B_{11} & B_{12} & \cdots \\ B_{21} & B_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \\ \rightarrow A + B = \begin{pmatrix} A_{11} + B_{11} & A_{12} + B_{12} & \cdots \\ A_{21} + B_{21} & A_{22} + B_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}, \end{aligned} \quad (22)$$

meaning that

$$(A + B)_{ij} = A_{ij} + B_{ij}. \quad (23)$$

All these rules applies to vectors when matrices are single-column or single-row.

■ Matrix Multiplication

Matrix multiplication is an *associative binary* operation:

$$\mathbb{C}^{m \times n} \times \mathbb{C}^{n \times l} \rightarrow \mathbb{C}^{m \times l}, \quad (24)$$

meaning that two matrices can multiply if and only if the *column* dimension of the *left* matrix matches the *row* dimension of the *right* matrix.

- Explicitly, when we write

$$\begin{pmatrix} A_{11} & A_{12} & \cdots \\ A_{21} & A_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} & \cdots \\ B_{21} & B_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} = \begin{pmatrix} C_{11} & C_{12} & \cdots \\ C_{21} & C_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}, \quad (25)$$

we mean that

$$C_{ij} = \sum_k A_{ik} B_{kj}, \quad (26)$$

where $k = 1, \dots, n$ is the index to be **contracted** (to be summed over).

- We can denote Eq. (25) on the matrix level simply as

$$A B = C. \quad (27)$$

Matrix-vector multiplication: If one of the matrix is reduced to a vector, the above rules still apply. A matrix can left-multiply a column vector or right-multiply a row vector, if their contracted dimensions matches.

- Left-multiplication

$$\begin{pmatrix} A_{11} & A_{12} & \cdots \\ A_{21} & A_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \vdots \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \end{pmatrix} \quad (28)$$

$$\Rightarrow v_i = \sum_j A_{ij} u_j,$$

- Right-multiplication

$$\begin{pmatrix} u_1 & u_2 & \cdots \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} & \cdots \\ A_{21} & A_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} = \begin{pmatrix} v_1 & v_2 & \cdots \end{pmatrix} \quad (29)$$

$$\Rightarrow v_j = \sum_i u_i A_{ij},$$

Vector-vector multiplication: If both matrices are reduced to vectors of the same dimension, we can define a **inner product** and a **outer product** between them.

- Inner product

$$\begin{pmatrix} u_1 & u_2 & \cdots \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \end{pmatrix} = \sum_i u_i v_i = \text{"a scalar (number)"}, \quad (30)$$

- Outer product

$$\begin{pmatrix} v_1 \\ v_2 \\ \vdots \end{pmatrix} \begin{pmatrix} u_1 & u_2 & \cdots \end{pmatrix} = \begin{pmatrix} v_1 u_1 & v_1 u_2 & \cdots \\ v_2 u_1 & v_2 u_2 & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}. \quad (31)$$

- Multiplying two row vectors or two column vectors are illegal (because dimensions do not match).

$$(u_1 \ u_2 \ \cdots)(u_1 \ u_2 \ \cdots) \rightarrow \text{No!}$$

$$\begin{pmatrix} v_1 \\ v_2 \\ \vdots \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \end{pmatrix} \rightarrow \text{No!} \quad (32)$$

■ Identity Matrix and Kronecker Symbol

Identity matrix: a special $n \times n$ square matrix whose *diagonal* are all 1's and *off-diagonal* are all 0's. It looks like

$$\mathbf{1} = \begin{pmatrix} 1 & 0 & 0 & \cdots \\ 0 & 1 & 0 & \ddots \\ 0 & 0 & 1 & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix}. \quad (33)$$

- The matrix element of an identity matrix can be expressed using the **Kronecker delta** symbol δ_{ij} ,

$$\mathbf{1}_{ij} = \delta_{ij} \equiv \begin{cases} 1 & i = j, \\ 0 & i \neq j. \end{cases} \quad (34)$$

- Identity matrix multiplying on any vector keeps the vector unchanged, i.e.
 $\forall \mathbf{u} \in \mathbb{C}^n : \mathbf{u} \mathbf{1} = \mathbf{1} \mathbf{u} = \mathbf{u}$. This implies that the Kronecker delta has the following property

$$\begin{aligned} \sum_i u_i \delta_{ij} &= u_j, \\ \sum_j \delta_{ij} u_j &= u_i. \end{aligned} \quad (35)$$

- Rule of thumb: when δ_{ij} appears in a summation of i (or j), it annihilates with the summation symbol and replaces summation index i by j (or j by i) in the summand.

■ Matrix Algebra in *Mathematica*

Construct two matrices

```
A = {{1, 2, 3}, {4, 5, 6}, {7, 8, 9}};
B = {{9, 8, 7}, {6, 5, 4}, {3, 2, 1}};
A // MatrixForm
B // MatrixForm
```

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

$$\begin{pmatrix} 9 & 8 & 7 \\ 6 & 5 & 4 \\ 3 & 2 & 1 \end{pmatrix}$$

- Linear combine them simply as

A + B // MatrixForm
 $\alpha A + \beta B$ // MatrixForm

$$\begin{pmatrix} 10 & 10 & 10 \\ 10 & 10 & 10 \\ 10 & 10 & 10 \end{pmatrix}$$

$$\begin{pmatrix} \alpha + 9\beta & 2\alpha + 8\beta & 3\alpha + 7\beta \\ 4\alpha + 6\beta & 5\alpha + 5\beta & 6\alpha + 4\beta \\ 7\alpha + 3\beta & 8\alpha + 2\beta & 9\alpha + \beta \end{pmatrix}$$

- Multiply them using “.” symbol, standing for the “dot product”.

A.B // MatrixForm

$$\begin{pmatrix} 30 & 24 & 18 \\ 84 & 69 & 54 \\ 138 & 114 & 90 \end{pmatrix}$$

B.A // MatrixForm

$$\begin{pmatrix} 90 & 114 & 138 \\ 54 & 69 & 84 \\ 18 & 24 & 30 \end{pmatrix}$$

Unlike multiplying two number ($a b = b a$, which is *commutative*), matrix multiplication is **non-commutative**, meaning that

$$A B \neq B A, \tag{36}$$

for two square matrices $A, B \in \mathbb{C}^{n \times n}$ in general.

■ Matrix as a Machine

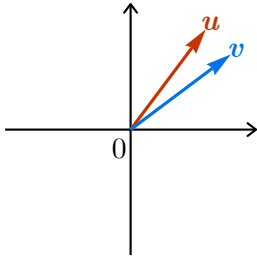
A *matrix* can be viewed as a machine that takes in a *vector*, acts (multiplies) on it, and returns a *new vector*.

Examples of 2×2 matrix M acting on 2-component vectors.

$$\mathbf{u} \xrightarrow{M} \mathbf{v} = M \mathbf{u}. \tag{37}$$

- Exchanging the two components in the vector by

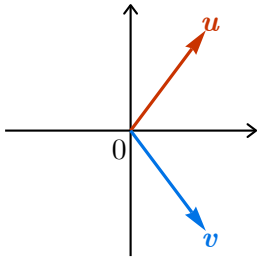
$$M = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{38}$$



$$\mathbf{u} = \begin{pmatrix} 0.6 \\ 0.8 \end{pmatrix} \xrightarrow{M} \mathbf{v} = \begin{pmatrix} 0.8 \\ 0.6 \end{pmatrix}$$

- Reflecting the vector with respect to an axis by

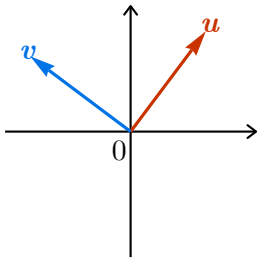
$$M = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (39)$$



$$\mathbf{u} = \begin{pmatrix} 0.6 \\ 0.8 \end{pmatrix} \xrightarrow{M} \mathbf{v} = \begin{pmatrix} 0.6 \\ -0.8 \end{pmatrix}$$

- Rotating the vector by 90° counterclockwise by

$$M = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \quad (40)$$



$$\mathbf{u} = \begin{pmatrix} 0.6 \\ 0.8 \end{pmatrix} \xrightarrow{M} \mathbf{v} = \begin{pmatrix} -0.8 \\ 0.6 \end{pmatrix}$$

■ Eigen System of a Matrix

An **eigen system** of a $n \times n$ square matrix M refers to the set of **eigenvalues** $\lambda^{(k)}$ and the corresponding **eigenvectors** $\mathbf{u}^{(k)}$ for $k = 1, \dots, n$ such that

$$M \mathbf{u}^{(k)} = \lambda^{(k)} \mathbf{u}^{(k)}, \quad (41)$$

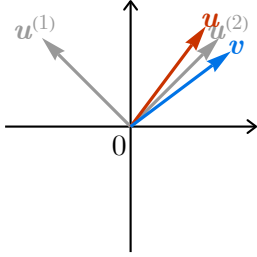
meaning that the multiplication of the matrix M to its eigenvector $\mathbf{u}^{(k)}$ can be reduced to the

scalar multiplication by $\lambda^{(k)}$. See [6] for more intuitive discussions.

Examples eigen systems of 2×2 matrices M .

- Exchanging the two components in the vector by

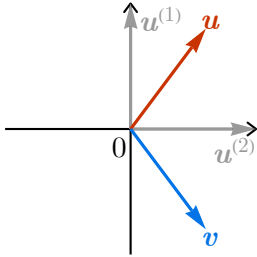
$$M = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \quad (42)$$



k	$\lambda^{(k)}$	$\mathbf{u}^{(k)}$
1	-1	$\begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$
2	1	$\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$

- Reflecting the vector with respect to an axis by

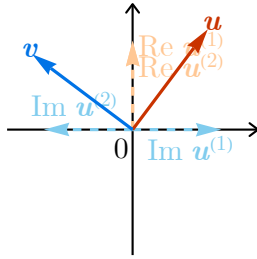
$$M = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (43)$$



k	$\lambda^{(k)}$	$\mathbf{u}^{(k)}$
1	-1	$\begin{pmatrix} 0 \\ 1 \end{pmatrix}$
2	1	$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$

- Rotating the vector by 90° counterclockwise by

$$M = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \quad (44)$$



k	$\lambda^{(k)}$	$u^{(k)}$
1	i	$\begin{pmatrix} \frac{i}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$
2	$-i$	$\begin{pmatrix} -\frac{i}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$

[6] 3Blue1Brown. Eigenvectors and eigenvalues | Chapter 14, Essence of linear algebra. (watch on YouTube).

■ Finding eigen systems in *Mathematica*

Choose a matrix

```
M = {{3, 1}, {0, 2}};
M // MatrixForm
```

$$\begin{pmatrix} 3 & 1 \\ 0 & 2 \end{pmatrix}$$

Find its eigenvalues and eigenvectors by `Eigensystem`, the result will be arranged as `{{val1, val2, ...}, {vec1, vec2, ...}}`.

```
Eigensystem[M]
```

$$\{\{3, 2\}, \{\{1, 0\}, \{-1, 1\}\}\}$$

Verify that the matrix multiplying on its eigenvector simply scales the eigenvector by the corresponding eigenvalue

```
M.{1, 0}
M.{-1, 1}
{3, 0}
{-2, 2}
```

Quantum States

■ Ket and Bra

■ Ket Vector

Postulate 1 (States): **States** of a quantum system are described as **vectors** in the associated Hilbert space.

In quantum mechanics, every **state** of a quantum system is described by a *complex column vector*, denoted by a **ket** (or ket state, ket vector) $|v\rangle$ in *Dirac's notation*,

$$|v\rangle \simeq \begin{pmatrix} v_1 \\ v_2 \\ \vdots \end{pmatrix}, \quad (45)$$

where $v_1, v_2, \dots \in \mathbb{C}$. The *length* of the vector depends on the **dimension** of the vector space.

- Note: “ $\simeq$ ” implies the vector representation is *basis dependent* and the values of vector components may change if we view the same state in a different basis.
- To write down the vector representation, we must specify a set of (orthonormal) **basis vectors** in the vector space, and represent them as one-hot unit vectors:

$$|1\rangle \simeq \begin{pmatrix} 1 \\ 0 \\ \vdots \end{pmatrix}, |2\rangle \simeq \begin{pmatrix} 0 \\ 1 \\ \vdots \end{pmatrix}, \dots \quad (46)$$

- Such that $|v\rangle$ can be expressed as a linear combination of basis vectors

$$\begin{aligned} |v\rangle &= v_1 |1\rangle + v_2 |2\rangle + \dots \\ &= \sum_i v_i |i\rangle. \end{aligned}$$

(47)

- The i th vector component v_i is the linear combination coefficient in front of the i th basis vector .

Superposition Principle: any linear combination of quantum states of a given quantum system is still a valid quantum state of the same system.

■ Bra Vector

Every **ket** $|v\rangle$ has a **dual vector**, called **bra** $\langle v|$, a *complex row vector*. They are related by *conjugate transpose* with respect to each other.

$$|v\rangle \simeq \begin{pmatrix} v_1 \\ v_2 \\ \vdots \end{pmatrix} \xleftrightarrow{\text{dual}} \langle v| \simeq (v_1^* \ v_2^* \ \cdots). \quad (48)$$

The name comes from the fact that they combine into a **bracket**, which represents a scalar product [to be introduced later].

- Every basis vector $|i\rangle$ also has a **dual basis vector** $\langle i|$, they are represented as

$$\begin{aligned} \langle 1| &\simeq (1 \ 0 \ \cdots), \\ \langle 2| &\simeq (0 \ 1 \ \cdots), \\ &\cdots. \end{aligned} \quad (49)$$

- The dual basis vectors form a set of basis for the bra vector. In terms of basis vectors,

$$\begin{aligned} \langle v| &= v_1^* \langle 1| + v_2^* \langle 2| + \cdots \\ &= \sum_i v_i^* \langle i|. \end{aligned} \quad (50)$$

- The i th vector component v_i^* is the linear combination coefficient in front of the i th dual basis vector $\langle i|$.

■ Qubit System

A **qubit** (or **quantum-bit**) is a quantum system that has two distinct states.

- The two distinct states are $|0\rangle$ and $|1\rangle$.
- We can *choose* $|0\rangle$ and $|1\rangle$ to be the **basis** vectors (like choosing a *coordinate system*) and write:

$$|0\rangle \simeq \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |1\rangle \simeq \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (51)$$

- The *vector representation* of a *quantum state* is also called a **state vector**.
- By saying that a qubit is a **two-state system**, its *state vector* has *two components*.
- A generic quantum state of a qubit is a complex *linear superposition* of the basis states

$$|\psi\rangle = \psi_0 |0\rangle + \psi_1 |1\rangle \simeq \begin{pmatrix} \psi_0 \\ \psi_1 \end{pmatrix}. \quad (52)$$

- $\psi_0, \psi_1 \in \mathbb{C}$ are complex numbers. They parameterize the state $|\psi\rangle$.
- Conversely, every two-component complex vector describes a qubit state.
- **Statistical interpretation:** $|\psi_0|^2$ and $|\psi_1|^2$ are respectively the probabilities to observe the qubit in the 0 and the 1 states.
- There is a dual bra vector $\langle\psi|$ associated with each ket vector $|\psi\rangle$,

$$\langle\psi| = \psi_0^* \langle 0| + \psi_1^* \langle 1| \simeq (\psi_0^* \ \psi_1^*), \quad (53)$$

which, the bra state, encodes the *same* information about the qubit as the ket state. They are equally good description of the quantum state of a qubit (but play different roles in defining the scalar product, as to be discussed soon).

■ Scalar Product

■ Definition

Scalar product (or **inner product**) is a function that takes two ket vectors, $|u\rangle$ and $|v\rangle$,

$$|u\rangle \simeq \begin{pmatrix} u_1 \\ u_2 \\ \vdots \end{pmatrix}, |v\rangle \simeq \begin{pmatrix} v_1 \\ v_2 \\ \vdots \end{pmatrix}, \quad (54)$$

and returns a complex number, denoted by the bracket $\langle u|v\rangle$,

$$\begin{aligned} \langle u|v\rangle &\simeq (u_1^* \ u_2^* \ \cdots) \begin{pmatrix} v_1 \\ v_2 \\ \vdots \end{pmatrix} \\ &= u_1^* v_1 + u_2^* v_2 + \dots \\ &= \sum_i u_i^* v_i. \end{aligned} \quad (55)$$

- Exchanging the two states in a scalar product lead to a complex conjugation of result

$$\langle v|u\rangle = \langle u|v\rangle^*. \quad (56)$$

- Scalar product of any vector $|v\rangle$ with itself is *real* and *positive definite*,

$$\langle v|v\rangle \geq 0. \quad (57)$$

More specifically,

$$\langle v|v\rangle \begin{cases} = 0 & \text{if } |v\rangle = 0 \\ > 0 & \text{otherwise} \end{cases}. \quad (58)$$

- This implies the **Cauchy-Schwarz inequality**

$$|\langle u|v\rangle|^2 \leq \langle u|u\rangle \langle v|v\rangle. \quad (59)$$

Exc
1

Prove Eq. (59).

Hilbert space: the space of all ket vectors together with their scalar product structure, denoted as $\mathcal{H}$, which is the space of all possible quantum states of a system.

■ Normalization

Squared norm of a vector $|v\rangle$ is the *scalar product* of the vector with itself, denoted as

$$\|v\|^2 = \langle v|v\rangle. \quad (60)$$

Taking off the square, $\|v\| = \sqrt{\langle v|v\rangle}$ is the **norm** of $|v\rangle$.

Normalized state: a state $|v\rangle$ is **normalized** $\Leftrightarrow$ Its *norm* is *one*, i.e.

$$\|v\|^2 = \langle v|v\rangle = \sum_i |v_i|^2 = 1. \quad (61)$$

- Example: Consider a qubit state

$$|v\rangle \simeq \begin{pmatrix} v_0 \\ v_1 \end{pmatrix}, \quad (62)$$

the normalization condition means

$$\langle v|v\rangle = v_0^* v_0 + v_1^* v_1 = 1. \quad (63)$$

- In general, the normalization condition means

$$\langle v|v\rangle = \sum_i |v_i|^2 = 1. \quad (64)$$

According to the statistical interpretation of quantum state, $|v_i|^2$ is the *probability* to observe the system in the i th basis state. The normalization condition is simply a requirement that the probabilities must *sum up to unity*.

- **Normalization** of a state: if a state was *not* normalized, it can be normalized by

$$|v\rangle \rightarrow \frac{|v\rangle}{\|v\|} = \frac{1}{\sqrt{\langle v|v\rangle}} |v\rangle, \quad (65)$$

unless $\|v\|$ is zero or infinity.

**Exc
2**

Normalize the vector .

■ Orthogonality

Orthogonal states: two states and are **orthogonal** to each other $\Leftrightarrow$ their *scalar product* is *zero*, i.e.

$$\langle u|v\rangle = \sum_i u_i^* v_i = 0. \quad (66)$$

- For example, the qubit states $|0\rangle$ and $|1\rangle$ (see Eq. (51)) are *orthogonal*, as

$$\langle 0|1\rangle = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0. \quad (67)$$

$|0\rangle$ and $|1\rangle$ are *orthogonal* for a good reason: they are **distinct** states of a qubit, i.e. if the qubit is in state 0, it is definitely not in state 1, vice versa.

■ Basis and Hilbert Space

■ Orthonormal Basis

Orthonormal basis: a (ordered) set of kets

$$\mathcal{B} = \{|i\rangle : i = 1, 2, \dots, n\}, \quad (68)$$

in which the vectors are **normalized** by themselves and **orthogonal** to each other:

$$\langle i|j\rangle = \delta_{ij} \equiv \begin{cases} 1 & i = j, \\ 0 & i \neq j. \end{cases} \quad (69)$$

- Each orthogonal basis state describes a distinct reality of the quantum system.
- Orthonormal basis states are *represented* by **one-hot vectors**, as they are normalized and orthogonal to each other

$$|1\rangle \simeq \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix}, |2\rangle \simeq \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix}, |3\rangle \simeq \begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \end{pmatrix}, \dots \quad (70)$$

Choosing a basis is always a helpful practice in quantum mechanics. But quantum mechanics can be formulated in a basis independent manner.

■ Hilbert Space

A set of orthonormal basis spans a **Hilbert space** (the vector space of kets), denoted as

$$\mathcal{H} = \text{span } \mathcal{B} = \text{span } \{|i\rangle : i = 1, 2, \dots, n\}. \quad (71)$$

- The **dimension** of the vector space $\dim \mathcal{H}$ = the number n of basis vectors = the *maximal* number of *linearly independent* vectors in the space.
- The Hilbert space dimension of a quantum system can be *finite* or *infinite*. Example: a qubit - $\dim \mathcal{H} = 2$, ten qubits - $\dim \mathcal{H} = 2^{10} = 1024$, a particle in a continuous space - $\dim \mathcal{H} = \infty$.
- Dimension of the Hilbert space is often a *choice*: we don't really know how many independent states are there in a quantum system. We only care about the states that are *relevant* to us.
- **Example:** $|0\rangle$ and $|1\rangle$ form an *orthonormal basis* of the qubit Hilbert space. They represent two distinct realities: if the qubit is in state $|0\rangle$, it is definitely not in state $|1\rangle$ (and vice versa).

■ State Basis Expansion

Completeness: Any *full* set of *distinct* states in the Hilbert space $\mathcal{H}$ forms a *complete* set of

orthonormal basis $\mathcal{B}$, such that *every* state $|v\rangle \in \mathcal{H}$ can be expanded as a *linear superposition* of the basis states,

$$|v\rangle = v_1 |1\rangle + v_2 |2\rangle + \dots = \sum_i v_i |i\rangle. \quad (72)$$

- Each basis state $|i\rangle$ describes a distinct reality that the quantum system can realize under observation.
- The *superposition coefficient* v_i are the **components** of the state vector, which can be extracted by the *scalar product* with the basis state,

$$v_i = \langle i | v \rangle. \quad (73)$$

Eq. (72) and Eq. (73) can be written in a more elegant form in terms of bras and kets only

$$|v\rangle = \sum_i |i\rangle \langle i | v \rangle. \quad (74)$$

- **Statistical interpretation:** If a quantum system is known to be in a superposition state $|v\rangle = \sum_i v_i |i\rangle$ of distinct realities described by orthogonal states $|i\rangle$ ($i = 1, 2, \dots$), an observation designed to discern which reality the system is really in will find the system in the state $|i\rangle$ with the probability

$$p(i | v) = |v_i|^2 = |\langle i | v \rangle|^2. \quad (75)$$

■ Born's Rule

■ Fidelity

The **fidelity** $F(u, v)$ between two quantum states $|u\rangle$ and $|v\rangle$ quantifies the similarity (overlap) between two states. It is given by the squared absolute value of their scalar product (assuming the normalization of state vectors)

$$F(u, v) = |\langle u | v \rangle|^2. \quad (76)$$

- Fidelity is *symmetric*: $F(u, v) = F(v, u)$.
- Fidelity takes values in the range of

$$0 \leq F(u, v) \leq 1. \quad (77)$$

This follows from the Cauchy-Schwarz inequality of scalar product Eq. (59) that $|\langle u | v \rangle|^2 \leq \langle u | u \rangle \langle v | v \rangle$.

■ Statistical Interpretation

Hypothesis testing: If a quantum system is prepared in a state $|v\rangle$, an observation

designed to check whether the system is in the state $|u\rangle$ will return a confirmative result with probability

$$p(u | v) = |\langle u | v \rangle|^2. \quad (78)$$

Detailed balance: the probability to observe one state given another is the same as the other way round, both are given by their fidelity

$$p(u | v) = p(v | u) = F(u, v) = |\langle u | v \rangle|^2. \quad (79)$$

- **Identical states.** Two states $|u\rangle$ and $|v\rangle$ are *identical* iff the fidelity between them is *one* (fully overlap)

$$|\langle u | v \rangle|^2 = 1. \quad (80)$$

- This is only achievable when

$$|u\rangle = e^{i\varphi} |v\rangle, \quad (81)$$

i.e. the two states are the same up to phase ambiguity.

- **Reality** must be *confirmable* by *repeated* observations: if a quantum system is known to be in a state $|v\rangle$, observing the system again will certainly confirm the state $|v\rangle$ (with probability 1).
- **Distinct states.** Two states $|u\rangle$ and $|v\rangle$ are *distinct* iff the fidelity between them is *zero* (no overlap)

$$|\langle u | v \rangle|^2 = 0. \quad (82)$$

- *Orthogonal* states $\Leftrightarrow$ *distinct* realities.
- **Distinct realities** are *distinguishable* by *repeated* observations: if a quantum system is known to be in a state $|v\rangle$, observing the system again will certainly not find the system in another orthogonal state $|u\rangle$.
- **Overlapping states.** In general, two different states $|u\rangle$ and $|v\rangle$ may have partial overlap (they don't need to be orthogonal), i.e. their fidelity falls between zero and one

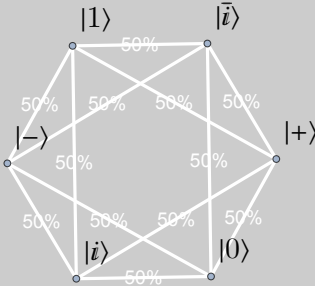
$$0 < |\langle u | v \rangle|^2 < 1. \quad (83)$$

- Realities can overlap: if two quantum states are more similar to (more overlapped with) each other, the probability to confuse them is higher.

$$|\"3\" \rangle \simeq \text{[3]}, \quad |\"5\" \rangle \simeq \text{[5]}.$$

HW
1

$\{|0\rangle, |1\rangle\}$, $\{|+\rangle, |-\rangle\}$, and $\{|i\rangle, |\bar{i}\rangle\}$ are three pairs of distinct states of a *qubit* (i.e. 2-dimensional Hilbert space). Within each pair, the two states are orthogonal. However, any two states from different pairs have 50% overlap (i.e. $1/2$ fidelity). Their overlapping relations can be visualized as the following graph.



Can you figure out an assignment of 2-component vector representation for these states that is consistent with their overlapping relations?

[Hint: read Lecture 2 of [7]]

[Comment: This result shows how it is possible to embed so many different realities just in a 2-dimensional Hilbert space.]

- [7] Leonard Susskind, Art Friedman. *Quantum Mechanics - the Theoretical Minimum*. Publisher: Basic Books (2014).

Quantum Operators

■ Matrix Representation

■ Definition

Operator: an operator acts on a state and returns a new state.

$$\begin{aligned}\hat{O} : \mathcal{H} &\rightarrow \mathcal{H} \\ |v\rangle &\mapsto |w\rangle = \hat{O} |v\rangle\end{aligned}\tag{84}$$

- **Linear operator:** action of the operator commutes with linear combination of the state

$$\hat{O} (\alpha |v\rangle + \beta |u\rangle) = \alpha \hat{O} |v\rangle + \beta \hat{O} |u\rangle.\tag{85}$$

- **Identity operator** is a special operator that maps any state to itself (the do-nothing operator), denoted as $\mathbb{1}$.

$$\forall |v\rangle : \mathbb{1} |v\rangle = |v\rangle.\tag{86}$$

■ Operator Acting on State

Recall: a matrix *multiplying* on a vector produces a new vector. If every **quantum state** is described by a **vector**, one may conjecture that every **quantum operator** should be described

by a (square) **matrix**. --- This is indeed a basic assumption of quantum mechanics: states are to be operated (transformed) *linearly*.

Applying an *operator* to a *state* $\simeq$ multiplying a *matrix* to a *vector*.

$$\begin{array}{ccc} |w\rangle & = & \hat{O} \quad |v\rangle \\ \downarrow \simeq & & \downarrow \simeq \quad \downarrow \simeq \\ \begin{pmatrix} w_1 \\ w_2 \\ \vdots \end{pmatrix} & = & \begin{pmatrix} O_{11} & O_{12} & \cdots \\ O_{21} & O_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \end{pmatrix} \end{array} \quad (87)$$

or equivalently

$$w_i = \sum_j O_{ij} v_j. \quad (88)$$

- The **matrix element** O_{ij} tells how the operator should act on each basis states: the operator $\hat{O}$ will turn a basis state $|j\rangle$ to a superposition state of basis states $|i\rangle$ with superposition coefficients O_{ij} .

$$\hat{O}|j\rangle = \sum_i O_{ij} |i\rangle. \quad (89)$$

Exc
3

Show Eq. (89) as a result of Eq. (87) using one-hot representation for basis vectors.

- In other words, O_{ij} is the **amplitude** to transform basis state $|j\rangle$ to basis state $|i\rangle$ under the action of the operator $\hat{O}$. It is sufficient to specify the operator by specifying its action on basis states, as all possible states are just linear combination of basis states, and the operator acts linearly.

■ Operator Basis Expansion

Given an *orthonormal* basis $\mathcal{B} = \{|i\rangle : i = 1, 2, \dots\}$ of the Hilbert space $\mathcal{H}$, every *operator* $\hat{O}$ acting in $\mathcal{H}$ can be expanded as a *linear combination* of **basis operators** $|i\rangle\langle j|$,

$$\hat{O} = \sum_{ij} |i\rangle O_{ij} \langle j|, \quad (90)$$

- $|i\rangle\langle j|$ denotes the operator that targets the state $|j\rangle$ and transforms it to the state $|i\rangle$, because

$$\begin{aligned} (|i\rangle\langle j|) |k\rangle &= |i\rangle \langle j|k\rangle = |i\rangle \delta_{jk} \\ &= \begin{cases} |i\rangle & \text{if } k = j, \\ 0 & \text{if } k \neq j. \end{cases} \end{aligned} \quad (91)$$

Thus Eq. (90) is consistent with the Eq. (89) in describing how the operator $\hat{O}$ acts on the state.

- $O_{ij} \in \mathbb{C}$ are *complex* coefficients, which can be extracted by

$$O_{ij} = \langle i | \hat{O} | j \rangle. \quad (92)$$

Exc
4

Prove Eq. (92) from Eq. (90) using the orthonormal property of the basis vectors, without representing them as on-hot vectors explicitly.

- Alternatively, $|i\rangle\langle j|$ can be represented as an **one-hot matrix** that is zero everywhere with a single 1 at the row- i column- j . For example, in a 2-dimensional Hilbert space [recall Eq. (31) for how to outer product two vectors]

$$\begin{aligned} |1\rangle\langle 1| &\simeq \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \\ |1\rangle\langle 2| &\simeq \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \\ |2\rangle\langle 1| &\simeq \begin{pmatrix} 0 \\ 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \\ |2\rangle\langle 2| &\simeq \begin{pmatrix} 0 \\ 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned} \quad (93)$$

Therefore, Eq. (90) indeed reconstructs the matrix representation

$$\begin{aligned} O_{11} |1\rangle\langle 1| + O_{12} |1\rangle\langle 2| + O_{21} |2\rangle\langle 1| + O_{22} |2\rangle\langle 2| \\ \simeq \begin{pmatrix} O_{11} & O_{12} \\ O_{21} & O_{22} \end{pmatrix}. \end{aligned} \quad (94)$$

The above can be generalized to larger matrices (higher dimensions).

Matrix representation. Every operator $\hat{O}$ can be represented as a matrix

$$\hat{O} \simeq \begin{pmatrix} O_{11} & O_{12} & \cdots \\ O_{21} & O_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}. \quad (95)$$

The i th row j th column matrix element O_{ij} describes:

- The linear combination coefficient in front of the basis operator $|i\rangle\langle j|$, as in Eq. (90).
- The amplitude to transform state $|j\rangle$ to state $|i\rangle$ under the action of the operator $\hat{O}$, as in Eq. (89).

■ Examples of Operators

Example I: Identity operator

Identity operator is universally represented by the **identity matrix** in any orthonormal basis (independent of the basis choice).

According to Eq. (92),

$$\mathbb{1}_{ij} = \langle i | \mathbb{1} | j \rangle = \langle i | j \rangle = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}. \quad (96)$$

- In matrix representation Eq. (95),

$$\mathbb{1} = \begin{pmatrix} 1 & & \\ & 1 & \\ & & \ddots \end{pmatrix}. \quad (97)$$

- Using Dirac notation Eq. (90),

$$\mathbb{1} = \sum_{ij} |i\rangle \mathbb{1}_{ij} \langle j| = \sum_i |i\rangle \langle i|. \quad (98)$$

This is also call the **resolution of identity**.

Example II: Pauli operators

Pauli operators are a set of operators acting on a qubit.

$$\begin{aligned} \hat{\sigma}^x &= |1\rangle \langle 0| + |0\rangle \langle 1|, \\ \hat{\sigma}^y &= i |1\rangle \langle 0| - i |0\rangle \langle 1|, \\ \hat{\sigma}^z &= |0\rangle \langle 0| - |1\rangle \langle 1|, \end{aligned} \quad (99)$$

Sometimes the identity operator

$$\mathbb{1} = |0\rangle \langle 0| + |1\rangle \langle 1|, \quad (100)$$

is also included as the 0th Pauli operator.

Pauli matrices - matrix representations of Pauli operators on the qubit basis $\{|0\rangle, |1\rangle\}$:

$$\mathbb{1} \simeq \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \hat{\sigma}^x \simeq \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \hat{\sigma}^y \simeq \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \hat{\sigma}^z \simeq \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (101)$$

■ Operator Algebra

■ Operator Product

Product (or **composition**) of two **operators** $\hat{O}$ and $\hat{P}$ is a combined operator $\hat{O} \hat{P}$ that first applies $\hat{P}$ to the state then applies $\hat{O}$ (from *right* to *left*):

$$(\hat{O} \hat{P}) |v\rangle = (\hat{O}(\hat{P} |v\rangle)). \quad (102)$$

- Composing two *operators* $\simeq$ multiplying two *matrices*.

$$\hat{O} \hat{P} \simeq \begin{pmatrix} O_{11} & O_{12} & \cdots \\ O_{21} & O_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} P_{11} & P_{12} & \cdots \\ P_{21} & P_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}. \quad (103)$$

Exc 5 | Prove Eq. (103) using Eq. (90).

- Operator product is *non-commutative* in general, i.e.

$$\hat{O} \hat{P} \neq \hat{P} \hat{O}. \quad (104)$$

■ Single-Qubit Pauli Operators

Example: product of Pauli operators

Multiplication table

	1	$\hat{\sigma}^x$	$\hat{\sigma}^y$	$\hat{\sigma}^z$
1	1	$\hat{\sigma}^x$	$\hat{\sigma}^y$	$\hat{\sigma}^z$
$\hat{\sigma}^x$	$\hat{\sigma}^x$	1	$i \hat{\sigma}^z$	$-i \hat{\sigma}^y$
$\hat{\sigma}^y$	$\hat{\sigma}^y$	$-i \hat{\sigma}^z$	1	$i \hat{\sigma}^x$
$\hat{\sigma}^z$	$\hat{\sigma}^z$	$i \hat{\sigma}^y$	$-i \hat{\sigma}^x$	1

(105)

Exc 6 | Verify Eq. (105) by multiplying Pauli matrices defined in Eq. (101).

- The table Eq. (105) can be summarized in a single formula: the product of Pauli matrices (as the defining property of Pauli matrices)

$$\hat{\sigma}^a \hat{\sigma}^b = \delta^{ab} \mathbb{1} + i \epsilon^{abc} \hat{\sigma}^c,$$

(106)

where $a, b, c = x, y, z$.

- δ^{ab} denotes the Kronecker delta symbol, defined as

$$\delta^{ab} = \begin{cases} 1 & \text{if } a = b \\ 0 & \text{if } a \neq b \end{cases} \quad (107)$$

- ϵ^{abc} denotes the Levi-Civita symbol, defined as

$$\epsilon^{abc} = \begin{cases} 1 & \text{if } (a \ b \ c) \text{ is a cyclic of } (x \ y \ z) \\ -1 & \text{if } (a \ b \ c) \text{ is a cyclic of } (z \ y \ x) \\ 0 & \text{otherwise} \end{cases} \quad (108)$$

- Another version of Eq. (106) using vector notation

$$(\mathbf{m} \cdot \hat{\sigma}) (\mathbf{n} \cdot \hat{\sigma}) = (\mathbf{m} \cdot \mathbf{n}) \mathbb{1} + i (\mathbf{m} \times \mathbf{n}) \cdot \hat{\sigma},$$

(109)

where $\mathbf{m}, \mathbf{n}$ are three-component vectors (each component is a scalar).

- The generalized vector $\hat{\sigma}$ should be understood as a vector of matrices, or as a three-dimensional tensor (shape: $3 \times 2 \times 2$).
- Here $\mathbf{m} \cdot \hat{\sigma}$ means

$$\begin{aligned}\mathbf{m} \cdot \hat{\boldsymbol{\sigma}} &= m_x \hat{\sigma}^x + m_y \hat{\sigma}^y + m_z \hat{\sigma}^z \\ &\simeq \begin{pmatrix} m_z & m_x - i m_y \\ m_x + i m_y & -m_z \end{pmatrix}.\end{aligned}\tag{110}$$

As we contract a 3-component vector $\mathbf{m}$ with a $3 \times 2 \times 2$ -component tensor $\hat{\boldsymbol{\sigma}}$ along the first index (the dimension 3 index), the result is a 2×2 matrix.

- Repeatedly applying Eq. (109) enables us to product more Pauli operators together. For example

$$(\mathbf{l} \cdot \hat{\boldsymbol{\sigma}})(\mathbf{m} \cdot \hat{\boldsymbol{\sigma}})(\mathbf{n} \cdot \hat{\boldsymbol{\sigma}}) = i \mathbf{l} \cdot (\mathbf{m} \times \mathbf{n}) \mathbb{1} + ((\mathbf{m} \cdot \mathbf{n}) \mathbf{l} - (\mathbf{l} \cdot \mathbf{n}) \mathbf{m} + (\mathbf{l} \cdot \mathbf{m}) \mathbf{n}) \cdot \hat{\boldsymbol{\sigma}}.\tag{111}$$

Exc
7

Derive Eq. (111).

■ Commutator

Commutator of two operators $\hat{O}$ and $\hat{P}$

$$[\hat{O}, \hat{P}] = \hat{O} \hat{P} - \hat{P} \hat{O}.\tag{112}$$

- Commutator is *antisymmetric*, $[\hat{O}, \hat{P}] = -[\hat{P}, \hat{O}]$.
- As a result, *commutator* of an operator with *itself* always *vanishes* $[\hat{O}, \hat{O}] = 0$.
- If the commutator vanishes $[\hat{O}, \hat{P}] = 0$, we say that the two operators $\hat{O}$ and $\hat{P}$ **commute**, i.e. $\hat{O} \hat{P} = \hat{P} \hat{O}$ (operators can *pass through* each other as if they were *numbers*) $\Rightarrow$ it does not matter which operator is applied first, the consequence will be the same.

Example: dressing up to school.

- A: put on the socks,
- B: put on the shoes,
- C: put on the hat,

A and B do *not commute* (changing the order leads to different result). But A and C *commute*, B and C also *commute* (changing the order does not affect the result).

Useful *rules* to evaluate commutators

- **Bi-linearity:**

$$\begin{aligned}[\hat{O}, \hat{P} + \hat{Q}] &= [\hat{O}, \hat{P}] + [\hat{O}, \hat{Q}], \\ [\hat{O} + \hat{P}, \hat{Q}] &= [\hat{O}, \hat{Q}] + [\hat{P}, \hat{Q}].\end{aligned}\tag{113}$$

Exc
8

Prove Eq. (113).

- **Product rules:**

$$\begin{aligned} [\hat{O}, \hat{P} \hat{Q}] &= [\hat{O}, \hat{P}] \hat{Q} + \hat{P} [\hat{O}, \hat{Q}], \\ [\hat{O} \hat{P}, \hat{Q}] &= [\hat{O}, \hat{Q}] \hat{P} + \hat{O} [\hat{P}, \hat{Q}]. \end{aligned} \quad (114)$$

**Exc
9**

Prove Eq. (114).

Example: Commutators of Pauli operators

$$\begin{aligned} [\hat{\sigma}^x, \hat{\sigma}^y] &= 2i \hat{\sigma}^z, \\ [\hat{\sigma}^y, \hat{\sigma}^z] &= 2i \hat{\sigma}^x, \\ [\hat{\sigma}^z, \hat{\sigma}^x] &= 2i \hat{\sigma}^y. \end{aligned} \quad (115)$$

Or more compactly as

$$[\hat{\sigma}^a, \hat{\sigma}^b] = 2i \epsilon^{abc} \hat{\sigma}^c, \quad (116)$$

for $a, b, c = x, y, z$, using the Levi-Civita symbol ϵ^{abc} defined in Eq. (108).

- Eq. (116) can be considered as the defining algebraic properties of *single-qubit operators* (Pauli matrices). Or even more compactly expressed using the **cross product** of vectors

$$\hat{\sigma} \times \hat{\sigma} = 2i \hat{\sigma}. \quad (117)$$

■ Operator Function

Operator power. n th power of an operator $\hat{O}$ is the composition of $\hat{O}$ by n times.

$$\hat{O}^n = \hat{O} \hat{O} \dots (n \text{ times}) \dots \hat{O}. \quad (118)$$

Operator function. Given a function $f(x)$ that admits Taylor expansion

$$f(x) = \sum_n c_n x^n, \quad (119)$$

the corresponding operator function is defined as

$$f(\hat{O}) = \sum_n c_n \hat{O}^n, \quad (120)$$

with the same set of coefficients c_n .

- $f(\hat{O})$ is still an operator that can act on states in $\mathcal{H}$.
- **Operator exponential.** Given the exponential function

$$e^x = 1 + x + \frac{x^2}{2!} + \dots = \sum_{n=0}^{\infty} \frac{1}{n!} x^n, \quad (121)$$

the exponential of an operator is defined as

$$e^{\hat{O}} = \mathbf{1} + \hat{O} + \frac{\hat{O}^2}{2!} + \dots = \sum_{n=0}^{\infty} \frac{1}{n!} \hat{O}^n, \quad (122)$$

- Note: exponentiating an matrix is NOT exponentiating each of the matrix element.

See [8] for a nice video about matrix exponential.

[8] 3Blue1Brown. How (and why) to raise e to the power of a matrix. (watch on YouTube)

Example: exponentiating a Pauli matrix

Exc
10

Given $\hat{\sigma}^y \simeq \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$,
show that the matrix representation of $e^{i\theta\hat{\sigma}^y}$ is
 $e^{i\theta\hat{\sigma}^y} \simeq \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}$.

HW
2

Use the definition Eq. (122) to prove that
 $\exp(i\theta \mathbf{n} \cdot \hat{\boldsymbol{\sigma}}) = \cos(\theta) \mathbf{1} + i \sin(\theta) \mathbf{n} \cdot \hat{\boldsymbol{\sigma}}$
given that $\mathbf{n}$ is a 3-component *real unit vector*.

■ Operator Trace

The **trace** of an operator $\hat{O}$ is defined as

$$\text{Tr } \hat{O} = \sum_i \langle i | \hat{O} | i \rangle. \quad (123)$$

The result is a scalar.

- On the matrix level, taking the trace is simply *summing* over *diagonal* matrix elements

$$\text{Tr} \begin{pmatrix} O_{11} & O_{12} & \cdots \\ O_{21} & O_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} = O_{11} + O_{22} + \dots = \sum_i O_{ii}. \quad (124)$$

- **Linear property:** trace is a *linear functional* of operators.

$$\text{Tr}(\alpha \hat{O} + \beta \hat{P}) = \alpha \text{Tr } \hat{O} + \beta \text{Tr } \hat{P}. \quad (125)$$

- **Cyclic property:** the trace of a product of operators is invariant under the cyclic permutation of the operators.

$$\begin{aligned} \text{Tr}(\hat{O} \hat{P}) &= \text{Tr}(\hat{P} \hat{O}), \\ \text{Tr}(\hat{O} \hat{P} \hat{Q}) &= \text{Tr}(\hat{P} \hat{Q} \hat{O}) = \text{Tr}(\hat{Q} \hat{O} \hat{P}), \end{aligned} \quad (126)$$

...

Exc 11 Prove Eq. (126).

The operator trace is *useful* in computing **scalar product** or **fidelity**:

- Scalar product

$$\langle u|v\rangle = \text{Tr } |v\rangle\langle u|. \quad (127)$$

- Fidelity

$$|\langle u|v\rangle|^2 = \langle u|v\rangle\langle v|u\rangle = \text{Tr } |v\rangle\langle v| |u\rangle\langle u|. \quad (128)$$

Example: trace of Pauli operators

Pauli operators are *traceless*.

$$\text{Tr } \hat{\sigma}^x = \text{Tr } \hat{\sigma}^y = \text{Tr } \hat{\sigma}^z = 0. \quad (129)$$

This is true for a Pauli operator along any direction

$$\text{Tr } \mathbf{n} \cdot \hat{\sigma} = 0. \quad (130)$$

Measurement

■ Hermitian Operators

■ Hermitian Conjugate

We have explained how an operator $\hat{O}$ acts on a *ket* state $|v\rangle$, what about its action on the *bra* state $\langle v|$?

	ket	bra (dual)
Hilbert space	$\mathcal{H}$	$\mathcal{H}^*$
basis	$\mathcal{B} = \{ i\rangle\}$	$\mathcal{B}^* = \{\langle i \}$
state	$ v\rangle = \sum_i v_i i\rangle$	$\langle v = \sum_i v_i^* \langle i $
vector	$\begin{pmatrix} v_1 \\ v_2 \\ \vdots \end{pmatrix}$	$(v_1^* \ v_2^* \ \cdots)$
component	$v_i = \langle i v\rangle$	$v_i^* = \langle v i\rangle$
operator	$\hat{O} = \sum_{ij} i\rangle O_{ij} \langle j $	$\hat{O}^\dagger = \sum_{ij} i\rangle O_{ji}^* \langle j $
matrix	$\begin{pmatrix} O_{11} & O_{12} & \cdots \\ O_{21} & O_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$	$\begin{pmatrix} O_{11}^* & O_{21}^* & \cdots \\ O_{12}^* & O_{22}^* & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$
component	$O_{ij} = \langle i \hat{O} j\rangle$	$O_{ij}^* = \langle j \hat{O}^\dagger i\rangle$

(131)

$$\text{action} \quad |w\rangle = \hat{O} |v\rangle \quad \langle w| = \langle v| \hat{O}^\dagger$$

- Just like the *bra* $\langle v|$ is the **dual** of the *ket* $|u\rangle$, the **Hermitian conjugate** operator $\hat{O}^\dagger$ is the **dual** of the original operator $\hat{O}$, such that

- if the operator $\hat{O}$ takes $|v\rangle$ to $|w\rangle$:

$$\begin{aligned} \hat{O} : \mathcal{H} &\rightarrow \mathcal{H} \\ |v\rangle &\mapsto |w\rangle = \hat{O} |v\rangle \end{aligned} \tag{132}$$

- then the operator $\hat{O}^\dagger$ takes $\langle v|$ to $\langle w|$:

$$\begin{aligned} \hat{O}^\dagger : \mathcal{H}^* &\rightarrow \mathcal{H}^* \\ \langle v| &\mapsto \langle w| = \langle v| \hat{O}^\dagger \end{aligned} \tag{133}$$

- Given an *orthonormal* basis $\mathcal{B} = \{|i\rangle : i = 1, 2, \dots\}$ of the Hilbert space $\mathcal{H}$, if $\hat{O}$ is given by

$$\hat{O} = \sum_{ij} |i\rangle O_{ij} \langle j|, \tag{134}$$

then $\hat{O}^\dagger$ should be given by

$$\hat{O}^\dagger = \sum_{ij} |i\rangle O_{ji}^* \langle j|. \tag{135}$$

**Exc
12**

Verify that Eq. (135) is consistent with the definition Eq. (133).

- In terms of *matrix* representation, the **Hermitian conjugate** acts as
 - **matrix transpose** (interchanges the rows and columns),
 - followed by **complex conjugation** of each matrix element.

$$\begin{pmatrix} O_{11} & O_{12} & \cdots \\ O_{21} & O_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}^\dagger = \begin{pmatrix} O_{11}^* & O_{21}^* & \cdots \\ O_{12}^* & O_{22}^* & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}. \tag{137}$$

How to think of it: Hermitian conjugate $\sim$ a generalization of *complex conjugate* from complex numbers to matrices.

Hermitian conjugate has the following properties:

- **Duality**: suppose $\hat{O}$ is an operator

$$\hat{O}^{\dagger\dagger} = \hat{O}. \tag{138}$$

- **Linearity**: suppose $\hat{O}$ and $\hat{P}$ are operators, α and β are complex numbers,

$$(\alpha \hat{O} + \beta \hat{P})^\dagger = \alpha^* \hat{O}^\dagger + \beta^* \hat{P}^\dagger. \tag{139}$$

- **Transpose Property:** suppose $\hat{O}$ and $\hat{P}$ are operators

$$(\hat{O} \hat{P})^\dagger = \hat{P}^\dagger \hat{O}^\dagger. \quad (140)$$

Exc
13

Prove the property Eq. (140).

■ Hermitian Operator

Real numbers play a special role in physics. The results of any measurements are real. If in quantum mechanics, physical observables are represented by *operators*, how do we impose the “real” condition on operators?

- A **real number** is a number whose *complex conjugation* is itself.

$$z = z^* \Leftrightarrow z \in \mathbb{R}. \quad (141)$$

- A ~~real operator~~ **Hermitian operator** is an linear operator whose *Hermitian conjugate* is itself.

An operator $\hat{O} = \sum_{ij} |i\rangle O_{ij} \langle j|$ is call **Hermitian**, if

$$\hat{O} = \hat{O}^\dagger, \quad (142)$$

or in terms of matrix elements,

$$O_{ij} = O_{ji}^*. \quad (143)$$

■ Eigensystem (General)

Given an operator $\hat{O}$, the **eigenvectors** $|O_k\rangle$ are a set of special vectors, on which the operator $\hat{O}$ acts as a *scalar* multiplication

$$\hat{O} |O_k\rangle = O_k |O_k\rangle, \quad (k = 1, 2, \dots) \quad (144)$$

and the corresponding scalars O_k are called the **eigenvalues** (of the corresponding eigenvectors).

- Eq. (144) is called the **eigen equation** of an operator $\hat{O}$.
- The eigenvalues can be found by solving the algebraic (polynomial) equation for O

$$\det(\hat{O} - O \mathbf{1}) = 0. \quad (145)$$

- For each solution of eigenvalue $O = O_k$, the corresponding eigenvector $|O_k\rangle$ is found by solving the linear equation

$$(\hat{O} - O_k \mathbf{1}) |O_k\rangle = 0. \quad (146)$$

- Use *Mathematica* to solve the eigen problem (recommended)

```
Eigensystem[{{0, 1}, {1, 0}}]
{{-1, 1}, {{-1, 1}, {1, 1}}}
```

■ Eigensystem (Hermitian Operators)

What is special about Hermitian operators?

Suppose $\hat{O} = \hat{O}^\dagger$ is a Hermitian operator and

$$\hat{O}|O_k\rangle = O_k|O_k\rangle, (k = 1, 2, \dots). \quad (147)$$

- **Eigenvalues** are **real**.

$$\hat{O} = \hat{O}^\dagger \Rightarrow O_k \in \mathbb{R}. \quad (148)$$

- **Eigenvectors** form a **complete** set of basis. (Any vector can be expanded as a sum of these eigenvectors.)

- Eigenvectors of *different* eigenvalues are *orthogonal* (automatically)

$$O_k \neq O_l \Rightarrow \langle O_k | O_l \rangle = 0. \quad (149)$$

- Eigenvectors of the *same* eigenvalue can be *made orthogonal* (by orthogonalization, e.g. Gram-Schmidt procedure).

```
Orthogonalize[{{1, 2}, {3, 4}}]
```

```
{{1/sqrt(5), 2/sqrt(5)}, {2/sqrt(5), -1/sqrt(5)}}
```

- For *bounded* Hermitian operators (e.g. finite matrices in finite dimensional Hilbert space), eigenvectors can be *normalized*.

Exc
14

Prove Eq. (148) and Eq. (149).

Therefore each **Hermitian operator** $\hat{O}$ generates a *complete* set of *orthonormal* basis $\{|O_k\rangle : k = 1, 2, \dots\}$ for the Hilbert space $\mathcal{H}$, also called the **eigenbasis** of $\hat{O}$.

- The completeness of the basis implies

$$\sum_k |O_k\rangle \langle O_k| = \mathbb{1}. \quad (151)$$

- Hermitian operator $\hat{O}$ can always be represented in its own eigenbasis, leading to the **spectral decomposition**

$$\hat{O} = \sum_k |O_k\rangle O_k \langle O_k|. \quad (152)$$

- Note: unlike a generic matrix representation $\hat{O} = \sum_{ij} |i\rangle O_{ij} \langle j|$, in the spectral decomposition Eq. (152), the summation only run through the eigenbasis once.
- In the eigenbasis, the Hermitian operator is represented as a **diagonal matrix**.

$$\hat{O} \simeq \begin{pmatrix} O_1 & & \\ & O_2 & \\ & & \ddots \end{pmatrix}. \quad (153)$$

So the procedure of bring the *matrix* representation to its *diagonal* form by transforming to its *eigenbasis* is called **diagonalization**. (We will discuss more about it later.)

Diagonalization is particularly useful in constructing the operator function. For example, the operator function $f(\hat{O})$ defined in Eq. (120) can be constructed by

$$f(\hat{O}) = \sum_k |O_k\rangle f(O_k) \langle O_k|, \quad (154)$$

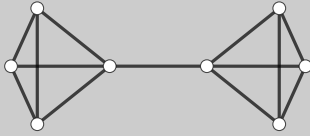
**Exc
15**

Prove Eq. (154).

or in the matrix form as

$$f(\hat{O}) \simeq \begin{pmatrix} f(O_1) & & \\ & f(O_2) & \\ & & \ddots \end{pmatrix}.$$

A particle can travel on a graph.



Let $|i\rangle$ denotes the state that the particle stays on the i th vertex of the graph. The following operator

$$\hat{H} = -\sum_{i \leftrightarrow j} (|i\rangle \langle j| + |j\rangle \langle i|)$$

describes the quantum process for the particle to tunnel from one vertex to the adjacent vertex (the summation sums over all links $i \leftrightarrow j$ on the graph).

- Represent the operator $\hat{H}$ as a matrix in the basis of $\{|i\rangle\}$.
- Write a computer program to compute the lowest and second lowest eigenvalues.
- Visualizing the corresponding eigen vectors by marking the vector components on the graph. What do you find?

[Comment: quantum mechanics can be applied to classify vertices on a graph --- an algorithm known as the **spectral clustering**.]

■ Eigensystem (Pauli Operators)

Example: Eigenvalues and eigenvectors of Pauli operators

Pauli matrices are 2×2 Hermitian matrices. Each one has two distinct eigenvalues, and two corresponding orthogonal eigenvectors.

operator	$\hat{\sigma}^x$	$\hat{\sigma}^y$	$\hat{\sigma}^z$
(matrix)	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

eigenvalue	+1	-1	+1	-1	+1	-1
eigenvector	$ +\rangle$	$ -\rangle$	$ i\rangle$	$ \bar{i}\rangle$	$ 0\rangle$	$ 1\rangle$
(vector)	$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$	$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$	$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}$	$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}$	$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 1 \end{pmatrix}$
projector	$ +\rangle\langle+ $	$ -\rangle\langle- $	$ i\rangle\langle i $	$ \bar{i}\rangle\langle\bar{i} $	$ 0\rangle\langle 0 $	$ 1\rangle\langle 1 $
(matrix)	$\frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$	$\frac{1}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$	$\frac{1}{2} \begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix}$	$\frac{1}{2} \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$

Spectral decompositions:

- Pauli- x

$$\hat{\sigma}^x = |+\rangle\langle+| - |-\rangle\langle-|, \quad (156)$$

with projection operators

$$\begin{aligned} |+\rangle\langle+| &= \frac{1 + \hat{\sigma}^x}{2}, \\ |-\rangle\langle-| &= \frac{1 - \hat{\sigma}^x}{2}. \end{aligned} \quad (157)$$

- Pauli- y

$$\hat{\sigma}^y = |i\rangle\langle i| - |\bar{i}\rangle\langle\bar{i}|, \quad (158)$$

with projection operators

$$\begin{aligned} |i\rangle\langle i| &= \frac{1 + \hat{\sigma}^y}{2}, \\ |\bar{i}\rangle\langle\bar{i}| &= \frac{1 - \hat{\sigma}^y}{2}. \end{aligned} \quad (159)$$

- Pauli- z

$$\hat{\sigma}^z = |0\rangle\langle 0| - |1\rangle\langle 1|, \quad (160)$$

with projection operators

$$\begin{aligned} |0\rangle\langle 0| &= \frac{1 + \hat{\sigma}^z}{2}, \\ |1\rangle\langle 1| &= \frac{1 - \hat{\sigma}^z}{2}. \end{aligned} \quad (161)$$

In general, the Pauli operator $\mathbf{n} \cdot \hat{\sigma}$ along the direction of the unit vector $\mathbf{n}$ has the following spectral decomposition

$$\mathbf{n} \cdot \hat{\sigma} = |\mathbf{n} \cdot \boldsymbol{\sigma} = +1\rangle\langle \mathbf{n} \cdot \boldsymbol{\sigma} = +1| - |\mathbf{n} \cdot \boldsymbol{\sigma} = -1\rangle\langle \mathbf{n} \cdot \boldsymbol{\sigma} = -1|, \quad (162)$$

with the projection operators

$$|n \cdot \sigma = \pm 1\rangle \langle n \cdot \sigma = \pm 1| = \frac{1 \pm n \cdot \hat{\sigma}}{2}. \quad (163)$$

Exc
16

Prove Eq. (162) and Eq. (163).

■ Observables

■ Physical Observable

Postulate 2 (Observables): **Physical observables** of a quantum system are described by **Hermitian operators** (represented as Hermitian matrices) acting on the associated Hilbert space.

Consider a Hermitian operator $\hat{O}$ with eigenvalues O_k and eigenvectors $|O_k\rangle$ ($m = 1, 2, \dots, g_k$), i.e.

$$\hat{O} = \sum_k |O_k\rangle O_k \langle O_k|. \quad (164)$$

The operator $\hat{O}$ corresponds to a physical observable O in the sense that

- All possible **measurement outcomes** (or **observation values**) of the observable O are given by (and only by) the *eigenvalues* O_k .
- The **measurement** projects (collapses) the quantum state to the eigenspace $\mathcal{H}_k$ spanned by the eigenstates of the corresponding measurement outcome O_k .

■ Measurement Postulate

Postulate 3 (Measurement): Given a quantum system in the **state** $|\psi\rangle$ and the **observable** O to be measured:

- (i) the **probability** to observe the measurement outcome O_k is $p(O_k | \psi) = |\langle O_k | \psi \rangle|^2$,
 - (ii) if O_k is observed, the state will **collapse** to $|O_k\rangle$.
-

- In quantum measurement, there is no way to tell for certain which outcome will be observed. There is only a **conditional probability** $p(O_k | \psi)$ that we can predict.
- Upon observing the measurement outcome O_k , the quantum state will be updated --- a process known as **quantum state collapse**.

$$|\psi\rangle \xrightarrow[\text{observe } O_k]{\text{measure } O} |O_k\rangle. \quad (165)$$

- $|\psi\rangle$ is called the **prior state** (pre-measurement state)
- $|O_k\rangle$ is called the **posterior state** (post-measurement state)

- **Bayesian view** of quantum state collapse:
 - The quantum state represents our *subjective* knowledge or belief about the system, not (necessarily) an *objective* physical reality.
 - Measurements provide new information that forces us to *update* our beliefs → the “collapse” happens in our knowledge.
 - The measurement postulate tells us how to update the quantum state given the observation, in a logically consistent manner.
- How to deal with **degeneracy**?

An eigenvalue O_k is n -fold degenerated $\Leftrightarrow$ there exists n orthonormal eigenstates (their choices are not unique) of $\hat{O}$ corresponding to the same eigenvalue:

$$\begin{aligned}
 \hat{O} |O_k, 1\rangle &= O_k |O_k, 1\rangle, \\
 \hat{O} |O_k, 2\rangle &= O_k |O_k, 2\rangle, \\
 &\dots \\
 \hat{O} |O_k, n\rangle &= O_k |O_k, n\rangle.
 \end{aligned} \tag{166}$$

Then if the measurement outcome O_k is observed in measuring O on state $|\psi\rangle$, how to compute $p(O_k | \psi)$ and the posterior state?

- Step I: Compute the scalar products $\alpha_m = \langle O_k, m | \psi \rangle$, meaning that

$$|\psi\rangle = \sum_{m=1}^n \alpha_m |O_k, m\rangle + \dots \text{(other states)}. \tag{167}$$

- Step II: Aggregate the probability:

$$p(O_k | \psi) = \sum_{m=1}^n |\alpha_m|^2 = \sum_{m=1}^n |\langle O_k, m | \psi \rangle|^2. \tag{168}$$

- Step III: Renormalize the amplitudes α_m

$$\tilde{\alpha}_m = \frac{\alpha_m}{\sqrt{p(O_k | \psi)}} = \frac{\langle O_k, m | \psi \rangle}{\sqrt{p(O_k | \psi)}}, \tag{169}$$

and reconstruct the posterior state

$$|\psi\rangle \xrightarrow[\text{observe } O_k]{\text{measure } O} |\psi'\rangle = \sum_{m=1}^n \tilde{\alpha}_m |O_k, m\rangle. \tag{170}$$

Note: it is always a good practice to normalize the state (i.e. ensuring $\langle \psi' | \psi' \rangle = 1$) after quantum state collapse.

HW
4

Let $\{|1\rangle, |2\rangle, |3\rangle\}$ be a set of orthonormal basis of a three-state system. Suppose the system is in the prior state $|\psi\rangle = \frac{1}{\sqrt{3}} (|1\rangle + |2\rangle + |3\rangle)$.

Consider measuring the observable $\hat{O} = |1\rangle\langle 2| + |2\rangle\langle 1| - |3\rangle\langle 3|$.

- (i) What are the possible measurement outcomes (observation values)?
- (ii) What are the probabilities to observe each outcome?
- (iii) What posterior states will the system collapse to after observing each outcome?

■ Expectation Value

The **expectation value** of an observable O , denoted as $\langle O \rangle$, is the *averaged* measurement outcome of O over many repeated experiments (with the same prior state $|\psi\rangle$ prepared each time).

According to the measurement postulate

$$\begin{aligned}\langle O \rangle &:= \sum_k O_k p(O_k | \psi) \\ &= \sum_k O_k |\langle O_k | \psi \rangle|^2 \\ &= \sum_k \langle \psi | O_k \rangle O_k \langle O_k | \psi \rangle\end{aligned}\tag{171}$$

Given $\hat{O} = \sum_k |O_k\rangle O_k \langle O_k|$, we conclude

$\langle O \rangle = \langle \psi | \hat{O} | \psi \rangle.$

(172)

- The answer is a *real* scalar (as $\hat{O}$ is Hermitian).
- Represented as *vectors* and *matrices*,

$$\langle O \rangle = (\psi_1^* \ \psi_2^* \ \cdots) \begin{pmatrix} O_{11} & O_{12} & \cdots \\ O_{21} & O_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \end{pmatrix}.\tag{173}$$

Alternatively, the expectation value can also be written as a trace of the product of the observable operator $\hat{O}$ and the state projector $|\psi\rangle\langle\psi|$

$$\langle O \rangle = \text{Tr } \hat{O} |\psi\rangle\langle\psi|.\tag{174}$$

- The advantage of this approach is to circumvent solving for $|\psi\rangle$ explicitly (sometimes the state projector is easier to construct than the state vector).

HW
5

Let $\mathbf{m}$ and $\mathbf{n}$ be three-component real unit vectors. For a qubit, consider measuring $\mathbf{n} \cdot \boldsymbol{\sigma}$ on the $|\mathbf{m} \cdot \boldsymbol{\sigma} = +1\rangle$ state.

- (i) What is the probability to observe $\mathbf{n} \cdot \boldsymbol{\sigma} = +1$?
 (ii) What is the expectation value of the operator $\mathbf{n} \cdot \hat{\boldsymbol{\sigma}}$ on the state $|\mathbf{m} \cdot \boldsymbol{\sigma} = +1\rangle$?
 [Express your results in terms of $\mathbf{m}$ and $\mathbf{n}$. Hint: using Eq. (174) and Eq. (163) can simplify the calculation.]

■ Variance

The **variance** of an observable O on a state $|\psi\rangle$ is defined as

$$\text{var } O = \langle (O - \langle O \rangle)^2 \rangle = \langle O^2 \rangle - \langle O \rangle^2. \quad (175)$$

where $\langle O^2 \rangle = \langle \psi | \hat{O}^2 | \psi \rangle$ and $\langle O \rangle = \langle \psi | \hat{O} | \psi \rangle$. The square root of the variance defines the **standard deviation**:

$$\text{std } O = \sqrt{\text{var } O}. \quad (176)$$

Uncertainty Relation: for any pair of *observables* A and B measured on any given *state* (repeatedly),

$$(\text{std } A)(\text{std } B) \geq \frac{1}{2} |\langle [A, B] \rangle|. \quad (177)$$

Exc
17

Prove Eq. (177).

- In words, the product of the *uncertainties* cannot be smaller than half of the magnitude of the expectation value of the *commutator*.
- For *commuting* observables ($[A, B] = 0$), $(\text{std } A)(\text{std } B) \geq 0$, it is possible to have $\text{std } A = \text{std } B = 0$ simultaneously, i.e. A and B can be jointly measured with perfect certainty.
- For *non-commuting* observables, there exists a state on which $|\langle [A, B] \rangle| \neq 0$. Then on such state, it is impossible to have $\text{std } A = \text{std } B = 0$ simultaneously, i.e. A and B can not be jointly measured with certainty.

Dynamics

■ Unitary Operators

■ Basis Transformation

Suppose we have two sets of orthonormal basis of the same Hilbert space $\mathcal{H}$

$$\begin{aligned}\mathcal{B} &= \{|i\rangle : i = 1, 2, \dots, \dim \mathcal{H}\}, \\ \mathcal{B}' &= \{|i'\rangle : i = 1, 2, \dots, \dim \mathcal{H}\}.\end{aligned}\tag{178}$$

For example, the eigen basis of $\hat{\sigma}^x$ v.s. that of $\hat{\sigma}^z$.

- The *same state* $|v\rangle$ can have *different vector* representations in different bases

$$v_i = \langle i | v \rangle, \quad v'_i = \langle i' | v \rangle.\tag{179}$$

- The *same operator* $\hat{O}$ can have *different matrix* representations in different bases

$$O_{ij} = \langle i | \hat{O} | j \rangle, \quad O'_{ij} = \langle i' | \hat{O} | j' \rangle.\tag{180}$$

How are representations in different bases related? - **Basis transformation.** Basis transformation from $\mathcal{B}$ to $\mathcal{B}'$ is describe by a matrix U with the matrix element

$$U_{ij} = \langle i' | j \rangle.\tag{181}$$

such that the representation in the new basis is related to that in the old basis by

$$\begin{aligned}v'_i &= \sum_j U_{ij} v_j, \\ O'_{ij} &= \sum_{kl} U_{ik} O_{kl} U_{jl}^*.\end{aligned}\tag{182}$$

Exc 18

Using Eq. (181) to prove that Eq. (182) is compatible with Eq. (179) and Eq. (180).

In quantum mechanics, every operator is a matrix, and every matrix is an operator. So does the basis transformation matrix.

$$\hat{U} = \sum_i |i\rangle \langle i'|.\tag{183}$$

Exc 19

Check that the matrix element of $\hat{U}$ in Eq. (183) is indeed given by Eq. (181), regardless of represented in the basis $\mathcal{B}$ or $\mathcal{B}'$.

$\hat{U}$ in Eq. (183) is an example of the **unitary operator**.

A operator $\hat{U}$ is **unitary**, iff

$$\hat{U}^\dagger \hat{U} = \hat{U} \hat{U}^\dagger = \mathbb{1}.\tag{184}$$

Exc 20

Check that Eq. (183) satisfies the defining property Eq. (184) for unitary operator.

- The *inverse* of a unitary operator is its *Hermitian conjugate*

$$\hat{U}^{-1} = \hat{U}^\dagger.\tag{185}$$

The operator (basis transformation) implemented by $\hat{U}$ is *reversed* by that of $\hat{U}^\dagger$, and vice versa.

- When the two sets of basis $|i\rangle$ and $|i'\rangle$ are identical, $U = \mathbb{1}$ becomes the identity operator (which is also unitary).

In terms of the unitary operator, the basis transformation Eq. (182) can be written as

$$\begin{aligned} \text{for ket state: } |v\rangle &\rightarrow \hat{U} |v\rangle, \\ \text{for bra state: } \langle v| &\rightarrow \langle v| \hat{U}^\dagger, \\ \text{for operator: } \hat{O} &\rightarrow \hat{U} \hat{O} \hat{U}^\dagger. \end{aligned} \tag{186}$$

- The operator $\hat{O}$ is also made of ket and bra states, so the unitary operator must be applied from both sides, when transforming an operator.
- The *expectation value* of an observable is *invariant* under *basis transformation*. (Physical reality should be *basis-independent*.)

$$\langle O \rangle = \langle \psi | \hat{O} | \psi \rangle \rightarrow \langle \psi | \hat{U}^\dagger \hat{U} \hat{O} \hat{U}^\dagger \hat{U} | \psi \rangle = \langle \psi | \mathbb{1} \hat{O} \mathbb{1} | \psi \rangle = \langle O \rangle. \tag{187}$$

■ Matrix Diagonalization

Diagonalization of a *Hermitian operator*: find a unitary operator $\hat{U}$ to bring the Hermitian operator $\hat{O}$ to *diagonal form* by transforming to its *eigenbasis*.

$$\begin{aligned} \hat{O} &= \sum_k |O_k\rangle O_k \langle O_k|, \\ \hat{U} &= \sum_k |k\rangle \langle O_k|, \end{aligned} \tag{188}$$

such that under $\hat{O} \rightarrow \hat{U} \hat{O} \hat{U}^\dagger$,

$$\hat{\Lambda} = \hat{U} \hat{O} \hat{U}^\dagger = \sum_k |k\rangle O_k \langle k| = \begin{pmatrix} O_1 & & \\ & O_2 & \\ & & \ddots \end{pmatrix} \tag{189}$$

is diagonal in the basis of **one-hot vectors** $|k\rangle$.

- Every *Hermitian* matrix can be written as

$$\hat{O} = \hat{U}^\dagger \hat{\Lambda} \hat{U}, \tag{190}$$

with $\hat{\Lambda}$ being *diagonal* and $\hat{U}$ being *unitary*.

- Or equivalently, the *unitary* transformation $\hat{U}$ brings the *Hermitian* matrix to its *diagonal* form,

$$\hat{U} \hat{O} \hat{U}^\dagger = \hat{\Lambda}. \tag{191}$$

Example: diagonalization of Pauli matrix

The Pauli matrix $\hat{\sigma}^x$ can be diagonalized by the following unitary transformation (whose row

vectors are bra eigenvectors of $\hat{\sigma}^x$)

$$\hat{U}_H = \begin{pmatrix} \langle + | \\ \langle - | \end{pmatrix} \simeq \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \quad (192)$$

- This unitary operation $\hat{U}_H$ is also known as the **Hadamard gate** in quantum information, an example of single-qubit gate.
- Under the unitary transformation, $\hat{\sigma}^x$ is brought to its diagonal form, which is $\hat{\sigma}^z$

$$\begin{aligned} \hat{U}_H \hat{\sigma}^x \hat{U}_H^\dagger &\simeq \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \simeq \hat{\sigma}^z. \end{aligned} \quad (193)$$

■ Hermitian Generators

If **Hermitian operators** are generalization of **real numbers**, then **unitary operators** are generalization of **phase factors**.

- A complex number $z \in \mathbb{C}$ is a phase factor, iff $|z| = 1$. Any phase factor can be written as $z = e^{i\theta}$, where $\theta \in \mathbb{R}$ is a *real* phase angle.

$$z^* z = z z^* = |z|^2 = 1 \Leftrightarrow z = e^{i\theta} \quad (194)$$

- Similar ideas apply to unitary operators: every **unitary operator** can be **generated** by a **Hermitian operator** $\hat{\Theta}$ in the form of

$$\hat{U} = e^{i\hat{\Theta}}. \quad (195)$$

Given a Hermitian operator $\hat{\Theta}$

$$\hat{\Theta} = \sum_k |\Theta_k\rangle \Theta_k \langle \Theta_k|, \quad (196)$$

by $e^{i\hat{\Theta}}$ we mean

- either by operator Taylor expansion (recall Eq. (122) on operator exponential)

$$e^{i\hat{\Theta}} = \mathbb{1} + i\hat{\Theta} + \frac{(i\hat{\Theta})^2}{2!} + \frac{(i\hat{\Theta})^3}{3!} + \dots \quad (197)$$

- or by spectral decomposition (HW 2)

$$e^{i\hat{\Theta}} = \sum_k |\Theta_k\rangle e^{i\Theta_k} \langle \Theta_k| \quad (198)$$

Don't do element-wise exponentiation on the matrix!

**Exc
21**

Use Eq. (198) to show that $\hat{U} = e^{i\hat{\Theta}}$ is unitary as long as $\hat{\Theta}$ is Hermitian.

Example: unitary generated by Pauli matrix. Recall $\hat{U}(\theta) = e^{i\theta\hat{\sigma}^y}$ in (Exc 10).

$$\hat{U}(\theta) = e^{i\theta\hat{\sigma}^y} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}. \quad (199)$$

It implements a **basis rotation** with θ being the **rotation angle**:

$$\hat{U}(\theta) |0\rangle = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos \theta \\ -\sin \theta \end{pmatrix}. \quad (200)$$

Special case: when $\theta = 0$, $\hat{U}(0) = \mathbb{1} \Rightarrow$ no rotation is performed.

More generally, let $\hat{U}(\theta)$ be the **unitary operator** that implements certain *basis rotation* by a real **angle** θ . When $\theta = \Delta\theta$ is **small**, we can Taylor expand

$$\hat{U}(\Delta\theta) = \hat{U}(0) + \hat{U}'(0) \Delta\theta + \dots = \mathbb{1} + \hat{U}'(0) \Delta\theta + \dots, \quad (201)$$

where $\hat{U}'(0)$ is $\partial_\theta \hat{U}(\theta)$ evaluated at $\theta = 0$.

- $\hat{U}'(0)$ is also an operator (matrix), usually denoted as $\hat{U}'(0) = i \hat{G}$. We call $\hat{G}$ the **generator** of the rotation/unitary operator, because it *generates* an **infinitesimal rotation**

$$\hat{U}(\Delta\theta) = \mathbb{1} + i \Delta\theta \hat{G} + \dots \quad (202)$$

- $\hat{U}(\Delta\theta)$ is **unitary** $\Rightarrow \hat{G}$ is **Hermitian**.

$$\begin{aligned} & \hat{U}(\Delta\theta)^\dagger \hat{U}(\Delta\theta) \\ &= (\mathbb{1} - i \Delta\theta \hat{G}^\dagger + \dots) (\mathbb{1} + i \Delta\theta \hat{G} + \dots) \\ &= \mathbb{1} + i \Delta\theta (\hat{G} - \hat{G}^\dagger) + \dots = \mathbb{1}. \end{aligned} \quad (203)$$

- *Large* rotations can be *accumulated* from *small* rotations.

$$\hat{U}(N \Delta\theta) = \hat{U}(\Delta\theta)^N = (\mathbb{1} + i \Delta\theta \hat{G})^N. \quad (204)$$

As $\Delta\theta$ is small (but N can be large, s.t. $\theta = N \Delta\theta$ is finite),

$$\ln \hat{U}(N \Delta\theta) = N \ln(\mathbb{1} + i \Delta\theta \hat{G}) = i N \Delta\theta \hat{G}, \quad (205)$$

So $\hat{U}(N \Delta\theta) = e^{i N \Delta\theta \hat{G}}$, we obtain the *exponential* form

$$\hat{U}(\theta) = e^{i\theta\hat{G}}.$$

(206)

Conclusion: every *Hermitian* operator $\hat{G} = \theta \hat{G}$ generates a *unitary* operator $e^{i\hat{G}}$ by the exponential map.

■ Time Evolution

■ Time-Evolution is Unitary

Unitarity: *information* is never lost!

Basic assumption: quantum *information* is preserved under quantum *dynamics*, i.e. two *identical* and isolated systems

- start out in **different** states $\Rightarrow$ **remains** in **different** states (towards both future and past).
- start out in the **same** state $\Rightarrow$ follow **identical evolution** (towards both future and past).

Although **measurement** seems to be **non-deterministic**, evolution of quantum **state** is **deterministic**: suppose you know the *state* at one time, then the quantum *equation of motion* tell you what it will be later.

$$|\psi(t)\rangle = \hat{U}(t) |\psi(0)\rangle, \quad (207)$$

$|\psi(0)\rangle$ is the initial state, and $|\psi(t)\rangle$ is the state at time t . $\hat{U}(t)$ is the **time-evolution operator** that takes $|\psi(0)\rangle$ to $|\psi(t)\rangle$. $\P$ We will show that $\hat{U}(t)$ should be *unitary*.

- *Distinct* states remain *distinct*:

$$\langle\phi(0)|\psi(0)\rangle = 0 \Rightarrow \langle\phi(t)|\psi(t)\rangle = \langle\phi(0)|\hat{U}(t)^\dagger \hat{U}(t)|\psi(0)\rangle = 0. \quad (208)$$

- *Identical* states remain the *identical*:

$$\langle\psi(0)|\psi(0)\rangle = 1 \Rightarrow \langle\psi(t)|\psi(t)\rangle = \langle\psi(0)|\hat{U}(t)^\dagger \hat{U}(t)|\psi(0)\rangle = 1. \quad (209)$$

Or, the fact that the normalization condition must be preserved.

Treat $|\psi(0)\rangle$ and $|\phi(0)\rangle$ as members of any orthonormal basis, then Eq. (208) and Eq. (209) implies

$$\langle i|\hat{U}(t)^\dagger \hat{U}(t)|j\rangle = \delta_{ij} \Rightarrow \hat{U}(t)^\dagger \hat{U}(t) = \mathbf{1}. \quad (210)$$

Therefore, the **time-evolution** operator $\hat{U}(t)$ is **unitary**.

■ Hamiltonian

Hamiltonian *generates* time-evolution!

As a *unitary* operator, the *time-evolution* operator is also *generated* by a *Hermitian* operator, called the **Hamiltonian**,

$$\hat{H} = i \hat{U}'(0) = i \partial_t \hat{U}(t)|_{t=0}. \quad (211)$$

For small Δt , *infinitesimal* evolution is given by

$$\hat{U}(\Delta t) = \mathbf{1} - i \hat{H} \Delta t + \dots, \quad (212)$$

therefore the state evolves as

$$|\psi(\Delta t)\rangle = \hat{U}(\Delta t) |\psi(0)\rangle = |\psi(0)\rangle - i \Delta t \hat{H} |\psi(0)\rangle, \quad (213)$$

meaning that

$$i \partial_t |\psi(0)\rangle = i \frac{|\psi(\Delta t)\rangle - |\psi(0)\rangle}{\Delta t} = \hat{H} |\psi(0)\rangle. \quad (214)$$

There is nothing special about $t = 0$. Eq. (214) should hold at any time.

$$i \partial_t |\psi(t)\rangle = \hat{H} |\psi(t)\rangle. \quad (215)$$

This is the **Schrödinger equation**, the *equation of motion* for the quantum state.

- The Hamiltonian $\hat{H}(t) = i \hat{U}'(t)$ can be **time-dependent** in general.
- But in many cases, we consider $\hat{H}$ to be **time-independent**, by assuming the **time-translation symmetry**.

What happens to Planck's constant?

$$\hbar = \frac{h}{2\pi} = 1.0545718(13) \times 10^{-34} \text{ J s}. \quad (216)$$

In quantum mechanics, the *observable* associated with the **Hamiltonian** is the **energy**. To balance the *dimensionality* across the Schrödinger equation, *Planck's constant* is inserted for Eq. (215):

$$i \hbar \partial_t |\psi(t)\rangle = \hat{H} |\psi(t)\rangle. \quad (217)$$

Why is $\hbar$ so small? Well, the answer has more to do with biology than with physics $\Rightarrow$ Why we are so big, heavy and slow? A natural choice for quantum mechanics is to set the units such that $\hbar = 1$. It is a common practice in theoretical physics (we will also use this convention sometimes).

■ Schrödinger Equation: State Dynamics

Postulate 4 (Dynamics): The **time-evolution** of the *state* of a quantum system is governed by the **Hamiltonian** of the system, according to the time-dependent **Schrödinger equation**.

$$i \hbar \partial_t |\psi(t)\rangle = \hat{H} |\psi(t)\rangle. \quad (218)$$

If the Hamiltonian $\hat{H}$ is **time-independent**, we can first find its eigenvalues (or **eigen energies**) and eigenvectors (or **energy eigenstates**).

$$\hat{H} |E_k\rangle = E_k |E_k\rangle. \quad (219)$$

This is also called the *time-independent Schrödinger equation*. Without solving a *differential equation*, we just need to *diagonalize* a *Hermitian matrix* in this case.

Each *energy eigenstate* will evolve in time simply by a *rotating overall phase*,

$$|E_k(t)\rangle = e^{-\frac{i}{\hbar} E_k t} |E_k\rangle. \quad (220)$$

- $|E_k\rangle$ form a complete set of orthonormal basis, called **energy eigenbasis**.

Exc 22 | Verify that Eq. (220) is a solution of Eq. (218):

Any initial state $|\psi(0)\rangle$ will evolve in time by first *representing* the initial state in the *energy eigenbasis*, and attaching to each energy eigenstate by its rotating overall phase,

$$\begin{aligned} |\psi(t)\rangle &= \sum_i e^{-\frac{i}{\hbar} E_i t} |E_i\rangle \langle E_i | \psi(0)\rangle \\ &= e^{-\frac{i}{\hbar} \hat{H} t} |\psi(0)\rangle. \end{aligned} \quad (221)$$

A *time-independent* Hamiltonian generates the time-evolution via *matrix exponentiation*

$$\hat{U}(t) = \exp\left(-\frac{i}{\hbar} \hat{H} t\right). \quad (222)$$

However, for *time-dependent* Hamiltonian, there no such a clean formula. Evolution must be carried out step by step, denoted as a *time-ordered* exponential

$$\hat{U}(t) = \mathcal{T} \exp\left(-\frac{i}{\hbar} \int_0^t \hat{H}(t') dt'\right). \quad (223)$$

□ Larmor Precession and Rabi Oscillation

How to write down a Hamiltonian?

- derive it from experiment,
- borrow it from some theory we like,
- pick one and see what happens. ☞

Hamiltonian must be *Hermitian* anyway. For a single spin (qubit), the most general Hamiltonian takes the form of

$$\begin{aligned} \hat{H} &= h_0 \mathbf{1} + h_x \hat{\sigma}^x + h_y \hat{\sigma}^y + h_z \hat{\sigma}^z \\ &= h_0 \mathbf{1} + \mathbf{h} \cdot \hat{\boldsymbol{\sigma}}, \end{aligned} \quad (224)$$

where $h_0, h_x, h_y, h_z \in \mathbb{R}$ are all *real* coefficients.

- The time-evolution operator (set $\hbar = 1$ in the following)

$$\begin{aligned} \hat{U}(t) &= e^{-i \hat{H} t} \\ &= e^{-i h_0 t} \left(\cos(|\mathbf{h}| t) \mathbf{1} - i \sin(|\mathbf{h}| t) \tilde{\mathbf{h}} \cdot \hat{\boldsymbol{\sigma}} \right), \end{aligned} \quad (225)$$

where $|\mathbf{h}| = \sqrt{\mathbf{h} \cdot \mathbf{h}}$ and $\tilde{\mathbf{h}} = \mathbf{h} / |\mathbf{h}|$.

Exc 23 | Derive Eq. (225) from Eq. (224).

- A state $|\psi(0)\rangle$ will evolve with time following

$$\begin{aligned} |\psi(t)\rangle &= \hat{U}(t) |\psi(0)\rangle \\ &= e^{-i h_0 t} (\cos(|\mathbf{h}| t) \mathbf{1} - i \sin(|\mathbf{h}| t) \tilde{\mathbf{h}} \cdot \hat{\boldsymbol{\sigma}}) |\psi(0)\rangle. \end{aligned} \quad (226)$$

- If we measure $\boldsymbol{\sigma}$ on the state $|\psi(t)\rangle$, the expectation value will be given by

$$\begin{aligned} \langle \boldsymbol{\sigma} \rangle_t &= \langle \psi(t) | \hat{\boldsymbol{\sigma}} | \psi(t) \rangle \\ &= \cos(2 |\mathbf{h}| t) \langle \boldsymbol{\sigma} \rangle_0 + \sin(2 |\mathbf{h}| t) \tilde{\mathbf{h}} \times \langle \boldsymbol{\sigma} \rangle_0 + (1 - \cos(2 |\mathbf{h}| t)) \tilde{\mathbf{h}} (\tilde{\mathbf{h}} \cdot \langle \boldsymbol{\sigma} \rangle_0). \end{aligned} \quad (227)$$

which also evolves with time.

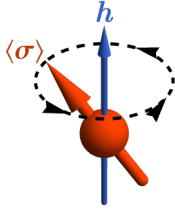
Exc 24 Derive Eq. (227) from Eq. (226).
Hint: Eq. (109) can make life much more easier.

Larmor precession: assume $\mathbf{h} = (0, 0, h_z)$ along the z -direction, and parameterize the expectation of the spin vector by $\langle \boldsymbol{\sigma} \rangle = (\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta)$.

$$\langle \boldsymbol{\sigma} \rangle_t = (\sin \theta_0 \cos (\varphi_0 + 2 h_z t), \sin \theta_0 \sin (\varphi_0 + 2 h_z t), \cos \theta_0), \quad (228)$$

where θ_0 and φ_0 are the initial azimuthal and polar angles.

- The solution describes the *spin* $\langle \boldsymbol{\sigma} \rangle$ *precessing* around the axis of the *Zeeman field* $\mathbf{h}$.



- The precession frequency $\omega = 2 |\mathbf{h}|$ is called the **Larmor frequency**. It can be used to probe the local Zeeman field strength, which has applications in nuclear magnetic resonance (NMR) and nitrogen-vacancy (NV) center.
- *Energy* of a spin in the Zeeman field is $\langle H \rangle = -\mathbf{h} \cdot \langle \boldsymbol{\sigma} \rangle$ (up to some constant energy shift h_0).

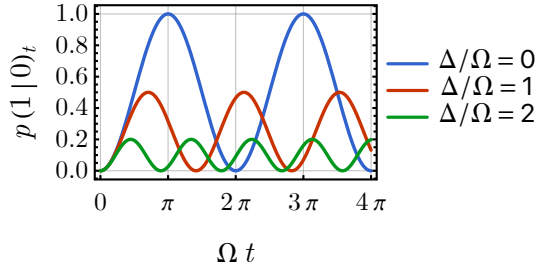
Rabi oscillation: a qubit initially prepared in state $|0\rangle$, evolved under the Hamiltonian

$$\hat{H} = \Omega \hat{\sigma}^x + \Delta \hat{\sigma}^z \simeq \begin{pmatrix} \Delta & \Omega \\ \Omega & -\Delta \end{pmatrix}, \quad (229)$$

where Ω is the *driving field* and Δ is called *detuning*. The probability to find the qubit in state $|1\rangle$ at time t is given by

$$p(1 | 0)_t = \frac{1 - \langle \sigma^z \rangle_t}{2} = \frac{\sin^2(\omega t / 2)}{1 + (\Delta / \Omega)^2}, \quad (230)$$

with the **Rabi frequency** $\omega = 2 \sqrt{\Omega^2 + \Delta^2}$.



- **Rabi π -Pulse:** flipping $|0\rangle$ to $|1\rangle$ (and vice versa) by a π -pulse (turn on the driving field Ω for time $t = \pi/\Omega$ and turn off) at resonance $\Delta = 0$. This implements a **NOT** gate (or X gate) on a single qubit.

■ Heisenberg Equation: Operator Dynamics

Two *pictures* of the **quantum dynamics**:

- **Schrödinger picture:** state evolves in time, operator remains fixed,

$$\langle O(t) \rangle = \langle \psi(t) | \hat{O} | \psi(t) \rangle. \quad (231)$$

- **Heisenberg picture:** operator evolves in time, state remains fixed,

$$\langle O(t) \rangle = \langle \psi | \hat{O}(t) | \psi \rangle. \quad (232)$$

The two pictures are consistent, if

$$|\psi(t)\rangle = \hat{U}(t) |\psi\rangle \Leftrightarrow \hat{O}(t) = \hat{U}(t)^\dagger \hat{O} \hat{U}(t), \quad (233)$$

such that Eq. (231) and Eq. (232) are consistent, as they both implies

$$\langle O(t) \rangle = \langle \psi | \hat{U}(t)^\dagger \hat{O} \hat{U}(t) | \psi \rangle. \quad (234)$$

Note: one should only apply one picture at a time, i.e. either the state or the operator is time-dependent, **but not both**.

In the *Heisenberg picture*, the time-evolution of an operator

$$\hat{O}(t) = \hat{U}(t)^\dagger \hat{O} \hat{U}(t), \quad (235)$$

described by the **Heisenberg equation**

$$i \hbar \partial_t \hat{O}(t) = [\hat{O}(t), \hat{H}]. \quad (236)$$

**Exc
25**

Derive Eq. (236) from Eq. (235).

Correspondingly, its expectation value evolves as

$$i \hbar \partial_t \langle O(t) \rangle = \langle [\hat{O}(t), \hat{H}] \rangle. \quad (237)$$

If $[\hat{O}, \hat{H}] = 0$, the *Heisenberg equation* Eq. (236) implies that $\partial_t \langle O \rangle = 0$, i.e. O will be invariant

in time. Conversely, the observable O is a **conserved quantity** (or an **integral of motion**) if $\hat{O}$ commutes with the Hamiltonian $\hat{H}$.

Use the Heisenberg equation to solve Rabi oscillation. Set $\hbar = 1$ for this problem. The Hamiltonian is given by

$$\hat{H} = \Omega \hat{\sigma}^x + \Delta \hat{\sigma}^z.$$

(i) Show that in the Heisenberg picture, the operator $\hat{\sigma}$ evolves by the following coupled linear differential equation

**HW
6**
$$\partial_t \begin{pmatrix} \hat{\sigma}^x \\ \hat{\sigma}^y \\ \hat{\sigma}^z \end{pmatrix} = 2 \begin{pmatrix} 0 & -\Delta & 0 \\ \Delta & 0 & -\Omega \\ 0 & \Omega & 0 \end{pmatrix} \begin{pmatrix} \hat{\sigma}^x \\ \hat{\sigma}^y \\ \hat{\sigma}^z \end{pmatrix}.$$

(ii) Starting with the initial condition $\langle \sigma(0) \rangle = (0, 0, 1)$ (i.e. the qubit was initially in the $|0\rangle$ state), find the time evolution of $\langle \sigma(t) \rangle$ by solving the differential equation derived in (i). [Hint: linear differential equation can be simply solved by matrix exponential.]

(iii) Compute the probability to find the qubit in state $|1\rangle$ at time t , which is given by $p(1|0)_t = \frac{1}{2} (1 - \langle \sigma^z(t) \rangle)$.