Select Solutions for "Quantum Computation and Quantum Information: 10th Anniversary Edition" by Nielsen and Chuang

Original author: goropikari
Extended by: tlesaul2

January 29, 2022

Repository

As of November, 2021, the original source LaTeX code, located at `https://github.com/goropikari/SolutionForQuantumComputationAndQuantumInformation` has not been updated since April 2020. The extended source LaTeX code is located at
`https://github.com/tlesaul2/SolutionQCQINielsenChuang` . It may be updated more actively.

For readers

This is an unofficial solution manual for "Quantum Computation and Quantum Information: 10th Anniversary Edition" (ISBN-13: 978-1107002173) by Michael A. Nielsen and Isaac L. Chuang.

From the original author:

I have studied quantum information theory as a hobby. And I'm not a researcher. So there is no guarantee that these solutions are correct. Especially because I'm not good at mathematics, proofs are often wrong. Don't trust me. Verify yourself!

If you find some mistake or have some comments, please feel free to open an issue or a PR.

goropikari

From the second author:

I'm a mathematician relatively new to quantum information theory as of the adoption of this repo, so hope to supplement the original author's work by checking and formalizing the mathematics, overly at times, while I use the task to learn the field. The original author's sentiments about self-verification are echoed.

tlesaul2

Contents

The exercise in this chapter is only interesting for it's mathematics, so it was moved to the end to avoid dissuading non-mathematicians from continuing to chapters more interesting for their quantum information theory.

Errata list

- p.101. eq (2.150) $\rho = \sum_m p(m)\rho_m$ should be $\rho' = \sum_m p(m)\rho_m$.

- p.408. eq (9.49) $\sum_i p_i D(\rho_i, \sigma_i) + D(p_i, q_i)$ should be $\sum_i p_i D(\rho_i, \sigma_i) + 2D(p_i, q_i)$.

$$
\begin{aligned}
\text{eqn (9.48)} &= \sum_i p_i \operatorname{Tr}(P(\rho_i - \sigma_i)) + \sum_i (p_i - q_i) \operatorname{Tr}(P\sigma_i) \\
&\leq \sum_i p_i \operatorname{Tr}(P(\rho_i - \sigma_i)) + \sum_i |p_i - q_i| \operatorname{Tr}(P\sigma_i) \quad (\because p_i - q_i \leq |p_i - q_i|) \\
&\leq \sum_i p_i \operatorname{Tr}(P(\rho_i - \sigma_i)) + \sum_i |p_i - q_i| \quad (\because \operatorname{Tr}(P\sigma_i) \leq 1) \\
&= \sum_i p_i \operatorname{Tr}(P(\rho_i - \sigma_i)) + 2\frac{\sum_i |p_i - q_i|}{2} \\
&= \sum_i p_i \operatorname{Tr}(P(\rho_i - \sigma_i)) + 2D(p_i, q_i)
\end{aligned}
$$

- p.409. Exercise 9.12. If $\rho = \sigma$, then $D(\rho, \sigma) = 0$. Furthermore trace distance is non-negative. Therefore $0 \leq D(\mathcal{E}(\rho), \mathcal{E}(\sigma)) \leq 0 \Rightarrow D(\mathcal{E}(\rho), \mathcal{E}(\sigma)) = 0$. So I think the map $\mathcal{E}$ is not strictly contractive. If $p \neq 1$ and $\rho \neq \sigma$, then $D(\mathcal{E}(\rho), \mathcal{E}(\sigma)) < D(\rho, \sigma)$ is satisfied.

- p.411. Exercise 9.16. eqn(9.73) $\operatorname{Tr}(A^\dagger B) = \langle m|A \otimes B|m\rangle$ should be $\operatorname{Tr}(A^T B) = \langle m|A \otimes B|m\rangle$.

  Simple counter example is the case that $A = \begin{bmatrix} i & 0 \\ 0 & 0 \end{bmatrix}$. $B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, In this case,

$$
\begin{aligned}
A^\dagger B &= \begin{bmatrix} -i & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} -i & 0 \\ 0 & 0 \end{bmatrix}, \\
\operatorname{Tr}(A^\dagger B) &= -i, \\
A \otimes B &= \begin{bmatrix} i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\
\langle m|A \otimes B|m\rangle &= (\langle 00| + \langle 11|)(A \otimes B)(|00\rangle + |11\rangle) = i.
\end{aligned}
$$

  Thus $\operatorname{Tr}(A^\dagger B) \neq \langle m|A \otimes B|m\rangle$.

  By using following relation, we can prove.

$$
\begin{aligned}
(I \otimes A)\,|m\rangle &= (A^T \otimes I)\,|m\rangle \\
\operatorname{Tr}(A) &= \langle m|I \otimes A|m\rangle
\end{aligned}
$$

$$\begin{aligned}\mathrm{Tr}(A^T B) = \mathrm{Tr}(BA^T) &= \langle m|I \otimes BA^T|m\rangle \\ &= \langle m|(I \otimes B)(I \otimes A^T)|m\rangle \\ &= \langle m|(I \otimes B)(A \otimes I)|m\rangle \\ &= \langle m|A \otimes B|m\rangle\,.\end{aligned}$$

- p.515. eqn (11.67) $S(\rho'||\rho)$ should be $S(\rho||\rho')$.

Chapter 2

Introduction to quantum mechanics

2.1) Show that $(1,-1),(1,2)$, and $(2,1)$ are linearly dependent.
Soln: It is enough to express $(0,0)$ as a linear combination of the specified vectors.

$$\begin{bmatrix}1\\-1\end{bmatrix}+\begin{bmatrix}1\\2\end{bmatrix}-\begin{bmatrix}2\\1\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}$$

2.2) 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 bais $|0\rangle,|1\rangle$, and the output basis $|0\rangle,|1\rangle$. Find input and output bases which give rise to a different matrix representation of A.
Soln: With specified operations, it is enough to solve for the entries of a 2x2 matrix which coverts the input vectors expressed as linear combinations of one basis, say $(|a_1\rangle,|a_2\rangle)$, into vectors expressed as linear combinations of another basis, say $(|b_1\rangle,|b_2\rangle)$.

$$A=\begin{array}{cc} & \begin{array}{cc}|b_1\rangle & |b_2\rangle\end{array}\\ \begin{array}{c}|a_1\rangle\\|a_2\rangle\end{array} & \begin{bmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{bmatrix}\end{array}$$

With $(|a_1\rangle,|a_2\rangle)=(|0\rangle,|1\rangle)$ and $(|b_1\rangle,|b_2\rangle)=(|0\rangle,|1\rangle)$, we have

$$A|0\rangle:=|1\rangle=0|0\rangle+1|1\rangle=A_{11}|b_1\rangle+A_{21}|b_2\rangle=A_{11}|0\rangle+A_{21}|1\rangle\Rightarrow A_{11}=0,\ A_{21}=1$$
$$A|1\rangle:=|0\rangle=1|0\rangle+0|1\rangle=A_{12}|b_1\rangle+A_{22}|b_2\rangle=A_{12}|0\rangle+A_{22}|1\rangle\Rightarrow A_{12}=1,\ A_{22}=0$$
$$\therefore A=\begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}$$

If the output basis was $(|b_1\rangle,|b_2\rangle)=(|1\rangle,|0\rangle)$ instead, then $A=I$. More formally:

$$A|0\rangle:=|1\rangle=1|1\rangle+0|0\rangle=A_{11}|b_1\rangle+A_{21}|b_2\rangle=A_{11}|1\rangle+A_{21}|0\rangle\Rightarrow A_{11}=1,\ A_{21}=0$$
$$A|1\rangle:=|0\rangle=0|1\rangle+1|0\rangle=A_{12}|b_1\rangle+A_{22}|b_2\rangle=A_{12}|1\rangle+A_{22}|0\rangle\Rightarrow A_{12}=0,\ A_{22}=1$$
$$\therefore A=\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$$

With a more interesting orthonormal output basis $(|b_1\rangle,|b_2\rangle)=(|+\rangle,|-\rangle)$:

$$A|0\rangle:=|1\rangle=\frac{\sqrt{2}}{2}|+\rangle-\frac{\sqrt{2}}{2}|-\rangle=A_{11}|b_1\rangle+A_{21}|b_2\rangle=A_{11}|+\rangle+A_{21}|-\rangle\Rightarrow A_{11}=\frac{\sqrt{2}}{2},\ A_2=-\frac{\sqrt{2}}{2}$$
$$A|1\rangle:=|0\rangle=\frac{\sqrt{2}}{2}|+\rangle+\frac{\sqrt{2}}{2}|-\rangle=A_{12}|b_1\rangle+A_{22}|b_2\rangle=A_{12}|+\rangle+A_{22}|-\rangle\Rightarrow A_{12}=\frac{\sqrt{2}}{2},\ A_{22}=\frac{\sqrt{2}}{2}$$
$$\therefore A=\frac{\sqrt{2}}{2}\begin{bmatrix}1 & -1\\1 & 1\end{bmatrix}$$

Note: This is similar, but not equal to **H**. Had A been the identity transformation when expressed with the same input and output bases, then the result would have been exactly **H**.

2.3) Suppose A is a linear operator from vector space V to vector space W, and B is a linear operator from vector space 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.
Soln: Fix i. We'll show that $(B \circ A)_{ki} = (B \cdot A)_{ki}$.

$$\begin{aligned}
(B \circ A)|v_i\rangle = \sum_k (B \circ A)_{ki} |x_k\rangle = B\left(\sum_j A_{ji} |w_j\rangle\right) & & \text{(Eqn 2.12, composition)} \\
= \sum_j A_{ji} B |w_j\rangle & & \text{(linearity)} \\
= \sum_{j,k} A_{ji} B_{kj} |x_k\rangle & & \text{(Eqn 2.12)} \\
= \sum_k \left(\sum_j B_{kj} A_{ji}\right) |x_k\rangle & & \text{(finiteness, commutativity)} \\
= \sum_k ((B \cdot A)_{ki}) |x_k\rangle & & \text{(definition)}
\end{aligned}$$

$$\therefore (B \circ A)_{ki} = (B \cdot A)_{ki}$$

2.4) 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. This matrix is known as the *identity matrix*
Soln: Let I be the matrix in question.

$$I|v_j\rangle := |v_j\rangle = \sum_i I_{ij} |v_i\rangle, \ \forall j.$$

$$\Rightarrow I_{ij} = \delta_{ij} := \begin{cases} 1 & i = j \\ 0 & o/w \end{cases}$$

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

$$((y_1, \cdots, y_n), (z_1, \cdots, z_n)) = \sum_i y_i^* z_i.$$

Equation (2.13.1), linearity in second argument:

$$\begin{aligned}
\left((y_1,\cdots,y_n),\sum_i \lambda_i(z_{i1},\cdots,z_{in})\right) &= \left((y_1,\cdots,y_n),\left(\sum_i \lambda_i z_{i1},\cdots,\sum_i \lambda_i z_{in}\right)\right) && \text{(definition)}\\
&= \sum_j y_j^*\left(\sum_i \lambda_i z_{ij}\right) && \text{(definition)}\\
&= \sum_j \left(\sum_i y_j^*\lambda_i z_{ij}\right) && \text{(linearity of multiplcation)}\\
&= \sum_j \left(\sum_i \lambda_i y_j^* z_{ij}\right) && \text{(associativity/commutativity)}\\
&= \sum_i \left(\sum_j \lambda_i y_j^* z_{ij}\right) && \text{(finiteness)}\\
&= \sum_i \lambda_i \left(\sum_j y_j^* z_{ij}\right) && \text{(linearity)}\\
&= \sum_i \lambda_i \left((y_1,\cdots,y_n),(z_{i1},\cdots,z_{in})\right) && \text{(definition)}
\end{aligned}$$

Equation (2.13.2), conjugate symmetry:

$$\begin{aligned}
\big((y_1,\cdots,y_n),(z_1,\cdots,z_n)\big)^* &= \left(\sum_i y_i^* z_i\right)^* && \text{(definition)}\\
&= \left(\sum_i y_i z_i^*\right) && \text{(conjugate symmetry in } \mathbb{C}^1)\\
&= \left(\sum_i z_i^* y_i\right) && \text{(commutativity in } \mathbb{C}^1)\\
&= \big((z_1,\cdots,z_n),(y_1,\cdots,y_n)\big) && \text{(definition)}
\end{aligned}$$

Equation (2.13.3), positive definiteness:

$$\begin{aligned}
\big((y_1,\cdots,y_n),(y_1,\cdots,y_n)\big) &= \sum_i y_i^* y_i && \text{(definition)}\\
&= \sum_i |y_i|^2 && \text{(definition)}\\
&\geq 0 && \text{(positive definiteness of } |\cdot|^2 \text{ over } \mathbb{C}^1)
\end{aligned}$$

Now:

$$\begin{aligned}
\big((y_1,\cdots,y_n),(y_1,\cdots,y_n)\big) = \sum_i |y_i|^2 &\stackrel{?}{=} 0 && \text{(hypothesis)}\\
\iff |y_i|^2 &= 0\ \forall i && \text{(positivity of } |\cdot|^2)\\
\iff y_i &= 0\ \forall i && \text{(positive definiteness of } |\cdot|^2 \text{ over } \mathbb{C}^1)\\
\iff (y_1,\cdots,y_n) &= \mathbf{0} && \text{(definition)}
\end{aligned}$$

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

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

Soln:

$$
\begin{aligned}
\left(\sum_i \lambda_i \left|w_i\right\rangle, \left|v\right\rangle\right) &= \big(\left|v\right\rangle, \sum_i \lambda_i \left|w_i\right\rangle\big)^* && \text{(conjugate symmetry)} \\
&= \left(\sum_i \lambda_i \big(\left|v\right\rangle, \left|w_i\right\rangle\big)\right)^* && \text{(linearlity in the 2nd arg.)} \\
&= \sum_i \lambda_i^* \big(\left|v\right\rangle, \left|w_i\right\rangle\big)^* && \text{(distributivity of complex conjugate)} \\
&= \sum_i \lambda_i^* (\left|w_i\right\rangle, \left|v\right\rangle) && \text{(conjugate symmetry)}
\end{aligned}
$$

2.7) Verify that $\left|w\right\rangle = (1,1)$ and $\left|v\right\rangle = (1,-1)$ are orthogonal. What are the normailized forms of these vectors?
Soln:

$$
\begin{aligned}
(\left|w\right\rangle, \left|v\right\rangle) &= \left\langle w|v\right\rangle && \text{(notation)} \\
&= \begin{bmatrix} 1^* & 1^* \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} && \text{(definition)} \\
&= 1^* \cdot 1 + 1^* \cdot (-1) && \text{(matrix multiplication)} \\
&= 1 \cdot 1 - 1 \cdot 1 = 0 && \text{(arithmetic)} \\
\frac{\left|w\right\rangle}{\|\left|w\right\rangle\|} &= \frac{\left|w\right\rangle}{\sqrt{\left\langle w|w\right\rangle}} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \left|+\right\rangle \\
\frac{\left|v\right\rangle}{\|\left|v\right\rangle\|} &= \frac{\left|v\right\rangle}{\sqrt{\left\langle v|v\right\rangle}} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \left|-\right\rangle
\end{aligned}
$$

2.8) Prove that the Gram-Schmidt procedure produces an orthonormal basis.
Soln: We prove inductively. For $d = 1$, the only requirement is that the procecure normalize $\left|w_d\right\rangle$, which it does by definition for all d. For $d = 2$, suppose $\left|v_1\right\rangle, \cdots, \left|v_{d-1}\right\rangle$ is a orthonormal basis for the subspace spanned by $\left|w_1\right\rangle, \cdots, \left|w_{d-1}\right\rangle$. Being a basis, the subspace spanned by $\left|v_1\right\rangle, \cdots, \left|v_{d-1}\right\rangle$ is the same. Linear independence of $\left|w_1\right\rangle, \cdots, \left|w_d\right\rangle$ implies that $\left|w_d\right\rangle$ is not in this subspace, so $\left|v_1\right\rangle, \cdots, \left|v_{d-1}\right\rangle, \left|w_d\right\rangle$ is easily seen to be linearly independent as well. It remains to be shown that $\left|v_d\right\rangle$ is linearly independent of $\left|v_1\right\rangle, \cdots, \left|v_{d-1}\right\rangle$, and is orthogonal to all such vectors. For independence, note that any dependence relation between $\left|v_1\right\rangle, \cdots, \left|v_d\right\rangle$ immediately induces one between $\left|v_1\right\rangle, \cdots, \left|v_{d-1}\right\rangle, \left|w_d\right\rangle$, violating their independence. For orthogonality, let $1 \le j \le d-1$. We show $\left\langle v_j|v_d\right\rangle = 0$, completing the proof.

$$
\begin{aligned}
\left\langle v_j|v_d\right\rangle &= \left\langle v_j\right| \left(\frac{\left|w_d\right\rangle - \sum_{i=1}^{d-1} \left\langle v_i|w_d\right\rangle \left|v_i\right\rangle}{\left\|\left|w_d\right\rangle - \sum_{i=1}^{d-1} \left\langle v_i|w_d\right\rangle \left|v_i\right\rangle\right\|}\right) && \text{(definition)} \\
&= \frac{\left\langle v_j|w_d\right\rangle - \sum_{i=1}^{d-1} \left\langle v_i|w_d\right\rangle \left\langle v_j|v_i\right\rangle}{\left\|\left|w_d\right\rangle - \sum_{i=1}^{d-1} \left\langle v_i|w_d\right\rangle \left|v_i\right\rangle\right\|} && \text{(linearity in the 2nd argument)} \\
&= \frac{\left\langle v_j|w_d\right\rangle - \sum_{i=1}^{d-1} \left\langle v_i|w_d\right\rangle \delta_{ij}}{\left\|\left|w_d\right\rangle - \sum_{i=1}^{d-1} \left\langle v_i|w_d\right\rangle \left|v_i\right\rangle\right\|} && \text{(orthonormality of } \left|v_1\right\rangle, \cdots, \left|v_{d-1}\right\rangle\text{)} \\
&= \frac{\left\langle v_j|w_d\right\rangle - \left\langle v_j|w_d\right\rangle}{\left\|\left|w_d\right\rangle - \sum_{i=1}^{d-1} \left\langle v_i|w_d\right\rangle \left|v_i\right\rangle\right\|} && \text{(definition of } \delta_{ij}\text{)} \\
&= 0. && \text{(arithmetic)}
\end{aligned}
$$

2.9) (Pauli operators and the outer product) The Pauli matrices 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 notation.

$$\begin{aligned} \sigma_0 = I \;\; &= |0\rangle\langle 0| + |1\rangle\langle 1| \\ \sigma_x = \sigma_1 = X &= |1\rangle\langle 0| + |0\rangle\langle 1| \\ \sigma_y = \sigma_2 = Y &= i\,|1\rangle\langle 0| - i\,|0\rangle\langle 1| \\ \sigma_z = \sigma_3 = Z &= |0\rangle\langle 0| - |1\rangle\langle 1| \end{aligned}$$

2.10) Suppose $|v_i\rangle$ is an orthonormal basis for an 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?
Soln:

$$\begin{aligned} |v_j\rangle\langle v_k| &= I_V\,|v_j\rangle\langle v_k|\,I_V && \text{(multiply by identity)} \\ &= \left(\sum_p |v_p\rangle\langle v_p|\right)|v_j\rangle\langle v_k|\left(\sum_q |v_q\rangle\langle v_q|\right) && \text{(completeness)} \\ &= \sum_{p,q} |v_p\rangle\langle v_p|v_j\rangle\langle v_k|v_q\rangle\langle v_q| && \text{(linearity and outer product definition)} \\ &= \sum_{p,q} \delta_{pj}\delta_{kq}\,|v_p\rangle\langle v_q| && \text{(orthonormality)} \end{aligned}$$

Thus

$$(|v_j\rangle\langle v_k|)_{pq} = \delta_{pj}\delta_{kq} = \begin{cases} 1 & p = j, k = q \\ 0 & o/w \end{cases}.$$

That is, $|v_j\rangle\langle v_k|$ is a square matrix with a 1 in row j, column k, and 0s everywhere else.

(Cauchy-Schwartz inequality) A brief expansion from a mathematician: in equation (2.26), other $|i\rangle$-basis vectors appear, but since $\langle i|v\rangle = \langle v|i\rangle^*$, $a \cdot a^* = \|a\| \geq 0$ for all $a \in \mathbb{C}$, and $\langle\cdot|\cdot\rangle$ is positive definite, all terms but the first constructed in terms of $|w\rangle$ are non-negative and can be removed, leaving the inequality.

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

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \; \det(X - \lambda I) = \det\left(\begin{bmatrix} -\lambda & 1 \\ 1 & -\lambda \end{bmatrix}\right) = \lambda^2 - 1 = 0 \Rightarrow \lambda = \pm 1$$

If $\lambda = 1$,

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} c_2 \\ c_1 \end{bmatrix} = 1\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \Rightarrow c_2 = c_1$$

The eigenspace corresponding to $\lambda = 1$ is the set of vectors $\begin{bmatrix} c \\ c \end{bmatrix}$. The vector $\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = |+\rangle$ is such a unit (normalized) vector. If $\lambda = -1$,

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} c_2 \\ c_1 \end{bmatrix} = -1\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \Rightarrow c_2 = -c_1$$

The eigenspace corresponding to $\lambda = -1$ is the set of vectors $\begin{bmatrix} c \\ -c \end{bmatrix}$. The vector $\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = |-\rangle$ is such a unit (normalized) vector. So, a diagonal representation of X (when expressed in terms of the computational basis) is $\big(|+\rangle\langle+|\big) - \big(|-\rangle\langle-|\big) = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} - \begin{bmatrix} \frac{1}{2} & \frac{-1}{2} \\ \frac{-1}{2} & \frac{1}{2} \end{bmatrix} \big(= X\big)$.

$$Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix},\ \det(Y - \lambda I) = \det\left(\begin{bmatrix} -\lambda & -i \\ i & -\lambda \end{bmatrix}\right) = \lambda^2 - (i)(-i) = \lambda^2 - 1 = 0 \Rightarrow \lambda = \pm 1$$

If $\lambda = 1$,

$$\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} -i \cdot c_2 \\ i \cdot c_1 \end{bmatrix} = 1\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \Rightarrow c_2 = i \cdot c_1$$

The eigenspace corresponding to $\lambda = 1$ is the set of vectors $\begin{bmatrix} c \\ i \cdot c \end{bmatrix}$. The vector $\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ i \end{bmatrix} \equiv |\psi_{y^+}\rangle$ is such a unit (normalized) vector. If $\lambda = -1$,

$$\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} -i \cdot c_2 \\ i \cdot c_1 \end{bmatrix} = -1\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \Rightarrow c_2 = -i \cdot c_1$$

The eigenspace corresponding to $\lambda = -1$ is the set of vectors $\begin{bmatrix} c \\ -i \cdot c \end{bmatrix}$. The vector $\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -i \end{bmatrix} \equiv |\psi_{y^-}\rangle$ is such a unit (normalized) vector. So, a diagonal representation of Y (when expressed in terms of the computational basis) is $\big(|\psi_{y^+}\rangle\langle\psi_{y^+}|\big) - \big(|\psi_{y^-}\rangle\langle\psi_{y^-}|\big) = \begin{bmatrix} \frac{1}{2} & \frac{-i}{2} \\ \frac{i}{2} & \frac{1}{2} \end{bmatrix} - \begin{bmatrix} \frac{1}{2} & \frac{i}{2} \\ \frac{-i}{2} & \frac{1}{2} \end{bmatrix} \big(= Y\big)$.

$$Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},\ \det(Z - \lambda I) = \det\left(\begin{bmatrix} 1-\lambda & 0 \\ 0 & -1-\lambda \end{bmatrix}\right) = (\lambda + 1)(\lambda - 1) = 0 \Rightarrow \lambda = \pm 1$$

If $\lambda = 1$,

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} c_1 \\ -c_2 \end{bmatrix} = 1\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \Rightarrow c_2 = -c_2 \Rightarrow c_2 = 0$$

The eigenspace corresponding to $\lambda = 1$ is the set of vectors $\begin{bmatrix} c \\ 0 \end{bmatrix}$. The vector $\begin{bmatrix} 1 \\ 0 \end{bmatrix} = |0\rangle$ is such a unit (normalized) vector. If $\lambda = -1$,

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} c_1 \\ -c_2 \end{bmatrix} = -1\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \Rightarrow c_1 = -c_1$$

The eigenspace corresponding to $\lambda = -1$ is the set of vectors $\begin{bmatrix} 0 \\ c \end{bmatrix}$. The vector $\begin{bmatrix} 0 \\ 1 \end{bmatrix} = |1\rangle$ is such a unit (normalized) vector. So, the computation basis *is* the eigenbasis for Z, and a diagonal representation of Z is $\big(|0\rangle\langle 0|\big) - \big(|1\rangle\langle 1|\big) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix} \big(= Z\big)$.

2.12) Prove that the matrix $\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ $(\equiv A)$ is not diagonalizable.

Soln:

$$\det(A - \lambda I) = \det\left(\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} - \lambda I\right) = \det\left(\begin{bmatrix} 1-\lambda & 0 \\ 1 & 1-\lambda \end{bmatrix}\right) = (1-\lambda)^2 = 0 \Rightarrow \lambda = 1 \text{ (with multiplicity 2)}$$

All eigenvectors $\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$ satisfy:

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} c_1 \\ c_1 + c_2 \end{bmatrix} = 1 \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \Rightarrow c_1 = 0$$

So, the eigenspace corresponding to eigenvalue 1 of A is 1-dimensional, with a single unit (normalized) vector $\begin{bmatrix} 0 \\ 1 \end{bmatrix} = |1\rangle$. The only possible diagonal representation of A would then be $A = |1\rangle \langle 1|$, but this equality does not hold. We conclude that A has no diagonal representation and is not diagonalizable.

2.13) If $|w\rangle$ and $|v\rangle$ are any two vectors, show that $\big(|w\rangle \langle v|\big)^\dagger = |v\rangle \langle w|$.
Soln: We show that $|v\rangle \langle w|$ has the defining property of $\big(|w\rangle \langle v|\big)^\dagger$, *i.e.* if $|\psi\rangle\,,\ |\phi\rangle$ are arbitrary vectors in V, then $\Big(|\psi\rangle\,, \big(|w\rangle \langle v|\big) |\phi\rangle\Big) = \Big(\big(|v\rangle \langle w|\big) |\psi\rangle\,, |\phi\rangle\Big)$. We do so by expanding $\Big(|\psi\rangle\,, \big(|w\rangle \langle v|\big) |\phi\rangle\Big)^*$ in two different ways.

$$\begin{aligned} \Big(|\psi\rangle\,,\ \big(|w\rangle \langle v|\big) |\phi\rangle\Big)^* &= \Big(\big(|w\rangle \langle v|\big)^\dagger |\psi\rangle\,,\ |\phi\rangle\Big)^* && \text{(defintion of } ^\dagger) \\ &= \Big(|\phi\rangle\,, \big(|w\rangle \langle v|\big)^\dagger |\psi\rangle\Big) && \text{(conjugate symmetry)} \end{aligned}$$

On the other hand,

$$\begin{aligned} \Big(|\psi\rangle\,, \big(|w\rangle \langle v|\big) |\phi\rangle\Big)^* &= \big(\langle \psi|w\rangle\,, \langle v|\phi\rangle\big)^* && \text{(associativity of } \langle\cdot|\,, |\cdot\rangle\,, \langle\cdot|\cdot\rangle\text{, and } |\cdot\rangle\langle\cdot|) \\ &= \big(\langle \phi|v\rangle\,, \langle w|\psi\rangle\big)^* && \text{(conjugate symmetry)} \\ &= \Big(\langle\phi|\,, \big(|v\rangle \langle w|\big) |\psi\rangle\Big). && \text{(notation)} \end{aligned}$$

Thus

$$\Big(|\phi\rangle\,, \big(|w\rangle \langle v|\big)^\dagger |\psi\rangle\Big) = \Big(\langle\phi|\,, \big(|v\rangle \langle w|\big) |\psi\rangle\Big) \text{ for arbitrary vectors } |\psi\rangle\,,\ |\phi\rangle$$

We conclude that $\big(|w\rangle \langle v|\big)^\dagger$ and $|v\rangle \langle w|$ are the same operator, so $\big(|w\rangle \langle v|\big)^\dagger = |v\rangle \langle w|$.

2.14) Anti-linearity of the adjoint: Show that the adjoint operation is anti-linear,

$$\left(\sum_i a_i A_i\right)^\dagger = \sum_i a_i^* A_i^\dagger$$

Soln: It is tempting to assume that $(\sum_i a_i A_i)^\dagger = \sum_i (a_i A_i)^\dagger$, *i.e.* that the † transformation is additive, but we don't yet know this. It will follow from the fact that $A^\dagger \equiv (A^*)^T$ given after problem 2.15, and that both * and ^{T} are linear. This itself is not hard to prove by observing that $(A^*)^T$ has the defining propery of $A^\dagger$, making use of the matrix formulation of the inner product. Without the assumption though, we must be careful to carry around the full sums until additivity (and in-fact full linearity) is known.

$$
\begin{aligned}
\left(\left(\sum_i a_i A_i\right)^\dagger |\phi\rangle\,,\ |\psi\rangle\right) &= \left(|\phi\rangle\,, \left(\sum_i a_i A_i\right) |\psi\rangle\right) && \text{(definition of } \dagger) \\
&= \left(|\phi\rangle\,, \sum_i a_i A_i\, |\psi\rangle\right) && \text{(distributivity of matrix multiplication)} \\
&= \sum_i a_i \big(|\phi\rangle\,,\ A_i\, |\psi\rangle\big) && \text{(linearity in the second argument)} \\
&= \sum_i a_i \left(A_i^\dagger\, |\phi\rangle\,,\ |\psi\rangle\right) && \text{(definition of } \dagger) \\
&= \sum_i \left(a_i^* A_i^\dagger\, |\phi\rangle\,,\ |\psi\rangle\right) && \text{(conjugate-linearity in the first argument)} \\
&= \left(\left(\sum_i a_i^* A_i^\dagger\right) |\phi\rangle\,,\ |\psi\rangle\right) && \text{(distributivity of matrix multiplication)} \\
\text{therefore } \left(\sum_i a_i A_i\right)^\dagger &= \sum_i a_i^* A_i^\dagger && \square
\end{aligned}
$$

2.15) Show that $\left(A^\dagger\right)^\dagger = A$.
Soln: We show that A has the defining property of the adjoint of $A^\dagger$.

$$
\begin{aligned}
\left(\left(A^\dagger\right)^\dagger |\psi\rangle\,,\ |\phi\rangle\right) &= \left(|\psi\rangle\,,\ A^\dagger\, |\phi\rangle\right) && \left(\text{definition of } \left(A^\dagger\right)^\dagger\right) \\
&= \left(A^\dagger\, |\phi\rangle\,,\ |\psi\rangle\right)^* && \text{(conjugate symmetry)} \\
&= (|\phi\rangle\,,\ A\, |\psi\rangle)^* && \text{(definition of } A^\dagger) \\
&= (A\, |\psi\rangle\,,\ |\phi\rangle) && \text{(conjugate symmetry)} \\
\text{therefore } \left(A^\dagger\right)^\dagger &= A && \square
\end{aligned}
$$

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

$$
\begin{aligned}
P &= \sum_i |i\rangle \langle i|\,. && \text{(definition)} \\
P^2 &= \left(\sum_i |i\rangle \langle i|\right) \left(\sum_j |j\rangle \langle j|\right) && \text{(square definition)} \\
&= \sum_{i,j} |i\rangle \langle i|j\rangle \langle j| && \text{(distributivity)} \\
&= \sum_{i,j} |i\rangle \langle j|\, \delta_{ij} && \big(\text{evaluate } \langle i|j\rangle\big) \\
&= \sum_i |i\rangle \langle i| && \text{(evaluate sum over } j) \\
&= P && \text{(definition)}
\end{aligned}
$$

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

Proof. ($\Rightarrow$) Suppose A is Hermitian. Then $A = A^\dagger$. Let λ be an eigenvalue of A with unit-eigenvector $|\lambda\rangle$. We have:

$$\begin{aligned} A|\lambda\rangle &= \lambda|\lambda\rangle && \text{(definition)} \\ \langle\lambda|A|\lambda\rangle &= \lambda\langle\lambda|\lambda\rangle && \big(\text{multiply by } \langle\lambda|\big) \\ &= \lambda. && (\lambda \text{ is a unit-vector}) \end{aligned}$$

Now:

$$\begin{aligned} \lambda^* &= \langle\lambda|A|\lambda\rangle^* && \text{(conjugate)} \\ &= \big(|\lambda\rangle, A|\lambda\rangle\big)^* && \text{(change notation)} \\ &= \big(A|\lambda\rangle, \lambda\big) && \text{(conjugate symmetry)} \\ &= \big(A^\dagger|\lambda\rangle, |\lambda\rangle\big) && \text{(hypothesis)} \\ &= \big(|\lambda\rangle, A|\lambda\rangle\big) && (\text{definiton of } ^\dagger) \\ &= \lambda && \text{(from above)} \end{aligned}$$

So the eigenvalue λ is real, since only real numbers are equal to their conjugates.
($\Leftarrow$) To prove the converse we make use of the spectral deomposition theorem. It's proof does *not* use the fact that a normal matrix is Hermition if and only if it's eigenvalues are real, so using it here does not make this proof circular. Suppose the eigenvalues of A are real. From the spectral decomposition theorem there exists a set of eigenvalues λ_i and a corresponding orthonormal basis $|\lambda_i\rangle$ such that

$$A = \sum_i \lambda_i |\lambda_i\rangle\langle\lambda_i| \qquad \text{(spectral decomposition)}$$

From this we have:

$$\begin{aligned} A^\dagger &= \left(\sum_i \lambda_i |\lambda_i\rangle\langle\lambda_i|\right)^\dagger && \text{(apply adjoint)} \\ &= \sum_i \lambda_i^* \big(|\lambda_i\rangle\langle\lambda_i|\big)^\dagger && \text{(anti-linearity)} \\ &= \sum_i \lambda_i |\lambda_i\rangle\langle\lambda_i| && (\lambda_i \text{ real, projectors are Hermitian}) \\ &= A && \text{(from spectral decomposition)} \end{aligned}$$

Thus A is Hermitian. □

2.18) Show that all eigenvalues of a unitary matrix have modulus 1, that is, can be written in the form $e^{i\theta}$ for some real θ.
Soln: Suppose λ is an eigenvalue with corresponding unit-eigenvector $|v\rangle$

$$\begin{aligned} 1 &= \langle v|v\rangle && (|v\rangle \text{ is a unit vector}) \\ &= \langle v|I|v\rangle && \text{(multiply by identity)} \\ &= \langle v|U^\dagger U|v\rangle && (U \text{ is unitary}) \\ &= \big(\langle v|U^\dagger\big)\big(U|v\rangle\big) && \text{(associativity of matrix multiplication)} \\ &= \big(U|v\rangle\big)^\dagger\big(U|v\rangle\big) && (\text{arithmetic properties of } ^\dagger) \\ &= \big(\lambda|v\rangle\big)^\dagger\big(\lambda|v\rangle\big) && (|v\rangle \text{ is an eigenvector}) \\ &= \lambda^*\lambda\langle v|v\rangle && (\text{re-apply } ^\dagger \text{ and simplify}) \\ &= \|\lambda\|^2 && (\text{definiton of } \|\cdot\|, |v\rangle \text{ is a unit-vector}) \end{aligned}$$

Now $\|\lambda\| = 1$, and all complex numbers with modulus 1 are located on the unit-circle in $\mathbb{C}$ and can be expressed as $e^{i\theta}$ for some real $\theta\big(\in [0, 2\pi)\big)$

2.19) Show that the Pauli matrices are Hermitian and unitary
Soln: It is easy to see that the Pauli matrices are Hermitian (self-adjoint) given the conjugate-transpose formula. We still must show that their squares are the identity:

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

$$Y^\dagger Y = Y^2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} = \begin{bmatrix} -i^2 & 0 \\ 0 & -i^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

$$Z^\dagger Z = Z^2 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & (-1)^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

2.20) 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''.
Soln:

$$U \equiv \sum_i |w_i\rangle\langle v_i|, \quad U^\dagger = \sum_j |v_j\rangle\langle w_j| \qquad \text{(construct a unitary operator and its adjoint)}$$

$$\begin{aligned}
A^{'}_{ij} &= \langle v_i|A|v_j\rangle && \text{(given)} \\
&= \langle v_i|UU^\dagger AUU^\dagger|v_j\rangle && (U \text{ is unitary; } UU^\dagger = I) \\
&= \sum_{p,q,r,s} \langle v_i|w_p\rangle\langle v_p|v_q\rangle\langle w_q|A|w_r\rangle\langle v_r|v_s\rangle\langle w_s|v_j\rangle && \text{(expand } U, U^\dagger\text{, apply linearity)} \\
&= \sum_{p,q,r,s} \langle v_i|w_p\rangle\,\delta_{pq}A^{''}_{qr}\delta_{rs}\,\langle w_s|v_j\rangle && (|v_i\rangle \text{ is orthonormal, apply given for } A'') \\
&= \sum_{p,r} \langle v_i|w_p\rangle\langle w_r|v_j\rangle\,A^{''}_{pr} && \text{(collect non-zero terms and re-index)}
\end{aligned}$$

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.

Theorem 2.1 **(Spectral decomposition)** A *Hermitian* operator M on a vector space V is diagonal with respect to some orthonormal basis for V.

Proof. We induct on the dimension of V, as in the boxed proof. Let λ be an eigenvalue of M, P be the projector onto the λ eigenspace, and Q the projector onto the orthogonal complement.

$$\begin{aligned}
M &= IMI && \text{(trivial)} \\
&= (P+Q)M(P+Q) && \text{(definition of } Q) \\
&= PMP + QMP + PMQ + QMQ && \text{(expand)}
\end{aligned}$$

Now $PMP = \lambda P$ and $QMP = 0$ as before. To show that $PMQ = 0$ is as easy as substituting $M^\dagger$:

$$\begin{aligned} PMQ &= PM^\dagger Q && (M \text{ is Heritian}) \\ &= P(M^{*T}Q) && (\dagger =^{*T}) \\ &= (QM^*P)^T && (\text{properties of } ^T) \\ &= ((QMP)^*)^T && (\text{properties of } ^*) \\ &= 0 && (QMP = 0) \end{aligned}$$

Thus $M = PMP + QMQ$. Next, we prove QMQ is normal.

$$\begin{aligned} QMQ(QMQ)^\dagger &= QMQQ^\dagger M^\dagger Q^\dagger && (\text{properties of } ^\dagger\text{, and symmetry}) \\ &= QMQQM^\dagger Q && (\text{projectors are Hermitian}) \\ &= QM^\dagger QQMQ && (M = M^\dagger) \\ &= Q^\dagger M^\dagger Q^\dagger QMQ && (\text{projectors are Hermitian}) \\ &= (QMQ)^\dagger QMQ && (\text{properties of } ^\dagger\text{, and symmetry}) \end{aligned}$$

Therefore QMQ is normal. By induction, QMQ is diagonal. The rest follows Box 2.2 identically. □

2.22) Prove that two eigenvectors of a Hermitian operator with different eigenvalues are necessarily orthogonal

Soln: Suppose A is a Hermitian operator and $|v_1\rangle, |v_2\rangle$ are eigenvectors of A with eigenvalues λ_1, λ_2, with $\lambda_1 \neq \lambda_2$. Then

$$\langle v_1|A|v_2\rangle = \lambda_2 \langle v_1|v_2\rangle . \qquad (\text{definition of } v_1\text{, linearity of } \langle\cdot|\cdot\rangle)$$

On the other hand,

$$\begin{aligned} \langle v_1|A|v_2\rangle &= \langle v_1|A^\dagger|v_2\rangle && (A \text{ is Hermitian}) \\ &= \langle v_2|A|v_1\rangle^* && (\text{properties of } ^\dagger\text{, Hermitian} \Rightarrow \text{self-transpose}) \\ &= \lambda_1 \langle v_2|v_1\rangle^* && (\text{definition of } v_1\text{, linearity of } \langle\cdot|\cdot\rangle) \\ &= \lambda_1 \langle v_1|v_2\rangle && (\text{properties of } ^*) \end{aligned}$$

Thus

$$(\lambda_1 - \lambda_2)\langle v_1|v_2\rangle = 0.$$

Since $\lambda_1 - \lambda_2 \neq 0$, we must have $\langle v_1|v_2\rangle = 0$, so v_1 and v_2 are orthogonal.

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

Soln: Suppose P is projector and $|v\rangle$ is an eigenvector of P with eigenvalue λ. By exercise 2.16, $P^2 = P$. We have $P|v\rangle = \lambda|v\rangle$ by hypothesis. Alternatively,

$$\begin{aligned} P|v\rangle &= P^2|\lambda\rangle && (\text{exercise 2.16}) \\ &= \lambda P|v\rangle && (\text{hypothesis, linearity}) \\ &= \lambda^2|v\rangle && (\text{hypothesis}) \end{aligned}$$

Therefore

$$\begin{aligned} \lambda &= \lambda^2 \\ \lambda^2 - \lambda &= 0 \\ \lambda(\lambda - 1) &= 0 \\ \lambda &= 0 \text{ or } 1. \end{aligned}$$

2.24) (Hermiticity of positive operators) Show that a positive operator is necessarily Hermitian.

Soln: Let A be a positive operator, that is, suppose $\langle v|A|v\rangle$ is real and ≥ 0 for all $|v\rangle$. Define $B = \frac{A+A^\dagger}{2}$ and $C = \frac{A-A^\dagger}{2i}$. Simple complex arithmetic will show that $A = B+iC$. B is clearly Hermitian by commutativity of operator addition. C is also Hermitian by linearity of the adjoint, noting that $\left(\frac{1}{2i}\right)^* = -\frac{1}{2i}$. There are two ways to proceed: one heuristic, and one mathematically rigorous. We'll start with a heuristic outline of the proof, then provide some mathematically rigorous detail after the fact.

Let v be a vector and note that it can be proven (below) that $\langle v|B|v\rangle$ and $\langle v|C|v\rangle$ are both real numbers. Now $\langle v|A|v\rangle = \langle v|B+iC|v\rangle = \langle v|B|v\rangle + i\langle v|C|v\rangle$ by the construction of B and C, and the linearity of $\langle\cdot|\cdot|\cdot\rangle$. By hypothesis, $\langle v|A|v\rangle$ is a non-negative real number, so $\langle v|C|v\rangle = 0$, since both $\langle v|B|v\rangle$ and $\langle v|C|v\rangle$ are real. This will be enough to show that $C = 0$ which yields $A = A^\dagger$ by the definition of C, that is, A is Hermiitian.

Now, to complete the proof, we need to rigorously show that both $\langle v|B|v\rangle$ and $\langle v|C|v\rangle$ are real numbers, and that if $\langle v|C|v\rangle = 0$ for all $|v\rangle$, then $C = 0$. Let W be Hermitian, thus normal, and note that by exercise 2.17, W has real eigenvalues, say ω_i. By the spectral decomposition theorem there is an orthnomal basis, say $|w_i\rangle$, such that $W = \sum_i \omega_i |w_i\rangle\langle w_i|$. Let $|v\rangle$ be an arbitrary vector, expressed in the orthonormal $|w_i\rangle$-basis as $\sum_i \alpha_i |w_i\rangle$.

$$\begin{aligned}
\langle v|W|v\rangle &= \left\langle \sum_j \alpha_j |w_j\rangle \middle| W \middle| \sum_i \alpha_i |w_i\rangle \right\rangle && \text{(by construction)}\\
&= \sum_i \sum_j \alpha_i \alpha_j^* \langle w_j|W|w_i\rangle && \text{((conjugate) linearity of } \langle\cdot|\cdot|\cdot\rangle)\\
&= \sum_i \sum_j \alpha_i \alpha_j^* \left\langle w_j \middle| \sum_k \omega_k |w_k\rangle\langle w_k| \middle| w_i \right\rangle && \text{(spectral decomposition)}\\
&= \sum_i \sum_j \sum_k \alpha_i \alpha_j^* \omega_k \langle w_j|w_k\rangle\langle w_k|w_i\rangle && \text{(linearity of } \langle\cdot|\cdot|\cdot\rangle)\\
&= \sum_i \sum_j \sum_k \alpha_i \alpha_j^* \omega_k \delta_{jk}\delta_{ki} && \text{(orthonomality of the } |w_i\rangle \text{ basis)}\\
&= \sum_k \alpha_k \alpha_k^* \omega_k && \text{(collecting non-zero terms)}\\
&= \sum_k \|\alpha_k\|^2 \omega_k && \text{(definition of } \|\cdot\|)
\end{aligned}$$

The ω_k are real numbers by exercise 2.17, and the $\|\alpha_k\|^2$ are real by the definition of $\|\cdot\|$, so $\langle v|W|v|\rangle$ is a sum of real number, and hence also real itself. Applying this to B and C above completes the first missing part. To finally complete the proof we'll require Theorem 2.0.1 below, more generally applicable to linear operators on complex vector spaces, without the assumption of Hermiticity. The proof follows an MIT 8.05 Quantum Physics II lecture note by Prof. Barton Zwiebach (`https://ocw.mit.edu/courses/physics/8-05-quantum-physics-ii-fall-2013/lecture-notes/MIT8_05F13_Chap_03.pdf`)

Proposition. 2.0.1. *Let T be a linear operator on a complex vector space V. If $\langle u|T|v\rangle = 0$ for all $|u\rangle, |v\rangle \in V$, then $T = 0$.*

Proof. Let $|u\rangle = T|v\rangle$. Then $\left\langle T|v\rangle \middle| T|v\right\rangle = \left\langle T|v\rangle \middle| T|v\rangle\right\rangle = \|T|v\rangle\|^2 = 0$, which implies $T|v\rangle = 0$ for all v by property 3 of the inner product (page 65). T is identically 0, so is the zero operator, i.e. $T = 0$. □

Theorem. 2.0.1. *Let T be a linear operator on a complex vector space V. If $\langle v|T|v\rangle = 0$ for all $|v\rangle \in V$, then $T = 0$.*

Proof. Note that the weakened hypothesis doesn't directly apply if $|u\rangle \neq |v\rangle$. We show that the "off-diagonal", distinct vector hypothesis of Proposition 2.0.1 can be derived from the weakened "diagonal" hypothesis' of this theorem, that is, if $\langle v|T|v\rangle = 0$ for all $|v\rangle$, then $\langle u|T|v\rangle = 0$ for all $|u\rangle, |v\rangle$. Then apply proposition 2.0.1

Suppose $|u\rangle, |v\rangle \in V$. Then note that by "foiling" the $\langle\cdot|\cdot|\cdot\rangle$'s, we can show a "polarization" identity, expressing $\langle u|T|v\rangle$ as follows

$$\begin{aligned}\frac{1}{4}\Big(\big\langle u+v|T|u+v\big\rangle - \big\langle u-v|T|u-v\big\rangle + \frac{1}{i}\big\langle u+iv|T|u+iv\big\rangle - \frac{1}{i}\big\langle u-iv|T|u-iv\big\rangle\Big) &= \\ \frac{1}{4}\Big(\big(\langle u|T|u\rangle + \langle u|T|v\rangle + \langle v|T|u\rangle + \langle v|T|v\rangle\big) - \big(\langle u|T|u\rangle - \langle u|T|v\rangle - \langle v|T|u\rangle + \langle v|T|v\rangle\big) + \dots \\ \frac{1}{i}\big(\langle u|T|u\rangle + i\langle u|T|v\rangle - i\langle v|T|u\rangle + \langle v|T|v\rangle\big) - \frac{1}{i}\big(\langle u|T|u\rangle - i\langle u|T|v\rangle + i\langle v|T|u\rangle + \langle v|T|v\rangle\big)\Big) &= \\ \frac{1}{4}\big(0\langle u|T|u\rangle + 4\langle u|T|v\rangle + 0\langle v|T|u\rangle + 0\langle v|T|v\rangle\big) &= \\ \langle u|T|v\rangle\end{aligned}$$

Applying the diagonal hypothesis to $|u+v\rangle, |u-v\rangle, |u+iv\rangle$, and $|u-iv\rangle$ in the first expression above gives that $\langle u|T|v\rangle = 0$ for all $|u\rangle, |v\rangle$, hence by Proposition 2.0.1, $T = 0$. □

Applying Theorem 2.0.1 to C from above finally completes the proof of the Hermiticity of positive operators.

2.25) Show that for any operator A, $A^\dagger A$ is positive.
Soln: Its enough to show that $\langle v|A^\dagger A|v\rangle \geq 0$ for all v, but note that $\langle v|A^\dagger A|v\rangle = \|Av\|^2$, which is non-negative, so $A^\dagger A$ is a positive operator.

2.26) Let $|\psi\rangle = (|0\rangle + |1\rangle)/\sqrt{2}(= |+\rangle)$. 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.
Soln:

$$\begin{aligned}|\psi\rangle^{\otimes 2} &= \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle \\ &= \frac{1}{2}(|00\rangle + |01\rangle + |10\rangle + |11\rangle) \\ &= \frac{1}{2}\begin{bmatrix}1\\1\\1\\1\end{bmatrix}\end{aligned}$$

$$\begin{aligned}|\psi\rangle^{\otimes 3} &= \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle \otimes \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle \\ &= \frac{1}{2\sqrt{2}}(|000\rangle + |001\rangle + |010\rangle + |011\rangle + |100\rangle + |101\rangle + |110\rangle + |111\rangle) \\ &= \frac{1}{2\sqrt{2}}\begin{bmatrix}1\\1\\1\\1\\1\\1\\1\\1\end{bmatrix}\end{aligned}$$

2.27) Calculate the matrix representations 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?

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

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

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

In general, the tensor product is not commutative.

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

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

Soln: Let A be $n_1 \times m_1$ and B be $n_2 \times m_2$, so that $A \otimes B$ is $n \times m$, where $n = n_1 n_2$ and $m = m_1 m_2$. The entries in $A \otimes B$ are products of a single entry in A and a single entry in B. Specifically, if $i = i_1 n_1 + i_2$ and $j = j_1 m_1 + j_2$, with $0 \le i_2 < n_1$ and $0 \le j_2 < m_1$, then $(A \otimes B)_{ij} = A_{i_1 j_1} B_{i_2 j_2}$.

$$\begin{aligned}
(A \otimes B)^* &= [A_{i_1 j_1} B_{i_2 j_2}]^* && \text{(from above)} \\
&= \left[A^*_{i_1 j_1} B^*_{i_2 j_2}\right] && \text{(piecewise conjugation)} \\
&= A^* \otimes B^* && \text{(consistent indexing)}
\end{aligned}$$

To see that $(A \otimes B)^T = A^T \otimes B^T$, note that $A^T \otimes B^T$ is $m \times n$, and $(A^T \otimes B^T)_{kl}$ is the product of a single entry in A^T and a single entry in B^T. Specifically, if $k = k_1 m_1 + k_2$ and $\ell = \ell_1 n_1 + \ell_2$, with $0 \le k_2 < m_2$ and $0 \le \ell_2 < n_2$, then $(A^T \otimes B^T)_{k\ell} = (A^T)_{k_1 \ell_1} (B^T)_{k_2 \ell_2} = A_{\ell_1 k_1} B_{\ell_2 k_2}$. Now, the hypotheses on k match the hypotheses on j above, and similarly for ℓ and i. This implies $(A^T \otimes B^T)_{k\ell} = (A \otimes B)_{\ell k} = (A \otimes B)^T_{k\ell}$. All entries in $A^T \otimes B^T$ and $(A \otimes B)^T$ are equal, so $(A \otimes B)^T = A^T \otimes B^T$.

Distributivity of † follows by applying distributivity of * and ^{T} in turn:

$$\begin{aligned}
(A \otimes B)^\dagger &= ((A \otimes B)^*)^T && \text{(definition of } \dagger\text{)} \\
&= (A^* \otimes B^*)^T && \text{(distribute } ^*\text{)} \\
&= (A^*)^T \otimes (B^*)^T && \text{(distribute } ^T\text{)} \\
&= A^\dagger \otimes B^\dagger. && \text{(definition of } ^\dagger\text{)}
\end{aligned}$$

2.29) Show that the tensor product of two unitary operators is unitary
Soln: Suppose U_1 and U_2 are unitary operators. To avoid implicit assumptions on mutliplication of tensor products, let $|v\rangle$ and $|w\rangle$ be vectors in the spaces on which U_1 and U_2 operate. Then:

$$\begin{aligned}
(U_1 \otimes U_2)(U_1 \otimes U_2)^\dagger(|v\rangle \otimes |w\rangle) &= (U_1 \otimes U_2)(U_1^\dagger \otimes U_2^\dagger)(|v\rangle \otimes |w\rangle) && \text{(distributivity of } ^\dagger\text{)} \\
&= (U_1 \otimes U_2)(U_1^\dagger |v\rangle \otimes U_2^\dagger |w\rangle) && \text{(definition of tensor product of operators)} \\
&= U_1 U_1^\dagger |v\rangle \otimes U_2 U_2^\dagger |w\rangle && \text{(definition of tensor product of operators)} \\
&= I|v\rangle \otimes I|w\rangle && (U_1 \text{ and } U_2 \text{ are unitary}) \\
&= (I \otimes I)(|v\rangle \otimes |w\rangle) && \text{(definition of tensor product of operators)} \\
&= I(|v\rangle \otimes |w\rangle) && (I \otimes I = I \text{ by construction})
\end{aligned}$$

So, $(U_1 \otimes U_2)(U_1 \otimes U_2)^\dagger = I$. Similarly, $(U_1 \otimes U_2)^\dagger (U_1 \otimes U_2) = I \otimes I = I$, so $U_1 \otimes U_2$ is unitary.

2.30) Show that the tensor product of two Hermitian operators is Hermitian.
Soln: Suppose A and B are Hermitian operators. Then by distributivity of † and Hermiticity:

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

Thus $A \otimes B$ is Hermitian.

2.31) Show that the tensor product of two positive operators is positive.
Soln: Suppose A and B are positive operators. Then

$$\begin{aligned}
\Big(|\psi\rangle \otimes |\phi\rangle, (A \otimes B)(|\psi\rangle \otimes |\phi\rangle)\Big) &= \Big(|\psi\rangle \otimes |\phi\rangle, A|\psi\rangle \otimes B|\phi\rangle\Big) && \text{(definition of } A \otimes B\text{)} \\
&= (|\psi\rangle \otimes |\phi\rangle)^\dagger (A|\psi\rangle \otimes B|\phi\rangle) && \text{(definition of inner-product)} \\
&= (\langle\psi| \otimes \langle\phi|)(A|\psi\rangle \otimes B|\phi\rangle) && \text{(distributivity of } ^\dagger\text{)} \\
&= \langle\psi|A|\psi\rangle \langle\phi|B|\phi\rangle .
\end{aligned}$$

Since A and B are positive operators, $\langle\psi|A|\psi\rangle \geq 0$ and $\langle\phi|B|\phi\rangle \geq 0$ for all $|\psi\rangle, |\phi\rangle$, so $\langle\psi|A|\psi\rangle \langle\phi|B|\phi\rangle \geq 0$, from which we conclude that $A \otimes B$ is positive.

2.32) Show that the tensor product of two projectors is a projector.
Soln: Suppose P_1 and P_2 are projectors. It is tempting to think that by applying exercise 2.16, which yields

$$\begin{aligned}(P_1 \otimes P_2)^2 &= P_1^2 \otimes P_2^2 && \text{(tensor product is multiplicative)}\\ &= P_1 \otimes P_2, && \text{(exercise 2.16)}\end{aligned}$$

exercise 2.16 would then imply that $P_1 \otimes P_2$ is also projector. However, this implication is the converse of exercise 2.16, which we have not proven. Instead, we need to prove that if $P_1 = \sum_{i=0}^{k} |v_i\rangle \langle v_i|$ and $P_2 = \sum_{j=0}^{\ell} |w_j\rangle \langle w_j|$ where $|v_i\rangle_{i=0}^{k}$ is a subset of an orthonormal basis $|v_i\rangle_i = 0^\kappa$, and $|w_j\rangle_{j=0}^{\ell}$ is a subset of an orthonormal basis $|w_j\rangle_{j=0}^{\chi}$, then $P_1 \otimes P_2 = \sum_{p=0}^{q} |r_p\rangle \langle r_p|$, where $|r_p\rangle_{p=0}^{q}$ is a subset of an orthonormal basis $|r_p\rangle_{p=0}^{\rho}$. ... to be continued.

2.33)

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

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

2.34)

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

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

Eigenvalues of A are $\lambda = 1,\ 7$. Corresponding eigenvectors are $|\lambda = 1\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, $|\lambda = 7\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

Thus

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

$$\begin{aligned}\sqrt{A} &= |\lambda = 1\rangle\langle\lambda = 1| + \sqrt{7}\,|\lambda = 7\rangle\langle\lambda = 7|\\ &= \frac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} + \frac{\sqrt{7}}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\\ &= \frac{1}{2} \begin{bmatrix} 1+\sqrt{7} & -1+\sqrt{7} \\ -1+\sqrt{7} & 1+\sqrt{7} \end{bmatrix}\end{aligned}$$

$$
\begin{aligned}
\log(A) &= \log(1) \ket{\lambda = 1}\bra{\lambda = 1} + \log(7) \ket{\lambda = 7}\bra{\lambda = 7} \\
&= \frac{\log(7)}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}
\end{aligned}
$$

2.35)

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

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

Eigenvalues are $\lambda = \pm 1$. Let $\ket{\lambda_{\pm 1}}$ be eigenvectors with eigenvalues ± 1.

Since $\vec{v} \cdot \vec{\sigma}$ is Hermitian, $\vec{v} \cdot \vec{\sigma}$ is diagonalizable. Then

$$
\vec{v} \cdot \vec{\sigma} = \ket{\lambda_1}\bra{\lambda_1} - \ket{\lambda_{-1}}\bra{\lambda_{-1}}
$$

Thus

$$
\begin{aligned}
\exp\left(i\theta \vec{v} \cdot \vec{\sigma}\right) &= e^{i\theta} \ket{\lambda_1}\bra{\lambda_1} + e^{-i\theta} \ket{\lambda_{-1}}\bra{\lambda_{-1}} \\
&= (\cos\theta + i\sin\theta) \ket{\lambda_1}\bra{\lambda_1} + (\cos\theta - i\sin\theta) \ket{\lambda_{-1}}\bra{\lambda_{-1}} \\
&= \cos\theta(\ket{\lambda_1}\bra{\lambda_1} + \ket{\lambda_{-1}}\bra{\lambda_{-1}}) + i\sin\theta(\ket{\lambda_1}\bra{\lambda_1} - \ket{\lambda_{-1}}\bra{\lambda_{-1}}) \\
&= \cos(\theta) I + i\sin(\theta) \vec{v} \cdot \vec{\sigma}.
\end{aligned}
$$

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

$$
\ket{\lambda_1}\bra{\lambda_1} + \ket{\lambda_{-1}}\bra{\lambda_{-1}} = I.
$$

2.36)

$$
\begin{aligned}
\mathrm{Tr}(\sigma_1) &= \mathrm{Tr}\left(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\right) = 0 \\
\mathrm{Tr}(\sigma_2) &= \mathrm{Tr}\left(\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}\right) = 0 \\
\mathrm{Tr}(\sigma_3) &= \mathrm{Tr}\left(\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\right) = 1 - 1 = 0
\end{aligned}
$$

2.37)

$$\begin{aligned}
\mathrm{Tr}(AB) &= \sum_i \langle i|AB|i\rangle \\
&= \sum_i \langle i|AIB|i\rangle \\
&= \sum_{i,j} \langle i|A|j\rangle \langle j|B|i\rangle \\
&= \sum_{i,j} \langle j|B|i\rangle \langle i|A|j\rangle \\
&= \sum_j \langle j|BA|j\rangle \\
&= \mathrm{Tr}(BA)
\end{aligned}$$

2.38)

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

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

2.39)

(1) $(A, B) \equiv \mathrm{Tr}(A^\dagger B)$.

(i)

$$\begin{aligned}
\left(A, \sum_i \lambda_i B_i\right) &= \mathrm{Tr}\left[A^\dagger\left(\sum_i \lambda_i B_i\right)\right] \\
&= \mathrm{Tr}(A^\dagger \lambda_1 B_1) + \cdots + \mathrm{Tr}(A^\dagger \lambda_n B_n) \quad (\because \text{Execise 2.38}) \\
&= \lambda_1 \mathrm{Tr}(A^\dagger B_1) + \cdots + \lambda_n \mathrm{Tr}(A^\dagger B_n) \\
&= \sum_i \lambda_i \mathrm{Tr}(A^\dagger B_i)
\end{aligned}$$

(ii)

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

(iii)

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

Since $A^\dagger A$ is positive, $\langle i|A^\dagger A|i\rangle \geq 0$ for all $|i\rangle$.

Let a_i be i-th column of A. If $\langle i|A^\dagger A|i\rangle = 0$, then

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

Therefore $(A,A) = 0$ iff $A = \mathbf{0}$.

(2)
(3)

2.40)

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

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

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

2.41)

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

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

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

$$
\begin{aligned}
\sigma_0^2 &= I^2 = I \\
\sigma_1^2 &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}^2 = I \\
\sigma_2^2 &= \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}^2 = I \\
\sigma_3^2 &= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}^2 = I
\end{aligned}
$$

2.42)

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

2.43)

From eq (2.75) and eq (2.76), $\{\sigma_j, \sigma_k\} = 2\delta_{jk} I$. From eq (2.77),

$$
\begin{aligned}
\sigma_j \sigma_k &= \frac{[\sigma_j, \sigma_k] + \{\sigma_j, \sigma_k\}}{2} \\
&= \frac{2i \sum_{l=1}^{3} \epsilon_{jkl} \sigma_l + 2\delta_{jk} I}{2} \\
&= \delta_{jk} I + i \sum_{l=1}^{3} \epsilon_{jkl} \sigma_l
\end{aligned}
$$

2.44)

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

$$\begin{aligned} A^{-1}AB &= 0 \\ IB &= 0 \\ B &= 0. \end{aligned}$$

2.45)

$$\begin{aligned} [A, B]^\dagger &= (AB - BA)^\dagger \\ &= B^\dagger A^\dagger - A^\dagger B^\dagger \\ &= \left[B^\dagger, A^\dagger\right] \end{aligned}$$

2.46)

$$\begin{aligned} [A, B] &= AB - BA \\ &= -(BA - AB) \\ &= -[B, A] \end{aligned}$$

2.47)

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

2.48)

(Positive)

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

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

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

(Unitary)

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 polar decomposition of U is $U = U$.

(Hermitian)

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, $H = \sum_i \lambda_i |i\rangle\langle i|$, $\lambda_i \in \mathbb{R}$.

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

2.49)

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)

Define $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$. $A^\dagger A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$.

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 associated eigenvectors are $|\lambda_\pm\rangle = \frac{1}{\sqrt{10\mp 2\sqrt{5}}} \begin{bmatrix} 2 \\ -1 \pm \sqrt{5} \end{bmatrix}$.

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

$$\begin{aligned} J = \sqrt{A^\dagger A} &= \sqrt{\lambda_+} |\lambda_+\rangle\langle\lambda_+| + \sqrt{\lambda_-} |\lambda_-\rangle\langle\lambda_-| \\ &= \sqrt{\frac{3+\sqrt{5}}{2}} \cdot \frac{5-\sqrt{5}}{40} \begin{bmatrix} 4 & 2\sqrt{5}-2 \\ 2\sqrt{5}-2 & 6-2\sqrt{5} \end{bmatrix} + \sqrt{\frac{3-\sqrt{5}}{2}} \cdot \frac{5+\sqrt{5}}{40} \begin{bmatrix} 4 & -2\sqrt{5}-2 \\ -2\sqrt{5}-2 & 6+2\sqrt{5} \end{bmatrix} \end{aligned}$$

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

$$U = AJ^{-1}$$

I'm tired.

2.51)

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

2.52)

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

Thus

$$H^2 = I.$$

2.53)

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

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

2.54)

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

$$\begin{aligned}\exp(A)\exp(B) &= \left(\sum_i \exp(a_i)\,|i\rangle\langle i|\right)\left(\sum_i \exp(b_i)\,|i\rangle\langle i|\right)\\ &= \sum_{i,j}\exp(a_i+b_j)\,|i\rangle\,\langle i|j\rangle\,\langle j|\\ &= \sum_{i,j}\exp(a_i+b_j)\,|i\rangle\langle j|\,\delta_{i,j}\\ &= \sum_i \exp(a_i+b_i)\,|i\rangle\langle i|\\ &= \exp(A+B)\end{aligned}$$

2.55)

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

$$\begin{aligned}U(t_2-t_1)U^\dagger(t_2-t_1) &= \exp\left(-\frac{iH(t_2-t_1)}{\hbar}\right)\exp\left(\frac{iH(t_2-t_1)}{\hbar}\right)\\ &= \sum_{E,E'}\left(\exp\left(-\frac{iE(t_2-t_1)}{\hbar}\right)|E\rangle\langle E|\right)\left(\exp\left(-\frac{iE'(t_2-t_1)}{\hbar}\right)|E'\rangle\langle E'|\right)\\ &= \sum_{E,E'}\left(\exp\left(-\frac{i(E-E')(t_2-t_1)}{\hbar}\right)|E\rangle\langle E'|\,\delta_{E,E'}\right)\\ &= \sum_E \exp(0)\,|E\rangle\langle E|\\ &= \sum_E |E\rangle\langle E|\\ &= I\end{aligned}$$

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

2.56)

$U = \sum_i \lambda_i |\lambda_i\rangle\langle\lambda_i| \quad (|\lambda_i| = 1)$.

$$\log(U) = \sum_j \log(\lambda_j) |\lambda_j\rangle\langle\lambda_j| = \sum_j i\theta_j |\lambda_j\rangle\langle\lambda_j| \text{ where } \theta_j = \arg(\lambda_j)$$
$$K = -i\log(U) = \sum_j \theta_j |\lambda_j\rangle\langle\lambda_j| .$$

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

2.57)

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

2.58)

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

2.59)

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

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

$$\begin{aligned}|\lambda_1\rangle\langle\lambda_1| &= \frac{1+v_3}{2}\begin{bmatrix}1\\ \frac{1-v_3}{v_1-iv_2}\end{bmatrix}\begin{bmatrix}1 & \frac{1-v_3}{v_1+iv_2}\end{bmatrix}\\ &= \frac{1+v_3}{2}\begin{bmatrix}1 & \frac{v_1-iv_2}{1+v_3}\\ \frac{v_1+iv_2}{1+v_3} & \frac{1-v_3}{1+v_3}\end{bmatrix}\\ &= \frac{1}{2}\begin{bmatrix}1+v_3 & v_1-iv_2\\ v_1+iv_2 & 1-v_3\end{bmatrix}\\ &= \frac{1}{2}\left(I+\begin{bmatrix}v_3 & v_1-iv_2\\ v_1+iv_2 & -v_3\end{bmatrix}\right)\\ &= \frac{1}{2}(I+\vec{v}\cdot\vec{\sigma})\end{aligned}$$

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

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

While I review my proof, I notice that my 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.

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

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

Proof. 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 $P_\pm = |\lambda_{\pm 1}\rangle\langle\lambda_{\pm 1}|$.

In order to prove above equation, all we have to do is prove following condition. (see Theorem 2.0.1)

$$\langle\psi|(P_\pm - |\lambda_{\pm 1}\rangle\langle\lambda_{\pm 1}|)|\psi\rangle = 0 \text{ for all } |\psi\rangle \in \mathbb{C}^2. \tag{2.1}$$

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

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

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

Therefore $\langle\psi|(P_\pm - |\lambda_{\pm1}\rangle\langle\lambda_{\pm1}|)|\psi\rangle = 0$ for all $|\psi\rangle \in \mathbb{C}^2$. Thus $P_\pm = |\lambda_{\pm1}\rangle\langle\lambda_{\pm1}|$. □

2.61)

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

Post-measurement state is

$$
\begin{aligned}
\frac{|\lambda_1\rangle\langle\lambda_1|0\rangle}{\sqrt{\langle 0|\lambda_1\rangle\langle\lambda_1|0\rangle}} &= \frac{1}{\sqrt{\frac{1}{2}(1+v_3)}} \cdot \frac{1}{2}\begin{bmatrix} 1 + v_3 \\ v_1 + iv_2 \end{bmatrix} \\
&= \sqrt{\frac{1}{2}(1+v_3)}\begin{bmatrix} 1 \\ \frac{v_1+iv_2}{1+v_3} \end{bmatrix} \\
&= \sqrt{\frac{1+v_3}{2}}\begin{bmatrix} 1 \\ \frac{1-v_3}{v_1-iv_2} \end{bmatrix} \\
&= |\lambda_1\rangle .
\end{aligned}
$$

2.62)

Suppose M_m is a measurement operator. From the assumption, $E_m = M_m^\dagger M_m = M_m$. Then

$$\langle\psi|E_m|\psi\rangle = \langle\psi|M_m|\psi\rangle \geq 0.$$

for all $|\psi\rangle$.

Since M_m is positive operator, M_m is Hermitian. Therefore,

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

Thus the measurement is a projective measurement.

2.63)

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

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

2.64) Read following paper:

- Lu-Ming Duan, Guang-Can Guo. Probabilistic cloning and identification of linearly independent quantum states. Phys. Rev. Lett.,80:4999-5002, 1998. arXiv:quant-ph/9804064
  `https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.80.4999`
  `https://arxiv.org/abs/quant-ph/9804064`
- Stephen M. Barnett, Sarah Croke, Quantum state discrimination, arXiv:0810.1970 [quant-ph]
  `https://arxiv.org/abs/0810.1970`
  `https://www.osapublishing.org/DirectPDFAccess/67EF4200-CBD2-8E68-1979E37886263936_176580/aop-1-2-238.pdf`

2.65)

$$|+\rangle \equiv \frac{|0\rangle + |1\rangle}{\sqrt{2}}, \quad |-\rangle \equiv \frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

2.66)

$$X_1 Z_2 \left(\frac{|00\rangle + |11\rangle}{\sqrt{2}} \right) = \frac{|10\rangle - |01\rangle}{\sqrt{2}}$$

$$\langle X_1 Z_2 \rangle = \left(\frac{\langle 00| + \langle 11|}{\sqrt{2}} \right) X_1 Z_2 \left(\frac{|00\rangle + |11\rangle}{\sqrt{2}} \right) = \frac{\langle 00| + \langle 11|}{\sqrt{2}} \cdot \frac{|10\rangle - |01\rangle}{\sqrt{2}} = 0$$

2.67)

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 $U' = \sum_i |u_i\rangle\langle w_i| + \sum_j |u'_j\rangle\langle w'_j|$, where $|u_i\rangle = U|w_i\rangle$.

Now

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

and

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

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

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

Therefore U' is an extension of U.

2.68)
$|\psi\rangle = \frac{|00\rangle+|11\rangle}{\sqrt{2}}$.
Suppose $|a\rangle = a_0 |0\rangle + a_1 |1\rangle$ and $|b\rangle = b_0 |0\rangle + b_1 |1\rangle$.

$$|a\rangle |b\rangle = a_0 b_0 |00\rangle + a_0 b_1 |01\rangle + a_1 b_0 |10\rangle + a_1 b_1 |11\rangle .$$

If $|\psi\rangle = |a\rangle |b\rangle$, then $a_0b_0 = 1,\ a_0b_1 = 0,\ a_1b_0 = 0,\ a_1b_1 = 1$ since $\{|ij\rangle\}$ is an orthonormal basis.
If $a_0b_1 = 0$, then $a_0 = 0$ or $b_1 = 0$.
When $a_0 = 0$, this is contradiction to $a_0b_0 = 1$. When $b_1 = 0$, this is contradiction to $a_1b_1 = 1$.
Thus $|\psi\rangle \neq |a\rangle |b\rangle$.

2.69) Define Bell states as follows.

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

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

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

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

First, we prove $\{|\psi_i\rangle\}$ is a linearly independent basis.

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

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

$$\therefore \begin{cases} a_1 + a_2 = 0 \\ a_3 + a_4 = 0 \\ a_3 - a_4 = 0 \\ a_1 - a_2 = 0 \end{cases}$$

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

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

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

2.70)

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

Suppose Eve measures the qubit Alice sent by measurement operators M_m. The probability that Eve gets 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)

From spectral decomposition,

$$\begin{aligned}
\rho &= \sum_i p_i |\psi_i\rangle\langle\psi_i|, \quad p_i \geq 0, \quad \sum_i p_i = 1. \\
\rho^2 &= \sum_{i,j} p_i p_j |i\rangle \langle i|j\rangle \langle j| \\
&= \sum_{i,j} p_i p_j |i\rangle\langle j| \delta_{ij} \\
&= \sum_i p_i^2 |i\rangle\langle i|
\end{aligned}$$

$$\mathrm{Tr}(\rho^2) = \mathrm{Tr}\left(\sum_i p_i^2 |i\rangle\langle i|\right) = \sum_i p_i^2 \,\mathrm{Tr}(|i\rangle\langle i|) = \sum_i p_i^2 \langle i|i\rangle = \sum_i p_i^2 \leq \sum_i p_i = 1 \quad (\because p_i^2 \leq p_i)$$

Suppose $\mathrm{Tr}(\rho^2) = 1$. 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 $\rho = |\psi\rangle\langle\psi|$.

$$\mathrm{Tr}(\rho^2) = \mathrm{Tr}(|\psi\rangle \langle\psi|\psi\rangle \langle\psi|) = \mathrm{Tr}(|\psi\rangle\langle\psi|) = \langle\psi|\psi\rangle = 1.$$

2.72)

(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, $\mathrm{Tr}(\rho) = a + d = 1$.

Define $a = (1 + r_3)/2$, $d = (1 - r_3)/2$ and $b = (r_1 - ir_2)/2$, $(r_i \in \mathbb{R})$.

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.

$$\begin{aligned}
\det(\rho - \lambda I) = (a-\lambda)(b-\lambda) - |b|^2 = \lambda^2 - (a+d)\lambda + ad - |b^2| = 0 \\
\lambda &= \frac{(a+d) \pm \sqrt{(a+d)^2 - 4(ad - |b|^2)}}{2} \\
&= \frac{1 \pm \sqrt{1 - 4\left(\frac{1-r_3^2}{4} - \frac{r_1^2+r_2^2}{4}\right)}}{2} \\
&= \frac{1 \pm \sqrt{1 - (1 - r_1^2 - r_2^2 - r_3^2)}}{2} \\
&= \frac{1 \pm \sqrt{|\vec{r}|^2}}{2} \\
&= \frac{1 \pm |\vec{r}|}{2}
\end{aligned}$$

Since ρ is positive, $\frac{1-|\vec{r}|}{2} \geq 0 \to |\vec{r}| \leq 1$.
Therefore an arbitrary density matrix for a mixed state qubit is written as $\rho = \frac{1}{2}(I + \vec{r}\cdot\vec{\sigma})$.

(2)

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

(3)

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

If ρ is pure, then $\mathrm{Tr}(\rho^2) = 1$.

$$\begin{aligned}
1 = \mathrm{Tr}(\rho^2) &= \frac{1}{4}(2 + 2|\vec{r}|^2) \\
&\therefore |\vec{r}| = 1.
\end{aligned}$$

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

2.73)

Theorem 2.6)

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

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

$$\sum_i |\tilde{\psi}_i\rangle\langle\tilde{\psi}_i| = \Big[|\tilde{\psi}_1\rangle \cdots |\tilde{\psi}_k\rangle\Big] \begin{bmatrix} \langle\tilde{\psi}_1| \\ \vdots \\ \langle\tilde{\psi}_k| \end{bmatrix} \tag{2.3}$$

$$= \Big[|\tilde{\varphi}_1\rangle \cdots |\tilde{\varphi}_k\rangle\Big] U^T U^* \begin{bmatrix} \langle\tilde{\varphi}_1| \\ \vdots \\ \langle\tilde{\varphi}_k| \end{bmatrix} \tag{2.4}$$

$$= \Big[|\tilde{\varphi}_1\rangle \cdots |\tilde{\varphi}_k\rangle\Big] \begin{bmatrix} \langle\tilde{\varphi}_1| \\ \vdots \\ \langle\tilde{\varphi}_k| \end{bmatrix} \tag{2.5}$$

$$= \sum_j |\tilde{\varphi}_j\rangle\langle\tilde{\varphi}_j|\,. \tag{2.6}$$

From spectral theorem, density matrix ρ is decomposed as $\rho = \sum_{k=1}^{d} \lambda_k \left|k\right\rangle\!\left\langle k\right|$ 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 $\rho = \sum_{k=1}^{l} p_k \left|k\right\rangle\!\left\langle k\right| = \sum_{k=1}^{l} |\tilde{k}\rangle\langle\tilde{k}|$, 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 $[u_{i1} \cdots u_{ik} \cdots u_{il}]$. Then we can define another ensemble such that

$$\Big[|\tilde{\psi}_1\rangle \cdots |\tilde{\psi}_i\rangle \cdots |\tilde{\psi}_l\rangle\Big] = \Big[|\tilde{k}_1\rangle \cdots |\tilde{k}_l\rangle\Big] U^T$$

[1] By Gram-Schmidt procedure construct an orthonormal basis $\{\boldsymbol{u}_j\}$ (row vector) with $\boldsymbol{u}_i = [u_{i1} \cdots u_{ik} \cdots u_{il}]$. Then define unitary $U = \begin{bmatrix} \boldsymbol{u}_1 \\ \vdots \\ \boldsymbol{u}_i \\ \vdots \\ \boldsymbol{u}_l \end{bmatrix}$.

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

$$\rho = \sum_k |\tilde{k}\rangle\langle\tilde{k}| = \sum_k |\tilde{\psi_k}\rangle\langle\tilde{\psi_k}|\,.$$

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

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

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

2.74)

$$\begin{aligned}
\rho_{AB} &= |a\rangle\langle a|_A \otimes |b\rangle\langle b|_B \\
\rho_A &= \mathrm{Tr}_B\,\rho_{AB} = |a\rangle\langle a|\,\mathrm{Tr}(|b\rangle\langle b|) = |a\rangle\langle a| \\
\mathrm{Tr}(\rho_A^2) &= 1
\end{aligned}$$

Thus ρ_A is pure.

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

$$\begin{aligned}
|\Phi_\pm\rangle\langle\Phi_\pm|_{AB} &= \frac{1}{2}(|00\rangle\langle 00| \pm |00\rangle\langle 11| \pm |11\rangle\langle 00| + |11\rangle\langle 11|) \\
\mathrm{Tr}_B(|\Phi_\pm\rangle\langle\Phi_\pm|_{AB}) &= \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|) = \frac{I}{2} \\
|\Psi_\pm\rangle\langle\Psi_\pm| &= \frac{1}{2}(|01\rangle\langle 01| \pm |01\rangle\langle 10| \pm |10\rangle\langle 01| + |10\rangle\langle 10|) \\
\mathrm{Tr}_B(|\Psi_\pm\rangle\langle\Psi_\pm|) &= \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|) = \frac{I}{2}
\end{aligned}$$

2.76)

Unsolved. ~~I think the polar decomposition can only apply to square matrix A, not arbitrary linear operators. Suppose A is $m \times n$ matrix. Then size of $A^\dagger A$ is $n \times n$. Thus the size of U should be $m \times n$. Maybe U is isometry, but I think it is not unitary.~~

I misunderstand linear operator.

> Quoted from "Advanced Liner Algebra" by Steven Roman, ISBN 0387247661.
>
> A linear transformation $\tau : V \to V$ is called a **linear operator** on V.[2]

Thus coordinate matrices of linear operator are square matrices. And Nielsen and Chaung say at Theorem 2.3, "Let A be a linear operator on a vector space V." Therefore A is a linear transformation such that $A : V \to V$.

2.77)

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

[2] According to Roman, some authors use the term linear operator for any linear transformation from V to W.

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)

Proof. Former 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.

($\Rightarrow$) Proved by exercise 2.74.

($\Leftarrow$) Let a pure state be $|\psi\rangle = \sum_i \lambda_i |i_A\rangle |i_B\rangle$. Then $\rho_A = \mathrm{Tr}_B(|\psi\rangle\langle\psi|) = \sum_i \lambda_i^2 |i\rangle\langle i|$. 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)

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)

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

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

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

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

$$\rho_A = \lambda_0 |\lambda_0\rangle\langle\lambda_0| + \lambda_1 |\lambda_1\rangle\langle\lambda_1| .$$

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

Then

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

(It's too tiresome to calculate $|a_i\rangle$)

2.80)

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)

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

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

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

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

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

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

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,

$$\mathrm{Tr}_R(|\psi_i\rangle|i\rangle) = |\psi_i\rangle .$$

(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

$$\mathrm{Tr}_R\left(|AR\rangle\langle AR|\right) = \sum_i \lambda_i |\phi_i^A\rangle\langle\phi_i^A| = \sum_i p_i|\psi_i\rangle\langle\psi_i| .$$

By theorem 2.6, there exits an unitary matrix u_{ij} such that $\sqrt{\lambda_i}\ket{\phi_i^A} = \sum_j u_{ij}\sqrt{p_j}\ket{\psi_j}$. Then

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

where $\ket{i} = \sum_k u_{ki}\ket{\phi_k^R}$.

About $\ket{i}$,

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

which implies $\ket{j}$ is an orthonormal basis for system R.

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

Problem 2.1)

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

$$\vec{n}\cdot\vec{\sigma} = \ket{\lambda_1}\bra{\lambda_1} - \ket{\lambda_{-1}}\bra{\lambda_{-1}}$$

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

Thus

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

Problem 2.2) Unsolved

Problem 2.3) Unsolved

Chapter 8

Quantum noise and quantum operations

8.1) Density operator of initial state is written by $|\psi\rangle\langle\psi|$ and final state is written by $U|\psi\rangle\langle\psi|U^\dagger$. Thus time development of $\rho = |\psi\rangle\langle\psi|$ can be written by $\mathcal{E}(\rho) = U\rho U^\dagger$.

8.2) From eqn (2.147) (on page 100),

$$\rho_m = \frac{M_m \rho M_m^\dagger}{\mathrm{Tr}(M_m^\dagger M_m \rho)} = \frac{M_m \rho M_m^\dagger}{\mathrm{Tr}(M_m \rho M_m^\dagger)} = \frac{\mathcal{E}_m(\rho)}{\mathrm{Tr}\,\mathcal{E}_m(\rho)}.$$

And from eqn (2.143) (on page 99), $p(m) = \mathrm{Tr}(M_m^\dagger M_m \rho) = \mathrm{Tr}(M_m \rho M_m^\dagger) = \mathrm{Tr}\,\mathcal{E}_m(\rho)$.

8.3)

8.4)
8.5)
8.6)
8.7)
8.8)
8.9)
8.10)
8.11)
8.12)
8.13)
8.14)
8.15)
8.16)
8.17)
8.18)
8.19)
8.20)
8.21)
8.22)
8.23)
8.24)
8.25)
8.26)
8.27)
8.28)
8.29)

8.30)
8.31)
8.32)
8.33)
8.34)
8.35)

Chapter 9

Distance measures for quantum information

9.1)

$$\begin{aligned} D((1,0),(1/2,1/2)) &= \frac{1}{2}\left(|1-1/2|+|0-1/2|\right) \\ &= \frac{1}{2}\left(\frac{1}{2}+\frac{1}{2}\right) \\ &= \frac{1}{2} \end{aligned}$$

$$\begin{aligned} D\left((1/2,1/3,1/6),(3/4,1/8,1/8)\right) &= \frac{1}{2}\left(|1/2-3/4|+|1/3-1/8|+|1/6-1/8|\right) \\ &= \frac{1}{2}\left(1/4+5/24+1/24\right) \\ &= \frac{1}{4} \end{aligned}$$

9.2)

$$\begin{aligned} D\left((p,1-p),(q,1-q)\right) &= \frac{1}{2}\left(|p-q|+|(1-p)-(1-q)|\right) \\ &= \frac{1}{2}\left(|p-q|+|-p+q|\right) \\ &= |p-q| \end{aligned}$$

9.3)

$$F((1,0),(1/2,1/2)) = \sqrt{1\cdot 1/2}+\sqrt{0\cdot 1/2} = \frac{1}{\sqrt{2}}$$

$$\begin{aligned} F\left((1/2,1/3,1/6),(3/4,1/8,1/8)\right) &= \sqrt{1/2\cdot 3/4}+\sqrt{1/3\cdot 1/8}+\sqrt{1/6\cdot 1/8} \\ &= \frac{4\sqrt{6}+\sqrt{3}}{12} \end{aligned}$$

9.4)

Define $r_x = p_x - q_x$. Let U be the whole index set.

$$\begin{aligned}\max_S |p(S) - q(S)| &= \max_S \left|\sum_{x\in S} p_x - \sum_{x\in S} q_x\right| \\ &= \max_S \left|\sum_{x\in S}(p_x - q_x)\right| \\ &= \max_S \left|\sum_{x\in S} r_x\right|\end{aligned}$$

Since $\sum_{x\in S} r_x$ is written as

$$\sum_{x\in S} r_x = \sum_{\substack{x\in S\\ r_x\geq 0}} r_x + \sum_{\substack{x\in S\\ r_x<0}} r_x, \tag{9.1}$$

$\left|\sum_{x\in S} r_x\right|$ is maximized when $S = \{x \in U | r_x \geq 0\}$ or $S = \{x \in U | r_x < 0\}$.
Define $S_+ = \{x \in U | r_x \geq 0\}$ and $S_- = \{x \in U | r_x < 0\}$.
Now the sum of all r_x is 0,

$$\begin{aligned}\sum_{x\in U} r_x = \sum_{x\in S_+} r_x + \sum_{x\in S_-} r_x = 0 \\ \therefore \sum_{x\in S_+} r_x = -\sum_{x\in S_-} r_x.\end{aligned}$$

Thus

$$\max_S \left|\sum_{x\in S} r_x\right| = \sum_{x\in S_+} r_x = -\sum_{x\in S_-} r_x. \tag{9.2}$$

On the other hand,

$$\begin{aligned}D(p_x, q_x) &= \frac{1}{2}\sum_{x\in U} |p_x - q_x| \\ &= \frac{1}{2}\sum_{x\in U} |r_x| \\ &= \frac{1}{2}\sum_{x\in S_+} |r_x| + \frac{1}{2}\sum_{x\in S_-} |r_x| \\ &= \frac{1}{2}\sum_{x\in S_+} r_x - \frac{1}{2}\sum_{x\in S_-} r_x \\ &= \frac{1}{2}\sum_{x\in S_+} r_x + \frac{1}{2}\sum_{x\in S_+} r_x \quad (\because \text{eqn}(9.2)) \\ &= \sum_{x\in S_+} r_x \\ &= \max_S \left|\sum_{x\in S} r_x\right|.\end{aligned}$$

Therefore $D(p_x, q_x) = \max_S \left|\sum_{x\in S} p_x - \sum_{x\in S} q_x\right| = \max_S |p(S) - q(S)|$.

9.5) From eqn (9.1) and (9.2), maximizing $\left|\sum_{x\in S} r_x\right|$ is equivalent to maximizing $\sum_{x\in S} r_x$.

Hence

$$D(p_x, q_x) = \max_S (p(S) - q(S)) = \max_S \left(\sum_{x \in S} p_x - \sum_{x \in S} q_x \right).$$

9.6)

Define $\rho = \frac{3}{4} |0\rangle\langle 0| + \frac{1}{4} |1\rangle\langle 1|$, $\sigma = \frac{2}{3} |1\rangle\langle 1| + \frac{1}{3} |1\rangle\langle 1|$.

$$\begin{aligned}
D(\rho, \sigma) &= \frac{1}{2} \operatorname{Tr} |\rho - \sigma| \\
&= D((3/4, 1/4), (2/3, 1/3)) \\
&= \frac{1}{2} \left(\left| \frac{3}{4} - \frac{2}{3} \right| + \left| \frac{1}{4} - \frac{1}{3} \right| \right) \\
&= \frac{1}{2} \left(\frac{1}{12} + \frac{1}{12} \right) \\
&= \frac{1}{12}
\end{aligned}$$

Define $\rho = \frac{3}{4} |0\rangle\langle 0| + \frac{1}{4} |1\rangle\langle 1|$, $\sigma = \frac{2}{3} |+\rangle\langle +| + \frac{1}{3} |-\rangle\langle -|$.

$$\begin{aligned}
|+\rangle\langle +| &= \frac{1}{2}(|0\rangle\langle 0| + |0\rangle\langle 1| + |1\rangle\langle 0| + |1\rangle\langle 1|) \\
|-\rangle\langle -| &= \frac{1}{2}(|0\rangle\langle 0| - |0\rangle\langle 1| - |1\rangle\langle 0| + |1\rangle\langle 1|)
\end{aligned}$$

$$\begin{aligned}
\rho - \sigma &= \left(\frac{3}{4} - \frac{1}{2} \right) |0\rangle\langle 0| - \frac{1}{6}(|0\rangle\langle 1| + |1\rangle\langle 0|) + \left(\frac{1}{4} - \frac{1}{2} \right) |1\rangle\langle 1| \\
&= \frac{1}{4} |0\rangle\langle 0| - \frac{1}{6}(|0\rangle\langle 1| + |1\rangle\langle 0|) - \frac{1}{4} |1\rangle\langle 1|
\end{aligned}$$

$$\begin{aligned}
(\rho - \sigma)^\dagger (\rho - \sigma) &= \frac{1}{4^2} |0\rangle\langle 0| - \frac{1}{4 \cdot 6} |0\rangle\langle 1| + \frac{1}{6^2} |0\rangle\langle 0| + \frac{1}{6 \cdot 4} |0\rangle\langle 1| - \frac{1}{4 \cdot 6} |1\rangle\langle 0| + \frac{1}{6^2} |1\rangle\langle 1| + \frac{1}{4 \cdot 6} |1\rangle\langle 0| + \frac{1}{4^2} |1\rangle\langle 1| \\
&= \left(\frac{1}{4^2} + \frac{1}{6^2} \right) (|0\rangle\langle 0| + |1\rangle\langle 1|)
\end{aligned}$$

$$\begin{aligned}
D(\rho, \sigma) &= \frac{1}{2} \operatorname{Tr} |\rho - \sigma| \\
&= \sqrt{\frac{1}{4^2} + \frac{1}{6^2}}
\end{aligned}$$

9.7)

Since $\rho - \sigma$ is Hermitian, we can apply spectral decomposition. Then $\rho - \sigma$ is written as

$$\rho - \sigma = \sum_{i=1}^{k} \lambda_i |i\rangle\langle i| + \sum_{i=k+1}^{n} \lambda_i |i\rangle\langle i|$$

where λ_i are positive eigenvalues for $i = 1, \cdots, k$ and negative eigenvalues for $i = k+1, \cdots, n$.

Define $Q = \sum_{i=1}^{k} \lambda_i |i\rangle\langle i|$ and $S = -\sum_{i=k+1}^{n} \lambda_i |i\rangle\langle i|$. Then P and S are positive operator. Therefore $\rho - \sigma = P - S$.

Proof of $|\rho - \sigma| = Q + S$.

$$\begin{aligned}
|\rho - \sigma| &= |Q - S| \\
&= \sqrt{(Q-S)^\dagger (Q-S)} \\
&= \sqrt{(Q-S)^2} \\
&= \sqrt{Q^2 - QS - SQ + S^2} \\
&= \sqrt{Q^2 + S^2} \\
&= \sqrt{\sum_i \lambda_i^2 |i\rangle\langle i|} \\
&= \sum_i |\lambda_i| \, |i\rangle\langle i| \\
&= Q + S
\end{aligned}$$

9.8)

Suppose $\sigma = \sigma_i$. Then $\sigma = \sum_i p_i \sigma_i$.

$$\begin{aligned}
D\left(\sum_i p_i \rho_i, \sigma\right) &= D\left(\sum_i p_i \rho_i, \sum_i p_i \sigma_i\right) && (9.3) \\
&\leq \sum_i p_i D(\rho_i, \sigma_i) \quad (\because \text{eqn}(9.50)) && (9.4) \\
&= \sum_i p_i D(\rho_i, \sigma). \quad (\because \text{assumption}). && (9.5)
\end{aligned}$$

9.9)
9.10)
9.11)
9.12)

Suppose $\rho = \frac{1}{2}(I + \vec{r} \cdot \vec{\sigma})$ and $\sigma = \frac{1}{2}(I + \vec{s} \cdot \vec{\sigma})$ where $\vec{v}$ and $\vec{s}$ are real vectors s.t. $|\vec{v}|, |\vec{s}| \leq 1$.

$$\mathcal{E}(\rho) = p\frac{I}{2} + (1-p)\rho, \quad \mathcal{E}(\sigma) = p\frac{I}{2} + (1-p)\sigma.$$

$$\begin{aligned}
D(\mathcal{E}(\rho), \mathcal{E}(\sigma)) &= \frac{1}{2} \operatorname{Tr} |\mathcal{E}(\rho) - \mathcal{E}(\sigma)| \\
&= \frac{1}{2} \operatorname{Tr} |(1-p)(\rho - \sigma)| \\
&= \frac{1}{2}(1-p) \operatorname{Tr} |\rho - \sigma| \\
&= (1-p) D(\rho, \sigma) \\
&= (1-p) \frac{|\vec{r} - \vec{s}|}{2}
\end{aligned}$$

Is this strictly contractive?

9.13)

Bit flip channel $E_0 = \sqrt{p}I$, $E_1 = \sqrt{1-p}\sigma_x$.

$$\begin{aligned}\mathcal{E}(\rho) &= E_0 \rho E_0^\dagger + E_1 \rho E_1^\dagger \\ &= p\rho + (1-p)\sigma_x \rho \sigma_x.\end{aligned}$$

Since $\sigma_x\sigma_x\sigma_x = \sigma_x$, $\sigma_x\sigma_y\sigma_x = -\sigma_y$ and $\sigma_x\sigma_z\sigma_x = -\sigma_z$, then $\sigma_x(\vec{r}\cdot\vec{\sigma}) = r_1\sigma_x - r_2\sigma_y - r_3\sigma_3$. Thus

$$\begin{aligned}D(\mathcal{E}(\rho), \mathcal{E}(\sigma)) &= \frac{1}{2}\operatorname{Tr}|\mathcal{E}(\rho) - \mathcal{E}(\sigma)| \\ &= \frac{1}{2}\operatorname{Tr}|p(\rho-\sigma) + (1-p)(\sigma_x\rho\sigma_x - \sigma_x\sigma\sigma_x)| \\ &\leq \frac{1}{2}p\operatorname{Tr}|\rho-\sigma| + \frac{1}{2}(1-p)\operatorname{Tr}|\sigma_x(\rho-\sigma)\sigma_x| \\ &= pD(\rho,\sigma) + (1-p)D(\sigma_x\rho\sigma_x, \sigma_x\sigma\sigma_x) \\ &= D(\rho,\sigma) \quad (\because \text{eqn}(9.21)).\end{aligned}$$

Suppose $\rho_0 = \frac{1}{2}(I + \vec{r}\cdot\vec{\sigma})$ is a fixed point. Then

$$\begin{aligned}&\rho_0 = \mathcal{E}(\rho_0) = p\rho_0 + (1-p)\sigma_x\rho_0\sigma_x \\ &\therefore (1-p)\rho_0 - (1-p)\sigma_x\rho_0\sigma_x = 0 \\ &\therefore (1-p)(\rho - \sigma_x\rho_0\sigma_x) = 0 \\ &\therefore \rho_0 = \sigma_x\rho_0\sigma_x \\ &\therefore \frac{1}{2}(I + r_1\sigma_x + r_2\sigma_y + r_3\sigma_z)\frac{1}{2}(I + r_1\sigma_x - r_2\sigma_y - r_3\sigma_z)\end{aligned}$$

Since $\{I, \sigma_x, \sigma_y, \sigma_z\}$ are linearly independent, $r_2 = -r_2$ and $r_3 = -r_3$. Thus $r_2 = r_3 = 0$. Therefore the set of fixed points for the bit flip channel is $\{\rho \mid \rho = \frac{1}{2}(I + r\sigma_x), |r| \leq 1, r \in \mathbb{R}\}$

9.14)

$$\begin{aligned}F(U\rho U^\dagger, U\sigma U^\dagger) &= \operatorname{Tr}\sqrt{(U\rho U^\dagger)^{1/2}\sigma(U\rho U^\dagger)} \\ &= \operatorname{Tr}\sqrt{U\rho^{1/2}\sigma\rho^{1/2}U^\dagger} \\ &= \operatorname{Tr}(U\sqrt{\rho^{1/2}\sigma\rho^{1/2}}U^\dagger) \\ &= \operatorname{Tr}(\sqrt{\rho^{1/2}\sigma\rho^{1/2}}U^\dagger U) \\ &= \operatorname{Tr}\sqrt{\rho^{1/2}\sigma\rho^{1/2}} \\ &= F(\rho,\sigma)\end{aligned}$$

I think the fact $\sqrt{UAU^\dagger} = U\sqrt{A}U^\dagger$ is not restricted for positive operator.
Suppose A is a normal matrix. From spectral theorem, it is decomposed as

$$A = \sum_i a_i |i\rangle\langle i| .$$

Let f be a function. Then

$$\begin{aligned}
f(UAU^\dagger) &= f(\sum_i a_i U |i\rangle\langle i| U^\dagger) \\
&= \sum_i f(a_i) U |i\rangle\langle i| U^\dagger \\
&= U(\sum_i f(a_i) U |i\rangle\langle i| U^\dagger) U^\dagger \\
&= U f(A) U^\dagger
\end{aligned}$$

9.15) $|\psi\rangle = (U_R \otimes \sqrt{\rho}U_Q)|m\rangle$ is any fixed purification of ρ, and $|\phi\rangle = (V_R \otimes \sqrt{\sigma}V_Q)|m\rangle$ is purification of σ. Suppose $\sqrt{\rho}\sqrt{\sigma} = |\sqrt{\rho}\sqrt{\sigma}|V$ is the polar decomposition of $\sqrt{\rho}\sqrt{\sigma}$. Then

$$\begin{aligned}
|\langle\psi|\phi\rangle| &= \left|\langle m| \left(U_R^\dagger V_R \otimes U_Q^\dagger \sqrt{\rho}\sqrt{\sigma}V_Q\right)|m\rangle\right| \\
&= \left|\mathrm{Tr}\left((U_R^\dagger V_R)^T U_Q^\dagger \sqrt{\rho}\sqrt{\sigma}V_Q\right)\right| \\
&= \left|\mathrm{Tr}\left(V_R^T U_R^* U_Q^\dagger \sqrt{\rho}\sqrt{\sigma}V_Q\right)\right| \\
&= \left|\mathrm{Tr}\left(V_Q V_R^T U_R^* U_Q^\dagger \sqrt{\rho}\sqrt{\sigma}\right)\right| \\
&= \left|\mathrm{Tr}\left(V_Q V_R^T U_R^* U_Q^\dagger |\sqrt{\rho}\sqrt{\sigma}|V\right)\right| \\
&= \left|\mathrm{Tr}\left(V V_Q V_R^T U_R^* U_Q^\dagger |\sqrt{\rho}\sqrt{\sigma}|\right)\right| \\
&\le \mathrm{Tr}\,|\sqrt{\rho}\sqrt{\sigma}| \\
&= F(\rho,\sigma)
\end{aligned}$$

Choosing $V_Q = V^\dagger$, $V_R^T = (U_Q^* U_R^\dagger)^\dagger$ we see that equality is attained.

9.16) I think eq (9.73) has a typo. $\mathrm{Tr}(A^\dagger B) = \langle m|A \otimes B|m\rangle$ should be $\mathrm{Tr}(A^{T} B) = \langle m|A \otimes B|m\rangle$. See errata list.

In order to show that this exercise, I will prove following two properties,

$$\mathrm{Tr}(A) = \langle m|(I \otimes A)|m\rangle\,, \quad (I \otimes A)|m\rangle = (A^T \otimes I)|m\rangle$$

where A is a linear operator and $|m\rangle$ is unnormalized maximally entangled state, $|m\rangle = \sum_i |ii\rangle$.

$$
\begin{aligned}
\langle m|I \otimes A|m\rangle &= \sum_{ij} \langle ii|(I \otimes A)|jj\rangle \\
&= \sum_{ij} \langle i|I|j\rangle \langle i|A|j\rangle \\
&= \sum_{ij} \delta_{ij} \langle i|A|j\rangle \\
&= \sum_{i} \langle i|A|i\rangle \\
&= \mathrm{Tr}(A)
\end{aligned}
$$

Suppose $A = \sum_{ij} a_{ij} |i\rangle\langle j|$.

$$
\begin{aligned}
(I \otimes A)|m\rangle &= \left(I \otimes \sum_{ij} a_{ij} |i\rangle\langle j|\right) \sum_k |kk\rangle \\
&= \sum_{ijk} a_{ij} |k\rangle \otimes |i\rangle \langle j|k\rangle \\
&= \sum_{ijk} a_{ij} |k\rangle \otimes |i\rangle \delta_{jk} \\
&= \sum_{ij} a_{ij} |j\rangle \otimes |i\rangle \\
&= \sum_{ij} a_{ji} |i\rangle \otimes |j\rangle
\end{aligned}
$$

$$
\begin{aligned}
(A^T \otimes I)|m\rangle &= \left(\sum_{ij} a_{ji} |i\rangle\langle j| \otimes I\right) \sum_k |kk\rangle \\
&= \sum_{ij} a_{ji} |i\rangle \langle j|k\rangle \otimes |k\rangle \\
&= \sum_{ij} a_{ji} |i\rangle \delta_{jk} \otimes |k\rangle \\
&= \sum_{ij} a_{ji} |ij\rangle \\
&= (I \otimes A)|m\rangle
\end{aligned}
$$

Thus

$$
\begin{aligned}
\mathrm{Tr}(A^T B) = \mathrm{Tr}(BA^T) &= \langle m|I \otimes BA^T|m\rangle \\
&= \langle m|(I \otimes B)(I \otimes A^T)|m\rangle \\
&= \langle m|(I \otimes B)(A \otimes I)|m\rangle \\
&= \langle m|A \otimes B|m\rangle .
\end{aligned}
$$

9.17) If $\rho = \sigma$, then $F(\rho, \sigma) = 1$. Thus $A(\rho, \sigma) = \arccos F(\rho, \sigma) = \arccos 1 = 0$.

If $A(\rho, \sigma) = 0$, then $\arccos F(\rho, \sigma) = 0 \Rightarrow \cos(\arccos F(\rho, \sigma)) = \cos(0) \Rightarrow F(\rho, \sigma) = 1$ ($\because$ text p.411, the fifth line from bottom).

9.18) For $0 \leq x \leq y \leq 1$, $\arccos(x) \geq \arccos(y)$. From $F(\mathcal{E}(\rho), \mathcal{E}(\sigma)) \geq F(\rho, \sigma)$ and $0 \leq F(\mathcal{E}(\rho), \mathcal{E}(\sigma)), F(\rho, \sigma) \leq 1$,

$$\begin{aligned} \arccos F(\mathcal{E}(\rho), \mathcal{E}(\sigma)) &\geq \arccos F(\rho, \sigma) \\ \therefore A(\mathcal{E}(\rho), \mathcal{E}(\sigma)) &\geq A(\rho, \sigma) \end{aligned}$$

9.19) From eq (9.92)

$$\begin{aligned} F\left(\sum_i p_i \rho_i, \sum_i p_i \sigma_i\right) &\geq \sum_i \sqrt{p_i p_i} F(\rho_i, \sigma_i) \\ &= \sum_i p_i F(\rho_i, \sigma_i). \end{aligned}$$

9.20) Suppose $\sigma_i = \sigma$. Then

$$\begin{aligned} F\left(\sum_i p_i \rho_i, \sigma\right) &= F\left(\sum_i p_i \rho_i, \sum_i p_i \sigma\right) \\ &= F\left(\sum_i p_i \rho_i, \sum_i p_i \sigma_i\right) \\ &\geq \sum_i p_i F(\rho_i, \sigma_i) \quad (\because \text{Exercise9.19}) \\ &= \sum_i p_i F(\rho_i, \sigma) \end{aligned}$$

9.21)

$$1 - F(|\psi\rangle, \sigma)^2 = 1 - \langle\psi|\sigma|\psi\rangle \quad (\because \text{eq(9.60)})$$

$$\begin{aligned} D(|\psi\rangle, \sigma) &= \max_P \mathrm{Tr}(P(\rho - \sigma)) \quad (\text{where } P \text{ is projector.}) \\ &\geq \mathrm{Tr}\left(|\psi\rangle\langle\psi| (\rho - \sigma)\right) \\ &= \langle\psi|(|\psi\rangle\langle\psi| - \sigma)|\psi\rangle \\ &= 1 - \langle\psi|\sigma|\psi\rangle \\ &= 1 - F(|\psi\rangle, \sigma)^2. \end{aligned}$$

9.22) (ref: QCQI Exercise Solutions (Chapter 9) - めもめも
`http://enakai00.hatenablog.com/entry/2018/04/12/134722`)

For all ρ, following inequality is satisfied,

$$\begin{aligned} d(VU\rho U^\dagger V^\dagger, \mathcal{F} \circ \mathcal{E}(\rho)) &\leq d(VU\rho U^\dagger V^\dagger, \mathcal{F}(U\rho U^\dagger)) + d(\mathcal{F}(U\rho U^\dagger), \mathcal{F} \circ \mathcal{E}(\rho)) \\ &\leq d(VU\rho U^\dagger V^\dagger) + d(U\rho U^\dagger, \mathcal{E}(\rho)) \\ &\leq E(V, \mathcal{F}) + E(U, \mathcal{E}). \end{aligned}$$

First inequality is triangular inequality, second is contractivity of the metric[1] and third is from definition of E.

[1] Trace distance and angle are satisfied with contractive (eq (9.35), eq (9.91)), but I don't assure that arbitrary metric satisfied with contractive.

Above inequality is hold for all ρ. Thus $E(VU, \mathcal{F} \circ \mathcal{E}) \leq E(V, \mathcal{F}) + E(U, \mathcal{E})$.

9.23) ($\Leftarrow$) If $\mathcal{E}(\rho_j) = \rho_j$ for all j such that $p_j > 0$, then

$$\bar{F} = \sum_j p_j F(\rho_j, \mathcal{E}(\rho_j))^2 = \sum_j p_j F(\rho_j, \rho_j)^2 = \sum_j p_j 1^2 = \sum_j p_j = 1.$$

($\Rightarrow$) Suppose $\mathcal{E}(\rho_j) \neq \rho_j$. Then $F(\rho_j, \mathcal{E}(\rho_j)) < 1$ ($\because$ text p.411, the fifth line from bottom). Thus

$$\bar{F} = \sum_j p_j F(\rho_j, \mathcal{E}(\rho_j))^2 < \sum_j p_j = 1.$$

Therefore if $\bar{F} = 1$, then $\mathcal{E}(\rho_j) = \rho_j$.

Problem 1)
Problem 2)

Problem 3) Theorem 5.3 of "Theory of Quantum Error Correction for General Noise", Emanuel Knill, Raymond Laflamme, and Lorenza Viola, Phys. Rev. Lett. 84, 2525 – Published 13 March 2000. arXiv:quant-ph/9604034 `https://arxiv.org/abs/quant-ph/9604034`

Chapter 11

Entropy and information

11.1) Fair coin:

$$H(1/2, 1/2) = \left(-\frac{1}{2}\log\frac{1}{2}\right) \times 2 = 1 \tag{11.1}$$

Fair die:

$$H(p) = \left(-\frac{1}{6}\log\frac{1}{6}\right) \times 6 = \log 6. \tag{11.2}$$

The entropy decreases if the coin or die is unfair.

11.2)
From assumption $I(pq) = I(p) + I(q)$.

$$\frac{\partial I(pq)}{\partial p} = \frac{\partial I(p)}{\partial p} + 0 = \frac{\partial I(p)}{\partial p} \tag{11.3}$$

$$\frac{\partial I(pq)}{\partial q} = 0 + \frac{\partial I(q)}{\partial q} = \frac{\partial I(q)}{\partial q} \tag{11.4}$$

$$\frac{\partial I(pq)}{\partial p} = \frac{\partial I(pq)}{\partial (pq)}\frac{\partial (pq)}{\partial p} = q\frac{\partial I(pq)}{\partial (pq)} \Rightarrow \frac{\partial I(pq)}{\partial (pq)} = \frac{1}{q}\frac{\partial I(p)}{\partial p} \tag{11.5}$$

$$\frac{\partial I(pq)}{\partial q} = \frac{\partial I(pq)}{\partial (pq)}\frac{\partial (pq)}{\partial q} = p\frac{\partial I(pq)}{\partial (pq)} \Rightarrow \frac{\partial I(pq)}{\partial (pq)} = \frac{1}{p}\frac{\partial I(q)}{\partial q} \tag{11.6}$$

Thus

$$\frac{1}{q}\frac{\partial I(p)}{\partial p} = \frac{1}{p}\frac{\partial I(q)}{\partial q} \tag{11.7}$$

$$\therefore\ p\frac{dI(p)}{dp} = q\frac{dI(q)}{dq} \qquad \text{for all } p, q \in [0,1]. \tag{11.8}$$

$$\tag{11.9}$$

Then $p(dI(p)/dp)$ is constant.
If $p(dI(p)/dp) = k$, $k \in \mathbb{R}$. Then $I(p) = k\ln p = k'\log p$ where $k' = k/\log e$.

11.3) $H_{\text{bin}}(p) = -p\log p - (1-p)\log(1-p)$.

$$\frac{dH_{\text{bin}}(p)}{dp} = \frac{1}{\ln 2}\left(-\log p - 1 + \log(1-p) + 1\right) \tag{11.10}$$
$$= \frac{1}{\ln 2}\ln\frac{1-p}{p} = 0 \tag{11.11}$$
$$\Rightarrow \frac{1-p}{p} = 1 \tag{11.12}$$
$$\Rightarrow p = 1/2. \tag{11.13}$$

11.4)
11.5)

$$H\left(p(x,y)||p(x)p(y)\right) = \sum_{x,y} p(x,y)\log\frac{p(x)p(y)}{p(x,y)} \tag{11.14}$$
$$= -H(p(x,y)) - \sum_{x,y} p(x,y)\log\left[p(x)p(y)\right] \tag{11.15}$$
$$= -H(p(x,y)) - \sum_{x,y} p(x,y)\left[\log p(x) + \log p(y)\right] \tag{11.16}$$
$$= -H(p(x,y)) - \sum_{x,y} p(x,y)\log p(x) - \sum_{x,y} p(x,y)\log p(y) \tag{11.17}$$
$$= -H(p(x,y)) - \sum_{x} p(x)\log p(x) - \sum_{y} p(y)\log p(y) \tag{11.18}$$
$$= -H(p(x,y)) + H(p(x)) + H(p(y)) \tag{11.19}$$
$$= -H(X,Y) + H(X) + H(Y). \tag{11.20}$$

From the non-negativity of the relative entropy,

$$H(X) + H(Y) - H(X,Y) \geq 0 \tag{11.21}$$
$$\therefore H(X) + H(Y) \geq H(X,Y). \tag{11.22}$$

11.6)

$$H(Y) + H(X,Y,Z) - H(X,Y) - H(Y,Z) = \sum_{x,y,z} p(x,y,z)\log\left(p(x,y)p(y,z)/p(y)p(x,y,z)\right) \tag{11.23}$$
$$\geq \frac{1}{\ln 2}\sum_{x,y,z} p(x,y,z)\left[1 - p(x,y)p(y,z)/p(y)p(x,y,z)\right] \tag{11.24}$$
$$= \frac{1-1}{\ln 2} = 0 \tag{11.25}$$

The equality occurs if and only if $p(x,y)p(y,z)/p(y)p(x,y,z) = 1$, which means a Markov chain condition of $Z \to Y \to X$; $p(x|y) = p(x|y,z)$
11.7)
11.8)
11.9)
11.10)
11.11)
11.12)

11.13)
11.14)
11.15)
11.16)
11.17)
11.18)
11.19)
11.20)
11.21)
11.22)
11.23)
11.24)
11.25)
11.26)

Problem 11.1)
Problem 11.2)
Problem 11.3)
Problem 11.4)
Problem 11.5)

12.31) Eve makes her qubits entangled with $|\beta_{00}\rangle$, and gets ρ^E.

$$|ABE\rangle = U\,|\beta_{00}^{\otimes n}\rangle\,|0\rangle_E \tag{11.26}$$

$$\rho^E = tr_{AB}(|ABE\rangle\,\langle ABE|) \tag{11.27}$$

Note that Eve's mutual information with Alice and Bob measurements does not depend on whether Eve measures ρ^E before Alice and Bob's measurement or after. So we can assume that Eve measures ρ^E after Alice and Bob's measurement. Alice and Bob measure their Bell state, getting binary string $\vec{k}$ as an outcome. Let ρ_k^E and p_k are the corresponding Eve's states and probabilities. Note,

$$\rho_E = \sum_k p_k \rho_k^E. \tag{11.28}$$

Let K is a variable of $\vec{k}$ and e is an outcom of a measurement of ρ^E, and E is its variable. From Holevo bound,

$$H(K:E) \leq S(\rho^E) - \sum_k p_k \rho_k^E \leq S(\rho^E) = S(\rho). \tag{11.29}$$

Chapter 1

Fundamental Concepts

1.1) Probabilistic Classical Deutsch-Jozsa Algorithm: 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?

Soln: To a mathematician, this problem is (*slightly*) under-specified. Missing is the probability that the function f in question is balanced, vice constant. We assume that both are **equally** likely, a priori. The results when all balanced or constant functions are chosen from randomly are significantly different, and likely less interesting. We describe *an* algorithm and analyze the error rate, but make no effort to show that it is the *best* algorithm, nor that this is the most effective analysis. Let C be the event that f is constant, and B be the event that it is balanced. By hypothesis $P(C) = P(B) = \frac{1}{2}$, a priori. Evaluating f provides information which can be used to update these prior probabilities. Classically evaluating the function once, say at x_0, provides no useful information, since comparison of values is at the heart of this problem. Evaluating f twice, say at x_0 and x_1, can unambiguously determine if f is balanced when their values disagree. So, let's assume they agree. We use Bayesian inference to iteratively update the probability that f is constant, given k successive measurements that agree. In a convenient abuse of notation, let $P(E \mid k) = P\big(E \mid f(x_0) = \cdots = f(x_{k-1})\big)$, $P(k \mid E) = P\big(f(x_0) = \cdots = f(x_{k-1}) \mid E\big)$, and $P(k) = P\big(f(x_0) = \cdots = f(x_{k-1})\big)$, for $E = B, C$, and $k \in \mathbb{N}$. We have $P(C \mid 0) = P(C \mid 1) = P(B \mid 0) = P(B \mid 1) = 1/2$. Note also that $P(k \mid C) = 1$, since if f is constant all evaluations (including the k in question) will agree. By Baye's theorem and the Law of Total Probability:

$$
\begin{aligned}
P(C \mid k) = &\frac{P(k \mid C) \cdot P(C \mid k-1)}{P(k)} \\
= &\frac{P(k \mid C) \cdot P(C \mid k-1)}{P(C \mid k-1) \cdot P(k \mid C) + P(B \mid k-1) \cdot P(k \mid B)}
\end{aligned}
$$

The formula above can be used to iteratively update $P(C, k)$, and hence $P(B, k) = 1 - P(C, k)$, but first we must discuss $P(k \mid B)$. It is important to note that when this quantity is used to update $P(C \mid k)$, it is already known with certainty that $f(x_0) = \cdots = f(x_{k-2})$, *i.e.* $P(k-1) = 1$. $P(k \mid B)$ is the probability that, given this information, evaluating f one more time, at x_{k-1}, yields another value in agreement with $f(x_0), \cdots, f(x_{k-2})$. We evaluate this by separating the two possible outcomes of evaluation and counting the number of balanced functions satisfying the hypotheses that would produce them. If $f(x_{k-1}) = f(x_0)$, then x_{k-1} is the k-th value on which f agrees. There are $\binom{n-k}{n/2-k}$ balanced functions which would produce this result, corresponding to the selections of $n/2 - k$ more of the remaining $n - k$ values on which f can agree. If $f(x_{k-1}) \neq f(x_0)$, then f must still agree on $n/2 - k + 1$ of the remaining $n - k$ values. There are $\binom{n-k}{n/2-k+1}$ balanced functions that would produce this result. So:

$$
P(k, B) = \frac{\binom{n-k}{n/2-k}}{\binom{n-k}{n/2-k} + \binom{n-k}{n/2-k+1}} = \frac{\binom{n-k}{n/2-k}}{\binom{n-k+1}{n/2-k+1}} = \frac{n/2 - k + 1}{n - k + 1} = \frac{n - 2k + 2}{2n - 2k + 2}
$$

We are finally in a position to calculate $P(C \mid k)$. Unfortunately, for fixed n, the machinery above does not produce formulas of bounded complexity as k grows. Each formula will be a rational function with equal degree in numerator and denominator, but those degrees seem to be $\lfloor k/2 \rfloor$. The coefficients of the leading terms show some structure that can be used for asymptotic analysis, which we do below. We illustrate the calculation of $P(C \mid 2)$, $P(C \mid 3)$, and $P(C \mid 4)$, and list formulas for $P(C \mid 5)$ through $P(C \mid 7)$, then discuss the results and some experimental confirmation.

$$
\begin{aligned}
P(C \mid 2) &= \frac{P(2 \mid C) \cdot P(C \mid 1)}{P(C \mid 1) \cdot P(2 \mid C) + P(B \mid 1) \cdot P(2 \mid B)} \\
&= \frac{1 \cdot \frac{1}{2}}{\frac{1}{2} \cdot 1 + \frac{1}{2} \cdot \frac{n-2}{2n-2}} \\
&= \frac{1}{1 + \frac{n-2}{2n-2}} \\
&= \frac{2n-2}{3n-4} \\
P(C \mid 3) &= \frac{P(3 \mid C) \cdot P(C \mid 2)}{P(C \mid 2) \cdot P(3 \mid C) + P(B \mid 2) \cdot P(3 \mid B)} \\
&= \frac{1 \cdot \frac{2n-2}{3n-4}}{\frac{2n-2}{3n-4} + \left(1 - \frac{2n-2}{3n-4}\right) \cdot \frac{n-4}{2n-4}} \\
&= \frac{4n-4}{5n-8} \\
P(C \mid 4) &= \frac{P(4 \mid C) \cdot P(C \mid 3)}{P(C \mid 3) \cdot P(4 \mid C) + P(B \mid 3) \cdot P(4 \mid B)} \\
&= \frac{1 \cdot \frac{4n-4}{5n-8}}{\frac{4n-4}{5n-8} + \left(1 - \frac{4n-4}{5n-8}\right) \cdot \frac{n-6}{2n-6}} \\
&= \frac{8n^2 - 32n + 24}{9n^2 - 42n + 48} \\
P(C \mid 5) &= \frac{16n^2 - 64n + 48}{17n^2 - 78n + 96} \\
P(C \mid 6) &= \frac{32n^3 - 288n^2 + 736n - 480}{33n^3 - 312n^2 + 924n - 960} \\
P(C \mid 7) &= \frac{64n^3 - 576n^2 + 1472n - 960}{65n^3 - 606n^2 + 1768n - 1920}
\end{aligned}
$$

There are clearly patterns, the most striking of which yields $P(C \mid k) \xrightarrow[n\to\infty]{} \frac{2^{k-1}}{2^{k-1}+1}$, that is, given $k \geq 2$ evaluations in agreement, the probability that f is constant is $\sim 1 - \frac{1}{2^{k-1}+1}$, at least for large n. In the quantum context, where n is likely to be exponential in the number of qubits, this asymptotic value would be approached rapidly. To confirm this analysis, a python script is included in the repo which experimentally calculates empirical values of $P(C \mid k)$ for specified values of n and k. It also calculates the theoretical values, recursing over k, for comparison. See `<git repo>/Python/Problem1.1.py`.

To answer the problem most directly, *i.e.*, "what is the performance of the best classical algorithm for this problem?", let n be fixed and $0 < \epsilon < \frac{1}{2}$ be specified. The "performance" of classically evaluating the function of n inputs in order to declare it constant with error less than ϵ is equivalent to determining the number k of evaluations in agreement after which the probability that f is constant is greater than $1 - \epsilon$. Note that in no case is this number less than two. The entries in the table below are such values, with rows indexed by n, and columns corresponding to exponentially decreasing values of ϵ. Specifically, column i lists the values of k corresponding to $\epsilon = 1/2^i$. The maximum value of k in each row is $n/2 + 1$, since this implies the function is constant.

$\epsilon = 1/2^i; i =$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
$n = 6$	2	3	3	4	→															
$n = 8$	2	3	4	4	4	5	→													
$n = 10$	2	3	4	4	5	5	6	→												
$n = 12$	2	3	4	4	5	5	6	6	7	→										
$n = 14$	2	3	4	5	5	6	6	7	7	7	8	→								
$n = 16$	2	3	4	5	5	6	6	7	7	8	8	8	9	→						
$n = 18$	2	3	4	5	5	6	7	7	8	8	8	9	9	9	10	→				
$n = 20$	2	3	4	5	6	6	7	7	8	8	9	9	9	10	10	10	11	→		
$n = 22$	2	3	4	5	6	6	7	7	8	9	9	9	10	10	11	11	11	11	12	→
$n = 24$	2	3	4	5	6	6	7	8	8	9	9	10	10	11	11	11	12	12	12	12
$n = 26$	2	3	4	5	6	6	7	8	8	9	9	10	10	11	11	12	12	12	13	13
$n = 28$	2	3	4	5	6	7	7	8	8	9	10	10	11	11	12	12	12	13	13	13
$n = 30$	2	3	4	5	6	7	7	8	9	9	10	10	11	11	12	12	13	13	13	14
$n = 32$	2	3	4	5	6	7	7	8	9	9	10	11	11	12	12	13	13	13	14	14
$n = 34$	2	3	4	5	6	7	7	8	9	9	10	11	11	12	12	13	13	14	14	15
$n = 36$	2	3	4	5	6	7	7	8	9	10	10	11	11	12	12	13	13	14	14	15
$n = 38$	2	3	4	5	6	7	8	8	9	10	10	11	12	12	13	13	14	14	15	15
$n = 40$	2	3	4	5	6	7	8	8	9	10	10	11	12	12	13	13	14	14	15	15
$n = 42$	2	3	4	5	6	7	8	8	9	10	10	11	12	12	13	14	14	15	15	16
$n = 44$	2	3	4	5	6	7	8	8	9	10	11	11	12	13	13	14	14	15	15	16
$n = 46$	2	3	4	5	6	7	8	8	9	10	11	11	12	13	13	14	14	15	16	16
$n = 48$	2	3	4	5	6	7	8	8	9	10	11	11	12	13	13	14	15	15	16	16
$n = 50$	2	3	4	5	6	7	8	9	9	10	11	11	12	13	13	14	15	15	16	16
$n = 52$	2	3	4	5	6	7	8	9	9	10	11	12	12	13	14	14	15	15	16	17
$n = 54$	2	3	4	5	6	7	8	9	9	10	11	12	12	13	14	14	15	16	16	17
$n = 56$	2	3	4	5	6	7	8	9	9	10	11	12	12	13	14	14	15	16	16	17
$n = 58$	2	3	4	5	6	7	8	9	9	10	11	12	12	13	14	15	15	16	16	17
$n = 60$	2	3	4	5	6	7	8	9	9	10	11	12	13	13	14	15	15	16	17	17
$n = 62$	2	3	4	5	6	7	8	9	10	10	11	12	13	13	14	15	15	16	17	17
$n = 64$	2	3	4	5	6	7	8	9	10	10	11	12	13	13	14	15	15	16	17	17

Loosely, for small numbers of evaluations and large n, *i.e* in the bottom left of the table, each exponential increase in the probability of being constant desired requires an additional evaluation. Eventually, the combinatorial reduction in the number of remaining balanced functions allows additonal evaluations to reduce the error with which the function can be declared constant by several powers of 2, as often seen in the top right. That is not to say that the probability is always reduced by at least a factor of 2. In fact, note that $P(C \mid 1) = 1/2$, and $P(C \mid 2) \xrightarrow[n\to\infty]{} 2/3$, so $\frac{1-P(C \mid 1)}{1-P(C \mid 2)} \xrightarrow[n\to\infty]{} 3/2$. The second evaluation only reduces the probability that the function is balanced by a factor of $\sim$1.5 for large n. Asymptotically, for $k \geq 2$, note that $\frac{1-P(C \mid k+1)}{1-P(C \mid k)} \xrightarrow[n\to\infty]{} \frac{\frac{1}{2^k+1}}{\frac{1}{2^{k-1}+1}} = \frac{2^{k-1}+1}{2^k+1} < 2$, so all k-th evaluations eventually reduce the probability of the function being balanced by less than a factor of 2, for large enough n. Theoretically, it is seemingly possible there's a case in which halving the probability of being balanced requires two additional evaluations. That is, there could exist n and an $\epsilon = \frac{1}{2^i}$ requiring k measurements to declare the function constant with error less than ϵ and at least $k+2$ measurements to declare the function constant with error less than $\epsilon/2$. However, attempts to search for such a pathological case have come up empty. The asymptotic short-fallings are overcome by the combinatorial reduction fast enough, before a power of 1/2 straddles two values of k. It is likely that more careful analysis could refute the possibility rigorously.

We finish discussion of this problem with a (perhaps unnecessary) table of values of $P(C \mid k)$ for fixed n and k (programmatically constructed with the python script mentioned above, as was the previous table.) Again, once $k = n/2 + 1$, the function must be constant, so all probabilities are 1.

k	2	3	4	5	6	7	8	9	10
$n=4$	$\frac{3}{4} \simeq 0.7500$	$\frac{1}{1} \simeq 1.0000$	$\rightarrow$						
$n=6$	$\frac{5}{7} \simeq 0.7143$	$\frac{10}{11} \simeq 0.9091$	$\frac{1}{1} \simeq 1.0000$	$\rightarrow$					
$n=8$	$\frac{7}{10} \simeq 0.7000$	$\frac{7}{8} \simeq 0.8750$	$\frac{35}{36} \simeq 0.9722$	$\frac{1}{1} \simeq 1.0000$	$\rightarrow$				
$n=10$	$\frac{9}{13} \simeq 0.6923$	$\frac{6}{7} \simeq 0.8571$	$\frac{21}{22} \simeq 0.9545$	$\frac{126}{127} \simeq 0.9921$	$\frac{1}{1} \simeq 1.0000$	$\rightarrow$			
$n=12$	$\frac{11}{16} \simeq 0.6875$	$\frac{11}{13} \simeq 0.8462$	$\frac{33}{35} \simeq 0.9429$	$\frac{66}{67} \simeq 0.9851$	$\frac{462}{463} \simeq 0.9978$	$\frac{1}{1} \simeq 1.0000$	$\rightarrow$		
$n=14$	$\frac{13}{19} \simeq 0.6842$	$\frac{26}{31} \simeq 0.8387$	$\frac{143}{153} \simeq 0.9346$	$\frac{143}{146} \simeq 0.9795$	$\frac{429}{431} \simeq 0.9954$	$\frac{1716}{1717} \simeq 0.9994$	$\frac{1}{1} \simeq 1.0000$	$\rightarrow$	
$n=16$	$\frac{15}{22} \simeq 0.6818$	$\frac{5}{6} \simeq 0.8333$	$\frac{13}{14} \simeq 0.9286$	$\frac{39}{40} \simeq 0.9750$	$\frac{143}{144} \simeq 0.9931$	$\frac{715}{716} \simeq 0.9986$	$\frac{6435}{6436} \simeq 0.9998$	$\frac{1}{1} \simeq 1.0000$	$\rightarrow$
$n=18$	$\frac{17}{25} \simeq 0.6800$	$\frac{34}{41} \simeq 0.8293$	$\frac{85}{92} \simeq 0.9239$	$\frac{34}{35} \simeq 0.9714$	$\frac{221}{223} \simeq 0.9910$	$\frac{442}{443} \simeq 0.9977$	$\frac{2431}{2432} \simeq 0.9996$	$\frac{24310}{24311} \simeq 1.0000$	$\frac{1}{1} \simeq 1.0000$
$n=20$	$\frac{19}{28} \simeq 0.6786$	$\frac{19}{23} \simeq 0.8261$	$\frac{323}{351} \simeq 0.9202$	$\frac{646}{667} \simeq 0.9685$	$\frac{646}{653} \simeq 0.9893$	$\frac{323}{324} \simeq 0.9969$	$\frac{4199}{4202} \simeq 0.9993$	$\frac{8398}{8399} \simeq 0.9999$	$\frac{92378}{92379} \simeq 1.0000$
$n=22$	$\frac{21}{31} \simeq 0.6774$	$\frac{14}{17} \simeq 0.8235$	$\frac{133}{145} \simeq 0.9172$	$\frac{57}{59} \simeq 0.9661$	$\frac{323}{327} \simeq 0.9878$	$\frac{1292}{1297} \simeq 0.9961$	$\frac{969}{970} \simeq 0.9990$	$\frac{4522}{4523} \simeq 0.9998$	$\frac{29393}{29394} \simeq 1.0000$
$n=24$	$\frac{23}{34} \simeq 0.6765$	$\frac{23}{28} \simeq 0.8214$	$\frac{161}{176} \simeq 0.9148$	$\frac{161}{167} \simeq 0.9641$	$\frac{437}{443} \simeq 0.9865$	$\frac{437}{439} \simeq 0.9954$	$\frac{7429}{7439} \simeq 0.9987$	$\frac{14858}{14863} \simeq 0.9997$	$\frac{14858}{14859} \simeq 0.9999$
$n=26$	$\frac{25}{37} \simeq 0.6757$	$\frac{50}{61} \simeq 0.8197$	$\frac{115}{126} \simeq 0.9127$	$\frac{230}{239} \simeq 0.9623$	$\frac{805}{817} \simeq 0.9853$	$\frac{575}{578} \simeq 0.9948$	$\frac{10925}{10943} \simeq 0.9984$	$\frac{2185}{2186} \simeq 0.9995$	$\frac{37145}{37149} \simeq 0.9999$
$n=28$	$\frac{27}{40} \simeq 0.6750$	$\frac{9}{11} \simeq 0.8182$	$\frac{225}{247} \simeq 0.9109$	$\frac{270}{281} \simeq 0.9609$	$\frac{690}{701} \simeq 0.9843$	$\frac{345}{347} \simeq 0.9942$	$\frac{1035}{1037} \simeq 0.9981$	$\frac{1725}{1726} \simeq 0.9994$	$\frac{6555}{6556} \simeq 0.9998$
$n=30$	$\frac{29}{43} \simeq 0.6744$	$\frac{58}{71} \simeq 0.8169$	$\frac{261}{287} \simeq 0.9094$	$\frac{261}{272} \simeq 0.9596$	$\frac{1305}{1327} \simeq 0.9834$	$\frac{1740}{1751} \simeq 0.9937$	$\frac{10005}{10027} \simeq 0.9978$	$\frac{10005}{10012} \simeq 0.9993$	$\frac{10005}{10007} \simeq 0.9998$
$n=32$	$\frac{31}{46} \simeq 0.6739$	$\frac{31}{38} \simeq 0.8158$	$\frac{899}{990} \simeq 0.9081$	$\frac{899}{938} \simeq 0.9584$	$\frac{8091}{8234} \simeq 0.9826$	$\frac{8091}{8146} \simeq 0.9932$	$\frac{4495}{4506} \simeq 0.9976$	$\frac{13485}{13496} \simeq 0.9992$	$\frac{310155}{310232} \simeq 0.9998$
$n=34$	$\frac{33}{49} \simeq 0.6735$	$\frac{22}{27} \simeq 0.8148$	$\frac{341}{376} \simeq 0.9069$	$\frac{2046}{2137} \simeq 0.9574$	$\frac{9889}{10071} \simeq 0.9819$	$\frac{1798}{1811} \simeq 0.9928$	$\frac{24273}{24338} \simeq 0.9973$	$\frac{5394}{5399} \simeq 0.9991$	$\frac{13485}{13489} \simeq 0.9997$
$n=36$	$\frac{35}{52} \simeq 0.6731$	$\frac{35}{43} \simeq 0.8140$	$\frac{77}{85} \simeq 0.9059$	$\frac{22}{23} \simeq 0.9565$	$\frac{682}{695} \simeq 0.9813$	$\frac{1705}{1718} \simeq 0.9924$	$\frac{4495}{4508} \simeq 0.9971$	$\frac{12586}{12599} \simeq 0.9990$	$\frac{37758}{37771} \simeq 0.9997$
$n=38$	$\frac{37}{55} \simeq 0.6727$	$\frac{74}{91} \simeq 0.8132$	$\frac{1295}{1431} \simeq 0.9050$	$\frac{259}{271} \simeq 0.9557$	$\frac{407}{415} \simeq 0.9807$	$\frac{1628}{1641} \simeq 0.9921$	$\frac{12617}{12656} \simeq 0.9969$	$\frac{11470}{11483} \simeq 0.9989$	$\frac{33263}{33276} \simeq 0.9996$
$n=40$	$\frac{39}{58} \simeq 0.6724$	$\frac{13}{16} \simeq 0.8125$	$\frac{481}{532} \simeq 0.9041$	$\frac{1443}{1511} \simeq 0.9550$	$\frac{3367}{3435} \simeq 0.9802$	$\frac{481}{485} \simeq 0.9918$	$\frac{1221}{1225} \simeq 0.9967$	$\frac{814}{815} \simeq 0.9988$	$\frac{2294}{2295} \simeq 0.9996$
$n=42$	$\frac{41}{61} \simeq 0.6721$	$\frac{82}{101} \simeq 0.8119$	$\frac{533}{590} \simeq 0.9034$	$\frac{1066}{1117} \simeq 0.9543$	$\frac{19721}{20129} \simeq 0.9797$	$\frac{19721}{19891} \simeq 0.9915$	$\frac{19721}{19789} \simeq 0.9966$	$\frac{1517}{1519} \simeq 0.9987$	$\frac{16687}{16695} \simeq 0.9995$
$n=44$	$\frac{43}{64} \simeq 0.6719$	$\frac{43}{53} \simeq 0.8113$	$\frac{1763}{1953} \simeq 0.9027$	$\frac{3526}{3697} \simeq 0.9537$	$\frac{45838}{46807} \simeq 0.9793$	$\frac{22919}{23123} \simeq 0.9912$	$\frac{848003}{851063} \simeq 0.9964$	$\frac{848003}{849193} \simeq 0.9986$	$\frac{65231}{65265} \simeq 0.9995$
$n=46$	$\frac{45}{67} \simeq 0.6716$	$\frac{30}{37} \simeq 0.8108$	$\frac{129}{143} \simeq 0.9021$	$\frac{387}{406} \simeq 0.9532$	$\frac{1763}{1801} \simeq 0.9789$	$\frac{35260}{35583} \simeq 0.9909$	$\frac{343785}{345077} \simeq 0.9963$	$\frac{22919}{22953} \simeq 0.9985$	$\frac{848003}{848479} \simeq 0.9994$
$n=48$	$\frac{47}{70} \simeq 0.6714$	$\frac{47}{58} \simeq 0.8103$	$\frac{705}{782} \simeq 0.9015$	$\frac{141}{148} \simeq 0.9527$	$\frac{6063}{6196} \simeq 0.9785$	$\frac{2021}{2040} \simeq 0.9907$	$\frac{82861}{83184} \simeq 0.9961$	$\frac{414305}{414951} \simeq 0.9984$	$\frac{1077193}{1077839} \simeq 0.9994$
$n=50$	$\frac{49}{73} \simeq 0.6712$	$\frac{98}{121} \simeq 0.8099$	$\frac{2303}{2556} \simeq 0.9010$	$\frac{658}{691} \simeq 0.9522$	$\frac{987}{1009} \simeq 0.9782$	$\frac{1974}{1993} \simeq 0.9905$	$\frac{14147}{14204} \simeq 0.9960$	$\frac{198058}{198381} \simeq 0.9984$	$\frac{4060189}{4062773} \simeq 0.9994$
$n=52$	$\frac{51}{76} \simeq 0.6711$	$\frac{17}{21} \simeq 0.8095$	$\frac{833}{925} \simeq 0.9005$	$\frac{4998}{5251} \simeq 0.9518$	$\frac{11186}{11439} \simeq 0.9779$	$\frac{5593}{5648} \simeq 0.9903$	$\frac{50337}{50546} \simeq 0.9959$	$\frac{11186}{11205} \simeq 0.9983$	$\frac{28294}{28313} \simeq 0.9993$
$n=54$	$\frac{53}{79} \simeq 0.6709$	$\frac{106}{131} \simeq 0.8092$	$\frac{901}{1001} \simeq 0.9001$	$\frac{901}{947} \simeq 0.9514$	$\frac{44149}{45161} \simeq 0.9776$	$\frac{25228}{25481} \simeq 0.9901$	$\frac{296429}{297694} \simeq 0.9958$	$\frac{592858}{593903} \simeq 0.9982$	$\frac{296429}{296638} \simeq 0.9993$
$n=56$	$\frac{55}{82} \simeq 0.6707$	$\frac{55}{68} \simeq 0.8088$	$\frac{583}{648} \simeq 0.8997$	$\frac{583}{613} \simeq 0.9511$	$\frac{9911}{10141} \simeq 0.9773$	$\frac{4505}{4551} \simeq 0.9899$	$\frac{31535}{31673} \simeq 0.9956$	$\frac{12614}{12637} \simeq 0.9982$	$\frac{592858}{593295} \simeq 0.9993$
$n=58$	$\frac{57}{85} \simeq 0.6706$	$\frac{38}{47} \simeq 0.8085$	$\frac{1045}{1162} \simeq 0.8993$	$\frac{1254}{1319} \simeq 0.9507$	$\frac{11077}{11337} \simeq 0.9771$	$\frac{11077}{11192} \simeq 0.9897$	$\frac{51357}{51587} \simeq 0.9955$	$\frac{85595}{85756} \simeq 0.9981$	$\frac{119833}{119925} \simeq 0.9992$
$n=60$	$\frac{59}{88} \simeq 0.6705$	$\frac{59}{73} \simeq 0.8082$	$\frac{1121}{1247} \simeq 0.8990$	$\frac{2242}{2359} \simeq 0.9504$	$\frac{24662}{25247} \simeq 0.9768$	$\frac{12331}{12461} \simeq 0.9896$	$\frac{653543}{656533} \simeq 0.9954$	$\frac{59413}{59528} \simeq 0.9981$	$\frac{1010021}{1010826} \simeq 0.9992$
$n=62$	$\frac{61}{91} \simeq 0.6703$	$\frac{122}{151} \simeq 0.8079$	$\frac{3599}{4005} \simeq 0.8986$	$\frac{3599}{3788} \simeq 0.9501$	$\frac{68381}{70019} \simeq 0.9766$	$\frac{273524}{276449} \simeq 0.9894$	$\frac{752191}{755701} \simeq 0.9954$	$\frac{752191}{753686} \simeq 0.9980$	$\frac{3624193}{3627183} \simeq 0.9992$
$n=64$	$\frac{63}{94} \simeq 0.6702$	$\frac{21}{26} \simeq 0.8077$	$\frac{1281}{1426} \simeq 0.8983$	$\frac{549}{578} \simeq 0.9498$	$\frac{3599}{3686} \simeq 0.9764$	$\frac{3599}{3638} \simeq 0.9893$	$\frac{68381}{68706} \simeq 0.9953$	$\frac{478667}{479642} \simeq 0.9980$	$\frac{5265337}{5269822} \simeq 0.9991$