Basic Notation and Background

Chi-Kwong Li
Department of Mathematics, The College of William and Mary,
Williamsburg, Virginia 23187, USA (ckli@math.wm.edu)
Department of Mathematics, Taiyuan University of Technology,
Taiyuan, Shanxi, P.R. of China.

We will give two lectures providing some basic notation, mathematics and physics background needed for the subsequent discussion in this summer school. Most material are adapted from [3, Chapter 2] and [1] (see also [2, Chapter 1]).

1 Hilbert spaces

The mathematical platform of quantum mechanics/computing is Hilbert space (complete inner product space) V. We mainly focus on finite dimensional complex inner product space $\mathbf{C}^n$, the set of $n \times 1$ column vectors. Let $\mathbf{C}^{n*}$ be the *dual* vector space of $\mathbf{C}^n$ consisting of $1 \times n$ row vectors

In physics, we use the *bra* and *ket* vector notation. (Dirac notation.) Let

$$|x\rangle = (x_1, \dots, x_n)^t = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbf{C}^n.$$

Then

$$\langle x| = (\bar{x}_1, \dots, \bar{x}_n) \in \mathbf{C}^{n*}$$

is the dual vector. The *norm* (*length*) of $|x\rangle$ is

$$\|x\| = \langle x|x\rangle^{1/2} = \{(\bar{x}_1, \dots, \bar{x}_n)(x_1, \dots, x_n)^t\}^{1/2} = \left\{\sum_{j=1}^n |x_j|\right\}^{1/2}.$$

(a) For $|x\rangle \in \mathbf{C}^n$, we can construct its transpose $|x\rangle^t$, the conjugate $|\overline{x}\rangle$, the conjugate transpose $|x\rangle^\dagger$ (instead of $|x\rangle^*$).

(b) The *inner product* of $|x\rangle, |y\rangle \in \mathbf{C}^n$ is $\langle x|y\rangle = \sum_{j=1}^n x_j^* y_j$. The vectors are *orthogonal* if their inner product is zero.

(c) Given a set of vectors S in $\mathbf{C}^n$, we can determine whether it is a *linearly independent set*, a *generating set*, an *orthonormal set*, a *basis*, or an *orthonormal basis*.

2 Matrices

Linear maps (transformations/functions) on finite dimensional vector spaces can be identified with matrices, namely, $A : \mathbf{C}^n \to \mathbf{C}^n$ so that $|x\rangle \mapsto A|x\rangle$.

Operations and Properties

Let M_n be the set (vector space/algebra) of $n \times n$ matrices.

(a) One can perform $A + B$, AB and μA for $A, B \in M_n$ and $\mu \in \mathbf{C}$.

(b) One can compute the eigenvalues and eigenvectors of $A \in M_n$.

Let $\{|e_1\rangle, \ldots, |e_n\rangle\}$ be the standard basis for $\mathbf{C}^n$. Then

$$A_{ij} = \langle e_i|A|e_j\rangle \quad \text{and} \quad A = \sum_{i,j} A_{ij}|e_i\rangle\langle e_j|.$$

The trace of A is defined by $\text{Tr}(A) = \sum_{j=1}^n A_{jj}$.

Exercise 1) If A is $m \times n$ and B is $n \times m$, then $\text{Tr}(AB) = \text{Tr}(BA)$.

2) If R is an $n \times n$ matrix, and $|\psi\rangle \in \mathbf{C}^n$, then $\langle\psi|R|\psi\rangle = \text{Tr}(R|\psi\rangle\langle\psi|)$.

Gram-Schmidt process and orthogonal projectors

If S is linearly independent set in $\mathbf{C}^n$, one can apply the Gram-Schmidt process to S to get an orthonormal set.

If $|e_k\rangle$ is a unit vector, then the projection of a vector $|v\rangle$ in the direction of $|e_k\rangle$ is $|v\rangle - P_k|v\rangle$, where $P = |e_k\rangle\langle e_k|$ is the *projection operator.* The vector $|v\rangle - P_k|v\rangle$ is orthogonal to $|e_k\rangle$.

If $\{|e_1\rangle, \ldots, |e_n\rangle\}$ is an orthonormal basis and $P_k = |e_k\rangle\langle e_k|$ for $k = 1, \ldots, n$, then

$$\text{(i) } P_k^2 = P_k, \quad \text{(ii) } P_jP_k = 0 \text{ for } j \neq k, \quad \text{(iii) } \textstyle\sum_{k=1}^n P_k = I_n.$$

More notation, definitions and examples

Let $A \in M_n$. One can compute its transpose A^t, the conjugate $\overline{A}$ and the conjugate transpose $A^\dagger$. The matrix is Hermitian if $A = A^\dagger$; it is skew-Hermitian if $A = -A^\dagger$; it is normal if $AA^\dagger = A^\dagger A$; it is unitary if $A^\dagger = A^{-1}$. If A is real and $A^t = A^{-1}$, the A is a real orthogonal matrix.

In quantum information science, the following Pauli matrices are useful:

$$\sigma_0 = I_2, \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Exercises 1) The Pauli matrices are trace zero Hermitian unitary matrices.

2) Let $\{i, j, k\} = \{x, y, z\}$, then

$$\sigma_i\sigma_j = i\gamma_{ij}\sigma_k = -\sigma_j\sigma_i, \quad \text{and} \quad [\sigma_i, \sigma_j] = 2i\gamma_{ij},$$

where $\gamma_{ij} = 1$ for $(i, j) = (x, y), (y, z), (z, x)$.

Some useful facts

Theorem (Schur Triangularization Lemma) Every matrix in M_n is unitarily similar to a matrix in upper or lower triangular form.

Theorem (Spectral Theorem) If $A \in M_n$ is normal, then there is a unitary U such that $UAU^\dagger = \text{diag}\,(\lambda_1, \ldots, \lambda_n)$; hence if U has columns $|u_1\rangle, \ldots, |u_n\rangle$ then

$$A = \lambda_1 |u_1\rangle\langle u_1| + \cdots + \lambda_n |u_n\rangle\langle u_n|.$$

(a) For any positive integer m,

$$A^m = \lambda_1^m |u_1\rangle\langle u_1| + \cdots + \lambda_n^m |u_n\rangle\langle u_n|.$$

The formula holds for negative integers m as well if A is invertible.

(b) If $f(z)$ is an analytic function, then

$$f(A) = \sum_{j=1}^{n} f(\lambda_j)|u_j\rangle\langle u_j|.$$

(c) In particular, if $f(z) = e^z$, then $f(A) = \sum_{j=1}^n e^{\lambda_j}|u_j\rangle\langle u_j|$.

Theorem (Singular Value Decomposition) For every $m \times n$ matrix A, there are unitary $U \in M_m$ and $V \in M_n$ so that $U^\dagger A V = D$ such that the (j, j) entries of D is s_j for $1 \le j \le \min\{m, n\}$, where $s_1^2 \ge s_2^2 \ge \cdots$ are the eigenvalues of $A^\dagger A$.

Theorem Every unitary matrix $U \in M_n$ is a product of no more than $n(n-1)/2$ tridiagonal unitary matrices, each of them differs from I_n by a 2×2 principal submatrix.

Remark Using the Gray code labeling of $2^m \times 2^m$, all the tridiagonal unitary matrices involve a change of two basic vectors with binary labels differ in one position.

Definition Given two real vectors $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$, we say that x is *majorized* by y if $\sum_{j=1}^n x_j = \sum_{j=1}^n y_j$ and the sum of the k largest entries of x is not larger than that of y for $k = 1, \ldots, n-1$.

Theorem The vector of diagonal entries $(d_1, \ldots, d_n)$ of a Hermitian matrix in M_n is *majorized* by the vector of its eigenvalues $(\lambda_1, \ldots, \lambda_n)$.

3 Tensor products

Let $A = (a_{ij})$ and B be two rectangular matrices or vectors. Then their tensor product (Kronecker product) is the matrix

$$A \otimes B = (a_{ij}B).$$

The following equalities hold:

$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD), \quad A \otimes (B + C) = A \otimes B + A \otimes C, \quad (A \otimes B)^\dagger = A^\dagger \otimes B^\dagger.$$

Note that A, B, C can be $|\psi_1\rangle, |\psi_2\rangle, |\psi_3\rangle$ vectors. If $A \in M_m$ and $B \in M_n$ are invertible, then

$$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}.$$

Every T on M_{mn} can be written as $T = \sum_{j=1} c_j A_j \otimes B_j$ with $c_j \in \mathbf{C}$ so that for any vectors $|u\rangle \in \mathbf{C}^m, |v\rangle \in \mathbf{C}^n$,

$$T|u\rangle|v\rangle = \left(\sum_j c_j A_j \otimes B_j\right) |u\rangle \otimes |v\rangle = \sum_j c_j A_j|u\rangle \otimes B_j|v\rangle.$$

Of course, every $|w\rangle \in \mathbf{C}^m \otimes \mathbf{C}^n$ has Schmidt decomposition $|w\rangle = \sum_j s_j|u_j\rangle \otimes |v_j\rangle$ so that

$$T|w\rangle = \sum_j \sum_k c_j s_k A_j|u_k\rangle \otimes B_j|v_k\rangle.$$

Remark We often use the abbreviation: $|x_1\rangle \otimes \cdots \otimes |x_n\rangle = |x_1\rangle \cdots |x_n\rangle = |x_1 \cdots x_n\rangle$.

Example Denote by $|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $|1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Let $H = (\frac{1}{\sqrt{2}}) \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$.

(1) $H|0\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$, $H|1\rangle) = (|0\rangle - |1\rangle)/\sqrt{2}$, and $H^2 = I_2$.

(2) Label the rows and columns of $A \in M_{2^n}$ by $(x_1 \cdots x_n)$ with $x_j \in \{0, 1\}$. Then

$$H_n = \overbrace{H \otimes \cdots \otimes H}^{n} = 2^{-n/2}((-1)^{x \cdot y}),$$

where $x \cdot y$ is the inner product of $(x_1, \dots, x_n)$ and $(y_1, \dots, y_n)$.

(3) We have $H_n|0 \cdots 0\rangle = 2^{-n/2} \sum_x |x\rangle$, where the summation ranges through all $x \in \{0, 1\}^n$.

Exercise Let $A \in M_m$ and $B \in M_n$. Prove the following.

- If $UAU^\dagger$ and $VBV^\dagger$ are in upper triangular form, then $(U \otimes V)(A \otimes B)(U \otimes V)^\dagger$ is in upper triangular form.
- $\det(A \otimes B) = \det(A)^n \det(B)^m$.
- $A \otimes B$ has eigenvalue $\lambda_i \mu_j$ corresponding to the eigenvector $|\lambda_i\rangle|\mu_j\rangle$ for $1 \le i \le m$ and $1 \le j \le n$, if $A \in M_m$ has eigenvectors $|\lambda_1\rangle, \dots, |\lambda_m\rangle$ corresponding to the eigenvalues $\lambda_1, \dots, \lambda_m$, and $B \in M_n$ has eigenvectors $|\mu_1\rangle, \dots, |\mu_n\rangle$ corresponding to the eigenvalues $\mu_1, \dots, \mu_n$.

4 Quantum Mechanics

The Copenhagen interpretation

A1 A state $|x\rangle$ is a unit vector in a Hilbert space $\mathcal{H}$ (usually $\mathbf{C}^n$). Linear combinations (*superposition*) of the physical states are allowed in the state space.

A2 Every physical quantity (*observable*) corresponds to a Hermitian operator (matrix) A. Suppose a state $|x\rangle = c_1|u_1\rangle + c_2|u_2\rangle$ such that $A|u_i\rangle = a_i|u_i\rangle$ for $i \in \{1, 2\}$. Then applying a measurement of $|x\rangle$ corresponding to A will cause a *wave function collapse* to $|u_1\rangle$ or $|u_2\rangle$ with probability of $|c_1|^2$ and $|c_2|^2$, respectively. Here c_1, c_2 are called the *probability amplitude* of the state $|x\rangle$.

A3 The time dependence of a state is governed by the Schrödinger equation

$$i\hbar \frac{\partial |x\rangle}{\partial t} = H(t)|x\rangle,$$

where $\hbar$ is the Planck constant, and H is a Hermitian operator (matrix) corresponding to the energy of the system known as the Hamiltonian.

Remarks

1. The phase of the state does not matter, i.e., $|x\rangle$ and $e^{i\alpha}|x\rangle$ represents the same states.

2. In the finite dimensional case, if the state and the observable are represented by

$$|x\rangle = \sum_{j=1}^{n} c_j|u_j\rangle \in \mathbf{C}^n \quad \text{and} \quad A = \sum_{j=1}^{n} \lambda_j |u_j\rangle\langle u_j| = \sum_{j=1}^{n} \lambda_j P_j,$$

   then the *projective measurement* of the state results in

$$\langle x|A|x\rangle = \sum_{j=1}^{n} \lambda_j |c_j|^2 \quad \text{and becomes} \quad \frac{P_i|x\rangle}{|c_i|^2}$$

   with a probability of $|c_i|^2$.

3. In the Schrödinger equation, if $H(t)$ does not depend on t, then

$$|x(t)\rangle = e^{-iHt/\hbar}|x(0)\rangle. \tag{1}$$

   Otherwise,

$$|x(t)\rangle = \exp\left(\frac{-i}{\hbar}\int_0^t H(s)ds\right)|x(0)\rangle. \tag{2}$$

Measurements

In connection to (A2), quantum measurements are described by a set of measurement operators $\{M_m : 1 \leq m \leq r\}$ such that $\sum_{j=1}^{r} M_r^\dagger M_r = I$. For each outcome m, construct a measurement operator so that the probability of obtaining outcome m in the state $|x\rangle$ is computed by

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

and the state immediately after the measurement is

$$|m\rangle = \frac{M_m|x\rangle}{\sqrt{p(m)}}.$$

If there are many copies of a state $|x\rangle$, we can let $M = \sum mP_m$. Then the expected value of M is

$$\text{Exp}_x(M) = \langle M\rangle = \sum_m mp(m) = \sum_m m\langle x|P_m|x\rangle = \langle x|M|x\rangle.$$

The variance (square of standard deviation) is

$$\langle (M - \langle M\rangle)^2\rangle = \langle x|M^2|x\rangle - \langle x|M|x\rangle^2.$$

The uncertainty principle

Let $\text{Exp}_x(A) = \langle x|A|x\rangle = \mu$ and

$$\text{Var}_x(A) = \text{Exp}_x((A - \mu I)^2) = \langle x|(A - \mu I)^2|x\rangle = \|(A - \mu I)|x\rangle\|^2.$$

Theorem For any observables A and B and for any state $|x\rangle$, we have

$$\text{Var}_x(A)\text{Var}_x(B) \geq \frac{1}{4}\langle x|[A, B]|x\rangle,$$

where $[A, B] = AB - BA$ is the commutator of A and B.

Remark There is no quantum measurement that can distinguish non-orthogonal states $|\psi_1\rangle, |\psi_2\rangle$, reliably. Suppose there is such a measurement. Let $E_1 = \sum M_j^\dagger M_j$ such that $\langle \psi_1|E_1|\psi_1\rangle = 1$ and $E_2 = \sum M_k^\dagger M_k$ such that $\langle \psi_2|E_2|\psi_1\rangle = 1$. Then $E_j - |\psi_j\rangle\langle\psi_j|$ is positive semidefinite for $j = 1, 2$. Since $I = \sum_i E_i$, we have

$$\langle\psi_1|I|\psi_1\rangle \geq \langle\psi_1|(E_1 + E_2)|\psi_1\rangle \geq \langle\psi_1|(|\psi_1\rangle\langle\psi_1|)|\psi_1\rangle + \langle\psi_1|(|\psi_2\rangle\langle\psi_2|)|\psi_1\rangle > 1,$$

which is a contradiction.

Positive Operator-Valued Measure (POVM) is a set of positive operators $E_j = M_j^\dagger M_j$ corresponding to the measuring operators M_j so that $\sum_j E_j = I_n$. The measurement(s) would allow Bob to identify correctly the state he receives or gets no information at all.

Note that a special case of POVM is the projective measurement $\{P_1, \ldots, P_r\}$, say, arising from an observable $A = \sum_{j=1}^{r} \lambda_j P_j$.

Example Alice sends Bob $|\psi_1\rangle = |0\rangle$ or $|\psi_2\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$. Consider the POVM $\{E_1, E_2, E_3\}$ with $E_1 = \alpha|1\rangle\langle 1|$, $E_2 = \beta(|0\rangle - |1\rangle)(\langle 0| - \langle 1|)$, and $E_3 = I - E_1 - E_2$.

- If Bob gets $|\psi_1\rangle$, there is zero probability to get E_1. Thus, getting E_1 measurement means that the received state is $|\psi_2\rangle$.
- Similarly, getting E_2 measurement means that the received state is $|\psi_1\rangle$.
- Getting E_3 measurement yields no information.

In connection to the evolution of a closed system, we have the following.

Example Recall that the differential equation $y' = ay$ has solution $y = e^{at}y_0$.

If $H = \sum_{j=1}^n \lambda_j P_j$ is Hermitian, then $e^{i\omega Ht} = \sum_{j=1}^n e^{i\omega\lambda_j t} P_j$.

If $H = -\hbar\omega\sigma_x/2$, then

$$|\psi(t)\rangle = \begin{pmatrix} \cos\omega t/2 & i\sin\omega t/2 \\ i\sin\omega t/2 & \cos\omega t/2 \end{pmatrix} |\psi(0)\rangle.$$

If $|\psi(0)\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, then $|\psi(t)\rangle = \begin{pmatrix} \cos\omega t/2 \\ i\sin\omega t/2 \end{pmatrix}$. If $|\psi(0)\rangle = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, then $|\psi(t)\rangle = \frac{e^{i\omega t/2}}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

General solution

Let $\mathbf{n} = (n_x, n_y, n_z)$ and $H = -\hbar\omega\mathbf{n}\cdot\sigma/2 = -\hbar\omega(n_x\sigma_x + n_y\sigma_y + n_z\sigma_z)/2$. Then

$$U(t) = \exp(-iHt/\hbar) = \cos(\omega/2)tI + i(\mathbf{n}\cdot\sigma)\sin(\omega/2)t.$$

A change of variable technique (using gauge functions)

Suppose

$$H = \frac{1}{2}\begin{pmatrix} -\omega_0 & \omega_1 e^{i\omega t} \\ \omega_1 e^{-\omega t t} & \omega_0 \end{pmatrix}.$$

Let

$$|\phi(t)\rangle = e^{-iw\sigma_z t/2}|\psi(t)\rangle.$$

Then

$$i\frac{d}{dt}|\phi(t)\rangle = \tilde{H}|\phi(t)\rangle, \qquad \text{where} \qquad \tilde{H} = [(w - w_0)\sigma_z + w_1\sigma_x]/2.$$

5 Multipartite system, tensor product and entangled states

A system may have two components described by two Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$. Then the *bipartite* system is represented by $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$. A general state in $\mathcal{H}$ has the form

$$|x\rangle = \sum_{i,j} c_{ij} |e_{1,i}\rangle \otimes |e_{2,j}\rangle \quad \text{with} \ \sum_{i,j} |c_{ij}|^2 = 1,$$

where $\{e_{r,1}, e_{r,2}, \ldots\}$ is an orthonormal basis for $\mathcal{H}_r$ with $r \in \{1, 2\}$.

A state of the form $|x\rangle = |x_1\rangle \otimes |x_2\rangle$ is a *separable state* or a *tensor product state.* Otherwise, it is an *entangled state.*

Proposition Every state $|x\rangle$ in $\mathcal{H}_1 \otimes \mathcal{H}_2$ admits a Schmidt decomposition

$$|x\rangle = \sum_{j=1}^{r} \sqrt{s_j} |u_j\rangle \otimes |v_j\rangle,$$

where $s_j > 0$ are the Schmidt coefficients satisfying $\sum_{j=1}^{r} s_j = 1$, r is the Schmidt number of $|x\rangle$, $\{|u_1\rangle, \ldots, |u_r\rangle\}$ is an orthonormal set of $\mathcal{H}_1$ and $\{|v_1\rangle, \ldots, |v_r\rangle\}$ is an orthonormal set of $\mathcal{H}_2$.

Remark In matrix theory, the Schmidt decomposition is just the singular value decomposition if one identify $\mathbf{C}^m \otimes \mathbf{C}^n$ with the space of $m \times n$ matrices. The following has a wide research interest in different branches of study.

Open problem Extend the Schmidt decomposition to $\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_k$ for $k \geq 3$.

Copying information

Theorem (No cloning) There is no unitary $\psi_0 \in \mathbf{C}^n$ and $U \in M_{n^2}$ such that $U|\psi\psi_0\rangle = |\psi\psi\rangle$ for every $|\psi\rangle \in \mathbf{C}^n$.

Remark There is unitary $U \in M_{n^2}$ such that $U|e_j\rangle|e_1\rangle = |e_j e_j\rangle$ for $j = 1, \ldots, n$, where $\{|e_1\rangle, \ldots, |e_n\rangle\}$ is an orthonormal set. So, classical information can be copied as we know!

Manipulation of multiple qubit states

Suppose two people, Alice and Bob, each possess one of the (maximally) entangled state, which is known as a Bell state:

$$|\psi_0\rangle = (|00\rangle + |11\rangle)/\sqrt{2}.$$

Each of them can manipulate her/his qubit.

For instance, in measuring each qubit, there is a 50-50 chance of seeing $|0\rangle$ and $|1\rangle$. The other qubit will be in the state of $|0\rangle$ and $|1\rangle$ accordingly.

One can apply $U \otimes I_2$, $I_2 \otimes U$ to a two qubit states. For example, Alice can apply the Hadamard gate $H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ to her qubit so that the entangled state $|\psi_0\rangle$ is changed to

$$\frac{1}{2}\{(|0\rangle + |1\rangle)|0\rangle + (|0\rangle - |1\rangle)|1\rangle\} = \frac{1}{2}\{|00\rangle + |10\rangle + |01\rangle - |11\rangle\}.$$

One can also apply a unitary $V \in M_4$ to a two qubit states if the two qubits are brought together. For example, the U_{CN} gate defined by

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

will have the effect change the basis of $\mathbf{C}^4 = \mathbf{C}^2 \otimes \mathbf{C}^2$ as follows.

$$|00\rangle \mapsto |00\rangle, \quad |01\rangle \mapsto |01\rangle, \quad |10\rangle \mapsto |11\rangle, \quad |11\rangle \mapsto |10\rangle.$$

This can be described as $|x\rangle|y\rangle \mapsto |x\rangle|x \oplus y\rangle$.

Applications of entanglement

Superdense coding

Suppose Alice and Bob share the entangled state $|\psi_0\rangle$. Alice can manipulate her qubit to encode one of the two bits of classical information, say, in $\{00, 01, 10, 11\}$, and send to Bob. Here is what she may do:

(1) she does nothing to send 00; (2) she applies a phase flip σ_z to send 01;
(3) she applies a not gate σ_x to send 10; (4) she applies $i\sigma_y$ to send 11.

The resulting state of Bob will be:

(1) $(|00\rangle + |11\rangle)/\sqrt{2}$, (2) $(|00\rangle - |11\rangle)/\sqrt{2}$, (3) $(|10\rangle + |01\rangle)/\sqrt{2}$, (4) $(|01\rangle - |10\rangle)/\sqrt{2}$,

which form the *Bell basis* of $\mathbf{C}^4$. By a suitable unitary gate $V \in M_4$, these will change to $|00\rangle, |01\rangle, |10\rangle, |11\rangle$ so that Bob can get the correct message upon measurement.

Teleportation

Suppose Alice put another state $|\xi\rangle = \alpha|0\rangle + \beta|1\rangle$ to the system, then Bob's qubit will be affected also. The resulting system with 3 qubits has quantum state

$$|\psi\rangle = |\xi\rangle|\psi_0\rangle = \frac{1}{\sqrt{2}} \{\alpha|0\rangle(|00\rangle + |11\rangle) + \beta|1\rangle(|00\rangle + |11\rangle\}.$$

Alice can apply a CNOT gate to her qubits to get

$$|\psi_1\rangle = \frac{1}{\sqrt{2}} \{\alpha(|0\rangle + |1\rangle)(|00\rangle + |11\rangle) + \beta|1\rangle(|10\rangle + |01\rangle\}.$$

Then apply a Hadamard gate to the first qubit to get

$$|\psi_2\rangle = \frac{1}{2} \{\alpha(|0\rangle + |1\rangle)(|00\rangle + |11\rangle) + \beta(|0\rangle - |1\rangle)(|10\rangle + |01\rangle\}.$$

Regrouping yields

$$|\psi_2\rangle = \frac{1}{2}\{|00\rangle(\alpha|0\rangle + \beta|1\rangle) + |01\rangle(\alpha|1\rangle + \beta|0\rangle) + |10\rangle(\alpha|0\rangle - \beta|1\rangle) + |11\rangle(\alpha|1\rangle - \beta|0\rangle)\}.$$

Now, measuring the two qubits of Alice gives one of the four possibilities: $|00\rangle, |01\rangle, |10\rangle, |11\rangle$. Bob can get this information from Alice and apply $\sigma_0, \sigma_x, \sigma_z, i\sigma_y$ to convert his qubit to $\alpha|0\rangle + \beta|1\rangle$.

6 Mixed States and Density Matrices

A system is in a *mixed state* if there is a probability p_i that the system is in state $|x_i\rangle$ for $i = 1, \dots, N$. If there is only one possible state, i.e., $p_1 = 1$, then the system is in *pure state*. The *mean value of the measurement* of the system corresponding to the observable described by the Hermitian matrix A is

$$\langle A \rangle = \sum_{j=1}^{N} p_i \langle x_j | A | x_j \rangle = \mathrm{Tr}(A\rho) \qquad \text{where} \qquad \rho = \sum_{j=1}^{N} p_j |x_j\rangle\langle x_j| \tag{3}$$

is a *density operator (matrix)*.

Description of a quantum system in mixed states

A1' A physical state is specified by a density matrix $\rho : \mathcal{H} \to \mathcal{H}$, which is positive semidefinite with trace equal to one.

A2' The mean value of an observable associate with the Hermitian matrix A is $\langle A \rangle = \mathrm{Tr}(\rho A)$.

A3' The temporal evolution of the density matrix is given by the Liouville-von Neumann equation

$$i\hbar \frac{d}{dt}\rho = [H, \rho] = H\rho - \rho H,$$

where H is the system Hamiltonian.

Remarks

(1) A density matrix $\rho = \sum_j p_j |x_j\rangle\langle x_j|$ corresponds to a mixed state, where the vector states $|x_1\rangle, \dots, |x_N\rangle$ need not be orthonormal.

(2) Each $|x_j\rangle$ satisfies the Schrödinger equation

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

One can derive the Liouville-von Neumann equation from these equations.

(3) The set of density matrices is compact and convex.

Exercises

1) The following conditions are equivalent for a given state (density matrix) ρ.

   (a) ρ is pure. (b) $\rho^2 = \rho$. (c) $\mathrm{Tr}(\rho^2) = 1$.

2) Show that every density matrix $\rho \in M_2$ has the form $\rho = \frac{1}{2}(\sigma_0 + x\sigma_x + y\sigma_y + z\sigma_z)$ with $x^2 + y^2 + z^2 \le 1$. The equality holds if and only if ρ is a pure state.

   Hence every density matrix in M_2 corresponds to a point in the unit sphere in $\mathbf{R}^3$, known as the Bloch sphere; ρ is a pure state if and only if it corresponds to a point on the sphere.

Definition Suppose $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$. A state ρ is *uncorrelated* if $\rho = \rho_1 \otimes \rho_2$; it is *separable* if it is a convex combination of uncorrelated states, i.e.,

$$\rho = \sum_{j=1}^{r} q_j \, \rho_{1,j} \otimes \rho_{2,j} \qquad \text{with } q_1, \ldots, q_r > 0,\ q_1 + \cdots + q_r = 1.$$

Otherwise, it is *inseparable.*

Remark Do not confuse this with the definitions of separability in the vector case. There will be discussion on this topic in depth.

Definition Let $\rho = \sum_{j=1}^{r} c_j \rho_{1,j} \otimes \rho_{2,j}$ act on $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$. The *partial transpose* of ρ with respect to $\mathcal{H}_2$ is

$$\rho^{\rm pt} = \sum_{j=1}^{r} \rho_{1,j} \otimes \rho_{2,j}^{t}.$$

Proposition If ρ is separable, then so is $\rho^{\rm pt}$. If $\rho^{\rm pt}$ has negative eigenvalues, then it is not physical and ρ is not separable. The converse holds if $\mathcal{H}$ has dimension at most 6.

Open problem Find effective way to determine separability.

Example Consider the Werner state and its partial transpose

$$\rho = \frac{1}{4}\begin{pmatrix} 1-p & 0 & 0 & 0 \\ 0 & 1+p & -2p & 0 \\ 0 & -2p & 1+p & 0 \\ 0 & 0 & 0 & 1-p \end{pmatrix}, \quad \rho^{\rm pt} = \frac{1}{4}\begin{pmatrix} 1-p & 0 & 0 & -2p \\ 0 & 1+p & 0 & 0 \\ 0 & 0 & 1+p & 0 \\ -2p & 0 & 0 & 1-p \end{pmatrix}.$$

Then the partial transpose has eigenvalues $(1+p)/4, (1+p)/4, (1+p)/2, (1-3p)/4$. So, it is not separable if and only if $p \in (1/2, 1]$.

The realignment matrices

For any $X = (x_{ij}) \in M_n$, let $\mathrm{vec}(X) = (x_{11}, x_{12}, \ldots, x_{nn})$. Suppose $\rho = (\rho_{ij})_{1 \le i,j \le m} \in M_{mn}$ is a density matrix such that $\rho_{ij} \in M_n$. The *realignment* matrix of ρ is the matrix

$$\rho^R = \begin{pmatrix} \mathrm{vec}(\rho_{11}) \\ \mathrm{vec}(\rho_{12}) \\ \vdots \\ \mathrm{vec}(\rho_{mm}) \end{pmatrix}.$$

Theorem Suppose $m \le n$ and $\rho \in M_{mn} = M_m \otimes M_n$ is a density matrix. If ρ is separable, then the sum of the singular values of ρ^R is at most one. In fact, the vector of singular values of ρ^R majorizes the vector $(\alpha, \beta, \ldots, \beta) \in \mathbf{R}^{1 \times m^2}$, where $\alpha = 1/\sqrt{mn}$ and $\beta = (1-\alpha)/m^2$.

Open problem If $\rho \in M_2 \otimes M_3$ is separable, then ρ^R cannot have its vector of singular values equal to $(1, (\sqrt{6}-1)/3, (\sqrt{6}-1)/3, (\sqrt{6}-1)/3)/\sqrt{6}$.

Partial Trace and Purification

Let A be an operator acting on $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$. The *partial trace* of A over $\mathcal{H}_2$ is an operator acting on $\mathcal{H}_1$ defined by

$$A_1 = \mathrm{Tr}_2 A = \sum_u (I \otimes \langle u|) A (I \otimes |u\rangle).$$

Remark The partial trace is the unique operation which gives rise to the correct description of observable quantities for subsystems of a composite system.

Theorem (Purification) Suppose $\rho_1 = \sum_{j=1}^n p_j |x_j\rangle\langle x_j| \in M_n$. Let $\{|y_1\rangle, \ldots, |y_n\rangle\}$ be an orthonormal basis of $\mathbf{C}^n$, and $|\psi\rangle = \sum_{j=1}^n \sqrt{p_j} |x_j\rangle \otimes |y_j\rangle$. Then $\mathrm{Tr}_2(|\psi\rangle\langle\psi|) = \rho_1$.

Exercise Suppose $\rho_1 = \mathrm{Tr}_2(|\psi_1\rangle\langle\psi_1|) = \mathrm{Tr}_2(|\psi_2\rangle\langle\psi_2|) \in M_m$ with $|\psi_1\rangle, |\psi_2\rangle \in \mathbf{C}^m \otimes \mathbf{C}^n$. Then there is a unitary $U \in M_n$ such that $|\psi_2\rangle = (I_m \otimes U)|\psi_1\rangle$.

An application: Quantum channels and quantum operations

A unitary time evolution of a closed system is determined by the quantum map $\mathcal{E}$ defined by

$$\mathcal{E}(\rho_S) = U(t)\rho_S U(t)^\dagger.$$

Here, ρ_S is the density matrix of a closed system at time $t = 0$ and $U(t)$ is the time evolution operator.

An open system is a system of interest (called the *principal system*) coupled with its environment. The total Hamiltonian is given by

$$H_T = H_S + H_E + H_{SE},$$

where H_S, H_E and H_{SE} are the system Hamiltonian, the environment Hamiltonian and their interaction Hamiltonian, respectively.

The state of the total system, which is assumed to be closed, will be described by ρ acting on the Hilbert space $\mathcal{H}_S \otimes \mathcal{H}_E$ such that one has the approximation $\rho(0) = \rho_S \otimes \rho_E$ and

$$\rho(t) = U(t)(\rho_S \otimes \rho_E)U(t)^\dagger \quad \text{for } t > 0.$$

For simplicity, we may assume that H_T is a constant matrix and $U(t) = e^{itH_T}$.

We study the system $(\mathcal{H}_S)$ by taking the *partial trace*

$$\rho_S(t) = \mathrm{Tr}_E[U(t)(\rho_S \otimes \rho_E)U(t)^\dagger] = \sum_{a \in J} (I_S \otimes \langle\varepsilon_a|)[U(t)(\rho_S \otimes \rho_E)U(t)^\dagger](I_S \otimes |\varepsilon_a\rangle)$$

for any complete orthonormal basis $\{|\varepsilon_a\rangle : a \in J\}$ for $\mathcal{H}_E$. We may assume $\rho_E = |\varepsilon_0\rangle\langle\varepsilon_0|$ by linearity or by purification. Let $E_a(t) = (I_S \otimes \langle\varepsilon_a|)U(t)(I_S \otimes |\varepsilon_0\rangle)$. Then

$$\rho_S(t) = \sum_a E_a(t)\rho_S E_a(t)^\dagger.$$

This is known as the *operator-sum representation* of the quantum operation. Note that

$$\begin{aligned}\sum_a E_a(t)^\dagger E_a(t) &= \sum_a (I_S \otimes \langle\varepsilon_0|)U(t)^\dagger (I_S \otimes |\varepsilon_a\rangle)(I_S \otimes \langle\varepsilon_a|)U(t)(I_S \otimes |\varepsilon_0\rangle) \\ &= (I_S \otimes \langle\varepsilon_0|)U(t)^\dagger (I_S \otimes I_E)U(t)(I_S \otimes |\varepsilon_0\rangle) = I_S.\end{aligned}$$

This is the *trace preserving* condition for the quantum operation. For certain quantum operations or channels, one may relax this condition.

7 EPR and the Bell inequality

Under the "real locality" theory of Einstein, Rosen, and Podolsky, Bell suggested the following inequality. Suppose a measurement of some quantity prepared by Charlie to that Alice measures Q or R, and Bob measures S and T, where Q, R, S, T each can assume the value 1 and -1, then $(Q + R)S = 0$ or $(R - Q)T = 0$ so that $QS + RS + RT - QT = \pm 2$ and

$$E(QS + RS + RT - QT) = E(QS) + E(RS) + E(RT) - E(QT) \leq 2.$$

However, in quantum world, Charlie prepares the state

$$|\psi\rangle = (|01\rangle - |10\rangle)/\sqrt{2}.$$

Alice performs the measurements $Q = \sigma_z$ and $R = \sigma_x$, where as Bob performs the measurements $S = -(\sigma_z + \sigma_x)/\sqrt{2}$ and $T = (\sigma_z - \sigma_x)/\sqrt{2}$. Then

$$\langle QS\rangle = 1/\sqrt{2}, \quad \langle RS\rangle = 1/\sqrt{2}, \quad \langle RT\rangle = 1/\sqrt{2}, \quad \langle QT\rangle = -1/\sqrt{2}$$

so that

$$\langle QS\rangle + \langle RS\rangle + \langle RT\rangle - \langle QT\rangle = 2\sqrt{2},$$

which is confirmed by experiment. So, the "real locality" theory does not apply to quantum mechanics.

References

[1] M. Nakahara and T. Ohmi, Quantum Computing: From Linear Algebra to Physical Realizations, CRS Press, New York, 2008.

[2] M. Nakahara, R. Rahimi, and A. SaiToh (editors), Mathematical Aspects of Quantum Computing 2007, World Scientific, Singapore, 2007.

[3] M.A. Nielsen and I.L. Chuang, Quantum Computation and Quantum Information, Cambridge University Press, Cambridge, 2000.