Chapter 0: Brief definitions and basic concepts

As the future version of me might forgot everything we have over the summer, as I did for now, I will make a review again from the simple definition to recall the necessary information to tell you why we are here and how we are going to proceed.

This section serve as reference for definitions, notations, and theorems that we will use later. This section can be safely ignored if you are already familiar with the definitions and theorems.

But for the future self who might have no idea what I'm talking about, we will provided detailed definitions to you to understand the concepts.

0.1 Complex vector spaces

The main vector space we are interested in is $\mathbb{C}^n$; therefore, all the linear operators we defined are from $\mathbb{C}^n$ to $\mathbb{C}^n$.

Definition 1. *We denote a vector in vector space as* $|\psi\rangle = (z_1, \ldots, z_n)$ *(might also be infinite dimensional, and* $z_i \in \mathbb{C}$*).*

Here ψ is just a label for the vector, and you don't need to worry about it too much. This is also called the ket, where the counterpart $\langle\psi|$ is called the bra, used to denote the vector dual to ψ; such an element is a linear functional if you really want to know what that is.

Few additional notation will be introduced, in this document, we will follows the notation used in mathematics literature [Axl23]

- $\langle\psi|\varphi\rangle$ is the inner product between two vectors, and $\langle\psi| A |\varphi\rangle$ is the inner product between $A|\varphi\rangle$ and $\langle\psi|$, or equivalently $A^\dagger \langle\psi|$ and $|\varphi\rangle$.
- Given a complex matrix $A = \mathbb{C}^{n\times n}$,
    1. $\overline{A}$ is the complex conjugate of A.

Example

$$A = \begin{bmatrix} 1+i & 2+i & 3+i \\ 4+i & 5+i & 6+i \\ 7+i & 8+i & 9+i \end{bmatrix}, \overline{A} = \begin{bmatrix} 1-i & 2-i & 3-i \\ 4-i & 5-i & 6-i \\ 7-i & 8-i & 9-i \end{bmatrix}$$

2. $A^\top$ denotes the transpose of A.

Example

$$A = \begin{bmatrix} 1+i & 2+i & 3+i \\ 4+i & 5+i & 6+i \\ 7+i & 8+i & 9+i \end{bmatrix}, A^\top = \begin{bmatrix} 1+i & 4+i & 7+i \\ 2+i & 5+i & 8+i \\ 3+i & 6+i & 9+i \end{bmatrix}$$

3. $A^* = \overline{(A^\top)}$ denotes the complex conjugate transpose, referred to as the adjoint, or Hermitian conjugate of A.

Example

$$A = \begin{bmatrix} 1+i & 2+i & 3+i \\ 4+i & 5+i & 6+i \\ 7+i & 8+i & 9+i \end{bmatrix}, A^* = \begin{bmatrix} 1-i & 4-i & 7-i \\ 2-i & 5-i & 8-i \\ 3-i & 6-i & 9-i \end{bmatrix}$$

4. A is unitary if $A^*A = AA^* = I$.
5. A is self-adjoint (hermitian in physics literature) if $A^* = A$.

Motivation of Tensor product

Recall from the traditional notation of product space of two vector spaces V and W, that is, $V \times W$, is the set of all ordered pairs $(|v\rangle, |w\rangle)$ where $|v\rangle \in V$ and $|w\rangle \in W$.

The space has dimension $\dim V + \dim W$.

We want to define a vector space with the notation of multiplication of two vectors from different vector spaces.

That is

$$(|v_1\rangle + |v_2\rangle) \otimes |w\rangle = (|v_1\rangle \otimes |w\rangle) + (|v_2\rangle \otimes |w\rangle)$$
$$|v\rangle \otimes (|w_1\rangle + |w_2\rangle) = (|v\rangle \otimes |w_1\rangle) + (|v\rangle \otimes |w_2\rangle)$$

and enables scalar multiplication by

$$\lambda(|v\rangle \otimes |w\rangle) = (\lambda |v\rangle) \otimes |w\rangle = |v\rangle \otimes (\lambda |w\rangle)$$

And we wish to build a way to associate the basis of V and W with the basis of $V \otimes W$. That makes the tensor product a vector space with dimension $\dim V \times \dim W$.

Definition 2. *Definition of linear functional*

A linear functional is a linear map from V to $\mathbb{F}$.

Note the difference between a linear functional and a linear map.

A generalized linear map is a function $f : V \to W$ satisfying the condition.

- $f(|u\rangle + |v\rangle) = f(|u\rangle) + f(|v\rangle)$
- $f(\lambda |v\rangle) = \lambda f(|v\rangle)$

Definition 3. *A bilinear functional is a bilinear function $\beta : V \times W \to \mathbb{F}$ satisfying the condition that $|v\rangle \to \beta(|v\rangle, |w\rangle)$ is a linear functional for all $|w\rangle \in W$ and $|w\rangle \to \beta(|v\rangle, |w\rangle)$ is a linear functional for all $|v\rangle \in V$.*

The vector space of all bilinear functionals is denoted by $\mathcal{B}(V, W)$.

Definition 4. *Let V, W be two vector spaces.*

Let V' and W' be the dual spaces of V and W, respectively, that is $V' = \{\psi : V \to \mathbb{F}\}$ and $W' = \{\phi : W \to \mathbb{F}\}$, ψ, ϕ are linear functionals.

The tensor product of vectors $v \in V$ and $w \in W$ is the bilinear functional defined by $\forall(\psi, \phi) \in V' \times W'$ given by the notation

$$(v \otimes w)(\psi, \phi) = \psi(v)\phi(w)$$

The tensor product of two vector spaces V and W is the vector space $\mathcal{B}(V', W')$

Notice that the basis of such vector space is the linear combination of the basis of V' and W', that is, if $\{e_i\}$ is the basis of V' and $\{f_j\}$ is the basis of W', then $\{e_i \otimes f_j\}$ is the basis of $\mathcal{B}(V', W')$.

That is, every element of $\mathcal{B}(V', W')$ can be written as a linear combination of the basis.

Since $\{e_i\}$ and $\{f_j\}$ are bases of V' and W', respectively, then we can always find a set of linear functionals $\{\phi_i\}$ and $\{\psi_j\}$ such that $\phi_i(e_j) = \delta_{ij}$ and $\psi_j(f_i) = \delta_{ij}$.

Here $\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$ *is the Kronecker delta.*

$$V \otimes W = \left\{ \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij} \phi_i(v) \psi_j(w) : \phi_i \in V', \psi_j \in W' \right\}$$

Note that $\sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij} \phi_i(v) \psi_j(w)$ is a bilinear functional that maps $V' \times W'$ to $\mathbb{F}$.

This enables basis-free construction of vector spaces with proper multiplication and scalar multiplication.

Examples of tensor product for vectors

Let $V = \mathbb{C}^2, W = \mathbb{C}^3$, choose bases $\{|0\rangle, |1\rangle\} \subset V, \{|0\rangle, |1\rangle, |2\rangle\} \subset W$.

$$v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = v_1 |0\rangle + v_2 |1\rangle \in V, w = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = w_1 |0\rangle + w_2 |1\rangle + w_3 |2\rangle \in W$$

.
Then the tensor product $v \otimes w$ is given by

$$v \otimes w = \begin{pmatrix} v_1 w_1 & v_1 w_2 & v_1 w_3 \\ v_2 w_1 & v_2 w_2 & v_2 w_3 \end{pmatrix} \in \mathbb{C}^6$$

Examples of tensor product for vector spaces

Let $V = \mathbb{C}^2, W = \mathbb{C}^3$, choose bases $\{|0\rangle, |1\rangle\} \subset V, \{|0\rangle, |1\rangle, |2\rangle\} \subset W$.
Then a basis of the tensor product is

$$\{|00\rangle, |01\rangle, |02\rangle, |10\rangle, |11\rangle, |12\rangle\},$$

where $|ij\rangle := |i\rangle \otimes |j\rangle$.
An example element of $V \otimes W$ is

$$|\psi\rangle = 2\ |0\rangle \otimes |1\rangle + (1+i)\ |1\rangle \otimes |0\rangle - i\ |1\rangle \otimes |2\rangle .$$

With respect to the ordered basis

$$(|00\rangle, |01\rangle, |02\rangle, |10\rangle, |11\rangle, |12\rangle),$$

this tensor corresponds to the coordinate vector

$$|\psi\rangle \longleftrightarrow \begin{pmatrix} 0 \\ 2 \\ 0 \\ 1+i \\ 0 \\ -i \end{pmatrix} \in \mathbb{C}^6.$$

Using the canonical identification

$$\mathbb{C}^2 \otimes \mathbb{C}^3 \cong \mathbb{C}^{2\times 3},$$

where

$$|i\rangle \otimes |j\rangle \longmapsto E_{ij},$$

the same tensor is represented by the matrix

$$|\psi\rangle \longleftrightarrow \begin{pmatrix} 0 & 2 & 0 \\ 1+i & 0 & -i \end{pmatrix}.$$

Definition 5. *The vector space defined by the tensor product is equipped with the unique inner*

product $\langle v \otimes w, u \otimes x \rangle_{V \otimes W} : V \otimes W \times V \otimes W \to \mathbb{F}$ *defined by*

$$\langle v \otimes w, u \otimes x \rangle = \langle v, u \rangle_V \langle w, x \rangle_W$$

In practice, we ignore the subscript of the vector space and just write $\langle v \otimes w, u \otimes x \rangle = \langle v, u \rangle \langle w, x \rangle$.
Partial trace

Definition 6. *Let* T *be a linear operator on* $\mathcal{H}$*,* $(e_1, e_2, \cdots, e_n)$ *be a basis of* $\mathcal{H}$ *and* $(\epsilon_1, \epsilon_2, \cdots, \epsilon_n)$ *be a basis of dual space* $\mathcal{H}^*$*. Then the trace of* T *is defined by*

$$\mathrm{Tr}(T) = \sum_{i=1}^{n} \epsilon_i(T(e_i)) = \sum_{i=1}^{n} \langle e_i, T(e_i) \rangle$$

This is equivalent to the sum of the diagonal elements of T.

Definition 7. *Let* T *be a linear operator on* $\mathcal{H} = \mathcal{A} \otimes \mathcal{B}$*, where* $\mathcal{A}$ *and* $\mathcal{B}$ *are finite-dimensional Hilbert spaces.*

An operator T *on* $\mathcal{H} = \mathcal{A} \otimes \mathcal{B}$ *can be written as*

$$T = \sum_{i=1}^{n} a_i A_i \otimes B_i$$

where A_i *is a linear operator on* $\mathcal{A}$ *and* B_i *is a linear operator on* $\mathcal{B}$*.*

The $\mathcal{B}$*-partial trace of* T *(*$\mathrm{Tr}_{\mathcal{B}}(T) : \mathcal{L}(\mathcal{A} \otimes \mathcal{B}) \to \mathcal{L}(\mathcal{A})$*) is the linear operator on* $\mathcal{A}$ *defined by*

$$\mathrm{Tr}_{\mathcal{B}}(T) = \sum_{i=1}^{n} a_i \,\mathrm{Tr}(B_i) A_i$$

Or we can define the map $L_v : \mathcal{A} \to \mathcal{A} \otimes \mathcal{B}$ by

$$L_v(u) = u \otimes v$$

Note that $\langle u, L_v^*(u') \otimes v' \rangle = \langle u, u' \rangle \langle v, v' \rangle = \langle u \otimes v, u' \otimes v' \rangle = \langle L_v(u), u' \otimes v' \rangle$.

Therefore, $L_v^* \sum_j u_j \otimes v_j = \sum_j \langle v, v_j \rangle u_j$.

Then the partial trace of T can also be defined by

Let $\{v_j\}$ be a set of orthonormal basis of $\mathcal{B}$.

$$\mathrm{Tr}_{\mathcal{B}}(T) = \sum_j L_{v_j}^*(T) L_{v_j}$$

Definition 8. *Let T be a linear operator on $\mathcal{H} = \mathcal{A} \otimes \mathcal{B}$, where $\mathcal{A}$ and $\mathcal{B}$ are finite-dimensional Hilbert spaces.*

Let ρ be a state on $\mathcal{B}$ consisting of orthonormal basis $\{v_j\}$ and eigenvalue $\{\lambda_j\}$.

The partial trace of T with respect to ρ is the linear operator on $\mathcal{A}$ defined by

$$\mathrm{Tr}_{\mathcal{A}}(T) = \sum_j \lambda_j L^*_{v_j}(T) L_{v_j}$$

This introduces a new model in mathematics explaining quantum mechanics: the non-commutative probability theory.

0.2 Non-commutative probability theory

The non-commutative probability theory is a branch of generalized probability theory that studies the probability of events in non-commutative algebras.

There are several main components of the generalized probability theory; let's see how we can formulate them, comparing with the classical probability theory.

First, we define the Hilbert space in case one did not make the step from the linear algebra courses like me.

Definition 9. *Hilbert space:*

A Hilbert space is a complete inner product space.

That is, a vector space equipped with an inner product that is complete (every Cauchy sequence converges to a limit).

Example

To introduce an example of Hilbert space we use when studying quantum mechanics, we need to introduce a common inner product used in $\mathbb{C}^n$.

Proposition 10. *The Hermitian inner product on the complex vector space $\mathbb{C}^n$ makes it a Hilbert space.*

Proof. We first verify that the Hermitian inner product

$$\langle u, v \rangle = \sum_{i=1}^{n} \overline{u_i} v_i$$

on $\mathbb{C}^n$ satisfies the axioms of an inner product:

1. **Conjugate symmetry:** For all $u, v \in \mathbb{C}^n$,

$$\langle u, v \rangle = \sum_{i=1}^{n} \overline{u_i} v_i = \overline{\sum_{i=1}^{n} \overline{v_i} u_i} = \overline{\langle v, u \rangle}.$$

2. **Linearity:** For any $u, v, w \in \mathbb{C}^n$ and scalars $a, b \in \mathbb{C}$, we have

$$\langle u, av + bw\rangle = \sum_{i=1}^{n} \overline{u_i}(av_i + bw_i) = a\langle u, v\rangle + b\langle u, w\rangle.$$

3. **Positive definiteness:** For every $u = (u_1, u_2, \cdots, u_n) \in \mathbb{C}^n$, let $u_j = a_j + b_j i$, where $a_j, b_j \in \mathbb{R}$.

$$\langle u, u\rangle = \sum_{j=1}^{n} \overline{u_j} u_j = \sum_{i=1}^{n} (a_i^2 + b_i^2) \geq 0,$$

with equality if and only if $u = 0$.

Therefore, the Hermitian inner product is an inner product.

Next, we show that $\mathbb{C}^n$ is complete with respect to the norm induced by this inner product:

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

Since $\mathbb{C}^n$ is finite-dimensional, every Cauchy sequence (with respect to any norm) converges in $\mathbb{C}^n$. This is a standard result in finite-dimensional normed spaces, which implies that $\mathbb{C}^n$ is indeed complete.

Therefore, since the Hermitian inner product fulfills the inner product axioms and $\mathbb{C}^n$ is complete, the complex vector space $\mathbb{C}^n$ with the Hermitian inner product is a Hilbert space. □

Another classical example of Hilbert space is $L^2(\Omega, \mathcal{F}, P)$, where $(\Omega, \mathcal{F}, P)$ is a measure space (Ω is a set, $\mathcal{F}$ is a σ-algebra on Ω, and P is a measure on $\mathcal{F}$). The L^2 space is the space of all function on Ω that is

1. **square integrable**: square integrable functions are the functions $f : \Omega \to \mathbb{C}$ such that

$$\int_\Omega |f(\omega)|^2 dP(\omega) < \infty$$

with inner product defined by

$$\langle f, g\rangle = \int_\Omega \overline{f(\omega)} g(\omega) dP(\omega)$$

2. **complex-valued**: functions are complex-valued measurable. $f = u + vi$ is complex-valued if u and v are real-valued measurable.

Example

Proposition 11. *$L^2(\Omega, \mathcal{F}, P)$ is a Hilbert space.*

Proof. We check the two conditions of the Hilbert space:

- Completeness: Let (f_n) be a Cauchy sequence in $L^2(\Omega, \mathcal{F}, P)$. Then for any $\epsilon > 0$, there exists an N such that for all $m, n \geq N$, we have

$$\int_\Omega |f_m(\omega) - f_n(\omega)|^2 dP(\omega) < \epsilon^2$$

This means that (f_n) is a Cauchy sequence in the norm of $L^2(\Omega, \mathcal{F}, P)$.

- Inner product: The inner product is defined by

$$\langle f, g\rangle = \int_\Omega \overline{f(\omega)} g(\omega) dP(\omega)$$

This is a well-defined inner product on $L^2(\Omega, \mathcal{F}, P)$. We can check the properties of the inner product:

  - Linearity:

$$\langle af + bg, h\rangle = a\langle f, h\rangle + b\langle g, h\rangle$$

  - Conjugate symmetry:

$$\langle f, g\rangle = \overline{\langle g, f\rangle}$$

  - Positive definiteness:

$$\langle f, f\rangle \geq 0$$

□

Let $\mathcal{H}$ be a Hilbert space. $\mathcal{H}$ consists of complex-valued functions on a finite set $\Omega = \{1, 2, \ldots, n\}$, and the functions $(e_1, e_2, \ldots, e_n)$ form an orthonormal basis of $\mathcal{H}$. (We use Dirac notation $|k\rangle$ to denote the basis vector e_k [Par92].)

As an analog to the classical probability space $(\Omega, \mathcal{F}, \mu)$, which consists of a sample space Ω and a probability measure μ on the state space $\mathcal{F}$, the non-commutative probability space $(\mathcal{H}, \mathcal{P}, \rho)$ consists of a Hilbert space $\mathcal{H}$ and a state ρ on the space of all orthogonal projections $\mathcal{P}$.

The detailed definition of the non-commutative probability space is given below:

Definition 12. *Non-commutative probability space:*

*A non-commutative probability space is a pair $(\mathcal{B}(\mathcal{H}), \mathcal{P})$, where $\mathcal{B}(\mathcal{H})$ is the set of all **bounded** linear operators on $\mathcal{H}$.*

*A linear operator on $\mathcal{H}$ is **bounded** if for all u such that $\|u\| \leq 1$, we have $\|Au\| \leq M$ for some $M > 0$.*

$\mathcal{P}$ is the set of all orthogonal projections on $\mathcal{B}(\mathcal{H})$.

The set $\mathcal{P} = \{P \in \mathcal{B}(\mathcal{H}) : P^* = P = P^2\}$ is the set of all orthogonal projections on $\mathcal{B}(\mathcal{H})$.

Recall from classical probability theory, we call the initial probability distribution for possible outcomes in the classical probability theory as our *state*, simillarly, we need to define the *state* in the non-commutative probability theory.

Definition 13. *Non-commutative probability state:*

Given a non-commutative probability space $(\mathcal{B}(\mathcal{H}), \mathcal{P})$,

A state is a unit vector $\langle\psi|$ in the Hilbert space $\mathcal{H}$, such that $\langle\psi|\,|\psi\rangle = 1$.

Every state uniquely defines a map $\rho : \mathcal{P} \to [0, 1]$, $\rho(P) = \langle\psi|\,P\,|\psi\rangle$ (commonly named as density operator) such that:

- *$\rho(O) = 0$, where O is the zero projection, and $\rho(I) = 1$, where I is the identity projection.*

- *If $P_1, P_2, \ldots, P_n$ are pairwise disjoint orthogonal projections, then $\rho(P_1 + P_2 + \cdots + P_n) = \sum_{i=1}^{n} \rho(P_i)$.*

Note that the pure states are the density operators that can be represented by a unit vector $\langle\psi|$ in the Hilbert space $\mathcal{H}$, whereas mixed states are the density operators that cannot be represented by a unit vector in the Hilbert space $\mathcal{H}$.

If $(|\psi_1\rangle, |\psi_2\rangle, \cdots, |\psi_n\rangle)$ is an orthonormal basis of $\mathcal{H}$ consisting of eigenvectors of ρ, for the eigenvalues $p_1, p_2, \cdots, p_n$, then $p_j \geq 0$ and $\sum_{j=1}^{n} p_j = 1$.

We can write ρ as

$$\rho = \sum_{j=1}^{n} p_j |\psi_j\rangle\langle\psi_j|$$

(Under basis $|\psi_j\rangle$, it is a diagonal matrix with p_j on the diagonal.)

The counterpart of the random variable in the non-commutative probability theory is called an observable, which is a Hermitian operator on $\mathcal{H}$ (for all ψ, ϕ in the domain of A, we have $\langle A\psi, \phi\rangle = \langle\psi, A\phi\rangle$. This kind of operator ensures that our outcome interpreted as probability is a real number).

Definition 14. *Observable:*

Let $\mathcal{B}(\mathbb{R})$ be the set of all Borel sets on $\mathbb{R}$.

A random variable on the Hilbert space $\mathcal{H}$ is a projection-valued map (measure) $P : \mathcal{B}(\mathbb{R}) \to \mathcal{P}$.

With the following properties:

- *$P(\emptyset) = O$ (the zero projection)*
- *$P(\mathbb{R}) = I$ (the identity projection)*
- *For any sequence $A_1, A_2, \cdots, A_n \in \mathcal{B}(\mathbb{R})$, the following holds:*
  - *$P(\bigcup_{i=1}^{n} A_i) = \bigvee_{i=1}^{n} P(A_i)$*
  - *$P(\bigcap_{i=1}^{n} A_i) = \bigwedge_{i=1}^{n} P(A_i)$*
  - *$P(A^c) = I - P(A)$*
  - *If A_j are mutually disjoint (that is $P(A_i)P(A_j) = P(A_j)P(A_i) = O$ for $i \neq j$), then $P(\bigcup_{j=1}^{n} A_j) = \sum_{j=1}^{n} P(A_j)$*

Definition 15. *Probability of a random variable:*

For a system prepared in state ρ, the probability that the random variable given by the projection-valued measure P is in the Borel set A is $\mathrm{Tr}(\rho P(A))$.

When operators commute, we recover classical probability measures.

Definition 16. *Definition of measurement:*

A measurement (observation) of a system prepared in a given state produces an outcome x, x is a physical event that is a subset of the set of all possible outcomes. For each x, we associate a measurement operator M_x on $\mathcal{H}$.

Given the initial state (pure state, unit vector) u, the probability of measurement outcome x is given by:

$$p(x) = \|M_x u\|^2$$

Note that to make sense of this definition, the collection of measurement operators $\{M_x\}$ must satisfy the completeness requirement:

$$1 = \sum_{x \in X} p(x) = \sum_{x \in X} \|M_x u\|^2 = \sum_{x \in X} \langle M_x u, M_x u \rangle = \langle u, (\sum_{x \in X} M_x^* M_x) u \rangle$$

So $\sum_{x \in X} M_x^* M_x = I$.

Here is Table 1 summarizing the analog of classical probability theory and non-commutative (*quantum*) probability theory [Fer]:

Table 1: Analog of classical probability theory and non-commutative (*quantum*) probability theory

Classical probability	**Non-commutative probability**
Sample space Ω, cardinality $\|\Omega\| = n$, example: $\Omega = \{0, 1\}$	Complex Hilbert space $\mathcal{H}$, dimension $\dim \mathcal{H} = n$, example: $\mathcal{H} = \mathbb{C}^2$
Common algebra of $\mathbb{C}$ valued functions	Algebra of bounded operators $\mathcal{B}(\mathcal{H})$
$f \mapsto \bar{f}$ complex conjugation	$P \mapsto P^*$ adjoint
Events: indicator functions of sets	Projections: space of orthogonal projections $\mathcal{P} \subseteq \mathcal{B}(\mathcal{H})$
functions f such that $f^2 = f = \overline{f}$	orthogonal projections P such that $P^* = P = P^2$
$\mathbb{R}$-valued functions $f = \overline{f}$	self-adjoint operators $A = A^*$
$\mathbb{I}_{f^{-1}(\{\lambda\})}$ is the indicator function of the set $f^{-1}(\{\lambda\})$	$P(\lambda)$ is the orthogonal projection to eigenspace
$f = \sum_{\lambda \in \text{Range}(f)} \lambda \mathbb{I}_{f^{-1}(\{\lambda\})}$	$A = \sum_{\lambda \in \text{sp}(A)} \lambda P(\lambda)$
Probability measure μ on Ω	Density operator ρ on $\mathcal{H}$
Delta measure δ_ω	Pure state $\rho = \|\psi\rangle\langle\psi\|$
μ is non-negative measure and $\sum_{i=1}^n \mu(\{i\}) = 1$	ρ is positive semi-definite and $\text{Tr}(\rho) = 1$
Expected value of random variable f is $\mathbb{E}_\mu(f) = \sum_{i=1}^n f(i)\mu(\{i\})$	Expected value of operator A is $\mathbb{E}_\rho(A) = \text{Tr}(\rho A)$
Variance of random variable f is $\text{Var}_\mu(f) = \sum_{i=1}^n (f(i) - \mathbb{E}_\mu(f))^2 \mu(\{i\})$	Variance of operator A is $\text{Var}_\rho(A) = \text{Tr}(\rho A^2) - \text{Tr}(\rho A)^2$
Covariance of random variables f and g is $\text{Cov}_\mu(f, g) = \sum_{i=1}^n (f(i) - \mathbb{E}_\mu(f))(g(i) - \mathbb{E}_\mu(g))\mu(\{i\})$	Covariance of operators A and B is $\text{Cov}_\rho(A, B) = \text{Tr}(\rho A \circ B) - \text{Tr}(\rho A)\,\text{Tr}(\rho B)$
Composite system is given by Cartesian product of the sample spaces $\Omega_1 \times \Omega_2$	Composite system is given by tensor product of the Hilbert spaces $\mathcal{H}_1 \otimes \mathcal{H}_2$
Product measure $\mu_1 \times \mu_2$ on $\Omega_1 \times \Omega_2$	Tensor product of space $\rho_1 \otimes \rho_2$ on $\mathcal{H}_1 \otimes \mathcal{H}_2$
Marginal distribution $\pi_* v$	Partial trace $\text{Tr}_2(\rho)$

0.2.1 Quantum physics and terminologies

In this section, we will introduce some terminologies and theorems used in quantum physics that are relevant to our study. Assuming no prior knowledge of quantum physics, we will provide brief definitions and explanations for each term.

One might ask, what is the fundamental difference between a quantum system and a classical system, and why can we not directly apply those theorems in classical computers to a quantum computer? It turns out that quantum error-correcting codes are hard due to the following definitions and features for quantum computing.

Definition 17. *All quantum operations can be constructed by composing four kinds of transformations: (adapted from Chapter 10 of [BZ17])*

1. *Unitary operations. $U(\cdot)$ for any quantum state. It is possible to apply a non-unitary operation for an open quantum system, but that is usually not the focus for quantum computing and usually leads to non-recoverable loss of information that we wish to obtain.*
2. *Extend the system. Given a quantum state $\rho \in \mathcal{H}^N$, we can extend it to a larger quantum system by "entangle" (For this report, you don't need to worry for how quantum entanglement works) it with some new states $\sigma \in \mathcal{H}^K$ (The space where the new state dwells is usually called ancilla system) and get $\rho' = \rho \otimes \sigma \in \mathcal{H}^N \otimes \mathcal{K}$.*
3. *Partial trace. Given a quantum state $\rho \in \mathcal{H}^N$ and some reference state $\sigma \in \mathcal{H}^K$, we can trace out some subsystems and get a new state $\rho' \in \mathcal{H}^{N-K}$.*
4. *Selective measurement. Given a quantum state, we measure it and get a classical bit; unlike the classical case, the measurement is a probabilistic operation. (More specifically, this is some projection to a reference state corresponding to a classical bit output. For this report, you don't need to worry about how such a result is obtained and how the reference state is constructed.)*

$U(n)$ is the group of all $n \times n$ **unitary matrices** over $\mathbb{C}$,

$$U(n) = \{A \in \mathbb{C}^{n \times n} : A^* A = A A^* = I_n\}$$

The uniqueness of such measurement came from the lemma below [Mec]

Lemma 18. *Let $(U(n), \|\cdot\|, \mu)$ be a metric measure space where $\|\cdot\|$ is the Hilbert-Schmidt norm and μ is the measure function.*

The Haar measure on $U(n)$ is the unique probability measure that is invariant under the action of $U(n)$ on itself.

That is, fixing $B \in U(n)$, $\forall A \in U(n)$, $\mu(A \cdot B) = \mu(B \cdot A) = \mu(B)$.

The Haar measure is the unique probability measure that is invariant under the action of $U(n)$ on itself.

Definition 19. *Pure state:*

A random pure state φ is any random variable distributed according to the unitarily invariant probability measure on the pure states $\mathcal{P}(A)$ of the system A, denoted by $\varphi \in_R \mathcal{P}(A)$.

It is trivial that for the space of pure state, we can easily apply the Haar measure as the unitarily invariant probability measure since the space of pure state is S^n for some n. However, for the case of mixed states, that is a bit complicated and we need to use partial tracing to defined the rank-s random states.

Definition 20. *Rank-s random state.*

For a system A and an integer $s \geq 1$, consider the distribution onn the mixed states $\mathcal{S}(A)$ of A induced by the partial trace over the second factor form the uniform distribution on pure states of $A \otimes \mathbb{C}^s$. Any random variable ρ distributed as such will be called a rank-s random states; denoted as $\rho \in_R \mathcal{S}_s(A)$. And $\mathcal{P}(A) = \mathcal{S}_1(A)$.

Proposition 21. *Proposition of indistinguishability:*

Suppose that we have two systems $u_1, u_2 \in \mathcal{H}_1$, the two states are distinguishable if and only if they are orthogonal.

Proof. Ways to distinguish the two states:

1. Set $X = \{0, 1, 2\}$ and $M_i = |u_i\rangle\langle u_i|$, $M_0 = I - M_1 - M_2$
2. Then $\{M_0, M_1, M_2\}$ is a complete collection of measurement operators on $\mathcal{H}$.
3. Suppose the prepared state is u_1, then $p(1) = \|M_1 u_1\|^2 = \|u_1\|^2 = 1$, $p(2) = \|M_2 u_1\|^2 = 0$, $p(0) = \|M_0 u_1\|^2 = 0$.

If they are not orthogonal, then there is no choice of measurement operators to perfectly distinguish the two states.

□

Intuitively, if the two states are not orthogonal, then for any measurement (projection) there exists non-zero probability of getting the same outcome for both states.

0.2.2 Random quantum states

First, we need to define what is a random state in a bipartite system.

References

[Axl23] S. Axler. *Linear Algebra Done Right*. Undergraduate Texts in Mathematics. Springer International Publishing, 2023. ISBN: 9783031410260. URL: https://books.google.com/books?id=0dnfEAAAQBAJ.

[BZ17] I. Bengtsson and K. Zyczkowski. *Geometry of Quantum States: An Introduction to Quantum Entanglement*. Cambridge University Press, 2017.

[Fer] R. Feres. *Math 444 Lecture notes – the mathematics of quantum theory*. URL: https://www.math.wustl.edu/~feres/Math444Spring25/Math444Spring25Syllabus.html.

[Mec] E. Meckes. *The Random Matrix Theory of the Classical Compact Groups*.

[Par92] K. R. Parthasarathy. *An Introduction to Quantum Stochastic Calculus*. Vol. 85. Monographs in Mathematics. Birkhäuser Basel, 1992, pp. XI, 292. ISBN: 978-3-0348-9711-2. DOI: 10.1007/978-3-0348-8641-3.