Solution Manual to

By Jia-Lin Chen, Nov, 2020

Quantum Computation & Quantum Information

By Micheal A. Nielsen & Isaac L. Chuang

Cambridge University Press, 10th Anniversary Edition

Contents

Chapter 1

Introduction and overview

1.1 Global perspectives

1.2 Quantum bits

1.3 Quantum computation

1.4 Quantum algorithm

Exercise 1.1. Suppose that the problem is not to distinguish between the constant and balanced functions *with certainty*, but rather, with some probability of error $\epsilon < 1/2$. What is the performance of the best classical algorithm for this problem?

Solution: Suppose Alice queries Bob j times, where $j \leq 2^{n-1}$. If she receives both 0(s) and 1(s), then she can determine $f(x)$ as balanced with certainty. If she receives all 0(s) (or all 1(s)), then according to Bayes' theorem and Law of total probability, the probability that $f(x)$ is constant is

$$\begin{aligned}
&P(f(x) \text{ is constant}|f(x) = 0\forall x = 0, 1, \cdots, j) = P(f(x) \text{ is constant}|f(x) = 1\forall x = 0, 1, \cdots, j) \\
=&\frac{P(f(x) \text{ is constant})P(f(x) = 0\forall j = 0, 1, \cdots, j|f(x) \text{ is constant})}{P(f(x) = 0\forall j = 0, 1, \cdots, j)} \\
=&\frac{P(f(x) \text{ is constant})P(f(x) = 0\forall j = 0, 1, \cdots, j|f(x) \text{ is constant})}{\begin{array}{l}P(f(x) \text{ is constant})P(f(x) = 0\forall j = 0, 1, \cdots, j|f(x) \text{ is constant}) \\ \quad + P(f(x) \text{ is balanced})P(f(x) = 0\forall j = 0, 1, \cdots, j|f(x) \text{ is balanced})\end{array}} \\
=&\frac{\frac{1}{2} \times \frac{1}{2}}{\frac{1}{2} \times \frac{1}{2} + \frac{1}{2} \times \frac{C_{2^{n-1}}^{2^n-k}}{C_{2^{n-1}}^{2^n}}} = \frac{1}{1 + 2\frac{(2^n-j)!(2^{n-1})!}{(2^n)!(2^{n-1}-j)!}} = \frac{(2^n-1)(2^n-2)\cdots(2^n-j+1)}{(2^n-1)(2^n-2)\cdots(2^n-j+1) + (2^{n-1}-1)(2^{n-1}-2)\cdots(2^{n-1}-j+1)}.
\end{aligned} \tag{1.1}$$

and the probability that $f(x)$ is balanced is

$$\begin{aligned}
&P(f(x) \text{ is balanced}|f(x) = 0\forall x = 0, 1, \cdots, j) = P(f(x) \text{ is balanced}|f(x) = 1\forall x = 0, 1, \cdots, j) \\
=&\frac{P(f(x) \text{ is balanced})P(f(x) = 0\forall j = 0, 1, \cdots, j|f(x) \text{ is balanced})}{P(f(x) = 0, 1, \cdots, j)} \\
=&\frac{P(f(x) \text{ is balanced})P(f(x) = 0\forall j = 0, 1, \cdots, j|f(x) \text{ is balanced})}{\begin{array}{l}P(f(x) \text{ is constant})P(f(x) = 0\forall x = 0, 1, \cdots, j|f(x) \text{ is constant}) \\ \quad + P(f(x) \text{ is balanced})P(f(x) = 0\forall x = 0, 1, \cdots, j|f(x) \text{ is balanced})\end{array}} \\
=&\frac{\frac{1}{2} \times \frac{C_{2^{n-1}}^{2^n-j}}{C_{2^{n-1}}^{2^n}}}{\frac{1}{2} \times \frac{1}{2} + \frac{1}{2} \times \frac{C_{2^{n-1}}^{2^n-k}}{C_{2^{n-1}}^{2^n}}} = \frac{2\frac{(2^n-k)!(2^{n-1})!}{(2^n)!(2^{n-1}-j)!}}{1 + 2\frac{(2^n-k)!(2^{n-1})!}{(2^n)!(2^{n-1}-j)!}} = \frac{(2^{n-1}-1)(2^{n-1}-2)\cdots(2^{n-1}-j+1)}{(2^n-1)(2^n-2)\cdots(2^n-j+1) + (2^{n-1}-1)(2^{n-1}-2)\cdots(2^{n-1}-j+1)}.
\end{aligned} \tag{1.2}$$

Note that

$$P(f(x) \text{ is constant}|f(x) = 0\forall x = 0, 1, \cdots, j) \geq P(f(x) \text{ is balanced}|f(x) = 0\forall x = 0, 1, \cdots, j) \tag{1.3}$$

with equality if and only if $j = 1$. Therefore, the best classical algorithm is:

1. Alice selects number $x = 0$ and queries Bob;
2. i. If Alice have received both 0(s) and 1(s), then she determines $f(x)$ as balanced;
   ii. If Alice have received either all 0(s) or 1(s), and $P(f(x) \text{ is constant}|f(x) = 0 \forall x = 0, 1, \cdots, j) > \frac{1}{2}$ or $P(f(x) \text{ is constant}|f(x) = 1 \forall x = 0, 1, \cdots, j) > \frac{1}{2}$, then she classifies $f(x)$ as constant;
   iii. Otherwise, let $x = x + 1$ and return to step 1.

In the worst case, Alice queries Bob (say) k times and receives k 0s. She classifies $f(x)$ as constant since the error probability of her judgement is just below 1/2:

$$\begin{aligned} P(\text{error}|f(x) = 0 \forall x = 0, 1, \cdots, k-1) =& P(f(x) \text{ is balanced}|f(x) = 0 \forall x = 0, 1, \cdots, k-1) \\ =& \frac{(2^{n-1}-1)(2^{n-1}-2)\cdots(2^{n-1}-k+1)}{(2^n-1)(2^n-2)\cdots(2^n-k+1)+(2^{n-1}-1)(2^{n-1}-2)\cdots(2^{n-1}-k+1)} < \frac{1}{2}. \end{aligned} \tag{1.4}$$

If Alice only queries Bob $(k-1)$ times and receives $(k-1)$ 0s, she can not make judgement with error probability less than 1/2:

$$\begin{aligned} P(\text{error}|f(x) = 0 \forall x = 0, 1, \cdots, k-2) =& P(f(x) \text{is balanced}|f(x) = 0 \forall x = 0, 1, \cdots, k-2) \\ \geq& \frac{(2^{n-1}-1)(2^{n-1}-2)\cdots(2^{n-1}-k+2)}{(2^n-1)(2^n-2)\cdots(2^n-k+2)+(2^{n-1}-1)(2^{n-1}-2)\cdots(2^{n-1}-k+2)} \geq \frac{1}{2}, \end{aligned} \tag{1.5}$$

so we have

$$k = 2. \tag{1.6}$$

i.e., the computation complexity of the best classical algorithm for this problem is $O(1)$. □

1.5 Experimental quantum information processing

Exercise 1.2. Explain how a device which, upon input of one of two non-orthogonal quantum states $|\psi\rangle$ or $|\varphi\rangle$ correctly identified the state, could be used to build a theorem which cloned the states $|\psi\rangle$ and $|\varphi\rangle$, in violation of the no-cloning theorem. Conversely, explain how device for cloning could be used to distinguish non-orthogonal quantum states.

Proof: The no-cloning theorem says that it is impossible for a cloning device to clone states that are not orthogonal to one another. If we have a device that can distinguish two non-orthogonal quantum states $|\psi\rangle$ and $|\varphi\rangle$, given a state either $|\phi\rangle$ or $|\varphi\rangle$, we can use this device to know whether the state is $|\phi\rangle$ and $|\varphi\rangle$, then find a another system and initialize it to the same state with not too much difficulty. In this way, we can clone states that are non-orthogonal, which violate the no-cloning theorem.

Conversely, if we have a device for cloning quantum states, given an unknown state, we can use this device to generate copies of the state and measure it as many times as we want, so that we will obtain enough information to determine the state. In this way, we can distinguish non-orthogonal quantum states. □

1.6 Quantum information

Chapter 2

Introduction to quantum mechanics

2.1 Linear algebra

Exercise 2.1 (Linear dependence: example). Show that $(1,-1)$, $(1,2)$ and $(2,1)$ are Linearly dependent.

Proof: Since

$$(1,-1)+(1,2)-(2,1)=0, \tag{2.1}$$

these three vectors are linearly dependent. □

Exercise 2.2 (Matrix representations: example). Suppose V is a vector space with basis vectors $|0\rangle$ and $|1\rangle$, and A is a linear operator from V to V such that $A|0\rangle = |1\rangle$ and $A|1\rangle = |0\rangle$. Give a matrix representation for A, with respect to the input basis $|0\rangle$, $|1\rangle$, and the output basis $|0\rangle$ and $|1\rangle$. Find input and output bases which give rise to a different matrix representation of A.

Solution: The matrix representation for A with respect to the input basis $|0\rangle$, $|1\rangle$ and the output basis $|0\rangle$ and $|1\rangle$ is

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}. \tag{2.2}$$

Keep $|0\rangle$ and $|1\rangle$ as the input basis and choose $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ and $|-\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$ as the output basis, then A can be regarded as a linear operator from V to V such that $A|0\rangle = (|+\rangle - |-\rangle)/\sqrt{2}$ and $A|1\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$. In this way, the matrix representation for A is

$$\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}. \tag{2.3}$$

□

Exercise 2.3 (Matrix representation for operator products). Suppose A is a linear operator from vector space V to vector space W, and B is a linear operator from vector spaces W to vector space X. Let $|v_i\rangle$, $|w_j\rangle$, and $|x_k\rangle$ be bases for the vector spaces V, W, and X, respectively. Show that the matrix representation for the linear transformation BA is the matrix product of the matrix representations for B and A, with respect to the appropriate bases.

Proof: Suppose the dimension of vector spaces V, W, and W are l, m, n, respectively. Since A is a linear operator from vector space V to vector space W, for each i in the range $1, \cdots, l$, there exist complex numbers A_{1i} through A_{mi} such that

$$A|v_i\rangle = \sum_j A_{ji}|w_j\rangle, \tag{2.4}$$

where A_{ji} is the entries of the matrix representation of the operator A. Since B is a linear operator from vector space W to vector space X, for each j in the range $1, \cdots, m$, there exist complex numbers A_{1j} through A_{nj} such that

$$B|w_j\rangle = \sum_k B_{kj}|x_k\rangle, \tag{2.5}$$

where B_{kj} is the entries of the matrix representation of the operator B. Putting the above two equations together, we have

$$BA|v_i\rangle = B\sum_j A_{ji}|w_j\rangle = \sum_j A_{ji}\sum_k B_{kj}|x_k\rangle = \sum_k \left(\sum_j B_{kj}A_{ji}\right)|x_k\rangle = \sum_k (BA)_{ki}|x_k\rangle. \tag{2.6}$$

where $(AB)_{ki}$ is the entries of the matrix representation of the operator BA. Therefore, the matrix representation for the linear transformation BA is the matrix product of the matrix representations for B and A, with respect to the appropriate bases. □

Exercise 2.4 (Matrix representation for identity). Show that the identity operator on a vector space V has a matrix representation which is one along the diagonal and zero everywhere else, if the matrix representation is taken with respect to the same input and output bases. The matrix is known as the *identity matrix*.

Proof: Suppose I is the identity operator on the vector space V and choose $|v_1\rangle, \cdots, |v_m\rangle$ as both the input basis and the output basis for V. Then for each j in the range $1, \cdots, m$, there exist complex numbers I_{1j} through I_{mj} such that

$$I|v_j\rangle = \sum_i A_{ij}|v_i\rangle = |v_j\rangle, \tag{2.7}$$

$$\Longrightarrow (A_{jj}-1)|v_j\rangle + \sum_{i\neq j} A_{ij}|v_i\rangle = 0, \tag{2.8}$$

where A_{ij} is the entries of the matrix representation of the operator I. Due to linear independence of the basis, there must be

$$v_{ij} = \begin{cases} 1, & i=j; \\ 0, & i\neq j, \end{cases} = \delta_{ij}, \quad \forall i,j = 1, \cdots, m. \tag{2.9}$$

i.e., the identity operator on the vector space V has a matrix representation which is one along the diagonal and zero everywhere else. □

Exercise 2.5. Verify that $(\cdot,\cdot)$ just defined is an inner product on $\mathbb{C}^n$.

Proof: $(\cdot,\cdot)$ just defined satisfies the requirements that:

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

$$\begin{aligned}\left((y_1,\cdots,y_n), \sum_j \lambda_j (z_1^{(j)},\cdots,z_n^{(j)})\right) &= \left((y_1,\cdots,y_n), (\sum_j \lambda_j z_1^{(j)},\cdots,\sum_j \lambda_j z_n^{(j)})\right) = \sum_i y_i^* \sum_j \lambda_j z_i^{(j)} \\ &= \sum_j \lambda_j \sum_i y_i^* z_i^{(j)} = \sum_j \lambda_j ((y_1,\cdots,y_n),(z_1^{(j)},\cdots,z_n^{(j)})).\end{aligned} \tag{2.10}$$

(2)

$$((y_1,\cdots,y_n),(z_1,\cdots,z_n)) = \sum_i y_i^* z_i = \left(\sum_i z_i^* y_i\right)^* = ((z_1,\cdots,z_n),(y_1,\cdots,y_n))^*. \tag{2.11}$$

(3)

$$((y_1,\cdots,y_n),(y_1,\cdots,y_n)) = \sum_i y_i^* y_i = \sum_i |y_i|^2 \geq 0, \tag{2.12}$$

with equality if and only if $(y_1,\cdots,y_n) = 0$.

Therefore, $(\cdot,\cdot)$ is an inner product on $\mathbb{C}^n$. □

Exercise 2.6. Show that any inner product $(\cdot,\cdot)$ is conjugate-linear in the first argument,

$$\left(\sum_i \lambda_i |w_i\rangle, |v\rangle\right) = \sum_i \lambda_i^* (|w_i\rangle, |v\rangle). \tag{2.13}$$

Proof:

$$\left(\sum_i \lambda_i|w_i\rangle, |v\rangle\right) = \left(|v\rangle, \sum_i \lambda_i|w_i\rangle\right)^* = \left(\sum_i \lambda_i(|v\rangle, |w_i\rangle)\right)^* = \sum_i \lambda_i^*(|w_i\rangle, |v\rangle). \tag{2.14}$$

Therefore, any inner product $(\cdot,\cdot)$ is conjugate-linear in the first argument. □

Exercise 2.7. Verify that $|w\rangle \equiv (1,1)$ and $|v\rangle \equiv (1,-1)$ are orthogonal. What are the normalized forms of these vectors?

Solution: Since

$$(|w\rangle, |v\rangle) = 1 \times 1 + 1 \times (-1) = 0, \tag{2.15}$$

these two vectors are orthogonal. The normalized form of $|w\rangle$ and $|v\rangle$ are

$$\frac{|w\rangle}{\||w\rangle\|} = \frac{1}{\sqrt{2}}(1,1), \tag{2.16}$$

and

$$\frac{|v\rangle}{\||v\rangle\|} = \frac{1}{\sqrt{2}}(1,-1), \tag{2.17}$$

respectively. □

Exercise 2.8. Prove that the Gram-Schmidt procedure produces an orthonormal basis for V.

Proof: Obviously, $|v_1\rangle, \cdots, |v_n\rangle$ are normalized, so we first prove that $|v_1\rangle, \cdots, |v_d\rangle$ are orthogonal with induction. For $k = 1$,

$$(|v_1\rangle, |v_{k+1}\rangle) = (|v_1, |v_2\rangle\rangle) = \left(|v_1\rangle, \frac{|w_2\rangle - \langle v_1|w_2\rangle|v_1\rangle}{\||v_1\rangle - \langle v_1|w_2\rangle|v_1\rangle\|}\right) = \frac{\langle v_1|w_2\rangle - \langle v_1|w_2\rangle\langle v_1|v_1\rangle}{\||v_1\rangle - \langle v_1|w_2\rangle|v_1\rangle\|} = \frac{\langle v_1|w_2\rangle - \langle v_1|w_2\rangle}{\||v_1\rangle - \langle v_1|w_2\rangle|v_1\rangle\|} = 0, \tag{2.18}$$

$$\tag{2.19}$$

so $|v_1\rangle, \cdots, |v_{k+1}\rangle$ are orthogonal for $k = 1$. For $k \geq 2$, if $|v_1\rangle, \cdots, |v_k\rangle$ are orthogonal, then

$$\begin{aligned}(|v_j\rangle, |v_{k+1}\rangle) &= \left(|v_j\rangle, \frac{|w_{k+1}\rangle - \sum_{i=1}^k \langle v_i|w_{k+1}\rangle|v_i\rangle}{\left\||w_{k+1}\rangle - \sum_{i=1}^k \langle v_i|w_{k+1}\rangle|v_i\rangle\right\|}\right) = \frac{\langle v_j|w_{k+1}\rangle - \sum_{i=1}^k \langle v_i|w_{k+1}\rangle\langle v_j|v_i\rangle}{\left\||w_{k+1}\rangle - \sum_{i=1}^k \langle v_i|w_{k+1}\rangle|v_i\rangle\right\|} \\ &= \frac{\langle v_j|w_{k+1}\rangle - \sum_{i=1}^k \langle v_i|w_{k+1}\rangle\delta_{ji}}{\left\||w_{k+1}\rangle - \sum_{i=1}^k \langle v_i|w_{k+1}\rangle|v_i\rangle\right\|} = \frac{\langle v_j|w_{k+1}\rangle - \langle v_j|w_{k+1}\rangle}{\left\||w_{k+1}\rangle - \sum_{i=1}^k \langle v_i|w_{k+1}\rangle|v_i\rangle\right\|} = 0, \quad \forall j = 1, \cdots, k,\end{aligned} \tag{2.20}$$

so $|v_1\rangle, \cdots, |v_{k+1}\rangle$ are orthogonal for $k \geq 2$, such as $k = d-1$. Till now, we proved that $|v_1\rangle, \cdots, |v_d\rangle$ are orthonormal.

We then prove that $|v_1\rangle, \cdots, |v_n\rangle$ are a basis for V. As proved above, $|v_1\rangle, \cdots, |v_d\rangle$ are orthonormal, and thus linear independent. Since $|w_1\rangle, \cdots, |w_d\rangle$ are a basis for V. Any vector $|v\rangle$ in V can be written as a linear combination of $|w_1\rangle, \cdots, |w_d\rangle$:

$$|v\rangle = \sum_{i=1}^d a_i|w_i\rangle. \tag{2.21}$$

Using

$$|v_1\rangle = \frac{|w_1\rangle}{\||w_1\rangle\|} \Longrightarrow |w_1\rangle = \||w_1\rangle\|\,|v_1\rangle, \tag{2.22}$$

$$\tag{2.23}$$

and

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

$$\Longrightarrow |w_{k+1}\rangle = \left\| |w_{k+1}\rangle - \sum_{i=1}^{k} \langle v_i|w_{k+1}\rangle |v_i\rangle \right\| |v_{k+1}\rangle + \sum_{i=1}^{k} \langle v_i|w_{k+1}\rangle |v_i\rangle, \quad \forall k = 1, \cdots, d-1 \tag{2.24}$$

we can rewrite $|v\rangle$ as a linear combination of $|v_1\rangle, \cdots, |v_d\rangle$:

$$|v\rangle = a_1 \||w_1\rangle\| \, |v_1\rangle + \sum_{k=1}^{d-1} a_{k+1} \left(\left\| |w_{k+1}\rangle - \sum_{i=1}^{k} \langle v_i|w_{k+1}\rangle |v_i\rangle \right\| |v_{k+1}\rangle + \sum_{i=1}^{k} \langle v_i|w_{k+1}\rangle |v_i\rangle \right), \tag{2.25}$$

so $|v_1\rangle, \cdots, |v_d\rangle$ span and form a basis for V.

Therefore, the Gram-Schmidt procedure produces an orthonormal basis for V. □

Exercise 2.9 (Pauli operators and the outer product). The Pauli matrices (Figure 2.2 on page 65) can be considered as operators with respect to an orthonormal basis $|0\rangle$, $|1\rangle$ for a two-dimensional Hilbert space. Express each of the Pauli operators in the outer product notion.

Solution: The Pauli operators in the outer product notion:

$$\sigma_0 = \sum_{m,n=1}^{2} \langle m|\sigma_0|n\rangle |m\rangle\langle n| = |0\rangle\langle 0| + |1\rangle\langle 1|, \tag{2.26}$$

$$\sigma_1 = \sum_{m,n=1}^{2} \langle m|\sigma_1|n\rangle |m\rangle\langle n| = |0\rangle\langle 1| + |1\rangle\langle 0|, \tag{2.27}$$

$$\sigma_2 = \sum_{m,n=1}^{2} \langle m|\sigma_2|n\rangle |m\rangle\langle n| = -i|0\rangle\langle 1| + i|1\rangle\langle 0|, \tag{2.28}$$

$$\sigma_3 = \sum_{m,n=1}^{2} \langle m|\sigma_3|n\rangle |m\rangle\langle n| = |0\rangle\langle 0| - |1\rangle\langle 1|. \tag{2.29}$$

□

Exercise 2.10. Suppose $|v_i\rangle$ is an orthonormal basis for and inner product space V. What is the matrix representation for the operator $|v_j\rangle\langle v_k|$, with respect to the $|v_i\rangle$ basis?

Solution: The matrix representation for the operator $|v_j\rangle\langle v_k|$ with respect to the $|v_i\rangle$ basis:

$$|v_j\rangle\langle v_k| = \begin{matrix} \\ \\ \\ \\ \\ \\ \\ \end{matrix} \overset{\begin{matrix} k\text{th column} \\ \downarrow \end{matrix}}{\begin{bmatrix} 0 & & & & & & \\ & \ddots & & & 0 & & \\ & & 0 & & & & \\ & & & 1 & & & \\ & & & & 0 & & \\ & 0 & & & & \ddots & \\ & & & & & & 0 \end{bmatrix}} \begin{matrix} \\ \\ \\ \leftarrow j\text{th row} \\ \\ \\ \\ \end{matrix} \tag{2.30}$$

(a matrix with all zeros expect a one at the jth row and kth column). □

Exercise 2.11 (Eigendecomposition of the Pauli matrices). Find the eigenvectors, eigenvalues, and diagonal representations of the Pauli matrices X, Y, and Z.

Solution: The characteristic equation of X

$$\det |X - \lambda I| = \begin{vmatrix} -\lambda & 1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 - 1 = 0, \tag{2.31}$$

gives the eigenvalues

$$\lambda_1 = 1, \quad \lambda_2 = -1. \tag{2.32}$$

The eigenequations of X

$$X|v_1\rangle = \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}a_1\\ b_1\end{bmatrix} = \begin{bmatrix}b_1\\ a_1\end{bmatrix} = \lambda_1|v_1\rangle = \begin{bmatrix}a_1\\ b_1\end{bmatrix}, \quad X|v_2\rangle = \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}a_2\\ b_2\end{bmatrix} = \begin{bmatrix}b_2\\ a_2\end{bmatrix} = \lambda_2|v_2\rangle = \begin{bmatrix}-a_2\\ -b_2\end{bmatrix}, \tag{2.33}$$

give the corresponding eigenvectors

$$|v_1\rangle = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\ 1\end{bmatrix} = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \quad |v_2\rangle = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\ -1\end{bmatrix} = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle). \tag{2.34}$$

The diagonal representation of X is

$$X = \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix} = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)\frac{1}{\sqrt{2}}(\langle 0| + \langle 1|) - \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)\frac{1}{\sqrt{2}}(\langle 0| - \langle 1|). \tag{2.35}$$

The characteristic equation of Y

$$\det|Y - \lambda I| = \begin{vmatrix}-\lambda & -i\\ i & -\lambda\end{vmatrix} = \lambda^2 - 1 = 0, \tag{2.36}$$

gives the eigenvalues

$$\lambda_1 = 1, \quad \lambda_2 = -1. \tag{2.37}$$

The eigenequations of Y

$$Y|v_1\rangle = \begin{bmatrix}0 & -i\\ i & 0\end{bmatrix}\begin{bmatrix}a_1\\ b_1\end{bmatrix} = \begin{bmatrix}-ib_1\\ ia_1\end{bmatrix} = \lambda_1|v_1\rangle = \begin{bmatrix}a_1\\ b_1\end{bmatrix}, \quad Y|v_2\rangle = \begin{bmatrix}0 & -i\\ i & 0\end{bmatrix}\begin{bmatrix}a_2\\ b_2\end{bmatrix} = \begin{bmatrix}-ib_2\\ ia_2\end{bmatrix} = \lambda_2|v_2\rangle = \begin{bmatrix}-a_2\\ -b_2\end{bmatrix}, \tag{2.38}$$

give the corresponding eigenvectors

$$|v_1\rangle = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\ i\end{bmatrix} = \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle), \quad |v_2\rangle = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\ -i\end{bmatrix} = \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle). \tag{2.39}$$

The diagonal representation of Y is

$$Y = \begin{bmatrix}0 & -i\\ i & 0\end{bmatrix} = \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle)\frac{1}{\sqrt{2}}(\langle 0| + i\langle 1|) - \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle)\frac{1}{\sqrt{2}}(\langle 0| - i\langle 1|). \tag{2.40}$$

The characteristic equation of Z

$$\det|Z - \lambda I| = \begin{vmatrix}1-\lambda & 0\\ 0 & -1-\lambda\end{vmatrix} = \lambda^2 - 1 = 0, \tag{2.41}$$

gives the eigenvalues

$$\lambda_1 = 1, \quad \lambda_2 = -1. \tag{2.42}$$

The eigenequations of Z

$$Z|v_1\rangle = \begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}\begin{bmatrix}a_1\\ b_1\end{bmatrix} = \begin{bmatrix}a_1\\ -b_1\end{bmatrix} = \lambda_1|v_1\rangle = \begin{bmatrix}a_1\\ b_1\end{bmatrix}, \quad Z|v_2\rangle = \begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}\begin{bmatrix}a_2\\ -b_2\end{bmatrix} = \lambda_2|v_2\rangle = \begin{bmatrix}-a_2\\ -b_2\end{bmatrix}, \tag{2.43}$$

give the corresponding eigenvectors

$$|v_1\rangle = \begin{bmatrix}1\\ 0\end{bmatrix} = |0\rangle, \quad |v_2\rangle = \begin{bmatrix}0\\ 1\end{bmatrix}. \tag{2.44}$$

The diagonal representation of Z is

$$Z = \begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix} = |0\rangle\langle 0| - |1\rangle\langle 1|. \tag{2.45}$$

□

Exercise 2.12. Prove that the matrix

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \tag{2.46}$$

is not diagonalizable.

Proof: The matrix is not normal,

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}^\dagger = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \neq \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}^\dagger \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}, \tag{2.47}$$

so it is not diagonalizable. □

Exercise 2.13. If $|w\rangle$ and $|v\rangle$ are any two vectors, show that $(|w\rangle\langle v|)^\dagger = |v\rangle\langle w|$.

Proof:

$$(|w\rangle\langle v|)^\dagger = \langle v|^\dagger |w\rangle^\dagger = |v\rangle\langle w|. \tag{2.48}$$

□

Exercise 2.14 (Anti-linearity of the adjoint). Show tha the adjoint operation is anti-linear,

$$\left(\sum_i a_i A_i\right)^\dagger = \sum_i a_i^* A_i^\dagger. \tag{2.49}$$

Proof: For two arbitrary vectors $|v\rangle$ and $|w\rangle$,

$$\left(\left(\sum_i a_i A_i\right)^\dagger |v\rangle, |w\rangle\right) = \left(|v\rangle, \sum_i a_i A_i |w\rangle\right) = \sum_i a_i (|v\rangle, A_i|w\rangle) = \sum_i a_i (A_i^\dagger |v\rangle, |w\rangle) = \left(\sum_i a_i^* A_i^\dagger |v\rangle, |w\rangle\right). \tag{2.50}$$

Due to the arbitrariness of $|v\rangle$ and $|w\rangle$,

$$\left(\sum_i a_i A_i\right)^\dagger = \sum_i a_i^* A_i^\dagger. \tag{2.51}$$

□

Exercise 2.15. Show that $(A^\dagger)^\dagger = A$.

Proof: For two arbitrary vectors $|v\rangle$ and $|w\rangle$,

$$((A^\dagger)^\dagger |v\rangle, |w\rangle) = (|v\rangle, A^\dagger |w\rangle) = (A^\dagger |w\rangle, |v\rangle)^* = (|w\rangle, A|v\rangle)^* = [(A|v\rangle, |w\rangle)^*]^* = (A|v\rangle, |w\rangle). \tag{2.52}$$

Due to the arbitrariness of $|v\rangle$ and $|w\rangle$,

$$(A^\dagger)^\dagger = A. \tag{2.53}$$

□

Exercise 2.16. Show that any projector P satisfies the equation $P^2 = P$.

Proof: For any orthonormal basis $|1\rangle, \cdots, |k\rangle$ for W,

$$P = \sum_{i=1}^k |i\rangle\langle i|, \tag{2.54}$$

and then

$$P^2 = \sum_{i=1}^k |i\rangle\langle i| \sum_{j=1}^k |j\rangle\langle j| = \sum_{i=1}^k \sum_{j=1}^k |i\rangle\langle i|j\rangle\langle j| = \sum_{i=1}^k \sum_{j=1}^k |i\rangle \delta_{ij} \langle j| = \sum_{i=1}^k |i\rangle\langle i| = P. \tag{2.55}$$

□

Exercise 2.17. Show that a normal matrix is Hermitian if and only if it has real eigenvalues.

Proof: *Sufficiency*: If normal matrix A has real eigenvalues $\lambda_1, \cdots, \lambda_n$ with corresponding eigenvectors $|1\rangle, \cdots, |n\rangle$. It can be written as

$$A = \sum_{i=1}^{n} \lambda_i |i\rangle\langle i|. \tag{2.56}$$

Since $\lambda_1, \cdots, \lambda_n$ are real,

$$A^\dagger = \sum_{i=1}^{n} \lambda_i^\dagger |i\rangle\langle i| = \sum_{i=1}^{n} \lambda_i |i\rangle\langle i| = A. \tag{2.57}$$

Therefore, A is Hermitian.

Necessity: Suppose normal and Hermitian matrix A has eigenvalues $\lambda_1, \cdots, \lambda_n$ with corresponding eigenvectors $|1\rangle, \cdots, |n\rangle$. For any $i = 1, \cdots, n$,

$$\begin{aligned}\lambda_i = \lambda_i(|i\rangle, |i\rangle) = (|i\rangle, \lambda_i|i\rangle) =& (|i\rangle, A|i\rangle) \\ =& (A^\dagger|i\rangle, |i\rangle) = (|i\rangle, A^\dagger|i\rangle)^* = ((A^\dagger)^\dagger|i\rangle, |i\rangle) = (A|i\rangle, |i\rangle) = (\lambda_i|i\rangle, |i\rangle) = \lambda_i^*(|i\rangle, |i\rangle) = \lambda_i^*.\end{aligned} \tag{2.58}$$

Therefore, all the eigenvalues $\lambda_1, \cdots, \lambda_n$ are real. □

Exercise 2.18. Show that all eigenvalues of a unitary matrix has modulus 1, that is can be written in the form $e^{i\theta}$ for some real θ.

Proof: Suppose unitary matrix A has eigenvalues $\lambda_1, \cdots, \lambda_n$ with corresponding eigenvectors $|1\rangle, \cdots, |n\rangle$. The eigenequations of A are

$$A|i\rangle = \lambda_i|i\rangle, \quad \forall i = 1, \cdots, n. \tag{2.59}$$

Taking Hermitian conjugate of the above equations,

$$\langle i|A^\dagger = (A|i\rangle)^\dagger = \lambda_i^*\langle i|, \quad \forall i = 1, \cdots, n. \tag{2.60}$$

Since A is unitary, for any $i = 1, \cdots, n$,

$$1 = \langle i|i\rangle = \langle i|I|i\rangle = \langle i|A^\dagger A|i\rangle = |\lambda_i|^2 \langle i|i\rangle = |\lambda_i|^2, \tag{2.61}$$

$$\implies |\lambda_i| = 1. \tag{2.62}$$

Therefore, all eigenvalues of a unitary matrix has modulus 1. □

Exercise 2.19 (Pauli matrices: Hermitian and unitary). Show that the Pauli matrices are Hermitian and unitary.

Proof: Since

$$\sigma_0^\dagger = I^\dagger = I = \sigma_0, \tag{2.63}$$

and

$$\sigma_0^\dagger \sigma_0 = I^\dagger I = II = I, \tag{2.64}$$

σ_0 is Hermitian and unitary. Since

$$\sigma_1^\dagger = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \sigma_1, \tag{2.65}$$

and

$$\sigma_1^\dagger \sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I, \tag{2.66}$$

σ_1 is Hermitian and unitary. Since

$$\sigma_2^\dagger = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} = \sigma_2, \tag{2.67}$$

and

$$\sigma_2^\dagger \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I, \tag{2.68}$$

σ_2 is Hermitian and unitary. Since

$$\sigma_3^\dagger = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \sigma_3, \tag{2.69}$$

and

$$\sigma_3^\dagger \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I, \tag{2.70}$$

σ_3 is Hermitian and unitary. □

Exercise 2.20 (Basis changes). Suppose A' and A'' are matrix representations of an operator A on a vector space V with respect to two different orthonormal bases, $|v_i\rangle$ and $|w_i\rangle$. Then the elements of A' and A'' are $A'_{ij} = \langle v_i|A|v_j\rangle$ and $A''_{ij} = \langle w_i|A|w_j\rangle$. characterize the relationship between A' and A''.

Solution: Define $U = \sum_i |w_i\rangle\langle v_i|$ so that $|w_i\rangle = U|v_i\rangle$ and $\langle w_i| = \langle v_i|U^\dagger \ \forall i$. Since $|v_i\rangle$ and $|w_i\rangle$ are two bases for V and

$$A''_{ij} = \langle w_i|A''|w_j\rangle = \langle v_i|U^\dagger A' U|v_j\rangle, \tag{2.71}$$

the relationship between A' and A'' is

$$A'' = U^\dagger A' U. \tag{2.72}$$

□

Exercise 2.21. Repeat the proof of the spectral decomposition in Box 2.2 for the case when M is Hermitian, simplifying the proof wherever possible.

Proof: *Forward*: Suppose vector space V is d-dimensional. The case $d = 1$ is trivial. For the case $d \geq 2$, let λ be an eigenvalue of M, P the projector onto the λ eigenspace, and Q the projector onto the orthogonal component. The $M = IMI = (P+Q)M(P+Q) = PMP + PMQ + QMP + QMQ = PMP + QMQ$, where $PMP = \lambda P$, i.e., PMP is diagonal, and $QMQQM^\dagger Q = QM^\dagger QQMQ$, i.e., QMQ is normal. By induction, QMQ is diagonal with respect to some orthonormal basis for the subspace Q. It follows that $M = PMP + QMQ$ is diagonal with respect to some orthonormal basis for the total vector space.

Converse: holds only if all the eigenvalues of M are real. Suppose M is diagonalizable with respect to an orthonormal basis $|i\rangle$ for V, i.e.,

$$M = \sum_i \lambda_i |i\rangle\langle i|, \tag{2.73}$$

where λ_i are the eigenvalues of M. Since M is Hermitian, all λ_i are real,

$$M^\dagger = \left(\sum_i \lambda_i |i\rangle\langle i| \right)^\dagger = \sum_i \lambda_i^* |i\rangle\langle i| = \sum_i \lambda_i |i\rangle\langle i|, \tag{2.74}$$

so M is Hermitian. □

Exercise 2.22. Prove that two eigenvectors of a Hermitian operators with different eigenvalues are necessarily orthogonal.

Proof: Suppose Hermitian operator A has two eigenvectors $|v_1\rangle$ and $|v_2\rangle$ corresponding to two different eigenvalues λ_1 and λ_2, i.e.,

$$A|v_1\rangle = \lambda_1 |v_1\rangle, \tag{2.75}$$
$$A|v_2\rangle = \lambda_2 |v_2\rangle. \tag{2.76}$$

Then,

$$\langle v_1|A|v_2\rangle = \lambda_1 \langle v_1|v_2\rangle = \lambda_2 \langle v_1|v_2\rangle. \tag{2.77}$$

Since $\lambda_1 \neq \lambda_2$, the above equation holds only if

$$\langle v_1|v_2\rangle = 0, \tag{2.78}$$

i.e., $|v_1\rangle$ and $|v_2\rangle$ are orthogonal. □

Exercise 2.23. Show that the eigenvalues of a projector P are all either 0 or 1.

Proof: Suppose $|1\rangle, \cdots, |k\rangle$ is an orthonormal basis for the subspace P and $|1\rangle, \cdots, |n\rangle$ is an orthonormal basis for the total vector space, where $k \leq n$. It is easy to see that $|1\rangle, \cdots, |n\rangle$ are eigenvectors of projector $P = \sum_{i=1}^{k} |i\rangle\langle i|$:

$$P|j\rangle = \sum_{i=1}^{k} |i\rangle\langle i|j\rangle = \sum_{i=1}^{k} |i\rangle\delta_{ij} = \begin{cases} |j\rangle, & \text{if } j = 1, \cdots, k; \\ 0, & \text{if } j = k+1, \cdots, n. \end{cases} \tag{2.79}$$

Therefore, the eigenvalues of a projector P are all either 0 or 1. □

Exercise 2.24 (Hermiticity of positive operators). Show that a positive operator is necessarily Hermitian. (*Hint*: Show that an arbitrary operator A can be written $A = B + iC$ where B and C are Hermitian.)

Proof: An arbitrary positive operator A can be written $A = B + iC$ where all the entries of $B = \frac{A+A^\dagger}{2}$ and $C = \frac{A-A^\dagger}{2i}$ are Hermitian. Since A is a positive operator, for any vector $|v\rangle$,

$$\langle v|A|v\rangle = \langle v|(B + iC)|v\rangle = \langle v|B|v\rangle + i\langle v|C|v\rangle \geq 0, \tag{2.80}$$

and is real. In this way, $\langle v|C|v\rangle$ can only be either purely imaginary or zero. Since C is Hermitian, it is diagonalizable and has a diagonal representation

$$C = \sum_i \lambda_i |i\rangle\langle i|, \tag{2.81}$$

where λ_i are the eigenvalues of C and real, and $|i\rangle$ is an orthonormal basis. Any vector $|v\rangle$ can be written as a linear combination of $|i\rangle$,

$$|v\rangle = \sum_i a_i |i\rangle. \tag{2.82}$$

Hence

$$\langle v|C|v\rangle = \sum_i a_i^* \langle i| \sum_j \lambda_j |j\rangle\langle j| \sum_k a_k |k\rangle = \sum_{i,j,k} \lambda_j a_i^* a_k \langle i|j\rangle\langle j|k\rangle = \sum_{ijk} \lambda_j a_i^* a_k \delta_{ij}\delta_{jk} = \sum_i \lambda_i |a_i|^2 \tag{2.83}$$

can not be purely imaginary and only be zero. Therefore, $A = B$ and is Hermitian. □

Exercise 2.25. Show that for any operator A, $A^\dagger A$ is positive.

Proof: For any vector $|v\rangle$,

$$(|v\rangle, A^\dagger A|v\rangle) = ((A^\dagger A)^\dagger |v\rangle, |v\rangle) = (A^\dagger A|v\rangle, |v\rangle) = (|v\rangle, A^\dagger A|v\rangle)^*, \tag{2.84}$$

so $(|v\rangle, A^\dagger A|v\rangle)$ is real. Besides,

$$(|v\rangle, A^\dagger A|v\rangle) = ((A^\dagger)^\dagger |v\rangle, A|v\rangle) = (A|v\rangle, A|v\rangle) \geq 0. \tag{2.85}$$

Therefore, for any operator A, $A^\dagger A$ is positive. □

Exercise 2.26. Let $|\psi\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$. Write out $|\psi\rangle^{\otimes 2}$ and $|\psi\rangle^{\otimes 3}$ explicitly, both in terms of tensor products like $|0\rangle|1\rangle$, and using the Kronecker product.

Solution:

$$|\psi\rangle^{\otimes 2} = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) = \frac{1}{2}(|0\rangle|0\rangle + |0\rangle|1\rangle + |1\rangle|0\rangle + |1\rangle|1\rangle) \tag{2.86}$$

$$= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \otimes \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \tag{2.87}$$

$$|\psi\rangle^{\otimes 3} = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) = \frac{1}{2\sqrt{2}}(|0\rangle|0\rangle|0\rangle + |0\rangle|0\rangle|1\rangle + |0\rangle|1\rangle|0\rangle + |0\rangle|1\rangle|1\rangle$$
$$+ |1\rangle|0\rangle|0\rangle + |1\rangle|0\rangle|1\rangle + |1\rangle|1\rangle|0\rangle + |1\rangle|1\rangle|1\rangle) \tag{2.88}$$

$$=\frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}\otimes\frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}\otimes\frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}=\frac{1}{2\sqrt{2}}\begin{bmatrix}1\\1\\1\\1\\1\\1\\1\\1\end{bmatrix}. \tag{2.89}$$

□

Exercise 2.27. Calculate the matrix representation of the tensor products of the Pauli operators (a) X and Z; (b) I and X; (c) X and I. Is the tensor product commutative?

Solution: (a) The matrix representation of the tensor product of X and Z:

$$X\otimes Z=\begin{bmatrix}0&1\\1&0\end{bmatrix}\otimes\begin{bmatrix}1&0\\0&-1\end{bmatrix}=\begin{bmatrix}0&0&1&0\\0&0&0&-1\\1&0&0&0\\0&-1&0&0\end{bmatrix}. \tag{2.90}$$

The matrix representation of the tensor product of Z and X:

$$Z\otimes X=\begin{bmatrix}1&0\\0&-1\end{bmatrix}\otimes\begin{bmatrix}0&1\\1&0\end{bmatrix}=\begin{bmatrix}0&1&0&0\\1&0&0&0\\0&0&0&-1\\0&0&-1&0\end{bmatrix}. \tag{2.91}$$

Therefore, the tensor product $X\otimes Z$ and $Z\otimes X$ are not commutative.

(b) The matrix representation of the tensor product of I and X:

$$I\otimes X=\begin{bmatrix}1&0\\0&1\end{bmatrix}\otimes\begin{bmatrix}0&1\\1&0\end{bmatrix}=\begin{bmatrix}0&1&0&0\\1&0&0&0\\0&0&0&1\\0&0&1&0\end{bmatrix}. \tag{2.92}$$

(c) The matrix representation of the tensor product of X and I:

$$X\otimes I=\begin{bmatrix}0&1\\1&0\end{bmatrix}\otimes\begin{bmatrix}1&0\\0&1\end{bmatrix}=\begin{bmatrix}0&0&1&0\\0&0&0&1\\1&0&0&0\\0&1&0&0\end{bmatrix}. \tag{2.93}$$

Therefore, the tensor product $I\otimes X$ and $X\otimes I$ are not commutative.

□

Exercise 2.28. Show that the transpose, complex conjugate, and joint operation distribute over the tensor product,

$$(A\otimes B)^*=A^*\otimes B^*;\quad (A\otimes B)^T=A^T\otimes B^T;\quad (A\otimes B)^\dagger=A^\dagger\otimes B^\dagger. \tag{2.94}$$

Proof: Suppose A is a m by n matrix, and B is a p by q matrix.

(a)

$$(A\otimes B)^*=\begin{bmatrix}A_{11}B&A_{12}B&\cdots&A_{1n}B\\A_{21}B&A_{22}B&\cdots&A_{2n}B\\\vdots&\vdots&\vdots&\vdots\\A_{m1}B&A_{m2}B&\cdots&A_{mn}B\end{bmatrix}^*=\begin{bmatrix}A_{11}^*B^*&A_{12}^*B^*&\cdots&A_{1n}^*B^*\\A_{21}^*B^*&A_{22}^*B^*&\cdots&A_{2n}^*B^*\\\vdots&\vdots&\vdots&\vdots\\A_{m1}^*B^*&A_{m2}^*B^*&\cdots&A_{mn}^*B^*\end{bmatrix}=A^*\otimes B^*. \tag{2.95}$$

(b)

$$(A\otimes B)^T=\begin{bmatrix}A_{11}B&A_{12}B&\cdots&A_{1n}B\\A_{21}B&A_{22}B&\cdots&A_{2n}B\\\vdots&\vdots&\vdots&\vdots\\A_{m1}B&A_{m2}B&\cdots&A_{mn}B\end{bmatrix}^T=\begin{bmatrix}A_{11}B^T&A_{21}B^T&\cdots&A_{m1}B^T\\A_{12}B^T&A_{22}B^T&\cdots&A_{m2}B^T\\\vdots&\vdots&\vdots&\vdots\\A_{1n}B^T&A_{2n}B^T&\cdots&A_{mn}B^T\end{bmatrix}=A^T\otimes B^T. \tag{2.96}$$

(c)

$$(A \otimes B)^\dagger = [(A \otimes B)^*]^T = (A^* \otimes B^*)^T = (A^*)^T \otimes (B^*)^T = A^\dagger \otimes B^\dagger. \tag{2.97}$$

□

Exercise 2.29. Show that the tensor product of two unitary operators is unitary.

Proof: Suppose A and B are two unitary operators,

$$A^\dagger A = I, \tag{2.98}$$
$$B^\dagger B = I. \tag{2.99}$$

Using the conclusion obtained in the previous exercise,

$$(A \otimes B)^\dagger (A \otimes B) = (A^\dagger \otimes B^\dagger)(A \otimes B) = (A^\dagger A) \otimes (B^\dagger B) = I \otimes I = I, \tag{2.100}$$

so the tensor product of two unitary operators is unitary. □

Exercise 2.30. Show that the tensor product of two Hermitian operators is Hermitian.

Proof: Suppose A and B are two Hermitian operators,

$$A^\dagger = A, \tag{2.101}$$
$$B^\dagger = B. \tag{2.102}$$

Using the conclusion obtained in Exercise 2.28,

$$(A \otimes B)^\dagger = A^\dagger \otimes B^\dagger = A \otimes B, \tag{2.103}$$

so the tensor product of two Hermitian operators is Hermitian. □

Exercise 2.31. Show that the tensor product of two positive operators is positive.

Proof: Suppose A and B are positive operators on vector spaces V and W respectively. For any vector $|v\rangle \in V$ and $|w\rangle \in W$,

$$(|v\rangle, A|v\rangle) \geq 0, \tag{2.104}$$
$$(|w\rangle, B|w\rangle) \geq 0, \tag{2.105}$$

so $A \otimes B$ is a positive operator,

$$(|v\rangle \otimes |w\rangle, (A \otimes B)(|v\rangle \otimes |w\rangle)) = (|v\rangle \otimes |w\rangle, (A|v\rangle) \otimes (B|w\rangle)) = (|v\rangle, A|v\rangle)(|w\rangle, B|w\rangle) \geq 0. \tag{2.106}$$

Therefore, the tensor product of two positive operators is positive. □

Exercise 2.32. Show that the tensor product of two projectors is a projector.

Proof: Suppose $|i\rangle_V$ is an orthonormal basis for vector space V, and $|j\rangle_W$ is an orthonormal basis for vector space W. The projector onto V is

$$P_V = \sum_i |i\rangle_V \langle i|_V, \tag{2.107}$$

and the projector onto W is

$$P_W = \sum_j |j\rangle_W \langle j|_W. \tag{2.108}$$

Their tensor product is

$$P_V \otimes P_W = \left(\sum_i |i\rangle_V \langle i|_V \right) \otimes \left(\sum_j |j\rangle_W \langle j|_W \right) = \sum_{i,j} (|i\rangle_V \otimes |j\rangle_W)(\langle i|_V \otimes \langle j|_W). \tag{2.109}$$

Since both $|i\rangle_V$ and $|j\rangle_W$ are independent, $|i\rangle \otimes |j\rangle$ are independent. Since both $|i\rangle_V$ and $|j\rangle_W$ are orthonormal, $|i\rangle \otimes |j\rangle$ are orthonormal,

$$(|i\rangle_V \otimes |j\rangle_W, |k\rangle_V \otimes |l\rangle_W) = (|i\rangle_V, |k\rangle_V)(|j\rangle_W, |l\rangle_W) = \delta_{ik}\delta_{jl}. \tag{2.110}$$

Since any vector $|v\rangle \in V$ can be written as a linear combination of $|i\rangle_V$,

$$|v\rangle = \sum_i a_i |i\rangle_V, \tag{2.111}$$

and any vector $|w\rangle \in W$ can be written as a linear combination of $|j\rangle_W$,

$$|w\rangle = \sum_j b_j |j\rangle_W, \tag{2.112}$$

the tensor product $|v \otimes |w\rangle \in V \otimes W$ can be written as a linear combination of $|i\rangle_V \otimes |j\rangle_W$,

$$|v\rangle \otimes |w\rangle = \left(\sum_i a_i |i\rangle_V\right) \otimes \left(\sum_j b_j |j\rangle_W\right) = \sum_{i,j} a_i b_j |i\rangle_V \otimes |j\rangle_W. \tag{2.113}$$

Hence $|i\rangle_V \otimes |j\rangle_W$ is an orthonormal basis for $V \otimes W$ and $P_V \otimes P_W$ is a projector onto $V \otimes W$. Therefore, the tensor product of two projectors is a projector. □

Exercise 2.33. The Hadamard operator on one qubit may be written as

$$H = \frac{1}{\sqrt{2}} \left[(|0\rangle + |1\rangle)\langle 0| + (|0\rangle - |1\rangle)\langle 1|\right]. \tag{2.114}$$

Show explicitly that the Hadamard transform on n qubits, $H^{\otimes n}$, may be written as

$$H^{\otimes n} = \frac{1}{\sqrt{2^n}} \sum_{x,y} (-1)^{x \cdot y} |x\rangle\langle y|. \tag{2.115}$$

Write out an explicit matrix representation for $H^{\otimes 2}$.

Solution: Hadamard transform on one qubit may be written as

$$H|0\rangle = \frac{1}{\sqrt{2}}(|0\rangle\langle 0| + |1\rangle\langle 0| + |0\rangle\langle 1| - |1\rangle\langle 1|) = \frac{1}{\sqrt{2}} \sum_{x,y=0}^{1} |x\rangle\langle y|. \tag{2.116}$$

Thus Hadamard transform on n qubits may be written as

$$\begin{aligned} H^{\otimes n} &= \left(\frac{1}{\sqrt{2}} \sum_{x_1,y_1=0}^{1} |x_1\rangle\langle y_1|\right) \otimes \cdots \otimes \left(\frac{1}{\sqrt{2}} \sum_{x_n,y_n=0}^{1} |x_n\rangle\langle y_n|\right) = \frac{1}{\sqrt{2^n}} \sum_{\substack{x_1,\cdots,x_n \\ y_1,\cdots,y_n=0}}^{1} |x_1, \cdots, x_n\rangle\langle y_1, \cdots, y_n| \\ &= \frac{1}{\sqrt{2^n}} \sum_{x,y} (-1)^{x \cdot y} |x\rangle\langle y|. \end{aligned} \tag{2.117}$$

The matrix representation for $H^{\otimes 2}$ is

$$H^{\otimes 2} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \otimes \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}. \tag{2.118}$$

□

Exercise 2.34. Find the square root and logarithm of the matrix

$$\begin{bmatrix} 4 & 3 \\ 3 & 4 \end{bmatrix}. \tag{2.119}$$

Solution: The square root of the matrix is

$$\begin{bmatrix} 2 & \sqrt{3} \\ \sqrt{3} & 2 \end{bmatrix}. \tag{2.120}$$

The logarithm of the matrix is

$$\begin{bmatrix} \ln 4 & \ln 3 \\ \ln 3 & \ln 4 \end{bmatrix}. \tag{2.121}$$

□

Exercise 2.35 (Exponential of the Pauli matrices). Let $\vec{v}$ be any real, three dimensional unit vector and θ a real number. Prove that

$$\exp(i\theta\vec{v}\cdot\vec{\sigma}) = \cos(\theta)I + i\sin(\theta)\vec{v}\cdot\vec{\sigma}, \tag{2.122}$$

where $\vec{v}\cdot\vec{\sigma} = \sum_{i=1}^{3} v_i\sigma_i$. This exercise is generalized in Problem 2.1 on page 117.

Proof: The left side of equation (2.122) is

$$\exp(i\theta\vec{v}\cdot\vec{\sigma}) = \sum_{n=0}^{\infty}\frac{1}{n!}(i\theta\vec{v}\cdot\vec{\sigma})^k = \sum_{n=1}^{\infty}\frac{(-1)^n}{(2n)!}(\theta\vec{v}\cdot\vec{\sigma})^{2n} + \sum_{n=0}^{\infty}\frac{i(-1)^n}{(2n+1)!}(\theta\vec{v}\cdot\vec{\sigma})^{2n+1}. \tag{2.123}$$

Note that

$$(\vec{v}\cdot\vec{\sigma})^2 = \left(\sum_{i=1}^{3} v_i\sigma_i\right)^2 = \sum_{i,j=1}^{3} v_iv_j\sigma_i\sigma_j. \tag{2.124}$$

Due to the anti-commutation relation between the Pauli matrices,

$$\{\sigma_i,\sigma_j\} = \sigma_i\sigma_j + \sigma_j\sigma_i = 2\delta_{ij}I = \begin{cases} 2I, & i = j; \\ 0, & i \neq j, \end{cases} \tag{2.125}$$

we have

$$(\vec{v}\cdot\vec{\sigma})^2 = \sum_i v_i^2 I = I. \tag{2.126}$$

Hence the left side of equation (2.122) can be written as

$$\exp(i\theta\vec{v}\cdot\vec{\sigma}) = \sum_{n=1}^{\infty}\frac{(-1)^n}{(2n)!}\theta^{2n} + \sum_{n=0}^{\infty}\frac{i(-1)^n}{(2n+1)}\theta^{2n+1}\vec{\sigma} = \cos(\theta)I + i\sin(\theta)\vec{v}\cdot\vec{\sigma}, \tag{2.127}$$

which equals the right side of equation (2.122). Therefore, equation (2.122) holds. □

Exercise 2.36. Show that the Pauli matrices except for I have trace zero.

Proof: The trace of I is

$$\mathrm{tr}\,(I) = 1 + 1 = 2. \tag{2.128}$$

The trace of X is

$$\mathrm{tr}\,(X) = 0 + 0 = 0. \tag{2.129}$$

The trace of Y is

$$\mathrm{tr}\,(Y) = 0 + 0 = 0. \tag{2.130}$$

The trace of Z is

$$\mathrm{tr}\,(Z) = 1 + (-1) = 0. \tag{2.131}$$

Therefore, the Pauli matrices except for I have trace zero. □

Exercise 2.37 (Cyclic property of the trace). If A and B are two linear operators show that

$$\mathrm{tr}\,(AB) = \mathrm{tr}\,(BA). \tag{2.132}$$

Proof:

$$\mathrm{tr}\,(AB) = \sum_i (AB)_{ii} = \sum_i \left(\sum_j A_{ij} B_{ji} \right) = \sum_j \left(\sum_i B_{ji} A_{ij} \right) = \sum_j (BA)_{jj} = \mathrm{tr}\,(BA). \tag{2.133}$$

□

Exercise 2.38 (Linearity of the trace). If A and B are two linear operators. show that

$$\mathrm{tr}\,(A+B) = \mathrm{tr}\,(A) + \mathrm{tr}\,(B) \tag{2.134}$$

and if z is an arbitrary complex number show that

$$\mathrm{tr}\,(zA) = z\,\mathrm{tr}\,(A). \tag{2.135}$$

Proof:

$$\mathrm{tr}\,(A+B) = \sum_i (A+B)_{ii} = \sum_i A_{ii} + \sum_i B_{ii} = \mathrm{tr}\,(A) + \mathrm{tr}\,(B). \tag{2.136}$$

$$\mathrm{tr}\,(zA) = \sum_i (zA)_{ii} = z \sum_i A_{ii} = z\,\mathrm{tr}\,(A). \tag{2.137}$$

□

Exercise 2.39 (The Hilbert-Schmidt inner product on operators). The set L_V of linear operators on a Hilbert space V is obviously a vector space – the sum of two linear operators is a Linear operator, zA is a linear operator if A is a linear operator and z is a complex number, and there is a zero element 0. An important additional result is that the vector space L_V can be given a natural inner product structure, turning it into a Hilbert space.

(1) Show that the function $(\cdot,\cdot)$ on $L_V \times L_V$ defined by

$$(A,B) \equiv \mathrm{tr}\,(A^\dagger B) \tag{2.138}$$

is an inner product function. This inner product is known as the *Hilbert-Schmidt* or *trace* inner product.

(2) If V has d dimensions show that L_V has dimension d^2.

(3) Find an orthonormal basis of Hermitian matrices for the Hilbert space L_V.

Solution: (1) The function $(\cdot,\cdot)$ on $L_V \times L_V$ satisfies the requirements that:

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

$$\left(A, \sum_j \lambda_j B^{(j)} \right) = \mathrm{tr}\,\left(A^\dagger \sum_j \lambda_j B^{(j)} \right) = \mathrm{tr}\,\left(\sum_j \lambda_j A^\dagger B^{(j)} \right) = \sum_j \lambda_j \,\mathrm{tr}\,(A^\dagger B^{(j)}) = \sum_j \lambda (A, B^\dagger). \tag{2.139}$$

(b)

$$\begin{aligned}(A,B) = \mathrm{tr}\,(A^\dagger B) = \sum_i (A^\dagger B)_{ii} = \sum_i \left[\sum_j (A^\dagger)_{ij} B_{ji} \right] = \sum_i \left(\sum_j A^*_{ji} B_{ji} \right) = \left[\sum_i \left(\sum_j A_{ji} B^*_{ji} \right) \right]^* \\ = \left\{ \sum_i \left[\sum_j (B^\dagger)_{ij} A_{ji} \right] \right\}^* = \left[\sum_i (B^\dagger A)_{ii} \right]^* = [\,\mathrm{tr}\,(B^\dagger A)]^* = (B,A)^*.\end{aligned} \tag{2.140}$$

(c)

$$(A,A) = \mathrm{tr}\,(A^\dagger A) = \sum_i (A^\dagger A)_{ii} = \sum_i \left[\sum_j (A^\dagger)_{ij} A_{ji} \right] = \sum_i \left(\sum_j A^*_{ji} A_{ji} \right) = \sum_{i,j} |A_{ji}|^2 \geq 0, \tag{2.141}$$

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

Therefore, the function $(\cdot,\cdot)$ defined on $L_V \times L_V$ is an inner product function.

(2) Suppose $|1\rangle, \cdots, |d\rangle$ form an orthonormal basis for V. Since

(a) any operator A in V can be written as

$$A = \sum_{i,j} A_{ij}|i\rangle\langle j|, \tag{2.142}$$

and

(b) $|i\rangle\langle j|$ are orthonormal,

$$(|i\rangle\langle j|, |m\rangle\langle n|) = \text{tr}\left[(|i\rangle\langle j|)^\dagger |m\rangle\langle n|\right] = \text{tr}\left(|j\rangle\langle i|m\rangle\langle n|\right) = \sum_k \langle k|j\rangle\langle i|m\rangle\langle n|k\rangle = \sum_k \delta_{kj}\delta_{im}\delta_{nk} = \delta_{im}\delta_{jn}, \tag{2.143}$$

and thus linearly independent, $|i\rangle\langle j|$ is an orthonormal basis for L_V. Since this basis has d^2 elements, L_V has dimension d^2.

(3) $\left\{|i\rangle\langle i|, \frac{|i\rangle\langle j|+|j\rangle\langle i|}{\sqrt{2}}, \frac{|i\rangle\langle j|-|j\rangle\langle i|}{i\sqrt{2}}; \forall 1 \le i < j \le d\right\}$ is an orthonormal basis of Hermitian matrices for the Hilbert space L_V. Here is the reason:

(a) For any Hermitian matrices A, since

$$A^\dagger = A \Longrightarrow A^*_{ji} = A_{ji}, \quad \forall i, j = 1, \cdots, d, \tag{2.144}$$

and A_{ii} is real $\forall i$, A can be written as

$$\begin{aligned}
A =& \sum_{i,j} A_{ij}|i\rangle\langle j| = \sum_i A_{ii}|i\rangle\langle i| + \sum_{i=2}^{d}\sum_{j=1}^{i-1} A_{ij}|i\rangle\langle j| + \sum_{j=2}^{d}\sum_{i}^{j-1} A_{ij}|i\rangle\langle j| \\
=& \sum_i A_{ii}|i\rangle\langle i| + \sum_{i=2}^{d}\sum_{j=1}^{i-1} A^*_{ji}|i\rangle\langle j| + \sum_{j=2}^{d}\sum_{i}^{j-1} A_{ij}|i\rangle\langle j| \\
=& \sum_i A_{ii}|i\rangle\langle i| + \sum_{j=2}^{d}\sum_{i=1}^{j-1} A^*_{ij}|j\rangle\langle i| + \sum_{j=2}^{d}\sum_{i}^{j-1} A_{ij}|i\rangle\langle j| \\
=& \sum_i A_{ii}|i\rangle\langle i| + \sum_{j=2}^{d}\sum_{i=1}^{j-1} \frac{A_{ij}+A^*_{ij}}{\sqrt{2}}\frac{|i\rangle\langle j|+|j\rangle\langle i|}{\sqrt{2}} + \sum_{j=2}^{d}\sum_{i}^{j-1} \frac{A_{ij}-A^*_{ij}}{-i\sqrt{2}}\frac{|i\rangle\langle j|-|j\rangle\langle i|}{i\sqrt{2}}.
\end{aligned}$$

(b) Elements in $\left\{|i\rangle\langle i|, \frac{|i\rangle\langle j|+|j\rangle\langle i|}{\sqrt{2}}, \frac{|i\rangle\langle j|-|j\rangle\langle i|}{i\sqrt{2}}; \forall 1 \le i < j \le d\right\}$ are orthonormal and thus linearly independent,

$$\begin{aligned}
\left(|i\rangle\langle i|, \frac{|m\rangle\langle n| + |n\rangle\langle m|}{\sqrt{2}}\right) =& \text{tr}\left[(|i\rangle\langle i|)^\dagger \left(\frac{|m\rangle\langle n| + |n\rangle\langle m|}{\sqrt{2}}\right)\right] = \frac{1}{\sqrt{2}}\,\text{tr}\,\left[|i\rangle\langle i|(|m\rangle\langle n| + |n\rangle\langle m|)\right] \\
=& \frac{1}{\sqrt{2}}\,\text{tr}\left(|i\rangle\langle i|m\rangle\langle n| + |i\rangle\langle i|n\rangle\langle m|\right) \\
=& \frac{1}{\sqrt{2}}\sum_k (\langle k|i\rangle\langle i|m\rangle\langle n|k\rangle + \langle k|i\rangle\langle i|n\rangle\langle m|k\rangle) \\
=& \frac{1}{\sqrt{2}}\sum_k (\delta_{ki}\delta_{im}\delta_{nk} + \delta_{ki}\delta_{in}\delta_{mk}) \\
=& \sqrt{2}\delta_{im}\delta_{in} = 0, \quad \forall i = 1, \cdots, d, 1 \le m < n \le d, \\
\left(|i\rangle\langle i|, \frac{|m\rangle\langle n| - |n\rangle\langle m|}{i\sqrt{2}}\right) =& \text{tr}\left[(|i\rangle\langle i|)^\dagger \left(\frac{|m\rangle\langle n| + |n\rangle\langle m|}{i\sqrt{2}}\right)\right] = \frac{1}{i\sqrt{2}}\,\text{tr}\,\left[|i\rangle\langle i|(|m\rangle\langle n| + |n\rangle\langle m|)\right] \\
=& \frac{1}{i\sqrt{2}}\,\text{tr}\left(|i\rangle\langle i|m\rangle\langle n| - |i\rangle\langle i|n\rangle\langle m|\right) \\
=& \frac{1}{i\sqrt{2}}\sum_k (\langle k|i\rangle\langle i|m\rangle\langle n|k\rangle - \langle k|i\rangle\langle i|n\rangle\langle m|k\rangle)
\end{aligned} \tag{2.145}$$

$$
\begin{aligned}
&= \frac{1}{i\sqrt{2}} \sum_k (\delta_{ki}\delta_{im}\delta_{nk} - \delta_{ki}\delta_{in}\delta_{mk}) \\
&= 0, \quad \forall i = 1, \cdots, d, 1 \le m < n \le d, \qquad (2.146)
\end{aligned}
$$

$$
\begin{aligned}
\left(\frac{|i\rangle\langle j| + |j\rangle\langle i|}{\sqrt{2}}, \frac{|m\rangle\langle n| - |n\rangle\langle m|}{i\sqrt{2}} \right) &= \mathrm{tr} \left[\left(\frac{|i\rangle\langle j| + |j\rangle\langle i|}{\sqrt{2}} \right)^\dagger \left(\frac{|m\rangle\langle n| - |n\rangle\langle m|}{i\sqrt{2}} \right) \right] \\
&= \frac{1}{2i} \mathrm{tr}\, [(|j\rangle\langle i| + |i\rangle\langle j|)(|m\rangle\langle n| - |n\rangle\langle m|)] \\
&= \frac{1}{2i} \mathrm{tr}\, (|j\rangle\langle i|m\rangle\langle n| - |j\rangle\langle i|n\rangle\langle m| + |i\rangle\langle j|m\rangle\langle n| - |i\rangle\langle j|n\rangle\langle m|) \\
&= \frac{1}{2i} \sum_k (\langle k|j\rangle\langle i|m\rangle\langle n|k\rangle - \langle k|j\rangle\langle i|n\rangle\langle m|k\rangle \\
&\quad + \langle k|i\rangle\langle j|m\rangle\langle n|k\rangle - \langle k|i\rangle\langle j|n\rangle\langle m|k\rangle) \\
&= \frac{1}{2i} \sum_k (\delta_{kj}\delta_{im}\delta_{nk} - \delta_{kj}\delta_{in}\delta_{mk} + \delta_{ki}\delta_{jm}\delta_{nk} - \delta_{ki}\delta_{jn}\delta_{mk}) \\
&= 0, \quad \forall 1 \le i < j \le d, \cdots, d, 1 \le m < n \le d, \qquad (2.147)
\end{aligned}
$$

$$
\begin{aligned}
(|i\rangle\langle i|, |j\rangle\langle j|) &= \mathrm{tr}\, [(|i\rangle\langle i|)^\dagger |j\rangle\langle j|] = \mathrm{tr}\, (|i\rangle\langle i|j\rangle\langle j|) = \sum_k \langle k|i\rangle\langle i|j\rangle\langle j|k\rangle = \sum_k \delta_{ki}\delta_{ij}\delta_{jk} \\
&= \delta_{ij}, \quad \forall i, j = 1, \cdots, d, \qquad (2.148)
\end{aligned}
$$

$$
\begin{aligned}
\left(\frac{|i\rangle\langle j| + |i\rangle\langle j|}{\sqrt{2}}, \frac{|m\rangle\langle n| + |n\rangle\langle m|}{\sqrt{2}} \right) &= \mathrm{tr} \left[\left(\frac{|i\rangle\langle j| + |j\rangle\langle i|}{\sqrt{2}} \right)^\dagger \frac{|m\rangle\langle n| + |n\rangle\langle m|}{\sqrt{2}} \right] \\
&= \frac{1}{2} \mathrm{tr}\, [(|j\rangle\langle i| + |i\rangle\langle j|)(|m\rangle\langle m| + |n\rangle\langle m|)] \\
&= \frac{1}{2} \mathrm{tr}\, (|j\rangle\langle i|m\rangle\langle n| + |j\rangle\langle i|n\rangle\langle m| + |i\rangle\langle j|m\rangle\langle n| + |i\rangle\langle j|n\rangle\langle m|) \\
&= \frac{1}{2} \sum_k (\langle k|j\rangle\langle i|m\rangle\langle n|k\rangle + \langle k|j\rangle\langle i|n\rangle\langle m|k\rangle \\
&\quad + \langle k|i\rangle\langle j|m\rangle\langle n|k\rangle + \langle k|i\rangle\langle j|n\rangle\langle m|k\rangle) \\
&= \frac{1}{2} \sum_k (\delta_{kj}\delta_{im}\delta_{nk} + \delta_{kj}\delta_{in}\delta_{mk} + \delta_{ki}\delta_{jm}\delta_{nk} + \delta_{ki}\delta_{jn}\delta_{mk}) \\
&= \delta_{im}\delta_{jn} + \delta_{in}\delta_{jm}, \quad \forall 1 \le i < j \le d, 1 \le m < n \le d, \qquad (2.149)
\end{aligned}
$$

$$
\begin{aligned}
\left(\frac{|i\rangle\langle j| - |j\rangle\langle i|}{i\sqrt{2}}, \frac{|m\rangle\langle n| - |n\rangle\langle m|}{i\sqrt{2}} \right) &= \mathrm{tr} \left[\left(\frac{|i\rangle\langle j| - |j\rangle\langle i|}{i\sqrt{2}} \right)^\dagger \frac{|m\rangle\langle n| - |n\rangle\langle m|}{i\sqrt{2}} \right] \\
&= \frac{1}{2} \mathrm{tr}\, [(|j\rangle\langle i| - |i\rangle\langle j|)(|m\rangle\langle n| - |n\rangle\langle m|)] \\
&= \frac{1}{2} \mathrm{tr}\, (|j\rangle\langle i|m\rangle\langle n| - |j\rangle\langle i|n\rangle\langle m| - |i\rangle\langle j|m\rangle\langle n| + |j\rangle\langle i|n\rangle\langle m|) \\
&= \frac{1}{2} \sum_k (\langle k|j\rangle\langle i|m\rangle\langle n|k\rangle - \langle k|j\rangle\langle i|n\rangle\langle m|k\rangle \\
&\quad - \langle k|i\rangle\langle j|m\rangle\langle n|k\rangle + \langle k|i\rangle\langle j|n\rangle\langle m|k\rangle) \\
&= \frac{1}{2} \sum_k (\delta_{kj}\delta_{im}\delta_{nk} - \delta_{kj}\delta_{in}\delta_{mk} - \delta_{ki}\delta_{jm}\delta_{nk} + \delta_{ki}\delta_{jn}\delta_{mk}) \\
&= \delta_{im}\delta_{jn} - \delta_{in}\delta_{jm} = \delta_{im}\delta_{jn}, \quad \forall 1 \le i < j \le d, 1 \le m < n \le d. \qquad (2.150)
\end{aligned}
$$

and thus linearly independent.

Therefore, $\left\{ |i\rangle\langle i|, \frac{|i\rangle\langle j| + |j\rangle\langle i|}{\sqrt{2}}, \frac{|i\rangle\langle j| - |j\rangle\langle i|}{i\sqrt{2}}; \forall 1 \le i < j \le d \right\}$ is an orthonormal basis of Hermitian matrices for the Hilbert space L_V.

□

Exercise 2.40 (commutation relation for the Pauli matrices). Verify the commutation relations

$$[X,Y] = 2iZ; \quad [Y,Z] = 2iX; \quad [Z,X] = 2iY. \tag{2.151}$$

There is an elegant way of writing this using ϵ_{jkl}, the alternative antisymmetric tensor on three indices for which $\epsilon_{jkl} = 0$ except for $\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = 1$, and $\epsilon_{321} = \epsilon_{213} = \epsilon_{132} = -1$:

$$[\sigma_j, \sigma_k] = 2i \sum_{l=1}^{3} \epsilon_{jkl} \sigma_l. \tag{2.152}$$

Proof: The commutation relations for the Pauli matrices are

$$[X,Y] = XY - YX = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} - \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} - \begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix} = 2i \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = 2iZ, \tag{2.153}$$

$$[Y,Z] = YZ - ZY = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} - \begin{bmatrix} 0 & -i \\ -i & 0 \end{bmatrix} = 2i \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = 2iX, \tag{2.154}$$

$$[Z,X] = ZX - XZ = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} - \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = 2i \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} = 2iY, \tag{2.155}$$

and

$$[\sigma_j, \sigma_j] = \sigma_l \sigma_l - \sigma_l \sigma_l = 0, \quad \forall j = 1,2,3, \tag{2.156}$$

$$[\sigma_j, \sigma_k] = \sigma_j \sigma_k - \sigma_k \sigma_j = -[\sigma_k, \sigma_j], \quad \forall j,k = 1,2,3. \tag{2.157}$$

Therefore, in general,

$$[\sigma_j, \sigma_k] = 2i \sum_{l=1}^{3} \epsilon_{jkl} \sigma_l. \tag{2.158}$$

□

Exercise 2.41 (Anti-commutation relation for the Pauli matrices). Verify the anti-commutation relations

$$\{\sigma_i, \sigma_j\} = 0 \tag{2.159}$$

where $i \neq j$ are both chosen from the set $1,2,3$. Also verify that $(i = 0,1,2,3)$

$$\sigma_i^2 = I. \tag{2.160}$$

Proof:

$$\{\sigma_1, \sigma_2\} = \sigma_1\sigma_2 + \sigma_2\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} + \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} + \begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix} = 0, \tag{2.161}$$

$$\{\sigma_2, \sigma_3\} = \sigma_2\sigma_3 + \sigma_3\sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} + \begin{bmatrix} 0 & -i \\ -i & 0 \end{bmatrix} = 0, \tag{2.162}$$

$$\{\sigma_3, \sigma_1\} = \sigma_3\sigma_1 + \sigma_1\sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = 0, \tag{2.163}$$

$$\sigma_0^2 = I^2 = I, \tag{2.164}$$

$$\sigma_1^2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I, \tag{2.165}$$

$$\sigma_2^2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I, \tag{2.166}$$

$$\sigma_3^2 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I. \tag{2.167}$$

Therefore, in general,

$$\{\sigma_i, \sigma_j\} = 0, \quad \forall i,j = 1,2,3, \text{ and } i \neq j, \tag{2.168}$$

and

$$\sigma_i^2 = I, \quad \forall i = 0,1,2,3. \tag{2.169}$$

□

Exercise 2.42. Verify that

$$AB = \frac{[A,B] + \{A,B\}}{2}. \tag{2.170}$$

Proof:

$$\frac{[A,B] + \{A,B\}}{2} = \frac{AB - BA + AB + BA}{2} = AB. \tag{2.171}$$

□

Exercise 2.43. Show that for $j,k = 1,2,3$,

$$\sigma_j \sigma_k = \delta_{jk} I + i \sum_{l=1}^{3} \epsilon_{jkl} \sigma_l. \tag{2.172}$$

Proof: As obtained in exercise 2.40, 2.41 and 2.42,

$$[\sigma_j, \sigma_k] = 2i \sum_{l=1}^{3} \epsilon_{jkl} \sigma_l, \quad \forall j,k = 1,2,3, \tag{2.173}$$

$$\{\sigma_j, \sigma_k\} = 2\delta_{jk} I, \quad \forall j,k = 1,2,3, \tag{2.174}$$

$$AB = \frac{[A,B] + \{A,B\}}{2}, \tag{2.175}$$

so

$$\sigma_j \sigma_k = \frac{[\sigma_j, \sigma_k] + \{\sigma_j, \sigma_k\}}{2} = \frac{2i \sum_{l=1}^{3} \epsilon_{jkl} \sigma_l + 2\delta_{jk} I}{2} = \delta_{jk} I + i \sum_{l=1}^{3} \epsilon_{jkl} \sigma_l, \quad \forall j,k = 1,2,3. \tag{2.176}$$

□

Exercise 2.44. Suppose $[A,B] = 0$, $\{A,B\} = 0$, and A is invertible. Show that B must be 0.

Proof:

$$AB = \frac{[A,B] + \{A,B\}}{2} = 0. \tag{2.177}$$

A is invertible and thus can not be 0, so B must be 0. □

Exercise 2.45. Show that $[A,B]^\dagger = [B^\dagger, A^\dagger]$.

Proof:

$$[A,B]^\dagger = (AB - BA)^\dagger = B^\dagger A^\dagger - A^\dagger B^\dagger = [B^\dagger, A^\dagger]. \tag{2.178}$$

□

Exercise 2.46. Show that $[A,B] = -[B,A]$.

Proof:

$$[A,B] = AB - BA = -(BA - AB) = -[B,A]. \tag{2.179}$$

□

Exercise 2.47. Suppose A and B are Hermitian. Show that $i[A,B]$ is Hermitian.

Proof: Since A and B are Hermitian,

$$A^\dagger = A, \tag{2.180}$$

$$B^\dagger = B. \tag{2.181}$$

Since

$$(i[A,B])^\dagger = [i(AB - BA)]^\dagger = -i(B^\dagger A^\dagger - A^\dagger B^\dagger) = i(AB - BA) = i[A,B], \tag{2.182}$$

$i[A,B]$ is Hermitian. □

Exercise 2.48. What is the polar decomposition of a positive matrix P? Of a unitary matrix U? Of a Hermitian matrix, H?

Solution: *Polar decomposition of positive matrix P*: There exists unitary U and positive operators J and K such that

$$P = UJ = KU, \tag{2.183}$$

where the unique positive operators $J = \sqrt{P^\dagger P}$ and $K = \sqrt{PP^\dagger}$. Since P is positive, it can be given a spectral decomposition, $P = \sum_i \lambda_i |i\rangle\langle i|$, where λ_i are real and $\lambda_i \geq 0\ \forall i$. In this way,

$$J = \sqrt{P^\dagger P} = \sum_i \sqrt{\lambda_i^* \lambda_i} |i\rangle\langle i| = \sum_i |\lambda_i|\, |i\rangle\langle i| = \sum_i \lambda_i |i\rangle\langle i| = P, \tag{2.184}$$

$$K = \sqrt{PP^\dagger} = \sum_i \sqrt{\lambda_i \lambda_i^*} |i\rangle\langle i| = \sum_i |\lambda_i|\, |i\rangle\langle i| = \sum_i \lambda_i |i\rangle\langle i| = P. \tag{2.185}$$

If P is invertible, i.e., P is positive definite, then so is J, so $U = PJ^{-1} = I$ is unique.
Polar decomposition of unitary matrix U: There exists unitary A and positive operators J and K such that

$$U = AJ = KA, \tag{2.186}$$

where the unique positive operators $J = \sqrt{U^\dagger U} = I$ and $K = \sqrt{KK^\dagger} = I$. Hence $A = UJ^{-1} = U$ is unique.
Polar decomposition of Hermitian matrix H: There exists unitary U and positive operators J and K such that

$$H = UJ = KU, \tag{2.187}$$

where the unique positive operators $J = \sqrt{H^\dagger H}$ and $K = \sqrt{HH^\dagger}$. Since H is Hermitian, it is normal. According to spectral decomposition theorem, H is diagonal with respect to some orthonormal basis, i.e.,

$$H = A\Lambda A^\dagger, \tag{2.188}$$

where Λ is the matrix of eigenvalues of H and A is unitary. In this way,

$$J = \sqrt{H^\dagger H} = \sqrt{HH} = A\sqrt{\Lambda^2} A^\dagger, \tag{2.189}$$

$$K = \sqrt{HH^\dagger} = \sqrt{HH} = A\sqrt{\Lambda^2} A^\dagger. \tag{2.190}$$

If H is invertible, then $U = HJ^{-1}$ is unique. □

Exercise 2.49. Express the polar decomposition of a normal matrix in the outer product representation.

Solution: *Polar decomposition of normal matrix in the outer products representation*: For a normal matrix M, there exists unitary U and positive operators J and K such that

$$M = UJ = KU, \tag{2.191}$$

where the unique operators $J = \sqrt{M^\dagger M}$ and $K = \sqrt{MM^\dagger}$. In the outer product representation,

$$M = \sum_i \lambda_i |i\rangle\langle i|, \tag{2.192}$$

so

$$J = \sum_i \sqrt{\lambda_i^* \lambda_i} |i\rangle\langle i| = \sum_i |\lambda_i|\, |i\rangle\langle i|, \tag{2.193}$$

$$K = \sum_i \sqrt{\lambda_i \lambda_i^*} |i\rangle\langle i| = \sum_i |\lambda_i|\, |i\rangle\langle i|. \tag{2.194}$$

If M is invertible, then

$$U = MJ^{-1} = \left(\sum_i \lambda_i |i\rangle\langle i| \right) \left(\sum_j |\lambda_j|^{-1} |j\rangle\langle j| \right) = \sum_i \frac{\lambda_i}{|\lambda_i|} |i\rangle\langle i| \tag{2.195}$$

is also unique. □

Exercise 2.50. Find the left and right polar decomposition of the matrix

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}. \tag{2.196}$$

Solution: Left polar decomposition:

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = UJ, \tag{2.197}$$

where

$$J = \sqrt{\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}^\dagger \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}}. \tag{2.198}$$

Since

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}^\dagger \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}, \tag{2.199}$$

whose characteristic equation

$$\begin{vmatrix} 2-\lambda & 1 \\ 1 & 1-\lambda \end{vmatrix} = \lambda^2 - 3\lambda + 1 = 0 \tag{2.200}$$

gives eigenvalues

$$\lambda_{1,2} = \frac{3 \pm \sqrt{5}}{2}, \tag{2.201}$$

and eigenequations

$$\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} |v_1\rangle = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ b_1 \end{bmatrix} = \begin{bmatrix} 2a_1 + b_1 \\ a_1 + b_1 \end{bmatrix} = \lambda_1 |v_1\rangle = \frac{3+\sqrt{5}}{2} \begin{bmatrix} a_1 \\ b_1 \end{bmatrix}, \tag{2.202}$$

$$\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} |v_2\rangle = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} a_2 \\ b_2 \end{bmatrix} = \begin{bmatrix} 2a_2 + b_2 \\ a_2 + b_2 \end{bmatrix} = \lambda_2 |v_2\rangle = \frac{3-\sqrt{5}}{2} \begin{bmatrix} a_2 \\ b_2 \end{bmatrix} \tag{2.203}$$

give the corresponding eigenvectors

$$|v_1\rangle = \begin{bmatrix} \sqrt{\frac{5+\sqrt{5}}{10}} \\ \sqrt{\frac{5-\sqrt{5}}{10}} \end{bmatrix}, \quad |v_2\rangle = \begin{bmatrix} \sqrt{\frac{5-\sqrt{5}}{10}} \\ -\sqrt{\frac{5+\sqrt{5}}{10}} \end{bmatrix}, \tag{2.204}$$

we have

$$J = \sqrt{\frac{3+\sqrt{5}}{2}} |v_1\rangle\langle v_1| + \sqrt{\frac{3-\sqrt{5}}{2}} |v_2\rangle\langle v_2| \approx \begin{bmatrix} 1.3416 & 0.4472 \\ 0.4472 & 0.8944 \end{bmatrix}, \tag{2.205}$$

and

$$U = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} J^{-1} \approx \begin{bmatrix} 0.8944 & -0.4472 \\ 0.4472 & 0.8944 \end{bmatrix}. \tag{2.206}$$

Right polar decomposition:

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = KU, \tag{2.207}$$

where

$$K = \sqrt{\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}^\dagger}. \tag{2.208}$$

Since

$$\begin{bmatrix}1 & 0\\ 1 & 1\end{bmatrix}\begin{bmatrix}1 & 0\\ 1 & 1\end{bmatrix}^\dagger = \begin{bmatrix}1 & 0\\ 1 & 1\end{bmatrix}\begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix} = \begin{bmatrix}1 & 1\\ 1 & 2\end{bmatrix}, \tag{2.209}$$

whose characteristic equation

$$\begin{vmatrix}1-\lambda & 1\\ 1 & 2-\lambda\end{vmatrix} = \lambda^2 - 3\lambda + 1 = 0 \tag{2.210}$$

gives eigenvalues

$$\lambda_{1,2} = \frac{3 \pm \sqrt{5}}{2}, \tag{2.211}$$

and eigenequations

$$\begin{bmatrix}1 & 1\\ 1 & 2\end{bmatrix}|v_1\rangle = \begin{bmatrix}1 & 1\\ 1 & 2\end{bmatrix}\begin{bmatrix}a_1\\ b_1\end{bmatrix} = \begin{bmatrix}a_1+b_1\\ a_1+2b_1\end{bmatrix} = \lambda_1|v_1\rangle = \frac{3+\sqrt{5}}{2}\begin{bmatrix}a_1\\ b_1\end{bmatrix}, \tag{2.212}$$

$$\begin{bmatrix}1 & 1\\ 1 & 2\end{bmatrix}|v_2\rangle = \begin{bmatrix}1 & 1\\ 1 & 2\end{bmatrix}\begin{bmatrix}a_2\\ b_2\end{bmatrix} = \begin{bmatrix}a_2+b_2\\ a_2+2b_2\end{bmatrix} = \lambda_2|v_2\rangle = \frac{3-\sqrt{5}}{2}\begin{bmatrix}a_2\\ b_2\end{bmatrix}, \tag{2.213}$$

give the corresponding eigenvectors

$$|v_1\rangle = \begin{bmatrix}\sqrt{\frac{5-\sqrt{5}}{10}}\\ \sqrt{\frac{5+\sqrt{5}}{10}}\end{bmatrix}, \quad |v_2\rangle = \begin{bmatrix}\sqrt{\frac{5+\sqrt{5}}{10}}\\ -\sqrt{\frac{5-\sqrt{5}}{10}}\end{bmatrix}, \tag{2.214}$$

we have

$$K = \sqrt{\frac{3+\sqrt{5}}{2}}|v_1\rangle\langle v_1| + \sqrt{\frac{3-\sqrt{5}}{2}}|v_2\rangle\langle v_2| \approx \begin{bmatrix}0.8944 & 0.4472\\ 0.4472 & 1.3416\end{bmatrix}. \tag{2.215}$$

□

2.2 The postulates of quantum mechanics

Exercise 2.51. Verify that the Hadamard gate H is unitary.

Proof: Since

$$H^\dagger H = \frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}^\dagger \frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix} = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} = I, \tag{2.216}$$

the Hadamard gate H is unitary. □

Exercise 2.52. Verify that $H^2 = I$.

Proof:

$$H^2 = \frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix} = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} = I. \tag{2.217}$$

□

Exercise 2.53. What are the eigenvalues and eigenvectors of H?

Solution: The characteristic equation of H

$$\begin{vmatrix}\frac{1}{\sqrt{2}}-\lambda & \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}}-\lambda\end{vmatrix} = \lambda^2 - 1 = 0 \tag{2.218}$$

gives the eigenvalues

$$\lambda_{1,2} = \pm 1. \tag{2.219}$$

The eigenequations of H

$$H|v_1\rangle = \frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}\begin{bmatrix}a_1\\ b_1\end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix}a_1+b_1\\ a_1-b_1\end{bmatrix} = \lambda_1|v_1\rangle = \begin{bmatrix}a_1\\ b_1\end{bmatrix}, \tag{2.220}$$

$$H|v_2\rangle = \frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}\begin{bmatrix}a_2\\ b_2\end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix}a_2+b_2\\ a_2-b_2\end{bmatrix} = \lambda_2|v_2\rangle = \begin{bmatrix}-a_2\\ -b_2\end{bmatrix}, \tag{2.221}$$

give the corresponding eigenvectors

$$|v_1\rangle = \begin{bmatrix}\sqrt{\frac{2+\sqrt{2}}{2}}\\ \sqrt{\frac{2-\sqrt{2}}{2}}\end{bmatrix}, \quad |v_2\rangle = \begin{bmatrix}\sqrt{\frac{2-\sqrt{2}}{4}}\\ -\sqrt{\frac{2+\sqrt{2}}{4}}\end{bmatrix}. \tag{2.222}$$

□

Exercise 2.54. Suppose A and B are commuting Hermitian operators. Prove that $\exp(A)\exp(B) = \exp(A+B)$. (*Hint*: Use the results of Section 2.1.9.)

Proof: Since A and B commute,

$$[A,B] = AB - BA = 0 \Longrightarrow AB = BA. \tag{2.223}$$

Hence

$$\begin{aligned}\exp(A+B) =& \sum_{n=0}^{\infty}\frac{(X+Y)^n}{n!} = \sum_{n=0}^{\infty}\sum_{m=0}^{n}\binom{n}{m}\frac{A^m B^{n-m}}{n!} = \sum_{n=0}^{\infty}\sum_{m=0}^{n}\frac{A^m B^{n-m}}{m!(n-m)!} = \sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\frac{A^j B^k}{j!k!}\\ =& \left(\sum_{j=0}^{\infty}\frac{A^j}{j!}\right)\left(\sum_{k=0}^{\infty}\frac{B^k}{k!}\right) = \exp(A)\exp(B).\end{aligned} \tag{2.224}$$

□

Exercise 2.55. Prove that $U(t_1,t_2)$ defined in Equation (2.91) is unitary.

Proof:

$$U(t_1,t_2) = \exp\left[\frac{-iH(t_2-t_1)}{\hbar}\right] = \sum_{n=0}^{\infty}\frac{1}{n!}\left[\frac{-iH(t_2-t_1)}{\hbar}\right]^n, \tag{2.225}$$

$$\begin{aligned}[U(t_1,t_2)]^\dagger =& \sum_{n=0}^{\infty}\frac{1}{n!}\left[\frac{iH^\dagger(t_2-t_1)}{\hbar}\right]^n = \sum_{n=0}^{\infty}\frac{1}{n!}\left[\frac{iH(t_2-t_1)}{\hbar}\right]^n = \exp\left[\frac{iH(t_2-t_1)}{\hbar}\right] = \left\{\exp\left[\frac{-iH(t_2-t_1)}{\hbar}\right]\right\}^{-1}\\ =& [U(t_1,t_2)]^{-1}.\end{aligned} \tag{2.226}$$

Therefore, $U(t_1,t_2)$ is unitary. □

Exercise 2.56. Use the spectral decomposition to show that $K = -i\log(U)$ is Hermitian for any unitary U, and thus $U = \exp(iK)$ for some Hermitian K.

Proof: The spectral decomposition of U in terms of outer product representation is

$$U = \sum_n \lambda_n |n\rangle\langle n|. \tag{2.227}$$

Since U is unitary,

$$U^\dagger = \sum_n \lambda_n^* |n\rangle\langle n| = \sum_n \lambda_n^{-1}|n\rangle\langle n| = U^{-1}, \tag{2.228}$$

we have

$$\lambda_n^* = \lambda_n^{-1}, \quad \forall n. \tag{2.229}$$

The spectral decomposition of K is

$$K = -i \log U = \sum_n -i \log \lambda_i |n\rangle\langle n|. \tag{2.230}$$

Since

$$K^\dagger = \sum_n (-i \log \lambda_n)^* |n\rangle\langle n| = \sum_n i \log(\lambda_n^*)|n\rangle\langle n| = \sum_n i \log(\lambda_n^{-1})|n\rangle\langle n| = \sum_n -i \log \lambda_n |n\rangle\langle n| = K, \tag{2.231}$$

K is Hermitian. Hence $U = \exp(iK)$ for some Hermitian K. □

Exercise 2.57 (Cascaded measurements are single measurements). Suppose $\{L_l\}$ and $\{M_m\}$ are two sets of measurement operators. Show that 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 $N_{ml} = M_m L_l$.

Proof: Suppose the state of the quantum state is $|\psi\rangle$ immediately before the measurement. Then the probability that the result m occurs in the first measurement is

$$p(l) = \langle\psi|L_l^\dagger L_l|\psi\rangle, \tag{2.232}$$

and the state of the system after the measurement is

$$\frac{L_l|\psi\rangle}{\sqrt{\langle\psi|L_l^\dagger L_l|\psi\rangle}}. \tag{2.233}$$

The probability that the result l occurs in the second measurement conditional on that the result m occurs in the first measurement is

$$p(m|l) = \frac{\langle\psi|L_l^\dagger}{\sqrt{\langle\psi|L_l^\dagger L_l|\psi\rangle}} M_m^\dagger M_m \frac{L_l|\psi\rangle}{\sqrt{\langle\psi|L_l^\dagger L_l|\psi\rangle}} = \frac{\langle\psi|L_l^\dagger M_m^\dagger M_m L_l|\psi\rangle}{\langle\psi|L_l^\dagger L_l|\psi\rangle}, \tag{2.234}$$

and the state of the system after the two measurements is

$$\frac{M_m L_l|\psi\rangle}{\sqrt{\frac{\langle\psi|L_l^\dagger M_m^\dagger M_m L_l|\psi\rangle}{\langle\psi|L_l^\dagger L_l|\psi\rangle}}} \tag{2.235}$$

Hence the probability that the result m occurs in the first measurement and the result l occurs in the second measurement is

$$p(l,m) = p(l)p(m|l) = \langle\psi|L_l^\dagger M_m^\dagger M_m L_l|\psi\rangle. \tag{2.236}$$

Therefore, a measurement defined by the measurement operators $\{L_l\}$ followed by a measurement defined by the measurement $\{M_m\}$ is physically equivalent to a single measurement defined by measurement operators $\{N_{lm}\}$ with the representation $N_{ml} = M_m L_l$. □

Exercise 2.58. Suppose we prepare a quantum system in an eigenstate $|\psi\rangle$ of some observable M, with corresponding eigenvalue m. What is the average observed value of M, and the standard deviation?

Solution: The average observed value of M is

$$\langle M\rangle = \langle\psi|M|\psi\rangle = \langle\psi|m|\rangle = m. \tag{2.237}$$

The average observed value of M^2 is

$$\langle M^2\rangle = \langle\psi|M^2|\psi\rangle = \langle\psi|m^2|\psi\rangle = m^2. \tag{2.238}$$

The standard deviation of M is

$$[\Delta(M)]^2 = \langle M^2\rangle - \langle M\rangle^2 = 0. \tag{2.239}$$

□

Exercise 2.59. Suppose we have qubit in the state $|0\rangle$, and we measure the observable X. What is the average value of X? What is the standard deviation of X?

Solution: The average value of X is

$$\langle 0|X|0\rangle = \begin{bmatrix}1 & 0\end{bmatrix}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}1\\ 0\end{bmatrix} = 0. \tag{2.240}$$

The average value of X^2 is

$$\langle 0|X^2|0\rangle = \begin{bmatrix}1 & 0\end{bmatrix}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}^2\begin{bmatrix}1\\ 0\end{bmatrix} = 1. \tag{2.241}$$

The standard deviation of X is

$$[\Delta(X)]^2 = \langle X^2\rangle - \langle X\rangle^2 = 1. \tag{2.242}$$

□

Exercise 2.60. Show that $\vec{v}\cdot\vec{\sigma}$ has eigenvalues ± 1, and that the projectors onto the corresponding eigenspaces are given by $P_\pm = (I \pm \vec{v}\cdot\vec{\sigma})/2$.

Proof:

$$\vec{v}\cdot\vec{\sigma} = v_1\sigma_1 + v_2\sigma_2 + v_3\sigma_3 = v_1\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix} + v_2\begin{bmatrix}0 & -i\\ i & 0\end{bmatrix} + v_3\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix} = \begin{bmatrix}v_3 & v_1 - iv_2\\ v_1 + iv_2 & -v_3\end{bmatrix}. \tag{2.243}$$

The characteristic equation of $\vec{v}\cdot\vec{\sigma}$

$$|\vec{v}\cdot\vec{\sigma} - \lambda I| = \begin{vmatrix}v_3 - \lambda & v_1 - iv_2\\ v_1 + iv_2 & -v_3 - \lambda\end{vmatrix} = \lambda^2 - v_1^2 - v_2^2 - v_3^2 = \lambda^2 - 1 = 0 \tag{2.244}$$

gives the eigenvalues

$$\lambda_{1,2} = \pm 1. \tag{2.245}$$

The eigenequations

$$(\vec{v}\cdot\vec{\sigma})|v_1\rangle = \begin{bmatrix}v_3 & v_1 - iv_2\\ v_1 + iv_2 & -v_3\end{bmatrix}\begin{bmatrix}a_1\\ b_1\end{bmatrix} = \begin{bmatrix}v_3a_1 + (v_1 - iv_2)b_1\\ (v_1 + iv_2)a_1 - v_3b_1\end{bmatrix} = \lambda_1|v_1\rangle = \begin{bmatrix}a_1\\ b_1\end{bmatrix}, \tag{2.246}$$

$$(\vec{v}\cdot\vec{\sigma})|v_2\rangle = \begin{bmatrix}v_3 & v_1 - iv_2\\ v_1 + iv_2 & -v_3\end{bmatrix}\begin{bmatrix}a_2\\ b_2\end{bmatrix} = \begin{bmatrix}v_3a_2 + (v_1 - iv_2)b_2\\ (v_1 + iv_2)a_2 - v_3b_2\end{bmatrix} = \lambda_2|v_2\rangle = \begin{bmatrix}-a_2\\ -b_2\end{bmatrix} \tag{2.247}$$

give the corresponding eigenvectors

$$|v_1\rangle = \frac{1}{\sqrt{2(1+v_3)}}\begin{bmatrix}1+v_3\\ v_1+iv_2\end{bmatrix}, \quad |v_2\rangle = \frac{1}{\sqrt{2(1-v_3)}}\begin{bmatrix}v_3-1\\ v_1+iv_2\end{bmatrix}. \tag{2.248}$$

The projectors onto the corresponding eigensapces are

$$\begin{aligned}
P_+ =& |v_1\rangle\langle v_1| = \frac{1}{\sqrt{2(1+v_3)}}\begin{bmatrix}1+v_3\\ v_1+iv_2\end{bmatrix}\frac{1}{\sqrt{2(1+v_3)}}\begin{bmatrix}1+v_3 & v_1-iv_2\end{bmatrix}\\
=& \frac{1}{2(1+v_3)}\begin{bmatrix}(1+v_3)^2 & (1+v_3)(v_1-iv_2)\\ (v_1+iv_2)(1+v_3) & (v_1+iv_2)(v_1-iv_2)\end{bmatrix} = \frac{1}{2(1+v_3)}\begin{bmatrix}(1+v_3)^2 & (1+v_3)(v_1-iv_2)\\ (v_1+iv_2)(1+v_3) & (1+v_3)(1-v_3)\end{bmatrix}\\
=& \frac{1}{2}\begin{bmatrix}1+v_3 & v_1-iv_2\\ v_1+iv_2 & 1-v_3\end{bmatrix} = (I+\vec{v}\cdot\vec{\sigma})/2,
\end{aligned} \tag{2.249}$$

$$\begin{aligned}
P_- =& |v_2\rangle\langle v_2| = \frac{1}{\sqrt{2(1-v_3)}}\begin{bmatrix}v_3-1\\ v_1+iv_2\end{bmatrix}\frac{1}{\sqrt{2(1-v_3)}}\begin{bmatrix}v_3-1 & v_1-iv_2\end{bmatrix}\\
=& \frac{1}{2(1-v_3)}\begin{bmatrix}(v_3-1)^2 & (v_3-1)(v_1-iv_2)\\ (v_1+iv_2)(v_3-1) & (v_1+iv_2)(v_1-iv_2)\end{bmatrix} = \frac{1}{2(1-v_3)}\begin{bmatrix}(1-v_3)^2 & -(1-v_3)(v_1-iv_2)\\ -(v_1+iv_2)(1-v_3) & (1+v_3)(1-v_3)\end{bmatrix}\\
=& \frac{1}{2}\begin{bmatrix}1-v_3 & -(v_1-iv_2)\\ v_1+iv_2 & 1+v_3\end{bmatrix} = (I-\vec{v}\cdot\vec{\sigma})/2.
\end{aligned} \tag{2.250}$$

□

Exercise 2.61. Calculate the probability of obtaining the result +1 for a measurement of $\vec{v}\cdot\vec{\sigma}$, given that the state priori to measurement is $|0\rangle$. What is the state of the system after the measurement if +1 is obtained?

Solution: The probability of obtaining +1 for a measurement of $\vec{v}\cdot\vec{\sigma}$ is

$$p(1) = \langle 0|P_+|0\rangle = \begin{bmatrix}1 & 0\end{bmatrix}\frac{1}{2}\begin{bmatrix}1-v_3 & -(v_1-iv_2)\\ v_1+iv_2 & 1+v_3\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix} = \frac{1-v_3}{2}. \tag{2.251}$$

The state of the system after the measurement if +1 is obtained is

$$|v_1\rangle = \frac{1}{\sqrt{2(1+v_3)}}\begin{bmatrix}1+v_3\\ v_1+iv_2\end{bmatrix}. \tag{2.252}$$

□

Exercise 2.62. Show that any measurement where the measurement operators and the POVM elements coincide is a projective measurement.

Proof: Consider a measurement whose measurement operators $\{M_m\}$ coincide with POVM elements $\{E_m = M_m^\dagger M_m\}$,

$$E_m = M_m^\dagger M_m = M_m. \tag{2.253}$$

Since

$$M_m = E_m^\dagger = M_m^\dagger M_m = E_m = M_m, \tag{2.254}$$

$\{M_m\}$ are Hermitian. Hence

$$M_m^\dagger M_m = M_m^2 = M_m, \tag{2.255}$$

$\{M_m\}$ are projectors and this measurement is a projective measurement. □

Exercise 2.63. Suppose a measurement is described by measurement operators M_m. Show that there exists unitary operators U_m such that $M_m = U_m\sqrt{E_m}$, where E_m is the POVM associated to the measurement.

Proof: Since U_m is unitary and the POVM measurement operators E_m is Hermitian,

$$M_m^\dagger M_m = \sqrt{E_m^\dagger}U_m^\dagger U_m\sqrt{E_m} = \sqrt{E_m}U_m^{-1}U_m\sqrt{E_m} = E_m. \tag{2.256}$$

Therefore, there exists unitary operators U_m such that $M_m = U_m\sqrt{E_m}$, where E_m is the POVM associated to the measurement. □

Exercise 2.64. Suppose Bob is given a quantum state chosen from a set $|\psi_1\rangle, \cdots, |\psi_m\rangle$ of linearly independent states. Construct a POVM $\{E_1, E_2, \cdots, E_{m+1}\}$ such that if outcome E_i occurs, $1 \le i \le m$, then Bob knows with certainty that he was given the state $|\psi_i\rangle$. (The POVM must be such that $\langle\psi_i|E_i|\psi_i\rangle > 0$ for each i.)

Solution: For each $1 \le i \le m$, using Gram-Schmidt procedure to produce the orthonormal basis set $|\phi_1^{(i)}\rangle, |\phi_2^{(i)}\rangle, \cdots, |\phi_{i-1}^{(i)}\rangle, |\phi_{i+1}^{(i)}\rangle$, from $|\psi_1\rangle, |\psi_2\rangle, \cdots, |\psi_{i-1}\rangle, |\psi_{i+1}\rangle, \cdots, |\psi_m\rangle$, and then

$$E_m = \frac{\left(|\psi_i\rangle - \sum_{j\neq i}\langle\psi_i|\phi_j^{(i)}\rangle|\phi_j^{(i)}\rangle\right)\left(\langle\psi_i| - \sum_{j\neq i}\langle\phi_j^{(i)}|\psi_i\rangle\langle\phi_j^{(i)}|\right)}{\left||\psi_i\rangle - \sum_{j\neq i}\langle\psi_i|\phi_j^{(i)}\rangle|\phi_j^{(i)}\rangle\right|^2}. \tag{2.257}$$

For $i = m+1$,

$$E_{m+1} = I - \sum_{i=1}^{m} E_m. \tag{2.258}$$

□

Exercise 2.65. Express the states $(|0\rangle + |1\rangle)/\sqrt{2}$ and $(|0\rangle - |1\rangle)/\sqrt{2}$ in a basis in which they are not the same up to a relative phase shift.

Solution: In the basis $\{|0\rangle, |1\rangle\}$,

$$\frac{|0\rangle + |1\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle, \tag{2.259}$$

$$\frac{|0\rangle - |1\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}e^{i\pi}|1\rangle, \tag{2.260}$$

which are not the same up to a relative phase shift. □

Exercise 2.66. Show that the average value of the observable X_1Z_2 for a two qubit system measured in the state $(|00\rangle + |11\rangle)/\sqrt{2}$ is zero.

Proof: The average value of the observable X_1Z_2 for a two qubit system measured in the state $(|00\rangle + |11\rangle)/\sqrt{2}$ is

$$\frac{\langle 00| + \langle 11|}{\sqrt{2}} X_1 Z_2 \frac{|00\rangle + |11\rangle}{\sqrt{2}} = \frac{\langle 00| + \langle 11|}{\sqrt{2}} X_1 \frac{|00\rangle - |11\rangle}{\sqrt{2}} = \frac{\langle 00| + \langle 11|}{\sqrt{2}} \frac{|10\rangle - |01\rangle}{\sqrt{2}} = 0. \tag{2.261}$$

□

Exercise 2.67. Suppose V is Hilbert space with a subspace W. Suppose $U : W \to V$ is linear operator which preserves inner products, that is, for any $|w_1\rangle$ and $|w_2\rangle$ in W,

$$\langle w_1|U^\dagger U|w_2\rangle = \langle w_1|w_2\rangle. \tag{2.262}$$

Prove that there exists a unitary operator $U' : V \to V$ which extends U. That is, $U'|w\rangle = U|w\rangle$ for all $|w\rangle$ in W, but U' is defined on the entire space V. Usually we omit the prime symbol $'$ and just write U to decode the extension.

Proof: Define

$$U' \equiv U \otimes A, \tag{2.263}$$

where the unitary operator $A : U - W \to \mathrm{Im}\,(U)^\perp$. Since both U and A are unitary, U' is unitary. For all $|w\rangle$ in V, we can express it as

$$|w\rangle = |w_1\rangle \otimes |w_2\rangle, \tag{2.264}$$

where $|w_1\rangle$ and $|w_2\rangle$ are in the image of U, $\mathrm{Im}\,(U)$, and the complement of the image of U, $\mathrm{Im}\,(U)^\perp = W - \mathrm{Im}\,(U)$, respectively, and then

$$U'|w\rangle = (U \otimes A)(|w_1\rangle \otimes |w_2\rangle) = (U|w_1\rangle) \otimes (A|w_2\rangle). \tag{2.265}$$

Specially, for all $|w\rangle$ in W, we can extend it as

$$|w'\rangle = |w\rangle \otimes |0\rangle \in V \tag{2.266}$$

where $\mathrm{Im}\,(U)^\perp$ is the zero vector in $V - W$, and then

$$U'|w'\rangle = (U \otimes A)(|w\rangle \otimes |0\rangle) = U|w\rangle \otimes |0\rangle \tag{2.267}$$

Therefore, there exists a unitary $U' : V \to V$ which extends U. □

Exercise 2.68. Prove that $|\psi\rangle \neq |a\rangle|b\rangle$ for all single qubit states $|a\rangle$ and $|b\rangle$.

Proof: Suppose

$$|a\rangle = a_0|0\rangle + a_1|1\rangle, \tag{2.268}$$

$$|b\rangle = b_0|0\rangle + b_1|1\rangle, \tag{2.269}$$

where the normalization condition requires that

$$|a_0|^2 + |a_1|^2 = 1, \tag{2.270}$$

$$|b_0|^2 + |b_1|^2 = 1. \tag{2.271}$$

If

$$\frac{|00\rangle + |11\rangle}{\sqrt{2}} = |\psi\rangle = |a\rangle|b\rangle = a_0b_0|00\rangle + a_0b_1|01\rangle + a_1b_0|10\rangle + a_1b_1|11\rangle, \tag{2.272}$$

then

$$a_0b_0 = \frac{1}{\sqrt{2}}, \tag{2.273}$$
$$a_0b_1 = 0, \tag{2.274}$$
$$a_1b_0 = 0, \tag{2.275}$$
$$a_1b_1 = \frac{1}{\sqrt{2}}. \tag{2.276}$$

Equation (2.273) and (2.276) means that none of a_0, a_1, b_0 and b_1 equals 0, which conflicts with equation (2.274) and (2.275). Therefore, equation (2.272) is impossible and $|\psi\rangle \neq |a\rangle|b\rangle$ for all single qubit states $|a\rangle$ and $|b\rangle$. □

2.3 Application: superdense coding

Exercise 2.69. Verify that the Bell basis forms an orthonormal basis for the two qubit state space.

Proof: The Bell basis satisfies that:

(a) The Bell basis is normalized,

$$||\beta_{00}\rangle| = \sqrt{\langle\beta_{00}|\beta_{00}\rangle} = \sqrt{\frac{\langle 00| + \langle 00|}{\sqrt{2}}\frac{|00\rangle + |11\rangle}{\sqrt{2}}} = 1, \tag{2.277}$$
$$||\beta_{01}\rangle| = \sqrt{\langle\beta_{01}|\beta_{01}\rangle} = \sqrt{\frac{\langle 01| + \langle 10|}{\sqrt{2}}\frac{|01\rangle + |10\rangle}{\sqrt{2}}} = 1, \tag{2.278}$$
$$||\beta_{10}\rangle| = \sqrt{\langle\beta_{10}|\beta_{10}\rangle} = \sqrt{\frac{\langle 00| - \langle 11|}{\sqrt{2}}\frac{|00\rangle - |11\rangle}{\sqrt{2}}} = 1, \tag{2.279}$$
$$||\beta_{11}\rangle| = \sqrt{\langle\beta_{11}|\beta_{11}\rangle} = \sqrt{\frac{\langle 01| - \langle 10|}{\sqrt{2}}\frac{|01\rangle - |10\rangle}{\sqrt{2}}} = 1. \tag{2.280}$$

(b) The Bell basis is orthogonal,

$$\langle\beta_{00}|\beta_{01}\rangle = \frac{\langle 00| + \langle 11|}{\sqrt{2}}\frac{|01\rangle + |10\rangle}{\sqrt{2}} = 0, \tag{2.281}$$
$$\langle\beta_{00}|\beta_{10}\rangle = \frac{\langle 00| + \langle 11|}{\sqrt{2}}\frac{|00\rangle - |11\rangle}{\sqrt{2}} = 0, \tag{2.282}$$
$$\langle\beta_{00}|\beta_{11}\rangle = \frac{\langle 00| + \langle 11|}{\sqrt{2}}\frac{|01\rangle - |10\rangle}{\sqrt{2}} = 0, \tag{2.283}$$
$$\langle\beta_{01}|\beta_{10}\rangle = \frac{\langle 01| + \langle 10|}{\sqrt{2}}\frac{|00\rangle - |11\rangle}{\sqrt{2}} = 0, \tag{2.284}$$
$$\langle\beta_{01}|\beta_{11}\rangle = \frac{\langle 01| + \langle 10|}{\sqrt{2}}\frac{|01\rangle - |10\rangle}{\sqrt{2}} = 0, \tag{2.285}$$
$$\langle\beta_{10}|\beta_{11}\rangle = \frac{\langle 00| - \langle 11|}{\sqrt{2}}\frac{|01\rangle - |10\rangle}{\sqrt{2}} = 0, \tag{2.286}$$

and thus also linearly independent.

(c) The dimension of the two qubit state space is $2 \times 2 = 4$, the Bell basis also has 4 independent elements, so the Bell basis can span the two qubit space.

Therefore, the Bell basis forms an orthonormal basis for the two qubit state space. □

Exercise 2.70. Suppose E is any positive operator acting on Alice's qubit. Show that $\langle\psi|E\otimes I|\psi\rangle$ *takes the same value* when $|\psi\rangle$ is any of the four Bell states. Suppose some malevolent third party ('Eve') intercepts Alice's qubit on the way to Bob in the superdense coding protocol. Can Eve infer anything about which of the four possible bit strings 00, 01, 10, 11 Alice is trying to send? If so, how, or if not, why not?

Proof: $\langle\psi|E\otimes I|\psi\rangle$ takes the same value when $|\psi\rangle$ is any of the four Bell states,

$$\begin{aligned}
\langle\beta_{00}|E\otimes I|\beta_{00}\rangle =& \frac{\langle 00|+\langle 11|}{\sqrt{2}}\frac{|00\rangle+|11\rangle}{\sqrt{2}} = \frac{1}{2}(\langle 00|E\otimes I|00\rangle + \langle 00|E\otimes I|11\rangle + \langle 11|E\otimes I|00\rangle + \langle 11|E\otimes I|11\rangle) \\
=& \frac{1}{2}(\langle 0|E|0\rangle\langle 0|I|0\rangle + \langle 0|E|1\rangle\langle 0|I|1\rangle + \langle 1|E|0\rangle\langle 1|I|0\rangle + \langle 1|E|1\rangle\langle 1|I|1\rangle) = \frac{1}{2}(\langle 0|E|0\rangle + \langle 1|E|1\rangle), \quad (2.287) \\
\langle\beta_{01}|E\otimes I|\beta_{01}\rangle =& \frac{\langle 01|+\langle 10|}{\sqrt{2}}\frac{|01\rangle+|10\rangle}{\sqrt{2}} = \frac{1}{2}(\langle 01|E\otimes I|01\rangle + \langle 01|E\otimes I|10\rangle + \langle 10|E\otimes I|01\rangle + \langle 10|E\otimes I|10\rangle) \\
=& \frac{1}{2}(\langle 0|E|0\rangle\langle 1|I|1\rangle + \langle 0|E|1\rangle\langle 1|I|0\rangle + \langle 1|E|0\rangle\langle 0|I|1\rangle + \langle 1|E|1\rangle\langle 0|I|0\rangle) = \frac{1}{2}(\langle 0|E|0\rangle + \langle 1|E|1\rangle), \quad (2.288) \\
\langle\beta_{10}|E\otimes I|\beta_{10}\rangle =& \frac{\langle 00|-\langle 11|}{\sqrt{2}}\frac{|00\rangle-|11\rangle}{\sqrt{2}} = \frac{1}{2}(\langle 00|E\otimes I|00\rangle - \langle 00|E\otimes I|11\rangle - \langle 11|E\otimes I|00\rangle + \langle 11|E\otimes I|11\rangle) \\
=& \frac{1}{2}(\langle 0|E|0\rangle\langle 0|I|0\rangle - \langle 0|E|1\rangle\langle 0|I|1\rangle - \langle 1|E|0\rangle\langle 1|I|0\rangle + \langle 1|E|1\rangle\langle 1|I|1\rangle) = \frac{1}{2}(\langle 0|E|0\rangle + \langle 1|E|1\rangle), \quad (2.289) \\
\langle\beta_{11}|E\otimes I|\beta_{11}\rangle =& \frac{\langle 01|-\langle 10|}{\sqrt{2}}\frac{|01\rangle-|10\rangle}{\sqrt{2}} = \frac{1}{2}(\langle 01|E\otimes I|01\rangle - \langle 01|E\otimes I|10\rangle - \langle 10|E\otimes I|01\rangle + \langle 10|E\otimes I|10\rangle) \\
=& \frac{1}{2}(\langle 0|E|0\rangle\langle 1|I|1\rangle - \langle 0|E|1\rangle\langle 1|I|0\rangle - \langle 1|E|0\rangle\langle 0|I|1\rangle + \langle 1|E|1\rangle\langle 0|I|0\rangle) = \frac{1}{2}(\langle 0|E|0\rangle + \langle 1|E|1\rangle). \quad (2.290)
\end{aligned}$$

Eve can not infer anything about which of the four possible bit string Alice is trying to send. Here is the reason: Suppose Eve intercepts Alice's qubit and try to do some measurement on it. Whichever of the four possible bit string 00, 01, 10, 11 Alice is trying to send, Eve will get result m with the same possibility

$$p(m) = \frac{1}{2}(\langle 0|E|0\rangle + \langle 1|E|1\rangle). \quad (2.291)$$

In this way, Eve can not obtain any knowledge about the bit string Alice is trying to send from the measurement result. □

2.4 The density operator

Exercise 2.71 (Criterion of decide if a state is mixed or pure). Let ρ be a density operator. Show that $\mathrm{tr}\,(\rho^2) \leq 1$, with equality if and only if ρ is a pure state.

Proof: The density operator is

$$\rho = \sum_i p_i|\psi_i\rangle\langle\psi_i|, \quad (2.292)$$

so

$$\begin{aligned}
\mathrm{tr}\,(\rho^2) =& \mathrm{tr}\left[\left(\sum_i p_i|\psi_i\rangle\langle\psi_i|\right)\left(\sum_j p_j|\psi_j\rangle\langle\psi_j|\right)\right] = \sum_i p_i\left[\sum_j p_j\,\mathrm{tr}\,(|\psi_i\rangle\langle\psi_i|\psi_j\rangle\langle\psi_j|)\right] \\
=& \sum_i p_i\left[\sum_j p_j\,\mathrm{tr}\,(\langle\psi_i|\psi_j\rangle\langle\psi_j|\psi_i\rangle)\right] = \sum_i p_i\left[\sum_j p_j\,\mathrm{tr}\,(|\langle\psi_i|\psi_j\rangle|^2)\right] = \sum_i p_i\left[\sum_j p_j\,|\langle\psi_i|\psi_j\rangle|^2\right] \\
\leq& \sum_i p_i\left(\sum_j p_j\right) = \sum_i p_i = 1. \quad (2.293)
\end{aligned}$$

If ρ is a pure state,

$$\rho = |\psi\rangle\langle\psi|, \quad (2.294)$$

then

$$\mathrm{tr}\,(\rho^2) = \mathrm{tr}\,(|\psi\rangle\langle\psi|\psi\rangle\langle\psi|) = \mathrm{tr}\,(|\psi\rangle\langle\psi|) = 1. \tag{2.295}$$

If $\mathrm{tr}\,(\rho^2) = 1$, then

$$\sum_i p_i \left[\sum_j p_j \left|\langle\psi_i|\psi_j\rangle\right|^2\right] = \sum_i p_i \left(\sum_j p_j\right), \tag{2.296}$$

$$\implies \left|\langle\psi_i|\psi_j\rangle\right|^2 = \left|\langle\psi_i|\psi_i\rangle\right|^2 = 1, \tag{2.297}$$

which means that, for all i and j, $|\psi_i\rangle$ and $|\psi_j\rangle$ is the same up to a phase difference

$$|\psi_j\rangle = e^{i\theta_{ij}}|\psi_i\rangle, \quad \forall i, j, \tag{2.298}$$

i.e., ρ must be a pure state

$$\rho = |\psi_1\rangle\langle\psi_1|. \tag{2.299}$$

Therefore, $\mathrm{tr}\,(\rho^2) \leq 1$, with equality if and only if ρ is a pure state. □

Exercise 2.72 (Bloch sphere for mixed states). The Bloch sphere picture for pure states of a single qubit was introduced in Section 1.2. This description has an important generalization to mixed states as follows.

(1) Show that an arbitrary density matrix for a mixed state qubit may be written as

$$\rho = \frac{I + \vec{r}\cdot\vec{\sigma}}{2}, \tag{2.300}$$

where $\vec{r}$ is a real three-dimensional vector such that $\|\vec{r}\| \leq 1$. This vector is known as the *Bloch vector* for the state ρ.

(2) What is the Bloch vector representation for the state $\rho = I/2$?

(3) Show that a state ρ is pure if and only if $\|\vec{r}\| = 1$.

(4) Show that for pure states the description of the Bloch vector we have given coincide with that in Section 1.2.

Solution:

(1) An arbitrary 2×2 matrix ρ can be expressed as a linear combination of I, σ_1, σ_2, σ_3,

$$\begin{aligned}\rho =& \begin{bmatrix}\rho_{11} & \rho_{12}\\ \rho_{21} & \rho_{22}\end{bmatrix} = \begin{bmatrix}\frac{1}{2}(\rho_{11}+\rho_{22}) + \frac{1}{2}(\rho_{11}-\rho_{22}) & \frac{1}{2}(\rho_{12}+\rho_{21}) + i(-i)\frac{1}{2}(\rho_{12}-\rho_{21})\\ \frac{1}{2}(\rho_{12}+\rho_{21}) + i\cdot i(\rho_{12}-\rho_{21}) & \frac{1}{2}(\rho_{11}+\rho_{22}) - \frac{1}{2}(\rho_{11}-\rho_{22})\end{bmatrix}\\ =& \frac{1}{2}(\rho_{11}+\rho_{22})\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} + \frac{1}{2}(\rho_{12}+\rho_{21})\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix} + \frac{i}{2}(\rho_{12}-\rho_{21})\begin{bmatrix}0 & -i\\ i & 0\end{bmatrix} + \frac{1}{2}(\rho_{11}-\rho_{22})\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}\\ =& \frac{1}{2}(\rho_{11}+\rho_{22})I + \frac{1}{2}(\rho_{12}+\rho_{21})\sigma_1 + \frac{i}{2}(\rho_{12}-\rho_{21})\sigma_2 + \frac{1}{2}(\rho_{11}-\rho_{22})\sigma_3\end{aligned} \tag{2.301}$$

If ρ is an arbitrary density matrix, then it has trace of 1,

$$\mathrm{tr}\,(\rho) = \rho_{11} + \rho_{22} = 1, \tag{2.302}$$

so

$$\rho = \frac{1}{2}I + \frac{1}{2}(\rho_{12}+\rho_{21})\sigma_1 + \frac{i}{2}(\rho_{12}-\rho_{21})\sigma_2 + \frac{1}{2}(\rho_{11}-\rho_{22})\sigma_3 = \frac{I + \vec{r}\cdot\vec{\sigma}}{2}, \tag{2.303}$$

where

$$\vec{r} = (r_1, r_2, r_3) = (\rho_{12}+\rho_{21}, i(\rho_{12}-\rho_{21}), \rho_{11}-\rho_{22}). \tag{2.304}$$

(2) For the state

$$\rho = \frac{I}{2} = \frac{I + \vec{r}\cdot\vec{\sigma}}{2}, \tag{2.305}$$

the Bloch vector representation is

$$\vec{r} = (0, 0, 0) = \vec{0}. \tag{2.306}$$

(3) *Necessity*: If $\|\vec{r}\| = 1$, then

$$\|\vec{r}\|^2 = r_1 r_1^* + r_2 r_2^* + r_3 r_3^* = (\rho_{12} + \rho_{21})(\rho_{12} + \rho_{21})^* + (\rho_{12} - \rho_{21})(\rho_{12} - \rho_{21})^* + (\rho_{11} - \rho_{22})(\rho_{11} - \rho_{22})^* = 1. \quad (2.307)$$

Since the density matrix ρ is positive, it is necessarily Hermitian, i.e., ρ_{11} and ρ_{22} are real and ρ_{12} and ρ_{21} are complex conjugate,

$$\rho_{12} = \rho_{21}^*. \quad (2.308)$$

Hence we can rewrite equation (2.307) as

$$(\rho_{12} + \rho_{21})(\rho_{21} + \rho_{12}) + (\rho_{12} - \rho_{21})(\rho_{21} - \rho_{12}) + (\rho_{11} - \rho_{22})(\rho_{11} - \rho_{22})^* = 4\rho_{12}\rho_{21} + \rho_{11}^2 + \rho_{22}^2 - 2\rho_{11}\rho_{22} = 1. \quad (2.309)$$

Since the trace of the density matrix ρ is 1,

$$\mathrm{tr}\,(\rho) = \rho_{11} + \rho_{22} = 1, \quad (2.310)$$

we can rewrite equation (2.309) as

$$4\rho_{12}\rho_{21} + 2(\rho_{11}^2 + \rho_{22}^2) - (\rho_{11}^2 + \rho_{22}^2 + 2\rho_{11}\rho_{22}) = 4\rho_{12}\rho_{21} + 2(\rho_{11}^2 + \rho_{22}^2) - (\rho_{11} + \rho_{22})^2 = 4\rho_{12}\rho_{21} + 2(\rho_{11}^2 + \rho_{22}^2) - 1 = 1, \quad (2.311)$$

$$\Longrightarrow 2\rho_{12}\rho_{21} + \rho_{11}^2 + \rho_{22}^2 = 1. \quad (2.312)$$

In this way, the trace of the square of the density matrix equals 1,

$$\mathrm{tr}\,(\rho^2) = \mathrm{tr}\left(\begin{bmatrix}\rho_{11} & \rho_{12}\\ \rho_{21} & \rho_{22}\end{bmatrix}\begin{bmatrix}\rho_{11} & \rho_{12}\\ \rho_{21} & \rho_{22}\end{bmatrix}\right) = \mathrm{tr}\begin{bmatrix}\rho_{11}^2 + \rho_{12}\rho_{21} & \rho_{11}\rho_{12} + \rho_{12}\rho_{22}\\ \rho_{21}\rho_{11} + \rho_{22}\rho_{21} & \rho_{21}\rho_{12} + \rho_{22}^2\end{bmatrix} = \rho_{11}^2 + \rho_{22}^2 + 2\rho_{12}\rho_{21} = 1. \quad (2.313)$$

Therefore, the state ρ is pure.
Sufficiency: If the state ρ is pure, the trace of its square is 1,

$$\mathrm{tr}\,(\rho^2) = \rho_{11}^2 + \rho_{22}^2 + 2\rho_{12}\rho_{21} = 1. \quad (2.314)$$

Then

$$\|\vec{r}\| = \sqrt{4\rho_{12}\rho_{21} + 2(\rho_{11}^2 + \rho_{22}^2) - 1} = 1. \quad (2.315)$$

Therefore, a state ρ is pure if and only if $\|\vec{r}\| = 1$.

(4) According to Section 1.2, we write the state of the pure state as

$$|\psi\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\varphi}\sin\frac{\theta}{2}|1\rangle, \quad (2.316)$$

and the Bloch vector is

$$\vec{r} = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta). \quad (2.317)$$

The density matrix of the pure state is

$$\begin{aligned}\rho =& \left(\cos\frac{\theta}{2}|0\rangle + e^{i\varphi}\sin\frac{\theta}{2}|1\rangle\right)\left(\cos\frac{\theta}{2}\langle 0| + e^{-i\varphi}\sin\frac{\theta}{2}\langle 1|\right) = \begin{bmatrix}\cos\frac{\theta}{2}\\ e^{i\varphi}\sin\frac{\theta}{2}\end{bmatrix}\begin{bmatrix}\cos\frac{\theta}{2} & e^{-i\varphi}\sin\frac{\theta}{2}\end{bmatrix}\\ =& \begin{bmatrix}\cos^2\frac{\theta}{2} & e^{-i\varphi}\sin\frac{\theta}{2}\cos\frac{\theta}{2}\\ e^{i\varphi}\sin\frac{\theta}{2}\cos\frac{\theta}{2} & \sin^2\frac{\theta}{2}\end{bmatrix}.\end{aligned} \quad (2.318)$$

According to the description we have given, the Bloch vector is

$$\begin{aligned}\vec{r} =& (\rho_{12} + \rho_{21}, i(\rho_{12} - \rho_{21}), \rho_{11} - \rho_{22})\\ =& \left(e^{-i\varphi}\sin\frac{\theta}{2}\cos\frac{\theta}{2} + e^{i\varphi}\sin\frac{\theta}{2}\cos\frac{\theta}{2}, i\left(e^{-i\varphi}\sin\frac{\theta}{2}\cos\frac{\theta}{2} - e^{i\varphi}\sin\frac{\theta}{2}\cos\frac{\theta}{2}\right), \cos^2\frac{\theta}{2} - \sin^2\frac{\theta}{2}\right)\\ =& (\sin\theta\cos\varphi, \sin\theta\sin\sin\varphi, \cos\theta).\end{aligned} \quad (2.319)$$

Therefore, for pure states, the description of the Bloch vector we have given coincide with that in Section 1.2. □

Exercise 2.73. 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 $|\psi_i\rangle$ be any state in the support of ρ. (The *support* of a Hermitian operator A is vector space spanned by the eigenvectors of A with non-zero eigenvalues.) Show that there is a minimal ensemble for ρ that contains $|\psi\rangle$, and moreover that in any such ensemble $|\psi_i\rangle$ must appear with probability

$$p_i = \frac{1}{\langle\psi_i|\rho^{-1}|\psi_i\rangle}, \tag{2.320}$$

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

Proof: Suppose the spectral decomposition of the density matrix ρ is

$$\rho = \sum_{k=1}^{N} q_k |k\rangle\langle k| = \sum_{k=1}^{N} |\tilde{k}\rangle\langle\tilde{k}|. \tag{2.321}$$

where q_k are the non-zero eigenvalues and $|k\rangle$ are the eigenvectors of ρ, $N = \text{rank}(\rho)$ and

$$|\tilde{k}\rangle = \sqrt{q_k}|k\rangle. \tag{2.322}$$

Since $|\psi_i\rangle$ is a state in the support of ρ, it can be written as a linear combination of the eigenvectors

$$|\psi_i\rangle = \sum_k \langle k|\psi_i\rangle|k\rangle = \sum_k c_{ik}|k\rangle, \tag{2.323}$$

where

$$c_{ik} = \langle k|\psi_i\rangle, \tag{2.324}$$

and

$$\sum_{k=1}^{N} |c_{ik}|^2 = 1. \tag{2.325}$$

Define

$$p_i = \frac{1}{\sum_{k=1}^{N} \frac{|c_{ik}|^2}{p_k}} \tag{2.326}$$

and

$$u_{ik} = \frac{\sqrt{p_i}c_{ik}}{\sqrt{q_k}}. \tag{2.327}$$

Since

$$\sum_{k=1}^{N} |u_{ik}|^2 = \sum_{k=1}^{N} \frac{p_i\,|c_{ik}|^2}{q_k} = p_i \sum_{k=1}^{N} \frac{|c_{ik}|^2}{q_k} = 1, \tag{2.328}$$

u_{ik} is a unitary matrix. According to Theorem 2.6, the set

$$|\tilde{\psi}_i\rangle = \sum_{k=1}^{N} u_{ik}|\tilde{k}\rangle = \sum_{k=1}^{N} \frac{\sqrt{p_i}c_{ik}}{\sqrt{q_k}}|\tilde{k}\rangle = \sqrt{p_i}\sum_{k=1}^{N} c_{ik}|k\rangle = \sqrt{p_i}|\psi_i\rangle, \quad i = 1, \cdots, N, \tag{2.329}$$

generate the same density matrix ρ,

$$\rho = \sum_{i=1}^{N} |\tilde{\psi}_i\rangle\langle\tilde{\psi}_i| = \sum_{i=1}^{N} p_i|\psi_i\rangle\langle\psi_i|. \tag{2.330}$$

Moreover,

$$\frac{1}{\langle\psi_i|\rho^{-1}|\psi_i\rangle} = \frac{1}{\langle\psi_i|\sum_{k=1}^{N} q_k^{-1}|k\rangle\langle k||\psi_i\rangle} = \frac{1}{\sum_{k=1}^{N} q_k^{-1}\,|\langle k|\psi_i\rangle|^2} = \frac{1}{\sum_{k=1}^{N} \frac{|c_{ik}|^2}{q_k}} = p_i. \tag{2.331}$$

Therefore, there is a minimal ensemble for ρ that contains $|\psi_i\rangle$, and moreover that in any such ensemble $|\psi_i\rangle$ must appear with probability

$$p_i = \frac{1}{\langle\psi_i|\rho^{-1}|\psi_i\rangle}. \tag{2.332}$$

□

Exercise 2.74. Suppose a composite of system A and B is in the state $|a\rangle|b\rangle$, where $|a\rangle$ is a pure state of system A, and $|b\rangle$ is a pure state of system B. Show that the reduced density operator of system A alone is a pure state.

Proof: The density operator of the composite system is

$$\rho^{AB} = |a\rangle\langle a| \otimes |b\rangle\langle b|. \tag{2.333}$$

The reduced density operator of system A is

$$\rho^A = \text{tr}_B(\rho^{AB}) = |a\rangle\langle a| \,\text{tr}_B(|b\rangle\langle b|) = |a\rangle\langle a|. \tag{2.334}$$

Since

$$\text{tr}\,[(\rho^A)^2] = \text{tr}\,[(|a\rangle\langle a|)(|a\rangle\langle a|)] = \text{tr}\,(|a\rangle\langle a|) = 1, \tag{2.335}$$

the reduced density operator of system A alone is a pure state. □

Exercise 2.75. For each of the four Bell states, find the reduced density operator for each qubit.

Solution: Suppose the first one of the two qubits described by the Bell states is A, and the other B. For the Bell state

$$|\beta_{00}\rangle = \frac{|00\rangle\langle 11|}{\sqrt{2}}, \tag{2.336}$$

its density operator is

$$\rho_{00}^{AB} = |\beta_{00}\rangle\langle\beta_{00}| = \frac{|00\rangle\langle 11|}{\sqrt{2}}\frac{\langle 00| + \langle 11|}{\sqrt{2}} = \frac{1}{2}(|00\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 00| + |11\rangle\langle 11|), \tag{2.337}$$

the reduced density operator of qubit A is

$$\begin{aligned}\rho_{00}^A =& \,\text{tr}_B(\rho_{00}^{AB}) = \frac{1}{2}[\,\text{tr}_B(|00\rangle\langle 00|) + \,\text{tr}_B(|00\rangle\langle 11|) + \,\text{tr}_B(|11\rangle\langle 00|) + \,\text{tr}_B(|11\rangle\langle 11|)] \\ =& \frac{1}{2}(|0\rangle\langle 0|\langle 0|0\rangle + |0\rangle\langle 1|\langle 0|1\rangle + |1\rangle\langle 0|\langle 1|0\rangle + |1\rangle\langle 1|\langle 1|1\rangle) = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|),\end{aligned}$$

and the reduced density operator of qubit B is

$$\begin{aligned}\rho_{00}^B =& \,\text{tr}_A(\rho_{00}^{AB}) = \frac{1}{2}[\,\text{tr}_A(|00\rangle\langle 00|) + \,\text{tr}_A(|00\rangle\langle 11|) + \,\text{tr}_A(|11\rangle\langle 00|) + \,\text{tr}_A(|11\rangle\langle 11|)] \\ =& \frac{1}{2}(\langle 0|0\rangle|0\rangle\langle 0| + \langle 0|1\rangle|0\rangle\langle 1| + \langle 1|0\rangle|1\rangle\langle 0| + \langle 1|1\rangle|1\rangle\langle 1|) = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|),\end{aligned} \tag{2.338}$$

For the Bell state

$$|\beta_{01}\rangle = \frac{|01\rangle + |10\rangle}{\sqrt{2}}, \tag{2.339}$$

its density operator is

$$\rho_{01}^{AB} = |\beta_{01}\rangle\langle\beta_{01}| = \frac{|01\rangle + |10\rangle}{\sqrt{2}}\frac{\langle 01| + \langle 10|}{\sqrt{2}} = \frac{1}{2}(|01\rangle\langle 01| + |01\rangle\langle 10| + |10\rangle\langle 01| + |10\rangle\langle 10|), \tag{2.340}$$

the reduced density operator of qubit A is

$$\begin{aligned}\rho_{01}^A =& \,\text{tr}_B(\rho_{01}^{AB}) = \frac{1}{2}[\,\text{tr}_B(|01\rangle\langle 01|) + \,\text{tr}_B(|01\rangle\langle 10|) + \,\text{tr}_B(|10\rangle\langle 01|) + \,\text{tr}_B(|10\rangle\langle 10|)] \\ =& \frac{1}{2}(|0\rangle\langle 0|\langle 1|1\rangle + |0\rangle\langle 1|\langle 1|0\rangle + |1\rangle\langle 0|\langle 0|1\rangle + |1\rangle\langle 1|\langle 0|0\rangle) = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|),\end{aligned} \tag{2.341}$$

and the reduced density operator of qubit B is

$$\begin{aligned}\rho_{01}^{B} =& \operatorname{tr}_A(\rho_{01}^{AB}) = \frac{1}{2}[\operatorname{tr}_A(|01\rangle\langle 01|) + \operatorname{tr}_A(|01\rangle\langle 10|) + \operatorname{tr}_A(|10\rangle\langle 01|) + \operatorname{tr}_A(|10\rangle\langle 10|)] \\ =& \frac{1}{2}(\langle 0|0\rangle|1\rangle\langle 1| + \langle 0|1\rangle|1\rangle\langle 0| + \langle 1|0\rangle|0\rangle\langle 1| + \langle 1|1\rangle|0\rangle\langle 0|) = \frac{1}{2}(|1\rangle\langle 1| + |0\rangle\langle 0|). \end{aligned} \tag{2.342}$$

For the Bell state

$$|\beta_{10}^{AB}\rangle = \frac{|00\rangle - |11\rangle}{\sqrt{2}}, \tag{2.343}$$

$$\rho_{10}^{AB} = |\beta_{10}^{AB}\rangle\langle\beta_{10}^{AB}| = \frac{|00\rangle - |11\rangle}{\sqrt{2}}\frac{\langle 00| - \langle 11|}{\sqrt{2}} = \frac{1}{2}(|00\rangle\langle 00| - |00\rangle\langle 11| - |11\rangle\langle 00| + |11\rangle\langle 11|), \tag{2.344}$$

the reduced density operator of qubit A is

$$\begin{aligned}\rho_{10}^{A} =& \operatorname{tr}_B(\rho_{10}^{AB}) = \frac{1}{2}[\operatorname{tr}_B(|00\rangle\langle 00|) - \operatorname{tr}_B(|00\rangle\langle 11|) - \operatorname{tr}_B(|11\rangle\langle 00|) + \operatorname{tr}_B(|11\rangle\langle 11|)] \\ =& \frac{1}{2}(|0\rangle\langle 0|\langle 0|0\rangle - |0\rangle\langle 1|\langle 0|1\rangle - |1\rangle\langle 0|\langle 1|0\rangle + |1\rangle\langle 1|\langle 1|1\rangle) = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|), \end{aligned} \tag{2.345}$$

and the reduced density operator of qubit B is

$$\begin{aligned}\rho_{10}^{B} =& \operatorname{tr}_A(\rho_{10}^{AB}) = \frac{1}{2}[\operatorname{tr}_A(|00\rangle\langle 00|) - \operatorname{tr}_A(|00\rangle\langle 11|) - \operatorname{tr}_A(|11\rangle\langle 00|) + \operatorname{tr}_A(|11\rangle\langle 11|)] \\ =& \frac{1}{2}(\langle 0|0\rangle|0\rangle\langle 0| - \langle 0|1\rangle|0\rangle\langle 1| - \langle 1|0\rangle|1\rangle\langle 0| + \langle 1|1\rangle|1\rangle\langle 1|) = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|). \end{aligned} \tag{2.346}$$

For the Bell state

$$|\beta_{11}^{AB}\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}}, \tag{2.347}$$

its density operator is

$$\rho_{11}^{AB} = |\beta_{11}^{AB}\rangle\langle\beta_{11}^{AB}| = \frac{|01\rangle - |10\rangle}{\sqrt{2}}\frac{\langle 01| - \langle 10|}{\sqrt{2}} = \frac{1}{2}(|01\rangle\langle 01| - |01\rangle\langle 10| - |10\rangle\langle 01| + |10\rangle\langle 10|), \tag{2.348}$$

the reduced density operator of qubit A is

$$\begin{aligned}\rho_{11}^{A} =& \operatorname{tr}_B(\rho_{11}^{AB}) = \frac{1}{2}[\operatorname{tr}_B(|01\rangle\langle 01|) - \operatorname{tr}_B(|01\rangle\langle 10|) - \operatorname{tr}_B(|10\rangle\langle 01|) + \operatorname{tr}_B(|10\rangle\langle 10|)] \\ =& \frac{1}{2}(|0\rangle\langle 0|\langle 1|1\rangle - |0\rangle\langle 1|\langle 1|0\rangle - |1\rangle\langle 0|\langle 0|1\rangle + |1\rangle\langle 1|\langle 0|0\rangle) = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|), \end{aligned} \tag{2.349}$$

and the reduced density operator of qubit B is

$$\begin{aligned}\rho_{11}^{B} =& \operatorname{tr}_A(\rho_{11}^{AB}) = \frac{1}{2}[\operatorname{tr}_A(|01\rangle\langle 01|) - \operatorname{tr}_A(|01\rangle\langle 10|) - \operatorname{tr}_A(|10\rangle\langle 01|) + \operatorname{tr}_A(|10\rangle\langle 10|)] \\ =& \frac{1}{2}(\langle 1|1\rangle|0\rangle\langle 0| - \langle 1|0\rangle|0\rangle\langle 1| - \langle 0|1\rangle|1\rangle\langle 0| + \langle 0|0\rangle|1\rangle\langle 1|) = \frac{1}{2}(|1\rangle\langle 1| + |0\rangle\langle 0|). \end{aligned} \tag{2.350}$$

□

2.5 The Schmidt decomposition and purification

Exercise 2.76. Extend the proof the Schmidt decomposition to the case where A and B may have the state spaces of different dimensionality.

Proof: Suppose the dimension of the state space for systems A and B are n and m, respectively, and let $|j\rangle$ and $|k\rangle$ be any fixed orthonormal bases for systems A and B, respectively. Then $|\psi\rangle$ can be written as

$$|\psi\rangle = \sum_{j=1}^{n}\sum_{k=1}^{m} a_{jk}|j\rangle|k\rangle, \tag{2.351}$$

for some matrix a of complex numbers a_{jk}. Without loss of generality, set $n > m$. By the singular value decomposition,

$$a = u \begin{bmatrix} d \\ 0 \end{bmatrix} v, \tag{2.352}$$

where u is an $n \times n$ unitary matrix, v is an $m \times m$ unitary matrix, d is an $m \times m$ diagonal matrix with non-negative real elements, and $\left[\begin{smallmatrix} d \\ 0 \end{smallmatrix}\right]$ is a $n \times m$ matrix whose $(m+1)$th to nth row are all zero. Thus

$$|\psi\rangle = \sum_{j=1}^{n} \sum_{i=1}^{m} \sum_{k=1}^{m} u_{ji} d_{ii} v_{ik} |j\rangle |k\rangle. \tag{2.353}$$

Define

$$|i_A\rangle \equiv \sum_{j=1}^{n} u_{ji} |j\rangle, \tag{2.354}$$

$$|i_B\rangle \equiv \sum_{k=1}^{m} u_{ik} |k\rangle, \tag{2.355}$$

and

$$\lambda_i \equiv d_{ii}. \tag{2.356}$$

We see that this gives

$$|\psi\rangle = \sum_{i=1}^{m} \lambda_i |i_A\rangle |i_B\rangle. \tag{2.357}$$

Due to the unitarity of u and v and the orthonormality of $|j\rangle$ and $|k\rangle$,

$$\langle i_A | l_A \rangle = \left(\sum_{j=1}^{n} u_{ji} |j\rangle \right)^{\dagger} \left(\sum_{j'=1}^{n} u_{j'l} |j'\rangle \right) = \sum_{j=1}^{n} \sum_{j'=1}^{n} u_{ij}^* u_{j'l} \langle j | j' \rangle = \sum_{j=1}^{n} \sum_{j'=1}^{n} u_{ij}^* u_{j'i} \delta_{jj'} = \sum_{j=1}^{n} u_{ij}^* u_{jl} = \delta_{il}, \tag{2.358}$$

$$\langle i_B | l_B \rangle = \left(\sum_{k=1}^{m} v_{ik} |k\rangle \right)^{\dagger} \left(\sum_{k'=1}^{m} v_{lk'} |k'\rangle \right) = \sum_{k=1}^{m} \sum_{k'=1}^{m} v_{ki}^* v_{lk'} \langle k | k' \rangle = \sum_{k=1}^{m} \sum_{k'=1}^{m} v_{ki}^* v_{lk'} \delta_{kk'} = \sum_{k=1}^{m} v_{ki}^* u_{lk} = \delta_{il}, \tag{2.359}$$

$|i_A\rangle$ and $|i_B\rangle$ form two orthonormal sets, respectively. Moreover, since $|\psi\rangle$ is a normalized pure state,

$$\langle \psi | \psi \rangle = \left(\sum_{i=1}^{m} \lambda_i |i_A\rangle |i_B\rangle \right)^{\dagger} \left(\sum_{l=1}^{m} \lambda_l |l_A\rangle |l_B\rangle \right) = \sum_{i=1}^{m} \sum_{l=1}^{m} \lambda_i^2 \langle i_A | l_A \rangle \langle i_B | l_B \rangle = \sum_{i=1}^{m} \sum_{l=1}^{m} \lambda_i \lambda_l \delta_{il} = \sum_{i=1}^{m} \lambda_i^2 = 1. \tag{2.360}$$

Therefore, there exists orthonormal states $|i_A\rangle$ for system A, and orthonormal states $|i_B\rangle$ of system B such that

$$|\psi\rangle = \sum_{i} \lambda_i |i_A\rangle |i_B\rangle, \tag{2.361}$$

where λ_i are non-negative real number satisfying $\sum_i \lambda_i^2 = 1$. □

Exercise 2.77. Suppose ABC is a three component quantum system. Show by example that there are quantum state $|\psi\rangle$ of such systems which can not be written in the form

$$|\psi\rangle = \sum_{i} \lambda_i |i_A\rangle |i_B\rangle |i_C\rangle, \tag{2.362}$$

where λ_i are real numbers, and $|i_A\rangle$, $|i_B\rangle$, $|i_C\rangle$ are orthonormal bases of the respective systems.

Solution: Prove this according to [1]. Suppose the dimension of the vector spaces of the three component systems are all d. Taking the normalization condition and the common phase factor taken into consideration, $2d^3 - 2$ real parameters are needed to determine the state of the composite system

$$|\psi\rangle = \sum_{j,k,l=1}^{d} A_{jkl} |j\rangle |k\rangle |l\rangle, \tag{2.363}$$

where $|j\rangle$, $|k\rangle$ and $|l\rangle$ are any fixed orthonormal bases for systems A, B and C, respectively. To rewrite the state into the form

$$|\psi\rangle = \sum_{i=1}^{d} |i_A\rangle|i_B\rangle|i_C\rangle, \tag{2.364}$$

where λ_i are real numbers, and $|i_A\rangle$, $|i_B\rangle$, $|i_C\rangle$ are orthonormal bases of the respective systems, we need take unitary transforms,

$$|i_A\rangle = \sum_{j=1}^{d} U_{ij}|j\rangle, \tag{2.365}$$

$$|i_B\rangle = \sum_{k=1}^{d} V_{ik}|k\rangle, \tag{2.366}$$

$$|i_C\rangle = \sum_{l=1}^{d} W_{il}|l\rangle. \tag{2.367}$$

Each $d \times d$ unitary matrix has $d(d-1)$ real parameters, so we only have $3d(d-1) + 2d - 2 = 3d^2 - d - 2$ real parameters in the rewriting process, which are not enough to solve the problem in general.

Therefore, there must be quantum state $|\psi\rangle$ of such systems which can not be written in the form

$$|\psi\rangle = \sum_{i} \lambda_i |i_A\rangle|i_B\rangle|i_C\rangle. \tag{2.368}$$

One of the examples is a quantum system consists of three qubits, A, B, and C, whose state is

$$|\psi\rangle = \frac{|001\rangle + |010\rangle + |100\rangle}{\sqrt{3}}. \tag{2.369}$$

□

Exercise 2.78. Prove that a state $|\psi\rangle$ of a composite system AB is a product state if and only if it has Schmidt number 1. Prove that $|\psi\rangle$ is a product state if and only if ρ^A (and thus ρ^B) are pure states.

Proof: *Sufficiency*: If the state $|\psi\rangle$ has Schmidt number 1, it can be written as

$$|\psi\rangle = \lambda_1 |1_A\rangle|1_B\rangle, \tag{2.370}$$

where the normalization condition requires that $|\lambda_1| = 1$ and we can actually absorb the phase factor of λ_1 into $|1_A\rangle$ or $|1_B\rangle$, so the phase $|\psi\rangle$ is a product state.
Necessity: If the state $|\psi\rangle$ is a product state, it can be written as the product of the state of A and the state of B,

$$|\psi\rangle = |1_A\rangle|1_B\rangle, \tag{2.371}$$

so its Schmidt number is 1.

Sufficiency: If ρ^A (and thus ρ^B) are pure states, they can be written as

$$\rho^A = |1_A\rangle\langle 1_A|, \tag{2.372}$$

$$\rho^B = |1_B\rangle\langle 1_B|, \tag{2.373}$$

$$\Longrightarrow \rho = \rho^A \otimes \rho^B = |1_A\rangle|1_B\rangle\langle 1_A|\langle 1_B| = |\psi\rangle\langle\psi|, \tag{2.374}$$

so $|\psi\rangle = |1_A\rangle|1_B\rangle$ is a product state.
Necessity: If $|\psi\rangle$ is a product state, it can be written as

$$|\psi\rangle = |1_A\rangle|1_B\rangle, \tag{2.375}$$

and its density matrix is

$$\rho = |\psi\rangle\langle\psi| = |1_A\rangle|1_B\rangle\langle 1_A|\langle 1_B|, \tag{2.376}$$

so

$$\rho^A = \mathrm{tr}_B(\rho) = |1_A\rangle\langle 1_A|, \tag{2.377}$$

$$\rho^B = \mathrm{tr}_A(\rho) = |1_B\rangle\langle 1_B|, \tag{2.378}$$

are pure states. □

Exercise 2.79. Consider a composite system consisting of two qubits. Find the Schmidt decomposition of the states

$$\frac{|00\rangle+|11\rangle}{\sqrt{2}};\quad \frac{|00\rangle+|01\rangle+|10\rangle+|11\rangle}{2};\quad \text{and}\quad \frac{|00\rangle+|01\rangle+|10\rangle}{\sqrt{3}}. \tag{2.379}$$

Proof: State $\frac{|00\rangle+|11\rangle}{\sqrt{2}}$ is already in the form of Schmidt decomposition,

$$\frac{|00\rangle+|11\rangle}{\sqrt{2}}=\frac{1}{\sqrt{2}}|0\rangle|0\rangle+\frac{1}{\sqrt{2}}|1\rangle|1\rangle. \tag{2.380}$$

For state $\frac{|00\rangle+|01\rangle+|10\rangle+|11\rangle}{2}$, by singular decomposition, we have

$$a=\frac{1}{2}\begin{bmatrix}1&1\\1&1\end{bmatrix}=\frac{1}{\sqrt{2}}\begin{bmatrix}-1&1\\1&1\end{bmatrix}\begin{bmatrix}0&0\\0&1\end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix}-1&1\\1&1\end{bmatrix}=udv, \tag{2.381}$$

where

$$u=v=\frac{1}{\sqrt{2}}\begin{bmatrix}-1&1\\1&1\end{bmatrix}, \tag{2.382}$$

$$d=\begin{bmatrix}0&0\\0&1\end{bmatrix}. \tag{2.383}$$

Hence we define

$$|1_A\rangle=u_{11}|0\rangle+u_{12}|1\rangle=\frac{1}{\sqrt{2}}(-|0\rangle+|1\rangle), \tag{2.384}$$

$$|2_A\rangle=u_{21}|1\rangle+u_{22}|1\rangle=\frac{1}{\sqrt{2}}(|0\rangle+|1\rangle), \tag{2.385}$$

$$|1_B\rangle=v_{11}|0\rangle+v_{12}|1\rangle=\frac{1}{\sqrt{2}}(-|0\rangle+|1\rangle), \tag{2.386}$$

$$|2_B\rangle=v_{21}|1\rangle+v_{22}|1\rangle=\frac{1}{\sqrt{2}}(|0\rangle+|1\rangle), \tag{2.387}$$

and have Schmidt decomposition

$$\frac{|00\rangle+|01\rangle+|10\rangle+|11\rangle}{2}=\sum_{i=1}^{2}d_{ii}|i_A\rangle|i_B\rangle=\frac{1}{\sqrt{2}}(|0\rangle+|1\rangle)\otimes\frac{1}{\sqrt{2}}(|0\rangle+|1\rangle). \tag{2.388}$$

For state $\frac{|00\rangle+|01\rangle+|10\rangle}{\sqrt{3}}$, by singular decomposition,

$$a=\frac{1}{\sqrt{3}}\begin{bmatrix}1&1\\1&0\end{bmatrix}=\begin{bmatrix}-\sqrt{\frac{5-\sqrt{5}}{10}}&\sqrt{\frac{5+\sqrt{5}}{10}}\\\sqrt{\frac{5+\sqrt{5}}{10}}&\sqrt{\frac{5-\sqrt{5}}{10}}\end{bmatrix}\begin{bmatrix}\frac{1-\sqrt{5}}{2\sqrt{3}}&0\\0&\frac{1+\sqrt{5}}{2\sqrt{3}}\end{bmatrix}\begin{bmatrix}-\sqrt{\frac{5-\sqrt{5}}{10}}&\sqrt{\frac{5+\sqrt{5}}{10}}\\\sqrt{\frac{5+\sqrt{5}}{10}}&\sqrt{\frac{5-\sqrt{5}}{10}}\end{bmatrix}=udv, \tag{2.389}$$

where

$$u=v=\begin{bmatrix}-\sqrt{\frac{5-\sqrt{5}}{10}}&\sqrt{\frac{5+\sqrt{5}}{10}}\\\sqrt{\frac{5+\sqrt{5}}{10}}&\sqrt{\frac{5-\sqrt{5}}{10}}\end{bmatrix}, \tag{2.390}$$

$$d=\begin{bmatrix}\frac{1-\sqrt{5}}{2\sqrt{3}}&0\\0&\frac{1+\sqrt{5}}{2\sqrt{3}}\end{bmatrix}. \tag{2.391}$$

Hence we define

$$|1_A\rangle=u_{11}|0\rangle+u_{12}|1\rangle=-\sqrt{\frac{5-\sqrt{5}}{10}}|0\rangle+\sqrt{\frac{5+\sqrt{5}}{10}}|1\rangle, \tag{2.392}$$

$$|2_A\rangle=u_{21}|0\rangle+u_{22}|1\rangle=\sqrt{\frac{5+\sqrt{5}}{10}}|0\rangle+\sqrt{\frac{5-\sqrt{5}}{10}}|1\rangle, \tag{2.393}$$

$$|1_B\rangle = v_{11}|0\rangle + v_{12}|1\rangle = -\sqrt{\frac{5-\sqrt{5}}{10}}|0\rangle + \sqrt{\frac{5+\sqrt{5}}{10}}|1\rangle, \tag{2.394}$$

$$|2_B\rangle = v_{21}|0\rangle + v_{22}|1\rangle = \sqrt{\frac{5+\sqrt{5}}{10}}|0\rangle + \sqrt{\frac{5-\sqrt{5}}{10}}|1\rangle, \tag{2.395}$$

and have Schmidt decomposition

$$\begin{aligned}\frac{|00\rangle + |01\rangle + |10\rangle}{\sqrt{3}} =& \frac{1-\sqrt{5}}{2\sqrt{3}} \left(-\sqrt{\frac{5-\sqrt{5}}{10}}|0\rangle + \sqrt{\frac{5+\sqrt{5}}{10}}|1\rangle\right) \otimes \left(-\sqrt{\frac{5-\sqrt{5}}{10}}|0\rangle + \sqrt{\frac{5+\sqrt{5}}{10}}|1\rangle\right) \\ &+ \frac{1+\sqrt{5}}{2\sqrt{3}} \left(\sqrt{\frac{5+\sqrt{5}}{10}}|0\rangle + \sqrt{\frac{5-\sqrt{5}}{10}}|1\rangle\right) \otimes \left(\sqrt{\frac{5+\sqrt{5}}{10}}|0\rangle + \sqrt{\frac{5-\sqrt{5}}{10}}|1\rangle\right).\end{aligned} \tag{2.396}$$

□

Exercise 2.80. Suppose $|\psi\rangle$ and $|\varphi\rangle$ are two pure states of a composite quantum system with components A and B, with identical Schmidt coefficients. Show that there are unitary transformations U on system A and V on system B such that $|\psi\rangle = (U \otimes V)|\varphi\rangle$.

Proof: Suppose the two pure states are respectively

$$|\psi\rangle = \sum_i \lambda_i |\psi_i^A\rangle|\psi_i^B\rangle, \tag{2.397}$$

$$|\varphi\rangle = \sum_i \lambda_i |\varphi_i^A\rangle|\varphi_i^B\rangle. \tag{2.398}$$

Define the two unitary matrices as

$$U = \sum_i |\psi_i^A\rangle\langle\varphi_i^A|, \tag{2.399}$$

$$V = \sum_i |\psi_i^B\rangle\langle\varphi_i^B|. \tag{2.400}$$

Then

$$\begin{aligned}(U \otimes V)|\varphi\rangle =& \sum_j |\psi_j^A\rangle\langle\varphi_j^A| \sum_k |\psi_k^B\rangle\langle\varphi_k^B| \sum_i \lambda_i |\varphi_i^A\rangle|\varphi_i^B\rangle = \sum_{i,j,k} \lambda_i |\psi_j^A\rangle\langle\varphi_j^A|\varphi_i^A\rangle|\psi_k^B\rangle\langle\varphi_k^A|\varphi_i^B\rangle \\ =& \sum_{i,j,k} \delta_{ij}\delta_{ik}\lambda_i |\psi_j^A\rangle|\psi_k^B\rangle = \sum_i \lambda_i |\psi_i^A\rangle|\psi_i^B\rangle = |\psi\rangle.\end{aligned} \tag{2.401}$$

□

Exercise 2.81 (Freedom in purifications). Let $|AR_1\rangle$ and $|AR_2\rangle$ be two purifications of a state ρ^A to a composite system AR. Prove that there exists a unitary transformation U_R acting on system R such that $|AR_1\rangle = (I_A \otimes U_R)|AR_2\rangle$.

Proof: Let the Schmidt decomposition of $|AR_1\rangle$ and $|AR_2\rangle$ be respectively

$$|AR_1\rangle = \sum_i \sqrt{p_i}|\psi_i^A\rangle|\psi_i^R\rangle, \tag{2.402}$$

$$|AR_2\rangle = \sum_i \sqrt{q_i}|\varphi_i^A\rangle|\varphi_i^R\rangle. \tag{2.403}$$

Since $|AR_1\rangle$ and $|AR_2\rangle$ are two purifications of state ρ^A,

$$\rho^A = \text{tr}_R(|AR_1\rangle\langle AR_1|) = \text{tr}_R(|AR_2\rangle\langle AR_2|), \tag{2.404}$$

$$\implies \sum_i p_i |\psi_i^A\rangle\langle\psi_i^A| = \sum_i q_i |\varphi_i^A\rangle\langle\varphi_i^A|. \tag{2.405}$$

Since both $|\psi_i^A\rangle$ and $|\varphi_i^A\rangle$ are orthogonal bases of system A and eigenvectors of ρ^A, without loss of generality, we can set

$$\sqrt{p_i} = \sqrt{q_i} \equiv \lambda_i, \tag{2.406}$$

$$|\psi_i^A\rangle = |\varphi_i^A\rangle \equiv |i^A\rangle. \tag{2.407}$$

In this way,

$$|AR_1\rangle = \sum_i \lambda_i |i^A\rangle|\psi_i^R\rangle, \tag{2.408}$$

$$|AR_2\rangle = \sum_i \lambda_i |i^A\rangle|\varphi_i^R\rangle \tag{2.409}$$

Define the unitary transformation as

$$U_R = \sum_i |\psi_i^R\rangle\langle\varphi_i^R|. \tag{2.410}$$

Then

$$\begin{aligned}(I_A \otimes U_R)|AR_2\rangle =& \left(I_A \otimes \sum_j |\psi_i^R\rangle\langle\varphi_i^R|\right) \sum_i \lambda_i |i^A\rangle|\varphi_i^R\rangle = \sum_{i,j} \lambda_i |i^A\rangle|\psi_j^R\rangle\langle\varphi_j^R|\varphi_i^R\rangle = \sum_{i,j} \delta_{ij}\lambda_i |i^A\rangle|\psi_i^R\rangle = \sum_i \lambda_i |i^A\rangle|\psi_i^R\rangle \\ =& |AR_2\rangle. \end{aligned} \tag{2.411}$$

□

Exercise 2.82. Suppose $\{p_i, |\psi_i\rangle\}$ is an ensemble of states generating a density matrix $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$ for a quantum system A. Introduce a system R with orthonormal basis $|i\rangle$.

(1) Show that $\sum_i \sqrt{p_i}|\psi_i\rangle|i\rangle$ is a purification of ρ.

(2) Suppose we measure R in the basis $|i\rangle$, obtaining outcome i. With what probability do we obtain the result i, and what is the corresponding state of system A.

(3) Let $|AR\rangle$ be any purification of ρ to the system AR. Show that there exists an orthonormal basis $|i\rangle$ in which R can be measured such that the corresponding post-measurement state for system A is $|\psi_i\rangle$ with probability p_i.

Solution: (1) Since

$$\mathrm{tr}_R \left[\left(\sum_i \sqrt{p_i}|\psi_i\rangle|i\rangle \right) \left(\sum_j \sqrt{p_j^*}\langle\psi_j|\langle j| \right) \right] = \sum_i p_i |\psi_i\rangle\langle\psi_i| = \rho, \tag{2.412}$$

$\sum_i \sqrt{p_i}|\psi_i\rangle|i\rangle$ is a purification of ρ.

(2) The probability of obtaining i is

$$\begin{aligned} p(i) =& \mathrm{tr} \left[(I^A \otimes M_i^R)^\dagger (I_A \otimes M_i^R) \left(\sum_j \sqrt{p_j}|\psi_j\rangle|j\rangle \right) \left(\sum_k \sqrt{p_k^*}\langle\psi_k|\langle k| \right) \right] \\ =& \mathrm{tr} \left[\sum_{j,k} \sqrt{p_j p_k^*}|\psi_j\rangle\langle\psi_k| \otimes |i\rangle\langle i|i\rangle\langle i|j\rangle\langle k| \right] = \mathrm{tr} \left[\sum_{j,k} \delta_{ij}\sqrt{p_j p_k^*}|\psi_j\rangle\langle\psi_k| \otimes |i\rangle\langle k| \right] \\ =& \mathrm{tr} \left[\sum_k \sqrt{p_i p_k^*}|\psi_i\rangle\langle\psi_k| \otimes |i\rangle\langle k| \right] = |p_i|. \end{aligned} \tag{2.413}$$

The state of the joint system after measurement is

$$\begin{aligned} \frac{(I^A \otimes M_i^R)^\dagger \left(\sum_j \sqrt{p_j}|\psi_j\rangle|j\rangle\right) \left(\sum_k \sqrt{p_k^*}\langle\psi_k|\langle k|\right) (I_A \otimes M_i^R)}{p(i)} =& \frac{\sum_{j,k} \sqrt{p_j p_k^*}|\psi_j\rangle\langle\psi_k||i\rangle\langle i|j\rangle\langle k|i\rangle\langle i|}{|p_i|} \\ =& \frac{\sum_{j,k} \delta_{ij}\delta_{ik}\sqrt{p_j p_k^*}|\psi_j\rangle\langle\psi_k| \otimes |i\rangle\langle i|}{|p_i|} \\ =& \frac{|p_i|\,|\psi_i\rangle\langle\psi_i| \otimes |i\rangle\langle i|}{|p_i|} = |\psi_j\rangle\langle\psi_k| \otimes |i\rangle\langle i|. \end{aligned} \tag{2.414}$$

The corresponding state of system A is

$$\mathrm{tr}_R(|\psi_j\rangle\langle\psi_k| \otimes |i\rangle\langle i|) = |\psi_i\rangle\langle\psi_i|, \tag{2.415}$$

i.e. $|\psi_i\rangle$.

(3) Suppose the Schmidt decomposition of the purification $|AR\rangle$ is

$$|AR\rangle = \sum_i \sqrt{q_i}|\varphi_i^A\rangle|\varphi_i^R\rangle. \tag{2.416}$$

According to Theorem 2.6, there exists a unitary matrix u_{ij} such that

$$\sqrt{q_i}|\varphi_i^A\rangle = \sum_j u_{ij}\sqrt{p_j}|\psi_i\rangle. \tag{2.417}$$

In this way,

$$|AR\rangle = \sum_i \left(\sum_j u_{ij}\sqrt{p_j}|\psi_i\rangle\right)|\varphi_i^R\rangle = \sum_j \sqrt{p_j}|\psi_j\rangle\left(\sum_i u_{ij}|\varphi_i^R\rangle\right) = \sum_j \sqrt{p_j}|\psi_j\rangle|j\rangle = \sum_i \sqrt{p_i}|\psi_i\rangle|i\rangle, \tag{2.418}$$

where

$$|i\rangle = \sum_j u_{ji}|\varphi_j^R\rangle. \tag{2.419}$$

Therefore, there exists an orthonormal basis $|i\rangle$ in which R can be measured such that the corresponding post-measurement state for system A is $|\psi_i\rangle$ with probability p_i.

□

2.6 EPR and the Bell inequality

Problem 2.1 (Functions of the Pauli matrices). Let $f(\cdot)$ be any function from complex numbers to complex numbers. Let $\vec{n}$ be a normalized vector in three dimensions, and let θ be real. Show that

$$f(\theta\vec{n}\cdot\vec{\sigma}) = \frac{f(\theta)+f(-\theta)}{2}I + \frac{f(\theta)-f(-\theta)}{2}\vec{n}\cdot\vec{\sigma}. \tag{2.420}$$

Proof: The left side of equation (2.420) is

$$f(\theta\vec{n}\cdot\vec{\sigma}) = \sum_{k=0}^{\infty}\frac{1}{k!}f^{(k)}(0)(\theta\vec{n}\cdot\vec{\sigma})^k = \sum_{k=0}^{\infty}\frac{1}{(2k)!}(\theta\vec{n}\cdot\vec{\sigma})^{2k} + \sum_{k=0}^{\infty}\frac{1}{(2k+1)!}(\theta\vec{n}\cdot\vec{\sigma})^{2k+1}.$$

Note that

$$(\vec{n}\cdot\vec{\sigma})^2 = \left(\sum_{i=1}^{3} n_i\sigma_i\right)^2 = \sum_{i,j=1}^{3} v_i v_j \sigma_i\sigma_j. \tag{2.421}$$

Using the anti commutation relation between the Pauli matrices,

$$\{\sigma_i,\sigma_j\} = \sigma_i\sigma_j + \sigma_j\sigma_i = 2\delta_{ij}I = \begin{cases} 2I, & i=j; \\ 0, & i\neq j, \end{cases} \tag{2.422}$$

we have

$$(\vec{n}\cdot\vec{\sigma})^2 = \sum_i n_i^2 I = I. \tag{2.423}$$

Hence the left side of equation (2.420) can be written as

$$f(\theta\vec{n}\cdot\vec{\sigma}) = \sum_{k=0}^{\infty}\frac{1}{(2k)!}f^{(2k)}(0)\theta^{2k} + \sum_{k=0}^{\infty}\frac{1}{(2k+1)!}f^{(2k+1)}(0)\theta^{2k+1}\vec{n}\cdot\vec{\sigma}. \tag{2.424}$$

Since

$$f(\theta)+f(-\theta)=\sum_{k=0}^{\infty}\frac{1}{k!}f^{(k)}(0)\theta^k+\sum_{k=0}^{\infty}\frac{1}{k!}f^{(k)}(0)(-\theta)^k=\sum_{k=0}^{\infty}\frac{1}{(2k)!}f^{(2k)}(0)\theta^{2k}, \tag{2.425}$$

$$f(\theta)-f(-\theta)=\sum_{k=0}^{\infty}\frac{1}{k!}f^{(k)}(0)\theta^k-\sum_{k=0}^{\infty}\frac{1}{k!}f^{(k)}(0)(-\theta)^k=\sum_{k=0}^{\infty}\frac{1}{(2k+1)!}f^{(2k+1)}(0)\theta^{2k+1}, \tag{2.426}$$

the left side of equation (2.420) equals its right side,

$$f(\theta\vec{n}\cdot\vec{\sigma})=\frac{f(\theta)-f(-\theta)}{2}I+\frac{f(\theta)-f(-\theta)}{2}\vec{n}\cdot\vec{\sigma}. \tag{2.427}$$

Therefore, equation (2.420) holds. □

Problem 2.2 (Properties of the Schmidt number). Suppose $|\psi\rangle$ is a pure state of a composite system with components A and B.

(1) Prove that the Schmidt number of $|\psi\rangle$ is equal to the rank of the reduced density matrix $\rho_A\equiv \mathrm{tr}_B(|\psi\rangle\langle\psi|)$. (Note that the rank of a Hermitian operator is equal to the dimension of its support.)

(2) Suppose $|\psi\rangle=\sum_j|\alpha_j\rangle|\beta_j\rangle$ is a representation for $|\psi\rangle$, where $|\alpha_i\rangle$ and $|\beta_j\rangle$ are (un-normalized) states for system A and B, respectively. Prove that the number of terms in such a decomposition is greater than or equal to the Schmidt number of $|\psi\rangle$, $\mathrm{Sch}(\psi)$.

(3) Suppose $|\psi\rangle=\alpha|\varphi\rangle+\beta|\gamma\rangle$. Prove that

$$\mathrm{Sch}(\psi)\geq\left|\mathrm{Sch}(\varphi)-\mathrm{Sch}(\gamma)\right|. \tag{2.428}$$

Proof: (1) Suppose the Schmidt decomposition of $|\psi\rangle$ is

$$|\psi\rangle=\sum_{i=1}^{\mathrm{Sch}(\psi)}\lambda_i|\psi_i^A\rangle|\psi_i^B\rangle. \tag{2.429}$$

The reduced density matrix is

$$\begin{aligned}\rho_A=&\mathrm{tr}_B(|\psi\rangle\langle\psi|)=\mathrm{tr}_B\left[\left(\sum_{i=1}^{\mathrm{Sch}(\psi)}\lambda_i|\psi_i^A\rangle|\psi_i^B\rangle\right)\left(\sum_{j=1}^{\mathrm{Sch}(\psi)}\lambda_j^*\langle\psi_j^A|\langle\psi_j^B|\right)\right]\\ =&\mathrm{tr}_B\left(\sum_{i,j=1}^{\mathrm{Sch}(\psi)}\lambda_i\lambda_j^*|\psi_i^A\rangle\langle\psi_j^A|\otimes|\psi_i^B\rangle\langle\psi_j^B|\right)=\sum_{i=1}^{\mathrm{Sch}(\psi)}|\lambda_i|^2|\psi_i^A\rangle\langle\psi_i^A|,\end{aligned} \tag{2.430}$$

whose rank is equal to the Schmidt number $\mathrm{Sch}(\psi)$.

(2) Suppose the number of terms in the decomposition, $|\psi\rangle=\sum_j|\alpha_j\rangle|\beta_j\rangle$, J is less than $\mathrm{Sch}(\psi)$. The reduced density matrix is

$$\begin{aligned}\rho_A=&\mathrm{tr}_B(|\psi\rangle\langle\psi|)=\mathrm{tr}_B\left[\left(\sum_{j=1}^{J}|\alpha_j\rangle|\beta_j\rangle\right)\left(\sum_{k=1}^{J}\langle\alpha_k|\langle\beta_k|\right)\right]=\mathrm{tr}_B\left(\sum_{j,k=1}^{J}|\alpha_j\rangle|\beta_j\rangle\langle\alpha_k|\langle\beta_k|\right)\\ =&\sum_{i=1}^{\mathrm{Sch}(\psi)}\sum_{j,k=1}^{J}|\alpha_j\rangle\langle\alpha_k|\langle\psi_i^B|\beta_j\rangle\langle\beta_k|\psi_i^B\rangle=\sum_{j,k=1}^{J}\left(\sum_{i=1}^{\mathrm{Sch}(\psi)}\langle\psi_i^B|\beta_j\rangle\langle\beta_k|\psi_i^B\rangle\right)|\alpha_j\rangle\langle\alpha_k|\end{aligned} \tag{2.431}$$

Even if $|\alpha_j\rangle$ are all independent, the rank of the reduced density matrix ρ^A is only J, which is less than $\mathrm{Sch}(\psi)$, and thus less than the rank of ρ^A. Therefore, the supposition is incorrect and the number of terms in such a decomposition is greater than or equal to the Schmidt number of $|\psi\rangle$, $\mathrm{Sch}(\psi)$.

(3) Suppose the Schmidt decompositions of $|\varphi\rangle$ and $|\gamma\rangle$ are respectively

$$|\varphi\rangle = \sum_{i=1}^{\mathrm{Sch}(\varphi)} |\varphi_i^A\rangle|\varphi_i^B\rangle, \tag{2.432}$$

$$|\gamma\rangle = \sum_{i=1}^{\mathrm{Sch}(\gamma)} |\gamma_i^A\rangle|\gamma_i^B\rangle. \tag{2.433}$$

Without loss of generality, we assume $\mathrm{Sch}(\varphi) \geq \mathrm{Sch}(\gamma)$ and thus the original problem is converted to proving that $\mathrm{Sch}(\varphi) + \mathrm{Sch}(\gamma) \geq \mathrm{Sch}(\psi)$. From

$$|\psi\rangle = \alpha|\varphi\rangle + \beta|\gamma\rangle, \tag{2.434}$$

we have

$$|\varphi\rangle = \frac{\beta}{\alpha}|\gamma\rangle - \frac{1}{\alpha}|\psi\rangle = \frac{\beta}{\alpha}\sum_{i=1}^{\mathrm{Sch}(\gamma)} |\gamma_i^A\rangle|\gamma_i^B\rangle - \frac{1}{\alpha}\sum_{i=1}^{\mathrm{Sch}(\psi)} |\psi_i^A\rangle|\psi_i^B\rangle = \sum_{i=1}^{\mathrm{Sch}(\gamma)+\mathrm{Sch}(\psi)} \delta_i|a_i\rangle|b_i\rangle, \tag{2.435}$$

where

$$\delta_i = \begin{cases} \frac{\beta}{\alpha}, & 1 \leq i \leq \mathrm{Sch}(\gamma); \\ -\frac{1}{\alpha}, & \mathrm{Sch}(\gamma) < i \leq \mathrm{Sch}(\gamma) + \mathrm{Sch}(\psi), \end{cases} \tag{2.436}$$

$$a_i = \begin{cases} \gamma_i^A, & 1 \leq i \leq \mathrm{Sch}(\gamma); \\ \psi_i^A, & \mathrm{Sch}(\gamma) < i \leq \mathrm{Sch}(\gamma) + \mathrm{Sch}(\psi), \end{cases} \tag{2.437}$$

$$b_i = \begin{cases} \gamma_i^B, & 1 \leq i \leq \mathrm{Sch}(\gamma); \\ \psi_i^B, & \mathrm{Sch}(\gamma) < i \leq \mathrm{Sch}(\gamma) + \mathrm{Sch}(\psi). \end{cases} \tag{2.438}$$

Using the conclusion obtained in (2), we have

$$\mathrm{Sch}(\gamma) + \mathrm{Sch}(\psi) \geq \mathrm{Sch}(\varphi). \tag{2.439}$$

Therefore,

$$\mathrm{Sch}(\psi) \geq |\mathrm{Sch}(\varphi) - \mathrm{Sch}(\gamma)|\,. \tag{2.440}$$

□

Problem 2.3 (Tsirelson's inequality). Suppose $Q = \vec{q}\cdot\vec{\sigma}$, $R = \vec{r}\cdot\vec{\sigma}$, $S = \vec{s}\cdot\vec{\sigma}$, $T = \vec{t}\cdot\vec{\sigma}$, where $\vec{q}$, $\vec{r}$, $\vec{s}$ and $\vec{t}$ are real unit vectors in three dimensions. Show that

$$(Q\otimes S + R\otimes S + R\otimes T - Q\otimes T)^2 = 4I + [Q,R]\otimes[S,T]. \tag{2.441}$$

Use this result to prove that

$$\langle Q\otimes S\rangle + \langle R\otimes S\rangle + \langle R\otimes T\rangle - \langle Q\otimes T\rangle \leq 2\sqrt{2}, \tag{2.442}$$

so the violation of the Bell inequality found in Equation (2.230) is the maximum possible in quantum mechanics.

Proof:

$$Q = \vec{q}\cdot\vec{\sigma} = q_1\sigma_1 + q_2\sigma_2 + q_3\sigma_3 = q_1\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix} + q_2\begin{bmatrix}0 & -i\\ i & 0\end{bmatrix} + q_3\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix} = \begin{bmatrix}q_3 & q_1 - iq_2\\ q_1 + iq_2 & -q_3\end{bmatrix}, \tag{2.443}$$

$$\Longrightarrow Q^2 = \begin{bmatrix}q_3 & q_1 - iq_2\\ q_1 + iq_2 & -q_3\end{bmatrix}\begin{bmatrix}q_3 & q_1 - iq_2\\ q_1 + iq_2 & -q_3\end{bmatrix} = \begin{bmatrix}q_1^2 + q_2^2 + q_3^2 & 0\\ 0 & q_1^2 + q_2^2 + q_3^2\end{bmatrix} = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} = I. \tag{2.444}$$

Similarly,

$$R^2 = S^2 = T^2 = I. \tag{2.445}$$

Hence

$$(Q\otimes S + R\otimes S + R\otimes T - Q\otimes T)^2 = QQ\otimes SS + QR\otimes SS + QR\otimes ST - QQ\otimes ST$$

$$
\begin{aligned}
&+ RQ \otimes SS + RR \otimes SS + RR \otimes ST - RQ \otimes ST \\
&+ RQ \otimes TS + RR \otimes TS + RR \otimes TT - RQ \otimes TT \\
&- QQ \otimes TS - QR \otimes TS - QR \otimes TT + QQ \otimes TT \\
=&I \otimes I + QR \otimes I + QR \otimes ST - I \otimes ST \\
&+ RQ \otimes I + I \otimes I + I \otimes ST - RQ \otimes ST \\
&+ RQ \otimes TS + I \otimes TS + I \otimes I - RQ \otimes I \\
&- I \otimes TS - QR \otimes TS - QR \otimes I + I \otimes I \\
=&4I + QR \otimes ST - RQ \otimes ST + RQ \otimes TS - QR \otimes TS \\
=&4I + (QR - RQ) \otimes (ST - TS) = 4I + [Q, R] \otimes [S, T]. \qquad (2.446)
\end{aligned}
$$

The average of the above equation is

$$
\langle (Q \otimes S + R \otimes S + R \otimes T - Q \otimes T)^2 \rangle = \langle 4I + [Q, R] \otimes [S, T] \rangle = 4\langle I \rangle + \langle [Q, S] \rangle \otimes \langle [S, T] \rangle = 4 + \langle [Q, R] \rangle \otimes \langle [S, T] \rangle, \qquad (2.447)
$$

where

$$
\langle [Q, R] \rangle = \langle \psi | [Q, R] | \psi \rangle = \langle \psi | QR - RQ | \psi \rangle = \langle \psi | QR | \rangle - \langle \psi | RQ | \psi \rangle = 2i \operatorname{Re} [\langle \psi | QR | \psi \rangle] = 2i \operatorname{Re} [\langle QR \rangle]. \qquad (2.448)
$$

Note that both Q and R are unitarity, so QR is unitary. Since unitary operators preserve the inner products and thus lengths of vectors, $QR|\psi\rangle$ is still a normalized vector and $|\langle [Q, R] \rangle| \leq 1 \Longrightarrow -1 \leq \operatorname{Re} [\langle QR \rangle] \leq 1$. Similarly,

$$
\langle [S, T] \rangle = 2i \operatorname{Re} [\langle ST \rangle]. \qquad (2.449)
$$

and $-1 \leq \operatorname{Re} [\langle ST \rangle] \leq 1$. Therefore,

$$
\begin{aligned}
\langle Q \otimes S \rangle + \langle R \otimes S \rangle + \langle R \otimes T \rangle - \langle Q \otimes T \rangle \leq& \sqrt{\langle (Q \otimes S + R \otimes S + R \otimes T - Q \otimes T)^2 \rangle} = \sqrt{4 + \langle [Q, R] \rangle \otimes \langle [S, T] \rangle} \\
=& \sqrt{4 - 4 \operatorname{Re} [\langle QR \rangle] \operatorname{Re} [\langle ST \rangle]} \leq 2\sqrt{2}. \qquad (2.450)
\end{aligned}
$$

□

Bibliography

[1] Asher Peres. Higher order schmidt decompositions. *arXiv preprint quant-ph/9504006*, 1995.