WiDS QRL - Linear Algebra and Postulates of Quantum Mechanics

Nilabha Saha

January 2023

1 Linear Algebra

1.1 Vector Space $\mathbb{C}^n$

A *vector space* of complex numbers is a set of n-tuples of complex numbers, whose elements are called *vectors*, which satisfy certain vector space axioms. The vectors are indicated using column matrix notation:

$$\begin{bmatrix} z_1 \\ ... \\ z_n \end{bmatrix}$$

In $\mathbb{C}^n$, addition for vectors is defined as:

$$\begin{bmatrix} z_1 \\ ... \\ z_n \end{bmatrix} + \begin{bmatrix} z_1' \\ ... \\ z_n' \end{bmatrix} = \begin{bmatrix} z_1 + z_1' \\ ... \\ z_n + z_n' \end{bmatrix}$$

Multiplication is defined as:

$$z \begin{bmatrix} z_1 \\ ... \\ z_n \end{bmatrix} = \begin{bmatrix} zz_1 \\ ... \\ zz_n \end{bmatrix}$$

The standard quantum mechanical notation for representing a vector in a vector space is $|\psi\rangle$, where ψ is a label for the vector, and the entire expression is called a *ket*.
The *zero vector* is represented by 0 (the ket notation is not used for the zero vector).
A *vector subspace* of a vector space V is a subset W of V such that W is also a vector space.

1.2 Bases and Linear Independence

A *spanning set* for a vector space is a set of vectors $|v_1\rangle, .., |v_n\rangle$ such that any vector $|v\rangle$ in the vector space can be written as a *linear combination* $|v\rangle =$

$\sum_i a_i|v_i\rangle$ of vectors in that set.
A set of non-zero vectors $|v_1\rangle, .., |v_n\rangle$ are *linearly dependent* if there exists a set of complex numbers $a_1, ..., a_n$ with $a_i \neq 0$ for at least one value of i, such that $a_1|v_1\rangle + a_2|v_2\rangle + ... + a_n|v_n\rangle = 0$. A set of vectors is *linearly independent* if it is not linearly dependent. Any two sets of linearly independent vectors which span a vector space V contain the same number of elements. Such a set is called a *basis* and the cardinality of a basis is called its *dimension*.

1.3 Linear Operators and Matrices

A *linear operator* between vector spaces V and W is defined to be any function $A : V \to W$ which is linear in its inputs:

$$A(\sum_i a_i|v_i\rangle) = \sum_i a_i A(|v_i\rangle)$$

An important linear operator on any vector space V is the *identity operator* I_V defined by $I_V|v\rangle = |v\rangle$ for all vectors $|v\rangle$. Another important linear operator is the *zero operator* denoted by 0 which maps all vectors to the zero vector, $0|v\rangle = 0$.
Once the action of a linear operator A on a basis is specified, the action of A is completely determined on all inputs.
Suppose V, W, and X are vector spaces, and $A : V \to W$ and $B : W \to X$ are linear operators. Then BA denotes the composition of B with A, defined as $BA|v\rangle = B(A(|v\rangle))$.
Linear operators can be understood in terms of their *matrix representations*. Matrices can be regarded as linear operators and linear operators can be represented as matrices, thereby making the two completely equivalent. Suppose $A : V \to W$ is a linear operator between vector spaces V and W. Suppose $|v_1\rangle, .., |v_m\rangle$ is a basis for V and $|w_1\rangle, .., |w_m\rangle$ is a basis for W. Then for each j in the range $1, .., m$, there exist complex numbers A_{1j} through A_{nj} such that

$$A|v_j\rangle = \sum_i A_{ij}|w_i\rangle$$

The matrix whose entries are the values A_{ij} is said to form a matrix representation of the operator A.

1.4 The Pauli Matrices

Four useful 2×2 matrices used occasionally are the *Pauli matrices*:

$$\sigma_0 = I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\sigma_1 = \sigma_x = X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$\sigma_2 = \sigma_y = Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$$

$$\sigma_3 = \sigma_z = Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

1.5 Inner Products

An *inner product* is a function which takes as input two vectors $|v\rangle$ and $|w\rangle$ from a vector space and produces a complex number as output.
A function $(\cdot, \cdot)$ from $V \times V$ to $\mathbb{C}$ is an inner product if it satisfies the requirements that:

1. $(\cdot, \cdot)$ is linear in the second argument,

$$(|v\rangle, \sum_i \lambda_i |w_i\rangle) = \sum_i \lambda_i (|v\rangle, |w_i\rangle).$$

2. $(|v\rangle, |w\rangle) = (|w\rangle, |v\rangle)^*$.
3. $(|v\rangle, |v\rangle) \geq 0$ with equality if and only if $|v\rangle = 0$.

The standard quantum mechanical notation for the inner product is $\langle v|w\rangle$ where $|v\rangle$ and $|w\rangle$ are vectors in the *inner product space*.
Discussions of quantum mechanics often refer to *Hilbert space*. In the finite dimensional complex vector spaces that come up in quantum computation and quantum information, a Hilbert space is exactly the same thing as an inner product space.
$\langle v|$ is used for the *dual vector* to the vector $|v\rangle$ which is a linear operator from the inner product space V to the complex numbers $\mathbb{C}$ defined by $\langle v|(|w\rangle) = \langle v|w\rangle$.
Vectors $|w\rangle$ and $|v\rangle$ are *orthogonal* if $\langle w|v\rangle = 0$.
The *norm* of a vector $|v\rangle$ is defined by $\||v\rangle\| = \sqrt{\langle v|v\rangle}$.
A *unit vector* is a vector $|v\rangle$ such that $\||v\rangle\| = 1$. The vector $|v\rangle / \||v\rangle\|$ is called the *normalised form* of $|v\rangle$, for any non-zero vector $|v\rangle$.
A set $|i\rangle$ of vectors with index i is *orthonormal* is each vector is a unit vector, and distinct vectors in the set are orthogonal, that is, $\langle i|j\rangle = \delta_{ij}$ where i and j are both chosen from the index set.
Suppose $|w_1\rangle, .., |w_d\rangle$ is a basis set for some inner product space V. The *Gram-Schmidt procedure* can be used to produce an orthonormal basis set $|v_1\rangle, ..., |v_d\rangle$ for the inner product space V. Define $|v_1\rangle = |w_1\rangle / \||w_1\rangle\|$, and for $1 \leq k \leq d-1$ define $|v_{k+1}\rangle$ inductively by

$$|v_{k+1}\rangle = \frac{|w_{k+1}\rangle - \sum_{i=1}^{k} \langle v_i|w_{k+1}\rangle |v_1\rangle}{\||w_{k+1}\rangle - \sum_{i=1}^{k} \langle v_i|w_{k+1}\rangle |v_1\rangle\|}$$

The vectors $|v_1\rangle, ..., |v_d\rangle$ form an orthonormal set which is also a basis for V.
Let $|w\rangle = \sum_i w_i |i\rangle$ and $|v\rangle = \sum_j v_j |j\rangle$ be representations of vectors $|v\rangle$ and $|w\rangle$

with respect to some orthonormal basis $|i\rangle$. Then

$$\langle v|w\rangle = \begin{bmatrix} v_1^* & ... & v_n^* \end{bmatrix} \begin{bmatrix} w_1 \\ ... \\ w_n \end{bmatrix}$$

Linear operators can also be represented using their *outer product* representation. Suppose $|v\rangle$ is a vector in an inner product space V, and $|w\rangle$ is a vector in an inner product space W. Define $|w\rangle\langle v|$ to be the linear operator from V to W whose action is defined by $(|w\rangle\langle v|)(|v^{'}\rangle) = \langle v|v^{'}\rangle|w\rangle$. The *completeness relation* is given by

$$\sum_i |i\rangle\langle i| = I$$

The completeness relation can be used to obtain a representation for any operator in the outer product notation. Suppose $A : V \to W$ is a linear operator, $|v_i\rangle$ is an orthonormal basis for V, and $|w_j\rangle$ an orthonormal basis for W. Using the completeness relation twice we obtain

$$A = I_W A I_V = \sum_{ij} |w_j\rangle\langle w_j|A|v_i\rangle\langle v_i| = \sum_{ij} \langle w_j|A|v_i\rangle|w_j\rangle\langle v_i|$$

which is the outer product representation for A.
The *Cauchy-Schwarz inequality* states that for any two vectors $|v\rangle$ and $|w\rangle$, $|\langle v|w\rangle|^2 \leq \langle v|v\rangle\langle w|w\rangle$.

1.6 Eigenvectors and Eigenvalues

An *eigenvector* of a linear operator A on a vector space is a non-zero vector $|v\rangle$ such that $A|v\rangle = v|v\rangle$, where v is a complex number known as the *eigenvalue* of A corresponding to $|v\rangle$. The *eigenspace* corresponding to an eigenvalue v is the set of vectors which have eigenvalue v. It is a vector subspace of the vector space on which A acts.
A *diagonal representation* for an operator A on a vector space V is a representation $A = \sum_i \lambda_i |i\rangle\langle i|$, where the vectors $|v\rangle$ for an orthonormal set of eigenvectors for A, with corresponding eigenvalues λ_i. An operator is said to be *diagonalisable* if it has a diagonal representation. Diagonal representations are also known as *orthonormal decompositions*. When an eigenspace is more than one dimensional, it is said to be *degenerate*.

1.7 Adjoints and Hermitian Operators

Suppose A is any linear operator on a Hilbert space V. The *adjoint* or *Hermitian conjugate* of the operator A is the unique linear operator $A^\dagger$ on V such that for all vectors $|v\rangle, |w\rangle \in V$,

$$(|v\rangle, A|w\rangle) = (A^\dagger|v\rangle, |w\rangle)$$

In the matrix representation of the operator A, the action of the Hermitian conjugation operation is to take the matrix of A to the conjugate-transpose matrix, $A^\dagger = (A^*)^T$.
An operator A whose adjoint is A is known as a *Hermitian* or *self-adjoint* operator.
Suppose W is a k-dimensional vector subspace of the d-dimensional vector space V. Using the Gram-Schmidt procedure, it is possible to construct an orthonormal basis $|1\rangle, ..., |d\rangle$ for V such that $|1\rangle, ..., |k\rangle$ is an orthonormal basis for W. By definition,

$$P = \sum_{i=1}^{k} |i\rangle\langle i|$$

is the *projector* onto the subspace W. This definition is independent of the orthornormal basis $|1\rangle, ..., |k\rangle$ used for W. P is Hermitian, $P^\dagger = P$. The *orthogonal complement* of P is the operator $Q = I - P$. Q is a projector onto the vector space spanned by $|k+1\rangle, ..., |d\rangle$.
An operator A is said to be *normal* if $AA^\dagger = A^\dagger A$.

> **Spectral Decomposition:** Any normal operator M on a vector space V is diagonal with respect to some orthonormal basis for V. Conversely, any diagonalisable operator is normal.

A matrix or operator U is said to be *unitary* if $UU^\dagger = I$. Unitary operators preserve inner products between vectors, that is, the inner product of $U|v\rangle$ and $U|w\rangle$ is the same as the inner product of $|v\rangle$ and $|w\rangle$:

$$(U|v\rangle, U|w\rangle) = \langle v|U^\dagger U|w\rangle = \langle v|I|w\rangle = \langle v|w\rangle.$$

Let $|v_i\rangle$ be any orthonormal basis set. Define $|w_i\rangle = U|v_i\rangle$, so $|u_i\rangle$ is also an orthonormal basis set, since unitary operators preserve inner products. Note that $U = \sum_i |w_i\rangle\langle v_i|$. Conversely, if $|v_i\rangle$ and $|w_i\rangle$ are any two orthonormal bases, then the operator $U = \sum_i |w_i\rangle\langle v_i|$ is a unitary operator.
A subclass of Hermitian operators is the *positive operators*. A positive operator A is defined to be an operator such that for any vector $|v\rangle$, $(|v\rangle, A|v\rangle)$ is a real, non-negative number. If $(|v\rangle, A|v\rangle)$ is strictly greater than for all $|v\rangle \neq 0$ then we say that A is *positive definite*.

1.8 Tensor Products

The *tensor product* is a way of putting vector spaces together to form larger vector spaces. Suppose V and W are Hilbert spaces of dimension m and n respectively. Then $V \otimes W$ is an mn dimensional vector space. The elements of $V \otimes W$ are linear combinations of 'tensor products' $|v\rangle \otimes |w\rangle$ of elements $|v\rangle$ of V and $|w\rangle$ of W. If $|i\rangle$ and $|j\rangle$ are orthonormal bases for the spaces V and W then $|i\rangle \otimes |j\rangle$ is a basis for $V \otimes W$.
By definition the tensor product satisfies the following basic properties:

1. For an arbitrary scalar z and elements $|v\rangle$ of V and $|w\rangle$ of W,

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

2. For arbitrary $|v_1\rangle$ and $|v_2\rangle$ in V and $|w\rangle$ in W,

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

3. For arbitrary $|v\rangle$ in V and $|w_1\rangle$ and $|w_2\rangle$ in W,

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

Suppose $|v\rangle$ and $|w\rangle$ are vectors in V and W, and A and B are linear operators on V and W, respectively. Then the linear operator $A \otimes B$ on $V \otimes W$ is defined as

$$(A \otimes B)(|v\rangle \otimes |w\rangle) = A|v\rangle \otimes B|w\rangle$$

The inner products on the spaces V and W can be used to define a natural inner product on $V \otimes W$ as

$$(\sum_i a_i |v_i\rangle \otimes |w_i\rangle, \sum_j b_j |v_j^{'}\rangle \otimes |w_j^{'}\rangle) = \sum_{ij} a_i^* b_j \langle v_i | v_j^{'}\rangle \langle w_i | w_j^{'}\rangle$$

A convenient matrix representation of the tensor product is called the *Kronecker product*. Suppose A is an m by n matrix, and B is p by q matrix. Then we have the matrix representation:

$$A \otimes B = \begin{bmatrix} A_{11}B & A_{12}B & \cdots & A_{1n}B \\ A_{21}B & A_{22}B & \cdots & A_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1}B & A_{m2}B & \cdots & A_{mn}B \end{bmatrix}$$

Note: $|\psi\rangle^{\otimes k}$ means $|\psi\rangle$ tensored with itself k times.

1.9 Operator Functions

Given a function f from the complex numbers to the complex numbers, it is possible to define a corresponding matrix function on normal matrices by the following construction: Let $A = \sum_a a|a\rangle\langle a|$ be a spectral decomposition for a normal operator A. Define $f(A) = \sum_a f(a)|a\rangle\langle a|$. $f(A)$ is uniquely defined.
The *trace* of A us defined to be the sum of its diagonal entries, that is, $tr(A) = \sum_i A_{ii}$. The trace is *cyclic*, $tr(AB) = tr(BA)$, and *linear*, $tr(A+B) = tr(A) + tr(B)$, $tr(zA) = ztr(A)$, where A and B are arbitrary matrices, and z is a complex number. The trace of a matrix is invariant under the *unitary similarity transformation* $A \rightarrow UAU^\dagger$, $tr(UAU^\dagger) = tr(A)$. We also have that $tr(A|\psi\rangle\langle\psi|) = \langle\psi|A|\psi\rangle$.

1.10 The Commutator and Anti-Commutator

The *commutator* between two operators A and B is defined to be $[A, B] = AB - BA$. If $[A, B] = 0$, then A *commutes* with B.
The *anti-commutator* of two operators A and B is defined by $\{A, B\} = AB + BA$. If $\{A, B\} = 0$, then A *anti-commutes* with B.

> **Simultaneous Diagonalisation Theorem:** Suppose A and B are Hermitian operators. Then $[A, B] = 0$ if and only if there exists an orthonormal basis such that both A and B are diagonal with respect to that basis. We say that A and B are *simultaneously diagonalisable* in this case.

The commutation relations for the Pauli matrices are

$$[X, Y] = 2iZ \quad [Y, Z] = 2iX \quad [Z, X] = 2iY$$

or more elegantly

$$[\sigma_j, \sigma_k] = 2i \sum_{l=1}^{3} \epsilon_{jkl} \sigma_l$$

1.11 The Polar and Singular Value Decompositions

> **Polar Decomposition:** Let A be a linear operator on a vector space V. Then there exists unitary U and positive operators J and K such that
>
> $$A = UJ = KU$$
>
> where the unique positive operators J and K satisfying these equations are defined by $J = \sqrt{A^\dagger A}$ and $K = \sqrt{AA^\dagger}$. Moreover, if A is invertible then U is unique.

The expressions $A = UJ$ is called the *left polar decomposition* of A, and $A = KU$ is called the *right polar decomposition* of A.

> **Singular Value Decomposition:** Let A be a square matrix. Then there exist unitary matrices U and V, and a diagonal matrix D with non-negative entries such that
>
> $$A = UDV$$
>
> The diagonal elements of D are called the *singular valuessingular value* of A.

1.12 Matrix Exponentials

Unitary transformations are often seen in the form $U = e^{i\gamma H}$ where H is a Hermitian matrix and γ is a real number. Such matrices U in this form are all unitary.
The *exponential of a matrix* is another matrix given by:

$$e^{i\gamma H} = \sum_{n=0}^{\infty} \frac{(i\gamma H)^n}{n!}$$

A matrix B is said to be *involutory* if $B^2 = I$. For an involutory matrix B, we have the identity:

$$e^{i\gamma B} = \cos(\gamma)I + i\sin(\gamma)B$$

The Pauli matrices are unitary, Hermitian, and involutory.
If a matrix M has eigenvector $|v\rangle$ corresponding to eigenvalue λ, $M|v\rangle = \lambda v$ then $|v\rangle$ is also an eigenvector of e^M with eigenvalue e^λ, $e^M|v\rangle = e^\lambda|v\rangle$.

2 The Postulates of Quantum Mechanics

2.1 State Space

Postulate 1: Associated to any physical system is a complex vector space with an inner product (that is, a Hilbert space) known as the *state space* of the system. The system is completely described by its *state vector*, which is a unit vector in the system's state space.

Any linear combination $\sum_i \alpha_i|\psi_i\rangle$ is called a *superposition* of the states $|\psi_i\rangle$ with *amplitude* α_i for the state $|\psi_i\rangle$.

2.2 Evolution

Postulate 2: The evolution of a *closed* quantum system is described by a *unitary transformation*. That is, the state $|\psi\rangle$ at a time t_1 is related to the state $|\psi^{'}\rangle$ of the system at time t_2 by a unitary operator U which depends only on the times t_1 and t_2,

$$|\psi^{'}\rangle = U|\psi\rangle$$

Postulate 2': The time evolution of the state of a closed quantum system is described by the *Schrödinger equation*,

$$i\hbar\frac{d|\psi\rangle}{dt} = H|\psi\rangle$$

H here is a fixed Hermitian operator known as the *Hamiltonian* of the closed system. Because the Hamiltonian is a Hermitian operator, it has a spectral decomposition

$$H = \sum_E E|E\rangle\langle E|$$

with eigenvalues E and corresponding normalised eigenvectors $|E\rangle$. The states $|E\rangle$ are conventionally referred to as *energy eigenstates*, or sometimes as *stationary states*, and E is the energy of the state $|E\rangle$. The lowest energy is known as the *ground state energy* for the system, and the corresponding energy eigenstate (or eigenspace) is known as the *ground state*. The states $|E\rangle$ are known as stationary states because their only change in time is to acquire an overall numerical factor,

$$|E\rangle \to e^{-iEt/\hbar}|E\rangle$$

The connection between the Hamiltonian picture of dynamics, Postulate 2', and the unitary operator picture, Postulate 2 is provided by writing down the solution to the Schrödinger equation, which is easily verified to be:

$$|\psi(t_2)\rangle = e^{-iH(t_2-t_1)/\hbar}|\psi(t_1)\rangle = U(t_1,t_2)|\psi(t_1)\rangle$$

It can be shown that any unitary operator U can be realised in the form $U = e^{iK}$ for some Hermitian operator K.

2.3 Quantum Measurement

Postulate 3: Quantum measurements are described by a collection $\{M_m\}$ of *measurement operators*. These operators acting on the state space of the system being measured. The index m refers to the measurement outcomes that may occur in the experiment. If the state of the quantum system is $|\psi\rangle$ immediately before the measurement then the probability that result m occurs is given by

$$p(m) = \langle\psi|M_m^\dagger M_m|\psi\rangle$$

and the state of the system after measurement is

$$\frac{M_m|\psi\rangle}{\sqrt{\langle\psi|M_m^\dagger M_m|\psi\rangle}}$$

The measurement operators satisfy the *completeness equation*:

$$\sum_m M_m^\dagger M_m = I$$

The completeness equation expresses the fact that probabilities sum to one

$$1 = \sum_m p(m) = \sum_m \langle\psi|M_m^\dagger M_m|\psi\rangle$$

The *measurement of a qubit in the computational basis* is defined by the two measurement operators $M_0 = |0\rangle\langle 0|$ and $M_1 = |1\rangle\langle 1|$.

> **Cascaded measurements are single measurements:** Suppose $\{L_l\}$ and $\{M_m\}$ are two sets of measurement operators. A measurement defined by the measurement operators $\{L_l\}$ followed by a measurement defined by the measurement operators $\{M_m\}$ is physically equivalent to a single measurement defined by measurement operators $\{N_{lm}\}$ with the representation $N_{lm} \equiv M_m L_l$.

2.4 Phase

The states $e^{i\theta}|\psi\rangle$ and $|\psi\rangle$ are said to be the same up to a *global phase factor*. The *statistics of measurement* for two such states are the same. To see it, suppose M_m is a measurement operator associated to some quantum measurement, and note that the respective probabilities for outcome m occurring are $\langle\psi|M_m^\dagger M_m|\psi\rangle$ and $\langle\psi|e^{-i\theta}M_m^\dagger M_m e^{i\theta}|\psi\rangle$. Therefore, from an observational point of view, both the states are identical.
Two amplitudes, a and b, are said to *differ by a relative phase* if there is a real θ such that $a = e^{i\theta}b$. Two states are said to *differ by a relative phase* in some basis if each of the amplitudes in that basis is related by such a phase factor.

2.5 Composite Systems

> **Postulate 4:** The state space of a *composite physical system* is the tensor product of the state spaces of the component physical systems. Moreover, if we have systems numbered 1 through n, and system number i is prepared in the state $|\psi_i\rangle$, then the joint state of the total system is $|\psi_1\rangle \otimes |\psi_2\rangle \otimes \cdots \otimes |\psi_n\rangle$.

A qubit state $|\psi\rangle$ for which there exist no single qubit states $|a\rangle$ and $|b\rangle$ such that $|\psi\rangle = |a\rangle|b\rangle$ is called an *entangled state*. In general, a state of a composite system which cannot be written as a product of states of its component systems is an entangled state.

3 The Density Operator

3.1 Ensembles of Quantum States

Suppose a quantum system is in one of a number of states $|\psi_i\rangle$, where i is an index, with respective probabilities p_i. We call $\{p_i, |\psi_i\rangle\}$ and *ensemble of pure states*. The *density operator* or *density matrix* for the system is defined by the equation

$$\rho \equiv \sum_i p_i|\psi_i\rangle\langle\psi_i|$$

The evolution of the density operator is described by the equation

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i| \xrightarrow{U} \sum_i p_i U|\psi_i\rangle\langle\psi_i|U^\dagger = U\rho U^\dagger$$

Suppose we perform a measurement described by measurement operators M_m. If the initial state was $|\psi_i\rangle$, then the probability of getting result m is

$$p(m|i) = \langle\psi_i|M_m^\dagger M_m|\psi_i\rangle = tr(M_m^\dagger M_m|\psi_i\rangle\langle\psi_i|)$$

By the law of total probability, the probability of obtaining result m is

$$\begin{aligned} p(m) &= \sum_i p(m|i)p_i & (1)\\ &= \sum_i p_i tr(M_m^\dagger M_m|\psi_i\rangle\langle\psi_i|) & (2)\\ &= tr(M_m^\dagger M_m \rho) & (3) \end{aligned}$$

If the initial state was $|\psi_i\rangle$ then the state after obtaining the result m is

$$|\psi_i^m\rangle = \frac{M_m|\psi_i\rangle}{\sqrt{\langle\psi_i|M_m^\dagger M_m|\psi_i\rangle}}$$

The corresponding density operator ρ_m is therefore

$$\rho_m = \sum_i p(i|m)|\psi_i^m\rangle\langle\psi_i^m| = \sum_i p(i|m)\frac{M_m|\psi_i\rangle\langle\psi_i|M_m^\dagger}{\langle\psi_i|M_m^\dagger M_m|\psi_i\rangle} = \frac{M_m\rho M_m^\dagger}{tr(M_m^\dagger M_m\rho)}$$

A quantum state whose state $|\psi\rangle$ is known exactly is said to be in a *pure state*. Otherwise, it is said to be in a *mixed state*.
Imagine a quantum system in the state ρ_i with probability p_i. This system may be described by the density matrix $\sum_i p_i\rho_i$. We say that ρ is a *mixture* of the states ρ_i with probabilities p_i.

3.2 General Properties of The Density Operator

Characterisation of density operators: An operator ρ is the density operator associated to some ensemble $\{p_i, |\psi_i\rangle\}$ if and only if it satisfies the conditions:

- **(Trace condition)** ρ has trace equal to one.
- **(Positivity condition)** ρ is a positive operator.

We can, in fact, instead define a density operator to be a positive operator ρ which has trace equal to one. The postulate of quantum mechanics can now be reformulated as follows:

Postulate 1: Associated to any isolated physical system is a complex vector space with inner product (that is, a Hilbert space) known as the *state space* of the system. The system is completely described by its *density operator*, which is a positive operator ρ with trace one, acting on the state space of the system. If a quantum system is in the state ρ_i with probability p_i, then the density operator of the system is $\sum_i p_i \rho_i$.
Postulate 2: The evolution of a *closed* quantum system is described by a *unitary transformation*. That is, the state ρ of a system at time t_1 is related to the state ρ' of the system at time t_2 by a unitary operator U which depends only on the times t_1 and t_2,

$$\rho' = U \rho U^\dagger.$$

Postulate 3: Quantum measurements are described by a collection $\{M_m\}$ of *measurement operators*. These are operators acting on the state space of the system being measured. The index m refers to the measurement outcomes that may occur in the experiment. If the state of the quantum system is ρ immediately before the measurement then the probability the result m occurs is given by

$$p(m) = tr(M_m^\dagger M_m \rho)$$

and the state of the system after the measurement is

$$\frac{M_m \rho M_m^\dagger}{tr(M_m^\dagger M_m \rho)}.$$

The measurement operators satisfy the *completeness equation*,

$$\sum_m M_m^\dagger M_m = I.$$

Postulate 4: The state space of a composite physical system is the tensor product of the state spaces of the component physical systems. Moreover, if we have systems numbered 1 through n, and system number i is prepared in the state ρ_i, then the joint state of the total system is $\rho_1 \otimes \rho_2 \otimes \ldots \otimes \rho_n$.

Criterion to decided if a state is mixed or pure: Let ρ be a density operator. Then $tr(\rho^2) \leq 1$, with equality holding if and only if ρ is a pure state.

We now turn to find out what class of ensembles give rise to a particular density matrix. For this, it is convenient to make use of the vectors $|\psi_i\rangle$ which may not be normalised to unit length. We say the set $|\psi_i\rangle$ *generates* the operator

$\rho \equiv \sum_i |\psi_i\rangle\langle\psi_i|$, and thus the connection to the usual ensemble picture is expressed by the equation $|\psi_i\rangle = \sqrt{p_i}|\psi_i\rangle$.

> **Unitary freedom in the ensemble for ensemble for density matrices:** The sets $|\psi_i\rangle$ and $|\phi_i\rangle$ generate the same density matrix if and only if
> $$|\psi_i\rangle = \sum_j u_{ij}|\phi_i\rangle$$
> where u_{ij} is a unitary matrix of complex numbers, with indices i and j, and we 'pad' whichever set of vectors $|\psi_i\rangle$ or $|\phi_i\rangle$ is smaller with additional vectors 0 so that the two sets have the same number of elements.

As a consequence of the above theorem, $\rho = \sum_i p_i|\psi_i\rangle\langle\psi_i| = \sum_i \sum_j q_j|\phi_j\rangle\langle\phi_j|$ for normalised states $|\psi_i\rangle$, $|\phi_i\rangle$ and probability distributions p_i and q_i if and only if

$$\sqrt{p_i}|\psi_i\rangle = \sum_j u_{ij}\sqrt{q_j}|\phi_j\rangle$$

for some unitary matrix u_{ij}, and we may pad the smaller ensemble with entries having probability zero in order to make the two ensembles the same size.

> **Bloch sphere for mixed states:** An arbitrary density matrix for a mixed state qubit may be written as
> $$\rho = \frac{I + \overrightarrow{r} \cdot \overrightarrow{\sigma}}{2}$$
> where $\overrightarrow{r}$ is a real three-dimensional vector such that $\|\overrightarrow{r}\| \leq 1$. This vector is known as the *Bloch vector* for the state ρ. A state ρ is pure if and only if $\|\overrightarrow{r}\| = 1$.

Let ρ be a density operator. A *minimal ensemble* for ρ is an ensemble $\{p_i, |\psi_i\rangle\}$ containing a number of elements equal to the rank of ρ. Let $|\psi\rangle$ be any state in the support of ρ. (The *support* of a Hermitian operator A is the vector space spanned by the eigenvectors of A with non-zero eigenvalues). Then there is a minimal ensemble for ρ that contains $|\psi\rangle$, and moreover that in any such ensemble $|\psi\rangle$ must appear with probability

$$p_i = \frac{1}{\langle\psi_i|\rho^{-1}|\psi_i\rangle}$$

where ρ^{-1} is defined to be the inverse of ρ, when ρ is considered to be an operator acting only on the support of ρ (This definition removes the problem that ρ may not have an inverse).

3.3 The Reduced Density Operator

The *reduced density operator* provides the description for *subsystems* of a composite quantum system. Suppose we have physical systems A and B, whose state is described by a density operator ρ^{AB}. The reduced density operator for system A is given by

$$\rho^A \equiv tr_B(\rho^{AB})$$

where tr_B is a map of operators known as the *partial trace* over system B. The partial trace is defined by

$$tr_B(|a_1\rangle\langle a_2| \otimes |b_1\rangle\langle b_2|) \equiv |a_1\rangle\langle a_2| tr(|b_1\rangle\langle b_2|)$$

where $|a_1\rangle$ and $|a_2\rangle$ are any two vectors in the state space of A, and $|b_1\rangle$ and $|b_2\rangle$ are any two vectors in the state space of B. As usual, $tr(|b_1\rangle\langle b_2|) = \langle b_2|b_1\rangle$. The partial trace also needs to be linear in its inputs.

Suppose a quantum system is in the product state $\rho^{AB} = \rho \otimes \sigma$, where ρ is a density operator for system A, and σ is a density operator for system B. Then

$$\rho^A = tr_B(\rho \otimes \sigma) = \rho tr(\sigma) = \rho$$

Similarly, $\rho^B = \sigma$ for this state.

3.4 The Schmidt Decomposition

Schmidt Decomposition: Suppose $|\psi\rangle$ is a pure state of a composite system AB. Then there exist orthonormal states $|i_A\rangle$ for system A, and orthonormal states $|i_B\rangle$ of system B such that

$$|\psi\rangle = \sum_i \lambda_i |i_A\rangle|i_B\rangle$$

where λ_i are non-negative real numbers satisfying $\sum_i \lambda_i^2 = 1$ known as *Schmidt coefficients*.

As a consequence, let $|\psi\rangle$ be a pure state of a composite system, AB. Then by Schmidt decomposition $\rho^A = \sum_i \lambda_i^2 |i_A\rangle\langle i_A|$ and $\rho^B = \sum_i \lambda_i^2 |i_B\rangle\langle i_B|$, so the eigenvalues of ρ^A and ρ^B are identical, namely λ_i^2 for both density operators. The bases $|i_A\rangle$ and $|i_B\rangle$ are called the *Schmidt bases* for A and B, respectively, and the number of non-zero values λ_i is called the *Schmidt number* for the state $|\psi\rangle$.

A state $|\psi\rangle$ of a composite system AB is a product state if and only if it has a Schmidt number 1. $|\psi\rangle$ is a product state if and only if ρ^A (and thus ρ^B) are pure states.

3.5 Purifications

Suppose we are given a state ρ_A of a quantum system A. It is possible to introduce another system, which we denote R, and define a pure state $|AR\rangle$ for the joint system AR such that $\rho_A = tr_R(|AR\rangle\langle AR|)$. That is, the pure state $|AR\rangle$ reduces to ρ_A when we look at system A alone. This is a purely mathematical procedure, known as *purification*, which allows us to associate pure states with mixed states. The system R is called a *reference system*.
Suppose ρ^A has orthonormal decomposition $\rho^A = \sum_i p_i|i^A\rangle\langle i^A|$. To purify ρ^A we introduce a system R which has the same state space as system A, with orthonormal basis states $|i^R\rangle$, and define a pure state for the combined system

$$|AR\rangle \equiv \sum_i \sqrt{p_i}|i^A\rangle|i^R\rangle$$

We now calculate the reduced density operator for system A corresponding to the state $|AR\rangle$:

$$\begin{aligned} tr_R(|AR\rangle\langle AR|) &= \sum_{ij} \sqrt{p_i p_j}|i^A\rangle\langle j^A|tr(|i^R\rangle\langle j^R|) && (4) \\ &= \sum_{ij} \sqrt{p_i p_j}|i^A\rangle\langle j^A|\delta_{ij} && (5) \\ &= \sum_i p_i|i^A\rangle\langle i^A| && (6) \\ &= \rho^A && (7) \end{aligned}$$

Thus $|AR\rangle$ is a purification of ρ^A.
The procedure used to purify a mixed state of system A is to define a pure state whose Schmidt basis for system A is just the basis in which the mixed state is diagonal, with the Schmidt coefficients being the square root of the eigenvalues of the density operator being purified.

> **Freedom in purifications:** Let $|AR_1\rangle$ and $|AR_2\rangle$ be two purifications of a state ρ^A corresponding to a composite system AR. Then there exists a unitary transformation U_R acting on system R such that $|AR_1\rangle = (I_A \otimes U_R)|AR_2\rangle$.