Chapter 2 solutions

2.19 Pauli Matrices are unitary and Hermitian

We know the first Pauli matrix is the identity matrix, which is by definition unitary and hermitian:

$$\sigma_0^\dagger = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \sigma_0 \tag{1}$$

So, we have :

$$\sigma_0^\dagger \sigma_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I \tag{2}$$

Similarly for the other Pauli Matrices, we have

$$\sigma_1^\dagger = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \sigma_1 \tag{3}$$

So, we have :

$$\sigma_1^\dagger \sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I \tag{4}$$

For σ_2 , we have:

$$\sigma_2^\dagger = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} = \sigma_2 \tag{5}$$

So, we have :

$$\sigma_2^\dagger \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I \tag{6}$$

For σ_3:

$$\sigma_3^\dagger = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \sigma_3 \tag{7}$$

So, we have :

$$\sigma_3^\dagger \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I \tag{8}$$

Hence, all Pauli matrices are unitary and hermitian.

2.20 Basis Change

Given operator A, with two matrix representations A' and A'', on a vector space V with two different orthonormal basis, $|v_i\rangle$ and $|w_i\rangle$. The relation between them:

$$A'_{ij} = \langle v_i|A|v_j\rangle$$

$$\Longrightarrow \sum_k \langle v_i|w_k\rangle\langle w_k|A|v_j\rangle$$

$$\Longrightarrow \sum_{k,l} \langle v_i|w_k\rangle\langle w_k|A|w_l\rangle\langle w_l|v_j\rangle$$

$$\Longrightarrow \sum_{k,l} \langle v_i|U|v_k\rangle\langle w_k|A|w_l\rangle\langle v_l|U^\dagger|v_j\rangle$$

$$\Longrightarrow \sum_{k,l} U_{ik}A''_{kl}U^\dagger_{lj}$$

where $U \equiv \sum_m |w_m\rangle\langle v_m|$

2.21 Spectral Decomposition

The spectral decomposition is an extremely useful representation theorem for normal operators. It states: Any normal operator M on a vector space V is diagonal with respect to some orthonormal basis for V. Conversely, any diagonalizable operator is normal.

We will solve this my the method of induction. We'll assume the $n = 1$ case is trivial. Now, let λ be an eigenvalue of M, P is the projector on to the λ eigenspace, and Q the projector onto the orthogonal component. Then,

$$M = (P+Q)M(P+Q)$$

$$M = PMP + QMP + PMQ + QMQ$$

Now, obviously $PMP = \lambda P, QMP = 0$, since M takes the P subspace onto itself.

Furthermore, we can take the Hermitian conjugate of both sides to show that:

$$0 = (QMP)^\dagger = PMQ$$

Showing that QMQ is normal is trivial, since $QMQ = (QMQ)^\dagger$. Since PMP is diagonal with respect to some vector space, and QMQ is normal , and thus diagonal with respect to another orthogonal vector space, $PMP + QMQ$ is diagonal with respect to some orthonormal basis for the total vector space.

Alternatively

Suppose M be hermitian. Then $M = M^\dagger$

$$M = IMI$$

$$M + (P + Q)M(P + Q)$$

$$M = PMP + QMP + PMQ + QMQ$$

Now, $PMP = \lambda P$, $QMP = 0$, $PMQ = PM^\dagger Q = (QMP)^* = 0$. Thus $M = PMP + QMQ$

Now, we need to prove QMQ is normal.

$$QMQ(QMQ)^\dagger = QMQQM^\dagger Q$$

$$\implies QM^\dagger QQMQ$$

$$\implies (QM^\dagger Q)QMQ$$

Therefore QMQ is normal. By induction, QMQ is diagonal.

2.22 Two Eigenvectors of a Hermitian operator with different eigenvalues are necessarily orthogonal

Let us assume M be the hermitian operator and $|v_i\rangle$ are the eigenvectors of M with eigenvalues λ_i. Then

$$\langle v_i|M|v_j\rangle = \lambda_j \langle v_i|v_j\rangle$$

, Similarly

$$\langle v_i|M|v_j\rangle = \langle v_i|M^\dagger|v_j\rangle = \langle v_j|M|v_i\rangle^* = \lambda_i^* \langle v_j|v_i\rangle^* = \lambda_i^* \langle v_i|v_j\rangle = \lambda_i \langle v_j|v_i\rangle$$

Thus,

$$(\lambda_i - \lambda_j)\langle v_i|v_j\rangle = 0$$

If $\lambda_i \neq \lambda_j$, then $\langle v_i|v_j\rangle = 0$

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

Let us assume P is the projector and $|\lambda\rangle$ are the eigenvectors of the projector, with eigen values λ. Since its a projector operator $P^2 = P$

$$P|\lambda\rangle = \lambda|\lambda\rangle \text{ and } P|\lambda\rangle = P^2|\lambda\rangle = \lambda P|\lambda\rangle = \lambda^2|\lambda\rangle$$

Therefore,

$$\lambda = \lambda^2$$

$$\lambda(\lambda - 1) = 0$$

$$\lambda = 0 \text{ or } 1$$

2.24 Hermiticity of positive operators

Suppose A be a positive operator, A can be decomposed as:

$$A = \frac{A + A^\dagger}{2} + \iota\frac{A - A^\dagger}{2i}$$

$$= B + \iota C \textbf{ where } B = \frac{A + A^\dagger}{2}, C = \frac{A - A^\dagger}{2i}$$

Now the operators B and C are hermitian.

$$\langle v|A|v\rangle = \langle v|B = \iota C|v\rangle$$

$$\implies = \langle v|B|v\rangle + \iota\langle v|C|v\rangle$$

$$\implies = \alpha + \iota\beta \text{ where } \alpha = \langle v|B|v\rangle, \beta = \langle v|C|v\rangle$$

Since B and C are hermitian, $\alpha, \beta \in \mathbb{R}$. From definition of positive operator , β should be vanished because $\langle v|A|v\rangle$ is real. Hence $\beta = \langle v|C|\rangle = 0$ for all $|v\rangle$, i.e $C = 0$.

Therefore $A = A^\dagger$.

2.25 For any operator A, $A^\dagger A$ is positive

$$\langle\psi|A^\dagger A|\psi\rangle = ||A|\psi\rangle||^2 \geq 0 \forall|\psi\rangle$$

Thus $A^\dagger A$ is positive.

2.26 Tensor Product

Given $|\psi\rangle = \frac{(|0\rangle+|1\rangle)}{\sqrt{2}}$, We need to find the tensor products, i.e:

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

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

$$= \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

Similarly for $|\psi\rangle^{\otimes 3}$:

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

$$= \frac{1}{2\sqrt{2}}(|000\rangle + |001\rangle + |010\rangle + |011\rangle + |100\rangle + |101\rangle + |110\rangle + |111\rangle)$$

$$= \frac{1}{2\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

2.27 Matrix representation of the tensor products of the Pauli operators

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

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

Similarly,

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

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

and, so for

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

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

In general, the tensor product is not commutable.

2.28 Show that the transpose, complex conjugation, and adjoint operations distribute over the tensor product

$$(A \otimes B)^* = \begin{bmatrix} A_{11}B & ... & A_{1n}B \\ . & . & . \\ . & . & . \\ . & . & . \\ A_{m1}B & ... & A_{mn}B \end{bmatrix}^*$$

$$= \begin{bmatrix} A_{11}^*B^* & ... & A_{1n}^*B^* \\ . & . & . \\ . & . & . \\ . & . & . \\ A_{m1}^*B^* & ... & A_{mn}^*B^* \end{bmatrix}$$

$$= A^* \otimes B^*$$

Now, for $(A \otimes B)^T$

$$(A \otimes B)^T = \begin{bmatrix} A_{11}B & ... & A_{1n}B \\ . & . & . \\ . & . & . \\ . & . & . \\ A_{m1}B & ... & A_{mn}B \end{bmatrix}^T$$

$$= \begin{bmatrix} A_{11}B^T & ... & A_{m1}B^T \\ . & . & . \\ . & . & . \\ . & . & . \\ A_{1n}B^T & ... & A_{mn}B^T \end{bmatrix}$$

$$= \begin{bmatrix} A_{11}B^T & ... & A_{1m}B^T \\ . & . & . \\ . & . & . \\ . & . & . \\ A_{n1}B^T & ... & A_{nm}B^T \end{bmatrix}$$

$$= A^T \otimes B^T$$

and for $(A \otimes B)^\dagger$

$$(A \otimes B)^\dagger = ((A \otimes B)^*)^T$$

$$= (A^* \otimes B^*)^T$$

$$= (A^*)^T \otimes (B^*)^T$$

$$A^\dagger \otimes B^\dagger$$

2.29 Tensor product of two unitary operators is unitary.

Suppose U_1 and U_2 are unitary operators. Then

$$(U_1 \otimes U_2)(U_1 \otimes U_2)^\dagger = U_1U_1^\dagger \otimes U_2U_2^\dagger$$

$$= I \otimes I$$

In a similar way:

$$(U_1 \otimes U_2)^\dagger(U_1 \otimes U_2) = I \otimes I$$

2.30 Tensor product of two Hermitian operators is Hermitian

Suppose A and B are Hermitian operators. Then

$$(A \otimes B)^\dagger = A^\dagger \otimes B^\dagger = A \otimes B$$

Thus we can say that $A \otimes B$ is Hermitian.

2.31 Tensor product of two positive operators is positive.

Suppose A and B are two positive operators. Then

$$\langle\psi| \otimes \langle\psi|(A \otimes B)|\psi\rangle \otimes |\phi\rangle = \langle\psi|A|\psi\rangle\langle\phi|B|\phi\rangle$$

Since A and B are positive operators, we have $\langle\psi|A|\psi\rangle \geq 0$ and $\langle\phi|B|\phi\rangle \geq 0 \forall |\psi\rangle, |\phi\rangle$. Then their product will also be greater than zero i.e $\langle\psi|A|\psi\rangle\langle\phi|B|\phi\rangle \geq 0$. Thus $A \otimes B$ is positive if A and B are positive.

2.32 Tensor product of two projectors is a projector

Suppose P_1 and P_2 are projectors. Then

$$\begin{aligned}(P_1 \otimes P_2)^2 &= P_1^2 \otimes P_2^2 \\ &= P_1 \otimes P_2\end{aligned}$$

Thus $P_1 \otimes P_2$ is also a projector.

2.33 Hadamard operator on one qubit may be written as

$$H = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

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

2.34 Square Root and Log of a Matrix

Given $A = \begin{bmatrix} 4 & 3 \\ 3 & 4 \end{bmatrix}$

We need to find the eigen values and the eigen vectors, i.e

$$\begin{aligned}\det\,(A - \lambda I) &= (4-\lambda)^2 - 3^2 \\ &= \lambda^2 - 8\lambda + 7 \\ &(\lambda - 1)(\lambda - 7)\end{aligned}$$

The eigenvalues for the A matrix are $\lambda = 1, 7$. Corresponding eigenvectors are :

$|\lambda = 1\rangle = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and $|\lambda = 7\rangle = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$
Thus the A matrix can be written as :

$$A = 1|\lambda = 1\rangle\langle\lambda = 1| + 7|\lambda = 7\rangle\langle\lambda = 7|$$

Now, it's easier to apply the operations, hence:

$$\sqrt{A} = \sqrt{1}|\lambda = 1\rangle\langle\lambda = 1| + \sqrt{7}|\lambda = 7\rangle\langle\lambda = 7|$$

$$= \frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} + \frac{\sqrt{7}}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

$$= \frac{1}{2}\begin{bmatrix} 1+\sqrt{7} & -1+\sqrt{7} \\ -1+\sqrt{7} & 1+\sqrt{7} \end{bmatrix}$$

Similarly for logarithm,

$$\log(A) = \log(1)\rangle\langle\lambda = 1| + \log(7)|\lambda = 7\rangle\langle\lambda = 7|$$

$$= \frac{\log(7)}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

2.35 Exponent of Pauli Matrices

We know :

$$\vec{v}\cdot\vec{\sigma} = \sum_{i=1}^{3} v_i\sigma_i$$

where σ_i is the i^{th} Pauli Matrix.

$$= v_1\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + v_2\begin{bmatrix} 0 & -\iota \\ \iota & 0 \end{bmatrix} + v_3\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} v_3 & v_1 - \iota v_2 \\ v_1 + \iota v_2 & -v_3 \end{bmatrix}$$

Now, for the eigen values:

$$\det\,(\vec{v}\cdot\vec{\sigma} - \lambda I) = (v_3 - \lambda)(-v_3\lambda) - (v_1 - \iota v_2)(v_1 + \iota v_2)$$

$$\lambda^2 - (v_1^2 + v_2^2 + v_3^2)$$
$$= \lambda^2 - 1 \text{ (Since } |\vec{v}| = 1)$$

Eigenvalues are $\lambda = \pm 1$. We let the eigenvectors be $|\lambda_{\pm 1}\rangle$ with the eigen values ± 1.

We know that $\vec{v}\cdot\vec{\sigma}$ is Hermitian, and hence diagonalizable. Then

$$\vec{v} \cdot \vec{\sigma} = |\lambda_1\rangle\langle\lambda_1| - |\lambda_{-1}\rangle\langle\lambda_{-1}|$$

Thus:

$$\exp \iota\theta\vec{v} \cdot \vec{\sigma} = \exp \iota\theta|\lambda_1\rangle\langle\lambda_1| + \exp -\iota\theta|\lambda_{-1}\rangle\langle\lambda_{-1}|$$
$$= (\cos\theta + \iota\sin\theta)|\lambda_1\rangle\langle\lambda_1| + (\cos\theta - \iota\sin\theta)|\lambda_{-1}\rangle\langle\lambda_{-1}|$$

$$= \cos\theta|\lambda_1\rangle\langle\lambda_1| + |\lambda_{-1}\rangle\langle\lambda_{-1}| + \iota\sin\theta|\lambda_1\rangle\langle\lambda_1| - |\lambda_{-1}\rangle\langle\lambda_{-1}|$$

$$= \cos\theta I + \iota\sin\theta\vec{v} \cdot \vec{\sigma}$$

Since $\vec{v} \cdot \vec{\sigma}$ is Hermitian, $|\lambda_1\rangle$ and $|\lambda_{-1}\rangle$ are orthogonal. Thus

$$|\lambda_1\rangle\langle\lambda_1| + |\lambda_{-1}\rangle\langle\lambda_{-1}| = I$$

2.36 Pauli Matrices except I has a trace 0

For σ_1:

$$\text{Tr}(\sigma_1) = \text{Tr}\left(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\right) = 0$$

For σ_2

$$\text{Tr}(\sigma_2) = \text{Tr}\left(\begin{bmatrix} 0 & -\iota \\ \iota & 0 \end{bmatrix}\right) = 0$$

For σ_3

$$\text{Tr}(\sigma_3) = \text{Tr}\left(\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\right) = 0$$

2.37 Cyclic Properties of Trace

$$\text{Tr}(AB) = \sum_i \langle i|AB|i\rangle$$
$$= \sum_i \langle i|AIB|i\rangle$$
$$= \sum_i \langle i|A|j\rangle\langle j|B|i\rangle$$
$$= \sum_i \langle j|B|i\rangle\langle i|A|j\rangle$$
$$= \sum_j \langle j|BA|j\rangle$$
$$\text{Tr}(BA)$$

2.38 Linearity of Trace

$$\begin{aligned}\mathrm{Tr}(A+B) &= \sum_i \langle i|A+B|i\rangle \\ &= \sum_i \langle i|A|i\rangle + \langle i|B|i\rangle \\ &= \sum_i \langle i|A|i\rangle + \sum_i \langle i|B|i\rangle \\ &= \mathrm{Tr}(A) + \mathrm{Tr}(B)\end{aligned}$$

$$\begin{aligned}\mathrm{Tr}(zA) &= \sum_i \langle i|zA|i\rangle \\ &\sum_i z\langle i|A|i\rangle \\ &= z\sum_i \langle i|A|i\rangle \\ &= z\mathrm{Tr}(A)\end{aligned}$$

2.39 Hilbert-Schmidt Inner Product on Operators

We have to show

$$(A, B) = \mathrm{Tr}(A^\dagger B)$$

$$\begin{aligned}\left(A, \sum_i \lambda_i B_i\right) =\ &\mathrm{Tr}\ \left[A^\dagger \left(\sum_i {}_i B_i\right)\right] \\ &= \mathrm{Tr}(A^\dagger \lambda_1 B_1) + ... + \mathrm{Tr}(A^\dagger \lambda_n B_n) \\ &\lambda_1 \mathrm{Tr}(A^\dagger B_1) + .. + \lambda_n \mathrm{Tr}(A^\dagger B_n) \\ &\sum_i \lambda_i \mathrm{Tr}(A^\dagger B_i)\end{aligned}$$

and for

$$\begin{aligned}(A, B)^* &= \Big(\mathrm{Tr}(A^\dagger B)\Big)^* \\ &= \left(\sum_{i,j} \langle i|A^\dagger|j\rangle\langle j|B|i\rangle\right)^* \\ &= \sum_{i,j} \langle i|A^\dagger|j\rangle^*\langle j|B|i\rangle^*\end{aligned}$$

$$
\begin{aligned}
&= \sum_{i,j} \langle j|B|i\rangle^* \langle i|A^\dagger|j\rangle^* \\
&= \sum_{i,j} \langle i|B^\dagger|j\rangle \langle j|A|i\rangle \\
&= \sum_{i,j} \langle i|B^\dagger A|i\rangle \\
&\mathrm{Tr}(B^\dagger A) \\
&= (B, A)
\end{aligned}
$$

and,

$$
\begin{aligned}
(A, A) &= \mathrm{Tr}(A^\dagger A) \\
&= \sum_i \langle i|A^\dagger A|i\rangle
\end{aligned}
$$

Since $A^\dagger A$ is positive, $\langle i|A^\dagger A|i\rangle \geq 0 \ \forall \ |i\rangle$ Let a_i be the i^{th} column of A. If $\langle i|A^\dagger A|i\rangle = 0$, then

$$\langle i|A^\dagger A|i\rangle = a_i^\dagger a_i = ||a_i||^2 = 0 \text{ iff } a_i = 0$$

Therefor $(A, A) = 0$ iff $A = 0$

ii). A linear transformation $T ::; \ V \rightarrow V$ where $\dim(V) = d$ can be represented as a $d \times d$ matrix. Since there are $d \times d = d^2$ matrices that are linearly independent, the dimension of L_V is d^2.

2.40 Commutation relation for the Pauli Matrices

We have

$$[X, Y] = XY - YX$$

$$
\begin{aligned}
[X, Y] &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & -\iota \\ \iota & 0 \end{bmatrix} - \begin{bmatrix} 0 & -\iota \\ \iota & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\
&= \begin{bmatrix} \iota & 0 \\ 0 & -\iota \end{bmatrix} - \begin{bmatrix} -\iota & 0 \\ 0 & \iota \end{bmatrix} \\
&= \begin{bmatrix} 2\iota & 0 \\ 0 & -2\iota \end{bmatrix} \\
&= 2\iota Z
\end{aligned}
$$

$$
\begin{aligned}
[Y, Z] &= \begin{bmatrix} 0 & -\iota \\ \iota & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 0 & -\iota \\ \iota & 0 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 2\iota \\ 2\iota & 0 \end{bmatrix} \\
&= 2\iota X
\end{aligned}
$$

Similarly,

$$[Z,X] = \begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}\begin{bmatrix}0 & 1\\1 & 0\end{bmatrix} - \begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}\begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}$$

$$= 2\iota\begin{bmatrix}0 & -\iota\\\iota & 0\end{bmatrix}$$

$$= 2\iota Y$$

2.41 Anti-Commutation relation for the Pauli Matrices

We have

$$\{\sigma_1, \sigma_2\} = \sigma_1\sigma_2 + \sigma_2\sigma_1$$

$$= \begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}\begin{bmatrix}0 & -\iota\\\iota & 0\end{bmatrix} + \begin{bmatrix}0 & -\iota\\\iota & 0\end{bmatrix}\begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}$$

$$= \begin{bmatrix}\iota & 0\\0 & -\iota\end{bmatrix} + \begin{bmatrix}-\iota & 0\\0 & \iota\end{bmatrix}$$

$$= 0$$

Similarly,

$$\{\sigma_2, \sigma_3\} == \begin{bmatrix}0 & -\iota\\\iota & 0\end{bmatrix}\begin{bmatrix}1 & 0\\0 & -1\end{bmatrix} - \begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}\begin{bmatrix}0 & -\iota\\\iota & 0\end{bmatrix}$$

$$= 0$$

and,

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

$$= 0$$

and, we have

$$\sigma_0^2 = I^2 = I$$

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

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

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

2.42 verify

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

2.43 Show $\sigma_j\sigma_k$

We know that $\{\sigma_j, \sigma_k\} = 2\delta_{jk}I$ using that

$$\sigma_j\sigma_k = \frac{[\sigma_j,\sigma_k]+\{\sigma_j,\sigma_k\}}{2}$$

$$= \frac{2\iota\sum_{i=1}^{3}\epsilon_{jkl}\sigma_l + 2\delta_{jk}I}{2}$$

$$\delta_{jk}I + \iota\sum_{l=1}^{3}\epsilon_{jkl}\sigma_l$$

2.44 B must be zero

By Assumption, $[A, B] = 0$ and $\{A, B\} = 0$, then $AB = 0$. Since A is invertible, multiply by A^{-1} from left,then

$$A^{-1}AB = 0$$

$$IB = 0$$

$$B = 0$$

2.45 Verify

$$[A,B]^\dagger = (AB-BA)^\dagger$$

$$= B^\dagger A^\dagger - A^\dagger B^\dagger$$

$$= [B^\dagger, A^\dagger]$$

2.46 Verify

$$[A,B] = AB - BA$$

$$= -(BA - AB)$$

$$= -[B,A]$$

2.47 Verify

$$\begin{aligned}(\iota[A,B])^\dagger &= \iota[A,B]^\dagger\\ &= \iota[B^\dagger, A^\dagger]\\ &-\iota[B,A]\\ &= \iota[A,B]\end{aligned}$$

2.48 Polar decomposition of P, U and H

Since P is positive, it is diagonalizable. Then $P = \sum_i \lambda_i |i\rangle\langle i|$, where $\lambda_i \geq 0$

$$J = \sqrt{P^\dagger P} = \sqrt{PP} = \sqrt{P^2} = \sum_i \sqrt{\lambda_i^2}|i\rangle\langle i| = \sum_i \lambda_i |i\rangle\langle i| = P$$

Therefore polar decomposition of P is $P = UP$ for all P. Thus $U == I$, then $P = P$

For the unitary U:

Suppose unitary U is decomposed by $U = WJ$ where W is unitary and J is positive, $J = \sqrt{U^\dagger U}$

$$J = \sqrt{U^\dagger U} = \sqrt{I} = I$$

Since unitary operators are invertible, $W = UJ^{-1} = UI^{-1} = UI = U$. Thus the polar decompositions of U is $U = U$

For the Hermitian H:

Suppose $H = UJ$

$$J = \sqrt{H^\dagger H} = \sqrt{HH} = \sqrt{H^2}$$

Thus $H = U\sqrt{H^2}$

In general $H \neq \sqrt{H^2}$ From spectral decomposition, $H = \sum_i \lambda_i |i\rangle\langle i|$, where $\lambda_i \in \mathbb{R}$

$$\sqrt{H^2} = \sqrt{\sum_i \lambda_i^2 |i\rangle\langle i|} = \sum_i \sqrt{\lambda_i^2}|i\rangle\langle i| = \sum_i |\lambda_i||i\rangle\langle i| \neq H$$

2.49 Polar decomposition of Normal Matrix

Normal matrix is diagonalizable, $A = \sum_i \lambda_i |i\rangle\langle i|$

$$J = \sqrt{A^\dagger A} = \sum_i |\lambda_i||i\rangle\langle i|$$

$$U = \sum_i |e_i\rangle\langle i|$$

$$A = UJ = \sum_i |\lambda_i||e_i\rangle\langle i|$$

2.50 Left and Right Polar Decomposition

We have $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ and so we have $A^\dagger A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$

Characteristic equation of $A^\dagger A$ is $det(A^\dagger A - \lambda I)) = \lambda^2 - 3\lambda + 1 = 0$. Eigenvalues of $A^\dagger A$ are $\lambda_\pm = \frac{3 \pm \sqrt{5}}{2}$ and the associated eigenvectors are $|\lambda_\pm\rangle = \frac{1}{\sqrt{10 \mp 2\sqrt{5}}} \begin{bmatrix} 2 \\ -1 \pm \sqrt{5} \end{bmatrix}$

$$A^\dagger A = \lambda_+ |\lambda_+\rangle\langle\lambda_+| + \lambda_- |\lambda_-\rangle\langle\lambda_-\rangle\langle\lambda_-|$$

$$J = \sqrt{A^\dagger A} = \sqrt{\lambda_+}|\lambda_+\rangle\langle\lambda_+| + \sqrt{\lambda_-}|\lambda_-\rangle\langle\lambda_-\rangle\langle\lambda_-|$$

Substituting all the values, you'll get:

$$J^{-1} = \frac{1}{\sqrt{\lambda_+}}|\lambda_+\rangle\langle\lambda_+| + \frac{1}{\sqrt{\lambda_-}}|\lambda_-\rangle\langle\lambda_-|$$

Which gives us $U = AJ^{-1}$

2.51 H is unitary

$$\begin{aligned} H^\dagger H &= \left(\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \right)^\dagger \left(\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \right) \\ &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \\ &= \frac{1}{2} \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = I \end{aligned}$$

2.52 H^2 is identity

$$\begin{aligned} H^\dagger &= \left(\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \right)^\dagger \\ &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = H \end{aligned}$$

Thus

$$H^2 = I$$

2.53 Eigen values and eigen vectors of H

We have $det(H - \lambda I) = \left(\frac{1}{\sqrt{2}} - \lambda\right)\left(\frac{-1}{\sqrt{2}} - \lambda\right) - \frac{1}{2}$

$$= \lambda^2 - \frac{1}{2} - \frac{1}{2}$$

$$= \lambda^2 - 1$$

Eigenvalues are $\lambda_\pm = \pm 1$ and associated eigenvectors are $|\lambda_\pm = \frac{1}{\sqrt{4 \mp 2\sqrt{2}}} \begin{bmatrix} 1 \\ -1 \pm \sqrt{2} \end{bmatrix}$

2.54 $\exp A \exp B = \exp A + B$
Since $[A, B] = 0$, A and B are simultaneously diagonlizable , $A = \sum_i a_i |i\rangle\langle i|$, and $B = \sum_i b_i |i\rangle\langle i|$.

So,

$$\exp A \exp B = \left(\sum_i \exp a_i |i\rangle\langle i|\right)\Bigg(\exp b_I |i\rangle\langle i|$$

$$= \sum_{i,j} \exp a_i + b_j |i\rangle\langle i|j\rangle\langle j|$$

$$= \sum_{i,j} \exp a_i + b_j |i\rangle\langle j|\delta_{i,j}$$

$$= \sum_i \exp a_i + b_i |i\rangle\langle i|$$

$$\exp A + B$$

2.55 $U(t_1, t_2)$ is unitary

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

$$U(t_2 - t_1)U^\dagger(t_2 - t_1) = \exp \frac{-\iota H(t_2 - t_1)}{\hbar} \exp \frac{\iota H(t_2 - t_1)}{\hbar}$$

$$= \sum_{E,E'} \left(\exp \frac{-\iota E(t_2 - t_1)}{\hbar} |E\rangle\langle E|\right)\left(\exp \frac{-\iota E'(t_2 - t_1)}{\hbar} |E'\rangle\langle E'|\right)$$

$$= \sum_E \exp 0 |E\rangle\langle E|$$

$$\sum_E |E\rangle\langle E|$$

$$= I$$

Similarly , $U^\dagger(t_2 - t_1)U(t_2 - t_1) = I$

2.56 Spectral Decomposition

$$U = \sum_i \lambda_i |\lambda_i\rangle\langle\lambda_i|$$

Now,

$$\log(U) = \sum_j \log(\lambda_j)|\lambda_j\rangle\langle\lambda_j| = \sum_j \iota\theta_j|\lambda_j\rangle\langle\lambda_j|$$

where $\theta_j = \text{ arg } (\lambda_j)$

$$K = \iota\log(U) = \sum_j \theta_j|\lambda_j\rangle\langle\lambda_j|$$

$$K^\dagger = (\iota\log U)^\dagger = \left(\sum_j \theta_j|\lambda_j\rangle\langle\lambda_j|\right)^\dagger = \sum_j \theta_j|\lambda_j\rangle\langle\lambda_j| = K$$

2.57 Cascaded Measurements are single measurements

We have

$$|\phi\rangle = \frac{L_l|\psi\rangle}{\sqrt{\langle\psi|L_l^\dagger L_l|\psi\rangle}}$$

$$\langle\phi|M_m^\dagger M_m|\phi\rangle = \frac{\langle\psi|L_l^\dagger M_m^\dagger M_m L_l|\psi\rangle}{\langle\psi|L_l^\dagger L_l|\psi\rangle}$$

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

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

$$= \frac{N_{lm}|\psi\rangle}{\sqrt{\langle\psi|N_{lm}^\dagger N_{lm}|\psi\rangle}}$$

2.58 Average Observed value of M

$$\langle M\rangle = \langle\psi|M|\psi\rangle = \langle\psi|m|\psi\rangle = m\langle\psi|\psi\rangle = m$$

$$\langle M^2\rangle = \langle\psi|M^2|\psi\rangle = \langle\psi|m^2|\psi\rangle = m^2\langle\psi|\psi\rangle = m^2$$

$$\text{deviation } = \langle M^2\rangle - \langle M\rangle^2 = m^2 - m^2 = 0$$

2.59 Average and Standard Deviation of X

$$\langle X \rangle = \langle 0|X|0 \rangle = \langle 0|1 \rangle = 0$$

$$\langle X^2 \rangle = \langle 0|X^2|0 \rangle = \langle 0|X|1 \rangle = \langle 0|0 \rangle = 1$$
$$\text{standard deviation } = \sqrt{\langle X^2 \rangle - \langle X \rangle^2} = 1$$

2.60

$$\begin{aligned}
\vec{v} \cdot \vec{\sigma} &= \sum_{i=1}^{3} v_i \sigma_i \\
&= v_1 \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + v_2 \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} + v_3 \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \\
&= \begin{bmatrix} v_3 & v_1 - iv_2 \\ v_1 + iv_2 & -v_3 \end{bmatrix}
\end{aligned}$$

$$\begin{aligned}
\det(\vec{v} \cdot \vec{\sigma} - \lambda I) &= (v_3 - \lambda)(-v_3 - \lambda) - (v_1 - iv_2)(v_1 + iv_2) \\
&= \lambda^2 - (v_1^2 + v_2^2 + v_3^2) \\
&= \lambda^2 - 1 \quad (\because |\vec{v}| = 1)
\end{aligned}$$

Eigenvalues are $\lambda = \pm 1$.

(i) if $\lambda = 1$

$$\begin{aligned}
\vec{v} \cdot \vec{\sigma} - \lambda I &= \vec{v} \cdot \vec{\sigma} - I \\
&= \begin{bmatrix} v_3 - 1 & v_1 - iv_2 \\ v_1 + iv_2 & -v_3 - 1 \end{bmatrix}
\end{aligned}$$

Normalized eigenvector is $|\lambda_1\rangle = \sqrt{\frac{1+v_3}{2}} \begin{bmatrix} 1 \\ \frac{1-v_3}{v_1 - iv_2} \end{bmatrix}$.

$$\lambda_1 = \frac{1+v_3}{2} \begin{bmatrix} 1 \\ \frac{1-v_3}{v_1-iv_2} \end{bmatrix} \begin{bmatrix} 1 & \frac{1-v_3}{v_1+iv_2} \end{bmatrix} = \frac{1+v_3}{2} \begin{bmatrix} 1 & \frac{v_1-iv_2}{1+v_3} \\ \frac{v_1+iv_2}{1+v_3} & \frac{1-v_3}{1+v_3} \end{bmatrix}$$

$$= \frac{1}{2} \begin{bmatrix} 1+v_3 & v_1 - iv_2 \\ v_1 + iv_2 & 1 - v_3 \end{bmatrix} = \frac{1}{2} \left(I + \begin{bmatrix} v_3 & v_1 - iv_2 \\ v_1 + iv_2 & -v_3 \end{bmatrix} \right) = \frac{1}{2}(I + \vec{v} \cdot \vec{\sigma})$$

(ii) If $\lambda = -1$.

$$\begin{aligned}
\vec{v} \cdot \vec{\sigma} - \lambda I &= \vec{v} \cdot \vec{\sigma} + I \\
&= \begin{bmatrix} v_3 + 1 & v_1 - iv_2 \\ v_1 + iv_2 & -v_3 + 1 \end{bmatrix}
\end{aligned}$$

Normalized eigenvalue is $|\lambda_{-1}\rangle = \sqrt{\frac{1-v_3}{2}} \begin{bmatrix} 1 \\ -\frac{1+v_3}{v_1-iv_2} \end{bmatrix}$.

$$|\lambda_{-1}\rangle = \frac{1-v_3}{2} \begin{bmatrix} 1 \\ -\frac{1+v_3}{v_1-iv_2} \end{bmatrix} \begin{bmatrix} 1 & -\frac{1+v_3}{v_1+iv_2} \end{bmatrix} = \frac{1-v_3}{2} \begin{bmatrix} 1 & -\frac{v_1-iv_2}{1-v_3} \\ -\frac{v_1+iv_2}{1-v_3} & \frac{1+v_3}{1-v_3} \end{bmatrix}$$

$$= \frac{1}{2} \begin{bmatrix} 1-v_3 & -(v_1-iv_2) \\ -(v_1+iv_2) & 1+v_3 \end{bmatrix} = \frac{1}{2} \left(I - \begin{bmatrix} v_3 & v_1 - iv_2 \\ (v_1+iv_2 & -v_3 \end{bmatrix} \right) = \frac{1}{2}(I - \vec{v} \cdot \vec{\sigma}).$$

This proof has a defect. The case $(v_1, v_2, v_3) = (0, 0, 1)$, second component of eigenstate, $\frac{1-v_3}{v_1-iv_2}$, diverges. So I implicitly assume $v_1 - iv_2 \neq 0$. Hence my proof is incomplete.

Since the exercise doesn't require explicit form of projector, we should prove the problem more abstractly. In order to prove, we use the following properties of $\vec{v} \cdot \vec{\sigma}$

- $\vec{v} \cdot \vec{\sigma}$ is Hermitian
- $(\vec{v} \cdot \vec{\sigma})^2 = I$ where $\vec{v}$ is a real unit vector.

We can easily check above conditions.

$$\begin{aligned} (\vec{v} \cdot \vec{\sigma})^\dagger &= (v_1\sigma_1 + v_2\sigma_2 + v_3\sigma_3)^\dagger \\ &= v_1\sigma_1^\dagger + v_2\sigma_2^\dagger + v_3\sigma_3^\dagger \\ &= v_1\sigma_1 + v_2\sigma_2 + v_3\sigma_3 \quad (\because \text{Pauli matrices are Hermitian.}) \\ &= \vec{v} \cdot \vec{\sigma} \end{aligned}$$

$$\begin{aligned} (\vec{v} \cdot \vec{\sigma})^2 &= \sum_{j,k=1}^{3} (v_j\sigma_j)(v_k\sigma_k) \\ &= \sum_{j,k=1}^{3} v_j v_k \sigma_j \sigma_k \\ &= \sum_{j,k=1}^{3} v_j v_k \left(\delta_{jk} I + i \sum_{l=1}^{3} \epsilon_{jkl} \sigma_l \right) \quad (\because \text{eqn(2.78) page78}) \\ &= \sum_{j,k=1}^{3} v_j v_k \delta_{jk} I + i \sum_{j,k,l=1}^{3} \epsilon_{jkl} v_j v_k \sigma_l \\ &= \sum_{j=1}^{3} v_j^2 I \\ &= I \quad \left(\because \sum_j v_j^2 = 1 \right) \end{aligned}$$

Suppose $|\lambda\rangle$ is an eigenstate of $\vec{v}\cdot\vec{\sigma}$ with eigenvalue λ. Then

$$\begin{aligned}\vec{v}\cdot\vec{\sigma}\,|\lambda\rangle &= \lambda\,|\lambda\rangle\\ (\vec{v}\cdot\vec{\sigma})^2\,|\lambda\rangle &= \lambda^2\,|\lambda\rangle\end{aligned}$$

On the other hand $(\vec{v}\cdot\vec{\sigma})^2 = I$,

$$\begin{aligned}(\vec{v}\cdot\vec{\sigma})^2\,|\lambda\rangle = I\,|\lambda\rangle &= |\lambda\rangle\\ \therefore \lambda^2\,|\lambda\rangle &= |\lambda\rangle\,.\end{aligned}$$

Thus $\lambda^2 = 1 \Rightarrow \lambda = \pm 1$. Therefore $\vec{v}\cdot\vec{\sigma}$ has eigenvalues ± 1.

Let $|\lambda_1\rangle$ and $|\lambda_{-1}\rangle$ are eigenvectors with eigenvalues 1 and -1, respectively. I will prove that $P_\pm = |\lambda_{\pm 1}\rangle$.

In order to prove above equation, all we have to do is prove following condition.

$$\langle\psi|(P_\pm - \lambda_{\pm 1})|\psi\rangle = 0 \text{ for all } |\psi\rangle \in \mathbb{C}^2$$

Since $\vec{v}\cdot\vec{\sigma}$ is Hermitian, $|\lambda_1\rangle$ and $|\lambda_{-1}\rangle$ are orthonormal vector ($\because$ Exercise 2.22). Let $|\psi\rangle \in \mathbb{C}^2$ be an arbitrary state. $|\psi\rangle$ can be written as

$$|\psi\rangle = \alpha\,|\lambda_1\rangle + \beta\,|\lambda_{\pm 1}\rangle \quad (|\alpha|^2 + |\beta|^2 = 1, \alpha, \beta \in \mathbb{C}).$$

$$\begin{aligned}\langle\psi|(P_\pm - \lambda_\pm)|\psi\rangle &= \langle\psi|P_\pm|\psi\rangle - \langle\psi|\lambda_\pm\rangle\,\langle\lambda_\pm|\psi\rangle\,.\\ \langle\psi|P_\pm|\psi\rangle &= \langle\psi|\frac{1}{2}(I \pm \vec{v}\cdot\vec{\sigma})|\psi\rangle\\ &= \frac{1}{2} \pm \frac{1}{2}\,\langle\psi|\vec{v}\cdot\vec{\sigma})|\psi\rangle\\ &= \frac{1}{2} \pm \frac{1}{2}(|\alpha|^2 - |\beta|^2)\\ &= \frac{1}{2} \pm \frac{1}{2}(2|\alpha|^2 - 1) \quad (\because |\alpha|^2 + |\beta|^2 = 1)\\ \langle\psi|\lambda_1\rangle\,\langle\lambda_1|\psi\rangle &= |\alpha|^2\\ \langle\psi|\lambda_{-1}\rangle\,\langle\lambda_{-1}|\psi\rangle &= |\beta|^2 = 1 - |\alpha|^2\end{aligned}$$

Therefore $\langle\psi|(P_\pm - \lambda_{\pm 1})|\psi\rangle = 0$ for all $|\psi\rangle \in \mathbb{C}^2$. Thus $P_\pm = \lambda_{\pm 1}$.

2.61 Calculate the Probability

$$\begin{gathered}\langle\lambda_1|0\rangle\langle 0|\lambda_1\rangle - \langle 0|\lambda_1\rangle\langle\lambda_1|0\rangle\\ = \langle 0|\frac{1}{2}(I + \vec{v}\cdot\vec{\sigma}|0\rangle\\ \frac{1}{2}(1 + v_3)\end{gathered}$$

and the state after measurement is

$$\begin{aligned}\frac{|\lambda_1\rangle\langle\lambda_1|0\rangle}{\sqrt{\langle 0|\lambda_1\rangle\langle\lambda_1|0\rangle}} &= \frac{1}{\sqrt{\frac{1}{2}(1+v_3)}}\cdot\frac{1}{2}\begin{bmatrix}1+v_3\\ v_1+\iota v_2\end{bmatrix}\\ &= \sqrt{\frac{1}{2}(1+v_3)}\cdot\begin{bmatrix}1\\ \frac{v_1+\iota v_2}{1+v_3}\end{bmatrix}\\ &= \sqrt{\frac{1+v_3}{2}}\begin{bmatrix}1\\ \frac{1-v_3}{v_1-\iota v_2}\end{bmatrix}\\ &= |\lambda_1\rangle\end{aligned}$$

2.62 Projective Measurement

Let's assume M_m is a measurement operator. From the assumption , we have $E_m = M_m^\dagger M_m = M_m$, Then

$$\langle\psi|E_m|\psi\rangle = \langle\psi|M_m|\psi\rangle \geq 0 \;\forall\; |\psi\rangle$$

Since M_m is a positive operator, M_m is Hermitian. Therfore,

$$E_m = M_m^\dagger M_m = M_m M_m = M_m^2 = M_m$$

Thus the measurement is a projective measurement.

2.63 Measurement Operators M_m

We have

$$\begin{aligned}M_m^\dagger M_m &= \sqrt{E_m}U_m^\dagger U_m\sqrt{E_m}\\ &= \sqrt{E_m}I\sqrt{E_m}\\ &= E_m\end{aligned}$$

Since E_m is POVM , for arbitrary unitary U, $M_m^\dagger M_m$ is a POVM

2.64 Construct a POVM

We can construct $|\psi_j^{'}\rangle$ that is orthogonal to all states except $|\psi_j\rangle$. That is

$$|\psi_j^{'}\rangle = |\psi_j\rangle - \sum_{k=1,k\neq j}^{m}\frac{\langle\psi_j|\psi_k\rangle|\psi_k\rangle}{||\psi_k\rangle|^2}$$

Then E_m is

$$E_m = A|\psi_m^{'}\rangle\langle\psi_m^{'}|$$

where A is chosen such that

$$E_{m+1} = I - \sum_{j=1}^{m} E_j$$

is positive.

2.65 Express the states in Composite system

The $|+\rangle$ and $|-\rangle$ states

$$|+\rangle = \frac{|0\rangle + |1\rangle}{2}$$

$$|-\rangle = \frac{|0\rangle - |1\rangle}{2}$$

2.66 Average Value

Let $|\psi\rangle = \frac{|00\rangle+|11\rangle}{\sqrt{2}}$. Then,

$$E(X_1Z_2) = \langle\psi|X_1Z_2|\psi\rangle$$

$$= \frac{1}{2}(\langle 00|X_1Z_2|00\rangle) + (\langle 00|X_1Z_2|11\rangle) + (\langle 11|X_1Z_2|00\rangle) + (\langle 11|X_1Z_2|11\rangle)$$

$$\frac{1}{2}(\langle 00|10\rangle - \langle 00|01\rangle + \langle 11|10\rangle - \langle 11|01\rangle$$

$$= 0$$

2.67 $V \approx$ Hilbert Space

Suppose $W^\perp$ is the orthogonal complement of W. Then $V = W \oplus W^\perp$. Let $|w_i\rangle, |w'_j\rangle, |u'_j\rangle$ be orthonormal bases for W, $W^\perp$, $(\text{image}(U))^\perp$, respectively.

Define $U' : V \to V$ as $U' = \sum_i u_i w_i + \sum_j u'_j w'_j$, where $|u_i\rangle = U|w_i\rangle$.

Now

$$\begin{aligned}(U')^\dagger U' &= \left(\sum_{i=1}^{\dim W} w_i u_i + \sum_{j=1}^{\dim W^\perp} w'_j u'_j\right)\left(\sum_i u_i w_i + \sum_j u'_j w'_j\right) \\ &= \sum_i w_i + \sum_j w'_j = I\end{aligned}$$

and

$$\begin{aligned}U'(U')^\dagger &= \left(\sum_i u_i w_i + \sum_j u'_j w'_j\right)\left(\sum_i w_i u_i + \sum_j w'_j u'_j\right) \\ &= \sum_i u_i + \sum_j u'_j = I.\end{aligned}$$

Thus U' is an unitary operator. Moreover, for all $|w\rangle \in W$,

$$\begin{aligned}
U' \left| w \right\rangle &= \left(\sum_i u_i w_i + \sum_j u'_j w'_j \right) \left| w \right\rangle \\
&= \sum_i \left| u_i \right\rangle \left\langle w_i | w \right\rangle + \sum_j \left| u'_j \right\rangle \left\langle w'_j | w \right\rangle \\
&= \sum_i \left| u_i \right\rangle \left\langle w_i | w \right\rangle \quad (\because \left| w'_j \right\rangle \perp \left| w \right\rangle) \\
&= \sum_i U \left| w_i \right\rangle \left\langle w_i | w \right\rangle \\
&= U \left| w \right\rangle .
\end{aligned}$$

Therefore U' is an extension of U.

2.68 $|\psi\rangle \neq |a\rangle|b\rangle$

Suppose one of the bell states:

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

can be written as a product of single particle states

$$|a\rangle = \alpha_a|0\rangle + \beta_a|1\rangle$$

$$|b\rangle = \alpha_b|0\rangle + \beta_b|1\rangle$$

Then

$$|a\rangle|b\rangle = \alpha_a\alpha_b|00\rangle + \alpha_a\beta_b|01\rangle + \beta_a\alpha_b|10\rangle + \beta_a\beta_b|11\rangle$$

Comparing this with the bell state $|\psi\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$ we have

$$\alpha_a\beta_b = 0$$

$$\beta_a\alpha_b = 0$$

This contradicts $|\psi\rangle = |a\rangle|b\rangle$. So $|\psi\rangle \neq |a\rangle|b\rangle$ for all single qubits states $|a\rangle$ and $|b\rangle$.

2.69 Bell Basis

We know the bell states are:

$$|\psi_1\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

$$|\psi_2\rangle = \frac{|00\rangle - |11\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\0\\0\\-1\end{bmatrix}$$

$$|\psi_3\rangle = \frac{|01\rangle + |10\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}}\begin{bmatrix}0\\1\\1\\0\end{bmatrix}$$

$$|\psi_4\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}}\begin{bmatrix}0\\1\\-1\\0\end{bmatrix}$$

To form a basis, it must be linearly independent, Hence:

$$a_1\,|\psi_1\rangle + a_2\,|\psi_2\rangle + a_3\,|\psi_3\rangle + a_4\,|\psi_4\rangle = 0$$

$$\implies \frac{1}{\sqrt{2}}\begin{bmatrix}a_1+a_2\\a_3+a_4\\a_3-a_4\\a_1-a_2\end{bmatrix} = 0$$

$$\implies a_1 = a_2 = a_3 = a_4 = 0$$

Thus $\{|\psi_i\rangle\}$ is a linearly independent basis.

Moreover $||\,|\psi_i\rangle\,|| = 1$ and $\langle\psi_i|\psi_j\rangle = \delta_{ij}$ for $i, j = 1, 2, 3, 4$. Thus it forms an orthonormal basis.

2.70 Bell States

For any Bell state we have $\langle\psi_i|E\otimes I|\psi_i\rangle = \frac{1}{2}(\langle 0|E|0\rangle + \langle 1|E|1\rangle)$

Suppose Eve measures the qubit Alice sent by measurement operators M_m. The probability that Eve get result m is $p_i(m) = \langle\psi_i|M_m^\dagger M_m \otimes I|\psi_i\rangle$ Since $M_m^\dagger M_m$ is positive , $p_i(m)$ are same values for all $|\psi_i\rangle$. Thus Eve can't distinguish Bell states.

2.71 Mixed or Pure

We know from spectral decomposition

$$\rho = \sum_i p_i|\psi_i\rangle\langle\psi_i|, \quad p_i \geq 0, \quad \sum_i p_i = 1$$

$$\begin{aligned}\rho^2 &= \sum_{i,j} p_i p_j |i\rangle\langle i||j\rangle\langle j| \\ &= \sum_{i,j} p_i p_j |i\rangle\langle i|\delta_{ij}\end{aligned}$$

$$= \sum_i p_i^2 |i\rangle\langle i|$$

$$Tr(\rho^2) = Tr(\sum_i p_i^2 |i\rangle\langle i|) = \sum_i p_i^2 Tr(|i\rangle\langle i|) = \sum_i p_i^2 |i\rangle\langle i| = \sum_i p_i^2 \geq \sum_i p_i = 1$$

Suppose $Tr(\rho^2) = 1$. Then $\sum_i p_i^2 = 1$. Since $p_i^2 < p_i$ for $0 < p_i < 1$, only single p_i should be 1 and otherwise have to vanish. Therefore $\rho = |\psi_i\rangle\langle\psi_i|$. It is a pure state.

Conversely if ρ is pure, then $\rho = |\psi_i\rangle\langle\psi_i|$

$$Tr(\rho^2) = Tr(|\psi_i\rangle\langle\psi_i||\psi_i\rangle\langle\psi_i|) = Tr(|\psi_i\rangle\langle\psi_i|) = |\psi_i\rangle\langle\psi_i| = 1$$

2.72 Bloch Sphere

(1) Since density matrix is Hermitian, matrix representation is $\rho = \begin{bmatrix} a & b \\ b^* & d \end{bmatrix}$, $a, d \in \mathbb{R}$ and $b \in \mathbb{C}$ w.r.t. standard basis. Because ρ is density matrix, $(\rho) = a + d = 1$.

Define $a = (1 + r_3)/2$, $d = (1 - r_3)/2$ and $b = (r_1 - \iota r_2)/2$, $(r_i \in \mathbb{R})$.

In this case,

$$\rho = \begin{bmatrix} a & b \\ b^* & d \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 + r_3 & r_1 - ir_2 \\ r_1 + ir_2 & 1 - r_3 \end{bmatrix} = \frac{1}{2}(I + \vec{r} \cdot \vec{\sigma}).$$

Thus for arbitrary density matrix ρ can be written as $\rho = \frac{1}{2}(I + \vec{r} \cdot \vec{\sigma})$.

Next, we derive the condition that ρ is positive.

If ρ is positive, all eigenvalues of ρ should be non-negative.

$$\det(\rho - \lambda I) = (a - \lambda)(b - \lambda) - |b|^2 = \lambda^2 - (a + d)\lambda + ad - |b^2| = 0$$

$$\begin{aligned} \lambda &= \frac{(a+d) \pm \sqrt{(a+d)^2 - 4(ad - |b|^2)}}{2} \\ &= \frac{1 \pm \sqrt{1 - 4\left(\frac{1 - r_3^2}{4} - \frac{r_1^2 + r_2^2}{4}\right)}}{2} \\ &= \frac{1 \pm \sqrt{1 - (1 - r_1^2 - r_2^2 - r_3^2)}}{2} = \frac{1 \pm \sqrt{|\vec{r}|^2}}{2} = \frac{1 \pm |\vec{r}|}{2} \end{aligned}$$

Since ρ is positive, $\frac{1 - |\vec{r}|}{2} \geq 0 \rightarrow |\vec{r}| \leq 1$.

Therefore an arbitrary density matrix for a mixed state qubit is written as $\rho = \frac{1}{2}(I + \vec{r} \cdot \vec{\sigma})$.

(2)

$\rho = I/2 \rightarrow \vec{r} = 0$. Thus $\rho = I/2$ corresponds to the origin of Bloch sphere.

(3)

$$\rho^2 = \frac{1}{2}(I + \vec{r}\cdot\vec{\sigma})\ \frac{1}{2}(I + \vec{r}\cdot\vec{\sigma})$$

$$= \frac{1}{4}\left[I + 2\vec{r}\cdot\vec{\sigma} + \sum_{j,k} r_j r_k \left(\delta_{jk} I + i \sum_{l=1}^{3} \epsilon_{jkl}\sigma_l \right) \right]$$

$$= \frac{1}{4}\left(I + 2\vec{r}\cdot\vec{\sigma} + |\vec{r}|^2 I \right)$$

$$(\rho^2) = \frac{1}{4}(2 + 2|\vec{r}|^2)$$

If ρ is pure, then $(\rho^2) = 1$.

$$1 = (\rho^2) = \frac{1}{4}(2 + 2|\vec{r}|^2)$$

$$\therefore |\vec{r}| = 1.$$

Conversely, if $|\vec{r}| = 1$, then $(\rho^2) = \frac{1}{4}(2 + 2|\vec{r}|^2) = 1$. Therefore ρ is pure.

2.73

Theorem 2.6

$$\rho = \sum_i p_i \psi_i = \sum_i \tilde{\psi}_i = \sum_j \tilde{\varphi}_j = \sum_j q_j \varphi_j \quad \Leftrightarrow \quad |\tilde{\psi}_i\rangle = \sum_j u_{ij} |\tilde{\varphi}_j\rangle$$

where u is unitary.
The-transformation in theorem 2.6, $|\tilde{\psi}_i\rangle = \sum_j u_{ij} |\tilde{\varphi}_j\rangle$, corresponds to

$$\Big[|\tilde{\psi}_1\rangle \cdots |\tilde{\psi}_k\rangle \Big] = \Big[\, |\tilde{\varphi}_1\rangle \cdots |\tilde{\varphi}_k\rangle \, \Big] U^T$$

where $k = \mathrm{rank}(\rho)$.

$$\begin{aligned}
\sum_i \tilde{\psi}_i &= \Big[|\tilde{\psi}_1\rangle \cdots |\tilde{\psi}_k\rangle \Big] \begin{bmatrix} \langle\tilde{\psi}_1| \\ \vdots \\ \langle\tilde{\psi}_k| \end{bmatrix} && (9) \\
&= \Big[\, |\tilde{\varphi}_1\rangle \cdots |\tilde{\varphi}_k\rangle \Big] U^T U^* \begin{bmatrix} \langle\tilde{\varphi}_1| \\ \vdots \\ \langle\tilde{\varphi}_k| \end{bmatrix} && (10) \\
&= \Big[\, |\tilde{\varphi}_1\rangle \cdots |\tilde{\varphi}_k\rangle \Big] \begin{bmatrix} \langle\tilde{\varphi}_1| \\ \vdots \\ \langle\tilde{\varphi}_k| \end{bmatrix} && (11) \\
&= \sum_j \tilde{\varphi}_j. && (12)
\end{aligned}$$

From spectral theorem, density matrix ρ is decomposed as $\rho = \sum_{k=1}^d \lambda_k k$ where $d = \dim \mathcal{H}$. Without loss of generality, we can assume $p_k > 0$ for $k = 1 \cdots, l$ where $l = \mathrm{rank}(\rho)$ and $p_k = 0$ for $k = l+1, \cdots, d$. Thus $\rho = \sum_{k=1}^l p_k k = \sum_{k=1}^l \tilde{k}$, where $|\tilde{k}\rangle = \sqrt{\lambda_k} |k\rangle$.

Suppose $|\psi_i\rangle$ is a state in support ρ. Then

$$|\psi_i\rangle = \sum_{k=1}^l c_{ik} |k\rangle, \quad \sum_k |c_{ik}|^2 = 1.$$

Define $p_i = \dfrac{1}{\sum_k \frac{|c_{ik}|^2}{\lambda_k}}$ and $u_{ik} = \dfrac{\sqrt{p_i} c_{ik}}{\sqrt{\lambda_k}}$.

Now

$$\sum_k |u_{ik}|^2 = \sum_k \frac{p_i |c_{ik}|^2}{\lambda_k} = p_i \sum_k \frac{|c_{ik}|^2}{\lambda_k} = 1.$$

Next prepare an unitary operator [1] such that ith row of U is $[u_{i1} \cdots u_{ik} \cdots u_{il}]$.

[1] By Gram-Schmidt procedure construct an orthonormal basis $\{\boldsymbol{u}_j\}$ (row vector) with $\boldsymbol{u}_i =$

Then we can define another ensemble such that

$$\Big[\, |\tilde{\psi_1}\rangle \cdots |\tilde{\psi_i}\rangle \cdots |\tilde{\psi_l}\rangle \,\Big] = \Big[\, |\tilde{k}_1\rangle \cdots |\tilde{k}_l\rangle \,\Big] U^T$$

where $|\tilde{\psi_i}\rangle = \sqrt{p_i}\,|\psi_i\rangle$. From theorem 2.6,

$$\rho = \sum_k \tilde{k} = \sum_k \tilde{\psi}_k.$$

Therefore we can obtain a minimal ensemble for ρ that contains $|\psi_i\rangle$. Moreover since $\rho^{-1} = \sum_k \frac{1}{\lambda_k} k$,

$$\langle\psi_i|\rho^{-1}|\psi_i\rangle = \sum_k \frac{1}{\lambda_k}\,\langle\psi_i|k\rangle\langle k|\psi_i\rangle = \sum_k \frac{|c_{ik}|^2}{\lambda_k} = \frac{1}{p_i}.$$

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

2.74

We know

$$\rho_{AB} = |a\rangle\langle a|_A \otimes |b\rangle\langle b|_B$$

$$\rho_A = Tr_B\rho_{AB} = |a\rangle\langle a| Tr(|b\rangle\langle b|) = |a\rangle\langle a|$$

$$Tr(\rho_A^2) = 1$$

Thus ρ_A is pure.

2.75 Reduced Density

Define $|\Phi_\pm\rangle = \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle)$ and $|\Psi_\pm\rangle = \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle)$.

$$|\Phi_\pm\rangle\,\langle\Phi_\pm|_{AB} = \frac{1}{2}(|00\rangle\,\langle 00| \pm |00\rangle\,\langle 11| \pm |11\rangle\,\langle 00| + |11\rangle\,\langle 11|)$$

$$Tr_B(|\Phi_\pm\rangle\,\langle\Phi_\pm|_{AB}) = \frac{1}{2}(|0\rangle\,\langle 0| + |1\rangle\,\langle 1|) = \frac{I}{2}$$

$$|\Psi_\pm\rangle\,\langle\Psi_\pm| = \frac{1}{2}(|01\rangle\,\langle 01| \pm |01\rangle\,\langle 10| \pm |10\rangle\,\langle 01| + |10\rangle\,\langle 10|)$$

$$Tr_B(|\Psi_\pm\rangle\,\langle\Psi_\pm|) = \frac{1}{2}(|0\rangle\,\langle 0| + |1\rangle\,\langle 1|) = \frac{I}{2}$$

$[u_{i1} \cdots u_{ik} \cdots u_{il}]$. Then define unitary $U = \begin{bmatrix} \boldsymbol{u}_1 \\ \vdots \\ \boldsymbol{u}_i \\ \vdots \\ \boldsymbol{u}_l \end{bmatrix}$.

2.76 Extend Proof

Click `https://en.wikipedia.org/wiki/Schmidt_decomposition#Proof` to check the proof

2.77 ABC three components

$$\begin{aligned}
|\psi\rangle &= |0\rangle |\Phi_+\rangle \\
&= |0\rangle \left[\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)\right] \\
&= (\alpha |\phi_0\rangle + \beta |\phi_1\rangle) \left[\frac{1}{\sqrt{2}}(|\phi_0\phi_0\rangle + |\phi_1\phi_1\rangle)\right]
\end{aligned}$$

where $|\phi_i\rangle$ are arbitrary orthonormal states and $\alpha, \beta \in \mathbb{C}$. We cannot vanish cross term. Therefore $|\psi\rangle$ cannot be written as $|\psi\rangle = \sum_i \lambda_i |i\rangle_A |i\rangle_B |i\rangle_C$.

2.78 Schmidt Number

Proof. First Part

If $|\psi\rangle$ is product, then there exist a state $|\phi_A\rangle$ for system A, and a state $|\phi_B\rangle$ for system B such that $|\psi\rangle = |\phi_A\rangle |\phi_B\rangle$.

Obviously, this Schmidt number is 1.

Conversely, if Schmidt number is 1, the state is written as $|\psi\rangle = |\phi_A\rangle |\phi_B\rangle$. Hence this is a product state. □

Proof. Later part.

(⇒) Proved by exercise 2.74.

(⇐) Let a pure state be $|\psi\rangle = \sum_i \lambda_i |i_A\rangle |i_B\rangle$. Then $\rho_A = Tr_B(|\psi\rangle \langle\psi|) = \sum_i \lambda_i^2 |i\rangle \langle i|$. If ρ_A is a pure state, then $\lambda_j = 1$ and otherwise 0 for some j. It follows that $|\psi_j\rangle = |j_A\rangle |j_B\rangle$. Thus $|\psi\rangle$ is a product state. □

2.79 Schmidt Decomposition

Procedure of Schmidt decomposition.
Goal: $|\psi\rangle = \sum_i \sqrt{\lambda_i} |i_A\rangle |i_B\rangle$

- Diagonalize reduced density matrix $\rho_A = \sum_i \lambda_i |i_A\rangle \langle i_A|$.
- Derive $|i_B\rangle$, $|i_B\rangle = \dfrac{(I \otimes \langle i_A|) |\psi\rangle}{\sqrt{\lambda_i}}$
- Construct $|\psi\rangle$.

(i)

$$\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \text{ This is already decomposed.}$$

(ii)

$$\frac{|00\rangle + |01\rangle + |10\rangle + |11\rangle}{2} = \left(\frac{|0\rangle + |1\rangle}{\sqrt{2}}\right) \otimes \left(\frac{|0\rangle + |1\rangle}{\sqrt{2}}\right) = |\psi\rangle\,|\psi\rangle \;\text{ where }\; |\psi\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}$$

(iii)

$$\begin{aligned} |\psi\rangle_{AB} &= \frac{1}{\sqrt{3}}(|00\rangle + |01\rangle + |10\rangle) \\ \rho_{AB} &= \psi_{AB} \end{aligned}$$

$$\begin{aligned} \rho_A =_B (\rho_{AB}) &= \frac{1}{3}\left(20 + 01 + 10 + 1\right) \\ \det(\rho_A - \lambda I) &= \left(\frac{2}{3} - \lambda\right)\left(\frac{1}{3} - \lambda\right) - \frac{1}{9} = 0 \\ \lambda^2 - \lambda + \frac{1}{9} &= 0 \\ \lambda &= \frac{1 \pm \sqrt{5}/3}{2} = \frac{3 \pm \sqrt{5}}{6} \end{aligned}$$

Eigenvector with eigenvalue $\lambda_0 \equiv \frac{3+\sqrt{5}}{6}$ is $|\lambda_0\rangle \equiv \frac{1}{\sqrt{\frac{5+\sqrt{5}}{2}}}\begin{bmatrix} \frac{1+\sqrt{5}}{2} \\ 1 \end{bmatrix}$.

Eigenvector with eigenvalue $\lambda_1 \equiv \frac{3-\sqrt{5}}{6}$ is $|\lambda_1\rangle \equiv \frac{1}{\sqrt{\frac{5-\sqrt{5}}{2}}}\begin{bmatrix} \frac{1-\sqrt{5}}{2} \\ 1 \end{bmatrix}$.

$$\rho_A = \lambda_0\lambda_0 + \lambda_1\lambda_1.$$

$$\begin{aligned} |a_0\rangle &\equiv \frac{(I \otimes \langle\lambda_0|)\,|\psi\rangle}{\sqrt{\lambda_0}} \\ |a_1\rangle &\equiv \frac{(I \otimes \langle\lambda_1|)\,|\psi\rangle}{\sqrt{\lambda_1}} \end{aligned}$$

Then

$$|\psi\rangle = \sum_{i=0}^{1} \sqrt{\lambda_i}\,|a_i\rangle\,|\lambda_i\rangle\,.$$

Calculate $|a_i\rangle$

2.80 Schmidt Coefficient

Let $|\psi\rangle = \sum_i \lambda_i |\psi_i\rangle_A |\psi_i\rangle_B$ and $|\varphi\rangle = \sum_i \lambda_i |\varphi_i\rangle_A |\varphi_i\rangle_B$.
Define $U = \sum_i |\psi_j\rangle \langle\varphi_j|_A$ and $V = \sum_j |\psi_j\rangle \langle\varphi_j|_B$.
Then

$$\begin{aligned}(U \otimes V) |\varphi\rangle &= \sum_i \lambda_i U |\varphi_i\rangle_A V |\varphi_i\rangle_B \\ &= \sum_i \lambda_i |\psi_i\rangle_A |\psi_i\rangle_B \\ &= |\psi\rangle .\end{aligned}$$

2.81 Purification

Let the Schmidt decomposition of $|AR_1\rangle$ be $|AR_1\rangle = \sum_i \sqrt{p_i} |\psi_i^A\rangle |\psi_i^R\rangle$ and let $|AR_2\rangle = \sum_i \sqrt{q_i} |\phi_i^A\rangle |\phi_i^R\rangle$.

Suppose ρ^A has orthonormal decomposition $\rho^A = \sum_i p_i |i\rangle \langle i|$.

Since $|AR_1\rangle$ and $|AR_2\rangle$ are purifications of the ρ^A, we have

$$\begin{aligned}Tr_R(|AR_1\rangle \langle AR_1|) = Tr_R(|AR_2\rangle \langle AR_2|) = \rho^A \\ \therefore \sum_i p_i |\psi_i^A\rangle \langle\psi_i^A| = \sum_i q_i |\phi_i^A\rangle \langle\phi_i^A| = \sum_i \lambda_i |i\rangle \langle i| .\end{aligned}$$

The $|i\rangle$, $|\psi_i^A\rangle$, and $|\psi_i^A\rangle$ are orthonormal bases and they are eigenvectors of ρ^A. Hence without loss of generality, we can consider

$$\lambda_i = p_i = q_i \text{ and } |i\rangle = |\psi_i^A\rangle = |\phi_i^A\rangle .$$

Then

$$\begin{aligned}|AR_1\rangle &= \sum_i \lambda_i |i\rangle |\psi_i^R\rangle \\ |AR_2\rangle &= \sum_i \lambda_i |i\rangle |\phi_i^R\rangle\end{aligned}$$

Since $|AR_1\rangle$ and $|AR_2\rangle$ have same Schmidt numbers, there are two unitary operators U and V such that $|AR_1\rangle = (U \otimes V) |AR_2\rangle$ from exercise 2.80.

Suppose $U = I$ and $V = \sum_i |\psi_i^R\rangle \langle\phi_i^R|$. Then

$$\begin{aligned}\left(I \otimes \sum_j |\psi_j^R\rangle \langle\phi_j^R|\right) |AR_2\rangle &= \sum_i \lambda_i |i\rangle \left(\sum_j |\psi_j^R\rangle \langle\phi_j^R| |\phi_i^R\rangle\right) \\ &= \sum_i \lambda_i |i\rangle |\psi_i^R\rangle \\ &= |AR_1\rangle .\end{aligned}$$

Therefore there exists a unitary transformation U_R acting on system R such that $|AR_1\rangle = (I \otimes U_R) |AR_2\rangle$.

2.82

(1)

Let $|\psi\rangle = \sum_i \sqrt{p_i} |\psi_i\rangle |i\rangle$.

$$\begin{aligned}
Tr_R(|\psi\rangle\langle\psi|) &= \sum_{i,j} \sqrt{p_i}\sqrt{p_j} |\psi_i\rangle\langle\psi_j| Tr_R(|i\rangle\langle j|) \\
&= \sum_{i,j} \sqrt{p_i}\sqrt{p_j} |\psi_i\rangle\langle\psi_j| \delta_{ij} \\
&= \sum_i p_i |\psi_i\rangle\langle\psi_i| = \rho.
\end{aligned}$$

Thus $|\psi\rangle$ is a purification of ρ.

(2)

Define the projector P by $P = I \otimes |i\rangle\langle i|$. The probability we get the result i is

$$Tr\left[P|\psi\rangle\langle\psi|\right] = \langle\psi|P|\psi\rangle = \langle\psi|(I \otimes i)|\psi\rangle = p_i \langle\psi_i|\psi_i\rangle = p_i.$$

The post-measurement state is

$$\frac{P|\psi\rangle}{\sqrt{p_i}} = \frac{(I \otimes |i\rangle\langle i|)|\psi\rangle}{\sqrt{p_i}} = \frac{\sqrt{p_i}|\psi_i\rangle|i\rangle}{\sqrt{p_i}} = |\psi_i\rangle|i\rangle.$$

If we only focus on the state on system A,

$$Tr_R(|\psi_i\rangle|i\rangle) = |\psi_i\rangle.$$

(3)

($\{|\psi_i\rangle\}$ is not necessary an orthonormal basis.)

Suppose $|AR\rangle$ is a purification of ρ and its Schmidt decomposition is $|AR\rangle = \sum_i \sqrt{\lambda_i} |\phi_i^A\rangle |\phi_i^R\rangle$.

From assumption

$$Tr_R\left(|AR\rangle\langle AR|\right) = \sum_i \lambda_i |\phi_i^A\rangle\langle\phi_i^A| = \sum_i p_i |\psi_i\rangle\langle\psi_i|.$$

By theorem 2.6, there exits an unitary matrix u_{ij} such that $\sqrt{\lambda_i}|\phi_i^A\rangle =$

$\sum_j u_{ij}\sqrt{p_j}\,|\psi_j\rangle$. Then

$$\begin{aligned}
|AR\rangle &= \sum_i \left(\sum_j u_{ij}\sqrt{p_j}\,|\psi_j\rangle\right)|\phi_i^R\rangle \\
&= \sum_j \sqrt{p_j}\,|\psi_j\rangle \otimes \left(\sum_i u_{ij}\,|\phi_i^R\rangle\right) \\
&= \sum_j \sqrt{p_j}\,|\psi_j\rangle\,|j\rangle \\
&= \sum_i \sqrt{p_i}\,|\psi_i\rangle\,|i\rangle
\end{aligned}$$

where $|i\rangle = \sum_k u_{ki}\,|\phi_k^R\rangle$.

About $|i\rangle$,

$$\begin{aligned}
\langle k|l\rangle &= \sum_{m,n} u_{mk}^* u_{nl}\,\langle\phi_m^R|\phi_n^R\rangle \\
&= \sum_{m,n} u_{mk}^* u_{nl}\delta_{mn} \\
&= \sum_m u_{mk}^* u_{ml} \\
s &= \delta_{kl}, \quad (\because u_{ij} \text{ is unitary.})
\end{aligned}$$

which implies $|j\rangle$ is an orthonormal basis for system R.

Therefore if we measure system R w.r.t $|j\rangle$, we obtain j with probability p_j and post-measurement state for A is $|\psi_j\rangle$ from (2). Thus for any purification $|AR\rangle$, there exists an orthonormal basis $|i\rangle$ which satisfies the assertion.

2.1

From Exercise 2.35, $\vec{n}\cdot\vec{\sigma}$ is decomposed as

$$\vec{n}\cdot\vec{\sigma} = |\lambda_1\rangle\langle\lambda_1| - |\lambda_1\rangle\langle\lambda_{-1}|$$

where $|\lambda_{\pm 1}\rangle$ are eigenvector of $\vec{n}\cdot\vec{\sigma}$ with eigenvalues ±1.

Thus

$$\begin{aligned}
f(\theta\vec{n}\cdot\vec{\sigma}) &= f(\theta)\,|\lambda_1\rangle\langle\lambda_1| + f(-\theta)\,|\lambda_{-1}\rangle\langle\lambda_{-1}| \\
&= \left(\frac{f(\theta)+f(-\theta)}{2} + \frac{f(\theta)-f(-\theta)}{2}\right)|\lambda_1\rangle\langle\lambda_1| + \left(\frac{f(\theta)+f(-\theta)}{2} - \frac{f(\theta)-f(-\theta)}{2}\right)|\lambda_{-1}\rangle\langle\lambda_{-1}| \\
&= \frac{f(\theta)+f(-\theta)}{2}\left(|\lambda_1\rangle\langle\lambda_1| + |\lambda_{-1}\rangle\langle\lambda_{-1}|\right) + \frac{f(\theta)-f(-\theta)}{2}\left(|\lambda_1\rangle\langle\lambda_1| - |\lambda_{-1}\rangle\langle\lambda_{-1}|\right) \\
&= \frac{f(\theta)+f(-\theta)}{2}I + \frac{f(\theta)-f(-\theta)}{2}\vec{n}\cdot\vec{\sigma}
\end{aligned}$$

2.2

Chapter 3 solutions

3.1 Non-computable processes in Nature

Since we know that a Turing Machine only maps from non-negative to non-negative numbers, and if any process in nature is found to map between different sets of values then that process can't be run on a Turing Machine.

3.2 Turing numbers

Taking the input of the Turing Machine $a_1, ..., a_k$ and then give the Turing machine the value $p_1^{a_1} \times p_2^{a_2} \times ..... \times p_k^{a_k}$ withe $p_1, p_2, p_3, ...p_k$ being the first k prime numbers, and thus all Turing machine with unique inputs will be given unique value identifiers since all numbers only have one kind of prime factorization.

3.3 Turing Machine to reverse a bit string

Our aim is to design a Turing Machine to reverse a binary string consisting of 0 and 1, we will call them a and b

- Move the last digit , replace x for a or x for b and move right to convert the corresponding B to a or b accordingly.
- Move left until the symbol left to x is reached.
- Perform Step 1 and Step 2 until B is reached while traversing left.
- Replace every x to B to make the cells empty since the reverse string is performed by the previous steps.

The transition diagram for the Turing machine looks like this

3.4 Turing Machine to add modulo 2

- Start in the initial state, with the tape head positioned over the leftmost digit of the first number.
- If the digit under the tape head is 0, move the tape head one position to the right and transition to the next state.
- If the digit under the tape head is 1, write a 0 at that position and move the tape head one position to the right.
- Repeat steps 2 and 3 until the tape head reaches the rightmost digit of the first number.

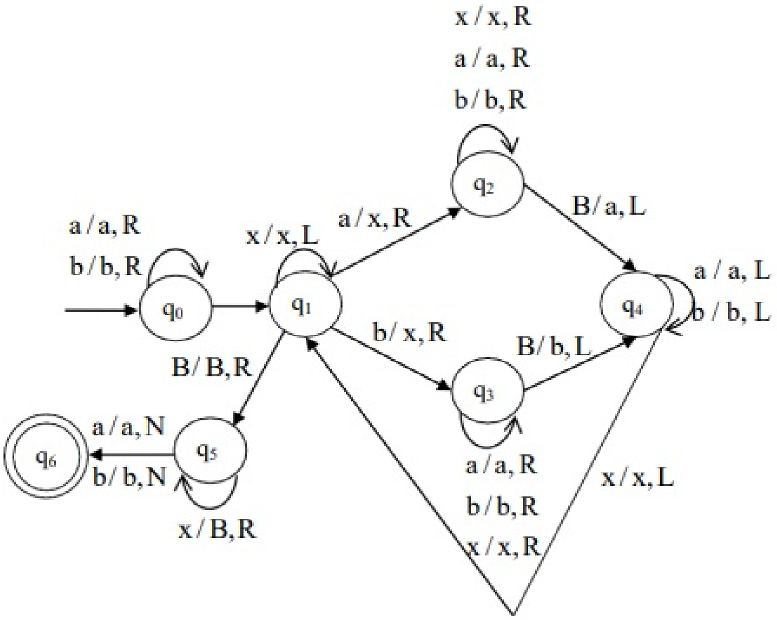

Figure 1: The Turing Machine for reverse string

- Transition to the next state and move the tape head to the leftmost digit of the second number.
- Repeat steps 2-4 for the second number.
- Move the tape head to the leftmost digit of the result and transition to the final state.

This Turing machine will add two binary numbers modulo 2, effectively performing a bitwise XOR operation on the input numbers.

3.5 Halting Problem with no inputs

The problem of determining whether a given Turing machine halts when the input to the machine is a blank tape is known as the halting problem. It was shown by Alan Turing in 1936 that there is no general algorithm that can determine whether an arbitrary Turing machine halts on a blank input.

The proof of this result relies on the concept of diagonalization, where a new machine is constructed that can determine whether any given machine halts on a blank input. If such an algorithm existed, then it could be encoded as a Turing machine, say H. It is possible to construct a new machine, say D, that takes as input the description of another machine M and runs both M and H on the blank tape. D compares the outputs of M and H for each machine it receives as input. If the outputs are different for any input machine M, D outputs ”M does not halt.” Otherwise, D outputs ”M halts.”

The problem is that D can be used to determine whether itself halts. If D halts, it will say ”D halts.” If D does not halt, it will say ”D does not halt.” This creates a contradiction, because D cannot simultaneously halt and not halt. Therefore, it is impossible for a machine to determine whether another machine halts on a blank input.

This proof was shown by Alan Turing in his paper ”On computable numbers, with an application to the Entscheidungsproblem” in 1936, showing that the halting problem is undecidable.

Define H2(M) = 0 if the machine doesn’t halt with blank input , 1 if the machine does halt with a blank input

We have the following algorithm

```
Turing(M)
Y = H2(M)
if y == 0
halt
else
‘Loop Forever
```

Here if M is blank, then $H2(M) = 1$ only if $y = H2(M) = 0$, thus we form a contradiction, meaning this machine cannot read input M, which is blank, meaning there is no algorithm to solve $H2(M)$ for this particular machine.

3.6 Probabilistic halting problem

Define Hp(x) = 0 if probability machine x halts in input $x < 1/2$, 1 if probability $> 1/2$ We have the following algorithm:

```
Turing(x)
Y = hp(x)
Y2 = flip an unbiased coin
if y==1 and y2 = heads
halt
Else
loop forever
End if
```

Here, assume hp(x) = 1 and thus corresponding probability $p > 1/2$. The probability program halts are $p \times 1/2 =$ at most $1/2$ since p is at most 1. Thus, this contradicts our original statement that hp(x) = 1, and thus there is no algorithm to correctly determine hp(x) for this machine.

3.7 Halting oracle

Yes. Before, the problem was no algorithm existed to compute HALT(x) for all x, but now that the black box exists that algorithm also exists, and since Turing machines can compute all algorithms, these Turing machines can compute this algorithm for all machines.

3.8 Universality of NAND

AND - input bit $[1, x_0, x_1]$. Apply NAND to x_0 and x_1 , then a NAND on $[1, \mathrm{NAND}(x_0, x_1)]$
NOT - NAND input bit x_0 and ancilla bit 1
XOR - input $[1, x_0, x_1, 1, 1]$ $\mathrm{NAND}[1, \mathrm{NAND}[x_0, x_1], \mathrm{NAND}[\mathrm{NAND}[1, x_0]\ \mathrm{NAND}[1, x_1]]]$

3.9 Prove that $f(n)$ is $O(g(n))$ iff $g(n)$ is $\Omega(f(n))$

$f(n)$ is $O(g(n)) \rightarrow f(n) \leq cg(n)$
$g(n)$ is $\Omega(f(n)) \rightarrow cf(n) \leq g(n)$
If $g(n)$ is $\Omega(f(n))$ then $cf(n) \leq g(n), f(n) \leq g(n)/c$
And thus $f(n) \leq cg(n)$ and $f(n)$ is $O(g(n))$

Thus $g(n)$ is $\Theta(f(n))$ and $f(n)$ is $O(g(n))$

3.10 Suppose $g(n)$ is a polynomial of degree k

$$g(n)\mathrm{is}O(n^1) \longrightarrow g(n) \leq cn^1$$

$$g(n) = An^k + Bn^{k-1} + ... + d \leq cn^1$$

Thus, because $An^k + Bn^{k-1} + ... + d \leq n^{k+1} \leq n^{k+2}$ and we design c such that $An^k + Bn^{k-1} + ... + d \leq cn^k$, thus $g(n)$ is $O(n^1)$.

3.11 Show that $\log$(n) is $O(n^k)$ for any $k > 0$

$\log n \leq cn^k$, $n \leq c10^{n^k}$, so we design c^k such that $n \leq c \times 10^{n^k}$ for all $k > 0$

3.12 $n^{\log n}$ is super polynomial

From 3.10 , we proved $\log n \geq k$ for sufficiently large n, thus $g(n) = n^k$ is in $O(n^{\log n})$. However, $\log n$ is not $\leq k$ for large n, so $g(n) = n^k$ is never in $O(n^k)$

3.13 $n^{\log n}$ is sub-exponential

c^n is $\Omega(n^{\log n})$, check graphically for sufficiently large n.
Since c^n is $\Omega(n^{\log n})$, $n^{\log n}$ can't be in $\Omega(c^n)$.

3.14 Suppose $e(n)$ is $O(f(n))$ and $g(n)$ is $O(h(n))$

$$e(n) \text{ is } O(f(n)) \rightarrow e(n) \leq cf(n)$$

$$g(n) \text{ is } O(h(n)) \rightarrow g(n) \leq c2 \times h(n)$$

$$e(n) \times g(n) \leq c \times c2 \times f(n) \times h(n) = c3 \times f(n) \times h(n)$$

3.15 Lower bound for compare and swap based sorts

After 1 swapping, there is only 2^1 ordering such that the swapping puts the whole thing in order. After 2 swapping, there are 2^2 orderings that were two swapping away from being sorted. After k swappings, it follows that 2^k initial orderings are now sorted.

$$2^n \log n = 2^{\log n^n = n^n} \geq n!$$

Thus, the lower bound on sorting is $n \log n$. You can see that asymptotic notations have the important effect of allowing us to find how efficient a particular algorithm can be on a process, but it doesn't necessarily tell us HOW to make such an algorithm, but it gives us an idea of how efficient the most efficient algorithm can be.

3.16 Hard to compute function exist

There exist Boolean functions on n inputs which require at least $2^n/\log(n)$ logic gates to compute. One such example is the function known as the "n-input majority function."

The majority function, denoted as $MAJ_n(x_1, x_2, ..., x_n)$, is a Boolean function that takes n binary inputs and outputs 1 if and only if more than $n/2$ of the inputs are 1. The function can be computed by creating n AND gates, one for each input, and $n - 1$ OR gates to combine the outputs of the AND gates. In the worst case, where all inputs are 1, the number of gates required is $(n-1)+(n-1) = 2n-2 = 2^n/log(n)$ gates.

Another example is the "n-input parity function," which is a Boolean function that takes n binary inputs and outputs 1 if and only if the number of 1s in the inputs is odd. The function can be computed by creating n XOR gates, one for each input, and then $n - 1$ XOR gates to combine the outputs of the previous XOR gates. In the worst-case scenario, the number of gates required is $(n-1)+(n-1) = 2n-2 = 2^n/log(n)$ gates.

This shows that there are Boolean functions that require at least $2^n/log(n)$ logic gates to compute, and that is a lower bound for the complexity of some Boolean functions.

3.17 Prove that a polynomial-time algorithm for finding the factors of a number m exists iff the factoring decision problem is in P

If this problem is in P, then a Turing machine exists for identifying if a value is a factor of a number m and less than L, and thus setting $L = m$ we have the factoring algorithm that can thus be done efficiently. If this problem isn't in P, the factoring obviously can't be done efficiently since if for any L it is inefficient then $L = m$ is definitely inefficient.

3.18 Prove that if $coNP \neq NP$, then $P \neq NP$

P is a subset of $CoNP$ so if $CoNP$ does not equal NP, there are some problems in P and $CoNP$, and the rest of the problems in P and NP, and thus there are some problems in P that are in $CoNP$ but not NP and thus P does not equal NP.

3.19 The REACHABILITY problem

Start at the first vertex, try to get to the 2nd. At most n vertices exist to visit, so its $O(n)$(algorithm would make sure not to visit a vertice more than once). Then use Reachability algorithm to form the following $O(n^2)$ algorithm:

```
For $i$ through all vertices
For $j$ through all vertices
Test Reachability(vertex(i), vertex(j)
End $j$
End $i$
```

3.20 Euler's Theorem

Euler's theorem is based on the following idea: If you visit a node, you can go on a new edge you haven't gone before, because the vertex has an even amount of incident edges so if a node is visited and then left, 2 edges have been traversed and used. Thus if you go through all nodes you will always have a new edge to move through for every node until there are no more edges left to visit, at which point you are back at the original node and have completed the cycle.

The procedure then would be to start at a node, go to all other nodes until you have reached back to the original node and there are no new edges to visit.

3.21 Transitive property of reduction

Since $L1->L2$ exists, there exists function $R(x)$ that gives a string in language $L2$ iff x is in $L1$. Since $L2->L3$ exists, there exists a function $R2(x_2)$ that gives a string in $L3$ iff x_2 is in $L2$. Thus, with $R(x)+R2(R(x))$ (polynomial time overhead), we reduce $L1$ to $L3$.

3.22 L is complete

Since all other problems can be reduced to L, and L can be reduced to L', all other problems can be reduced to L'.

3.23 SAT is NP-complete problem

SAT (Satisfiability) is a problem of determining whether a given Boolean formula, in conjunctive normal form (CNF) has a satisfying assignment.

First, to show that SAT is in NP, we need to show that a solution to the problem can be verified in polynomial time. Given a Boolean formula in CNF and an assignment, we can check in polynomial time if the assignment satisfies the formula, by checking each clause of the formula. If all the clauses are true for the given assignment, then the assignment is a satisfying assignment, otherwise not. Since the size of the input does not affect the time complexity of the verification, this means that SAT is in NP.

Next, to show that SAT is NP-complete, we need to show that there exists a reduction from an arbitrary problem in NP to SAT. One such problem is known as CSAT (Circuit SAT), which is the problem of determining whether a given Boolean circuit has a satisfying assignment.

To show that CSAT reduces to SAT, we can construct a Boolean formula in CNF that represents the given Boolean circuit. Each gate in the circuit can be represented by a clause in the CNF formula, and each input to the circuit can be represented by a variable in the formula. We can then use the variables to represent the inputs and the clauses to represent the gates, such that the

Boolean formula in CNF is satisfied if and only if the circuit has a satisfying assignment.

Since CSAT is in NP and it is possible to reduce it to SAT, it implies that SAT is NP-complete. This means that SAT is at least as hard as any problem in NP, and that any problem in NP can be reduced to SAT in polynomial time.

3.25 PSPACE subset EXP

With $lm^{p(n)}$ different states, all problems in PSPACE can be solved by going through all possible states in $lm^{p(n)}$ time, or exponential time. Thus PSPACE is a subset of EXP

3.26 L subset P

If you can solve a problem in L, then you can solve the problem by going through around $c \times log(n)$ spaces, thus you can solve it in time $c \times log(n)$, which is polynomial. Thus L is a subset of P.

3.27 Approximation algorithm for VERTEX COVER

In this algorithm, at worst the min vector span is made up of all the alpha's and none of the Beta's, in which case at worst this algorithm calculates the a vector span that is 2× the space of the min vector span.

3.29 Fredkin gate is self inverse

A Fredkin gate is a reversible gate that acts on three qubits, and it is also known as a Controlled-SWAP gate. The Fredkin gate is defined as:

$$|x{>}|y{>}|z{>}\text{ß}|x{>}|y{>}|z{>}if\text{x} = 0$$

$$|x{>}|y{>}|z{>}\text{ß}|x{>}|z{>}|y{>}\text{if}x = 1$$

Where $|x{>}$, $|y{>}$ and $|z{>}$ are the input qubits and —x⟩, —y⟩, —z⟩ are the output qubits. The qubit x is the control qubit, and the qubits y and z are the target qubits.

Now, let's consider applying two consecutive Fredkin gates, where the first one acts on qubits —x⟩, —y⟩, —z⟩ and the second one acts on qubits —x⟩, —z⟩, —w⟩.

If x = 0, the first Fredkin gate doesn't change the state of the qubits and the second Fredkin gate also don't change the state of the qubits, and the output state is —x⟩—y⟩—z⟩—w⟩ which is the same as the input state —x⟩—y⟩—z⟩—w⟩

If x = 1, the first Fredkin gate swaps the qubits —y⟩ and —z⟩, and the second Fredkin gate swaps the qubits —z⟩ and —w⟩. The net effect is that the qubits —y⟩ and —w⟩ are swapped and the output state is —x⟩—w⟩—z⟩—y⟩ which is the same as the input state —x⟩—y⟩—z⟩—w⟩

Therefore, applying two consecutive Fredkin gates gives the same outputs as inputs

3.32 From fredkin to toffoli

e a Fredkin gate? The Toffoli gate, also known as the Controlled-Controlled-NOT gate, is a reversible gate that acts on three qubits. The Toffoli gate is defined as:

$$|x{>}|y{>}|z{>}\text{ß}|x{>}|y{>}|z{>} \text{ if } x = 0 \text{ and } y = 0$$

$$|x{>}|y{>}|z{>}\text{ß}|x{>}|y{>}|z \otimes 1{>} \text{ if } x = 0 \text{ and } y = 1$$

where —x⟩, —y⟩, and —z⟩ are the input qubits and —x⟩, —y⟩, —z⟩ are the output qubits. The qubits x and y are the control qubits, and the qubit z is the target qubit.

The smallest number of Fredkin gates needed to simulate a Toffoli gate is 3. To simulate a Toffoli gate with three Fredkin gates, we can use a technique called 'controlled-SWAP-controlled-SWAP', where the first two Fredkin gates act on —x⟩, —y⟩, —z⟩ and the third Fredkin gate acts on —x⟩, —y⟩, —t⟩ where —t⟩ is an auxiliary qubit.

On the other hand, Toffoli gate is a more powerful gate than Fredkin gate and it can simulate Fredkin gate, but it requires 2 Toffoli gates to simulate 1 Fredkin gate. To simulate a Fredkin gate with two Toffoli gates, we can use a technique called "controlled-controlled-NOT-controlled-NOT" where the first Toffoli gate is applied on —x⟩, —y⟩, —t1⟩ where —t1⟩ is an auxiliary qubit, and the second Toffoli gate is applied on —x⟩, —t1⟩, —z⟩

It's important to note that while it's possible to simulate a Toffoli gate using Fredkin gates, or vice versa, it's not always the most efficient way to perform the operation, as the number of gates required for the simulation is higher than using the native gate.