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

Original author: goropikari
Extended by: tlesaul2

June 24, 2022

Copylight Notice:

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.

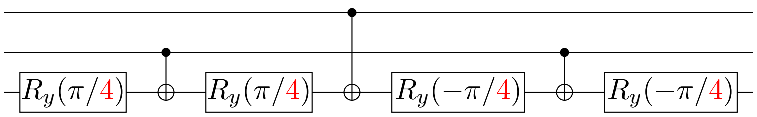
tlesaul2

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.103. Exercise 2.26. Show that the circuit:

  differs from a Toffoli gate only by relative phases.
- 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 $\mathrm{tr}(A^\dagger B) \neq \langle m|A \otimes B|m\rangle$.

By using following relation, we can prove.

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

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

- 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 basis $|0\rangle, |1\rangle$, 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

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

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

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

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

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

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

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:

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

Equation (2.13.2), conjugate symmetry:

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

Equation (2.13.3), positive definiteness:

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

Now:

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

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

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

Soln:

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

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

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

$$\frac{|w\rangle}{\||w\rangle\|} = \frac{|w\rangle}{\sqrt{\langle w|w\rangle}} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = |+\rangle$$

$$\frac{|v\rangle}{\||v\rangle\|} = \frac{|v\rangle}{\sqrt{\langle v|v\rangle}} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = |-\rangle$$

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 procedure normalize $|w_d\rangle$, 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 $|w_d\rangle$ 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 $|v_d\rangle$ 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 $1 \leq j \leq d - 1$. We show $\langle v_j|v_d\rangle = 0$, completing the proof.

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

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

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

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 $(|\psi_{y^+}\rangle \langle\psi_{y^+}|) - (|\psi_{y^-}\rangle \langle\psi_{y^-}|) = \begin{bmatrix} \frac{1}{2} & \frac{-i}{2} \\ \frac{i}{2} & \frac{1}{2} \end{bmatrix} - \begin{bmatrix} \frac{1}{2} & \frac{i}{2} \\ \frac{-i}{2} & \frac{1}{2} \end{bmatrix} (= Y)$.

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

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

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

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

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

On the other hand,

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

Thus

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

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

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 property 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.

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

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

Soln: We show that A has the defining property of the adjoint of $A^\dagger$.

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

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

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

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:

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

Now:

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

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 decomposition theorem. It's proof does *not* use the fact that a normal matrix is Hermitian 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

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

From this we have:

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

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 $|v\rangle$

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

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 $\theta (\in [0, 2\pi))$

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:

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

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

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 $M^\dagger$:

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

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 Hermitian.

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 orthonormal 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 $\sum_i \alpha_i\,|w_i\rangle$.

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

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

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

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?
Soln:

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

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

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

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 $i = i_1n_1 + i_2$ and $j = j_1m_1 + j_2$, with $0 \le i_2 < n_1$ and $0 \le j_2 < m_1$, then $(A\otimes B)_{ij} = A_{i_1j_1}B_{i_2j_2}$.

$$\begin{aligned}
(A\otimes B)^* &= [A_{i_1j_1}B_{i_2j_2}]^* && \text{(from above)}\\
&= \left[A^*_{i_1j_1}B^*_{i_2j_2}\right] && \text{(piecewise conjugation)}\\
&= A^*\otimes B^* && \text{(consistent indexing)}
\end{aligned}$$

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 $k = k_1m_1 + k_2$ and $\ell = \ell_1n_1 + \ell_2$, with $0 \le k_2 < m_2$ and $0 \le \ell_2 < n_2$, then $(A^T\otimes B^T)_{k\ell} = (A^T)_{k_1\ell_1}(B^T)_{k_2\ell_2} = A_{\ell_1k_1}B_{\ell_2k_2}$. 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:

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

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 multiplication of tensor products, let $|v\rangle$ and $|w\rangle$ be vectors in the spaces on which U_1 and U_2 operate. Then:

$$\begin{aligned}
(U_1\otimes U_2)(U_1\otimes U_2)^\dagger(|v\rangle\otimes|w\rangle) &= (U_1\otimes U_2)(U_1^\dagger\otimes U_2^\dagger)(|v\rangle\otimes|w\rangle) && \text{(distributivity of } ^\dagger)\\
&= (U_1\otimes U_2)(U_1^\dagger|v\rangle\otimes U_2^\dagger|w\rangle) && \text{(definition of tensor product of operators)}\\
&= U_1U_1^\dagger|v\rangle\otimes U_2U_2^\dagger|w\rangle && \text{(definition of tensor product of operators)}\\
&= I|v\rangle\otimes I|w\rangle && (U_1 \text{ and } U_2 \text{ are unitary})\\
&= (I\otimes I)(|v\rangle\otimes|w\rangle) && \text{(definition of tensor product of operators)}\\
&= I(|v\rangle\otimes|w\rangle) && (I\otimes I = I \text{ by construction})
\end{aligned}$$

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

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

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 $P_2 = \sum_{j=0}^{t} |w_j\rangle \langle w_j|$, where $|v_i\rangle_{i=0}^{k}$ is a subset of an orthonormal basis $|v_i\rangle_{i=0}^{\kappa}$, and $|w_j\rangle_{j=0}^{t}$ is a subset of an orthonormal basis $|w_j\rangle_{j=0}^{\tau}$, then $P_1 \otimes P_2 = \sum_{q=0}^{s} |r_q\rangle \langle r_q|$, where $|r_q\rangle_{q=0}^{s}$ is a subset of an orthonormal basis $|r_q\rangle_{q=0}^{\sigma}$. First, the fact that $P_1 \otimes P_2 = \sum\limits_{\substack{0\leq i\leq k\\0\leq j\leq t}} \big(|v_i\rangle \langle v_i|\big) \otimes \big(|w_j\rangle \langle w_j|\big)$ follows easily from distributivity of operator tensor products, having illustrated how easily that follows from distributivity of tensor products of vectors in exercise 2.29. We need to show that $\Big\{\big(|v_i\rangle \langle v_i|\big) \otimes \big(|w_j\rangle \langle w_j|\big)\Big\}_{\substack{0\leq i\leq k\\0\leq j\leq t}}$ is a subset of an orthonormal basis. It is automatically a subset of the set of vector tensor products resulting from loosening the restrictions on i and j to $0 \leq i \leq \kappa$ and $0 \leq j \leq \tau$, which we may assume is a basis, as stated on page 72. We need only show that the inner product of tensor products is multiplicative so that orthonormality is preserved. Let v_1, v_2, w_1 and w_2 be vectors.

$$\begin{aligned}\langle v_1 \otimes w_1|v_2 \otimes w_2\rangle &= |v_1 \otimes w_1\rangle^\dagger |v_2 \otimes w_2\rangle && \text{(definition of } \langle\cdot|\cdot\rangle)\\ &= \big(|v_1\rangle^\dagger \otimes |w_1\rangle^\dagger\big)\big(|v_2\rangle \otimes |w_2\rangle\big) && \text{(distributivity of } ^\dagger \text{ over } \otimes)\\ &= \big(|v_1\rangle^\dagger |v_2\rangle\big) \otimes \big(|w_1\rangle^\dagger |w_2\rangle\big) && \text{(mixed-product property of Kronecker product)}\\ &= \langle v_1|v_2\rangle \otimes \langle w_1|w_2\rangle && \text{(definition of } \langle\cdot|\cdot\rangle)\\ &= \langle v_1|v_2\rangle \langle w_1|w_2\rangle && (\langle\cdot|\cdot\rangle \text{ is a scalar})\end{aligned}$$

So, in the basis $\Big\{\big(|v_i\rangle \langle v_i|\big) \otimes \big(|w_j\rangle \langle w_j|\big)\Big\}_{\substack{0\leq i\leq \kappa\\0\leq j\leq \tau}}$, the inner product of two vectors $|v_{i_1}\rangle \otimes |w_{j_1}\rangle$ and $|v_{i_2}\rangle \otimes |w_{j_2}\rangle$ is $\langle v_{i_1} \otimes w_{j_1}|v_{i_2} \otimes w_{j_2}\rangle = \langle v_{i_1}|v_{i_2}\rangle \langle w_{j_1}|w_{j_2}\rangle = \delta_{i_1 i_2}\delta_{j_1 j_2}$, from which it follows that this basis is orthonormal, completing the proof.

2.33) The Hadamard operator on one qubit may be written as

$$H = \frac{1}{\sqrt{2}}\Big[\big(|0\rangle + |1\rangle\big) \langle 0| + \big(|0\rangle - |1\rangle\big) \langle 1|\Big].$$

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

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

Write out an explicit matrix representation for $H^{\otimes 2}$.
Soln: It is important to note what is meant by $x \cdot y$ in this formula. Here $\cdot$ does **not** mean integer multiplication. It can be taken to mean popparity of the binary AND of x and y. Note that this property is multiplicative across dimensions, when 1 is used for even popparity, and -1 for odd. We proceed by

induction on n. Assume the preceding formula for $n-1$. We must prove the formula for n.

$$\begin{aligned}
H^{\otimes n} &= H \otimes H^{\otimes n-1} && \text{(notation)}\\
&= \left(\frac{1}{\sqrt{2}}\Big[(|0\rangle + |1\rangle)\langle 0| + (|0\rangle - |1\rangle)\langle 1|\Big]\right) \otimes \left(\frac{1}{\sqrt{2^{n-1}}}\sum_{x,y}(-1)^{x\cdot y}|x\rangle\langle y|\right) && \text{(hypothesis)}\\
&= \frac{1}{\sqrt{2^n}}\sum_{x,y}(-1)^{x\cdot y}(|0\rangle\langle 0| + |1\rangle\langle 0| + |0\rangle\langle 1| - |1\rangle\langle 1|) \otimes |x\rangle\langle y| && \text{(rearrange)}\\
&= \frac{1}{\sqrt{2^n}}\sum_{x',y'}(-1)^{x'\cdot y'}|x'\rangle\langle y'| && \text{(multiplicativity of $\cdot$ across dimensions)}
\end{aligned}$$

Now, explcitly

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

and

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

2.34) Find the square root and logarithm of the matrix

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

Soln: Suppose $A = \begin{bmatrix} 4 & 3 \\ 3 & 4 \end{bmatrix}$. We will need the "spectral" decomposition A. First, to find the eigenvalues of A:

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

So, the eigenvalues of A are $\lambda = 1$, and $\lambda = 7$. The corresponding eigenvectors can easily be seen to be $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$, corresponding to $\lambda = 1$, and $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, corresponding to $\lambda = 7$. To construct an orthonormal basis from these eigenvectors, we can scale both by $\frac{1}{\sqrt{2}}$, and denote these scaled vectors by $|\lambda = 1\rangle = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and $|\lambda = 7\rangle = \frac{1}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Now, seeing as A is real and self-transpose/adjoint, it is a normal matrix/operator, and as such, A can be "spectrally decomposed"/diagonalized as:

$$\begin{aligned}
A &= |\lambda = 1\rangle\langle\lambda = 1| + 7\,|\lambda = 7\rangle\langle\lambda = 7|\\
&= \frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} + 7\left(\frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\right).
\end{aligned}$$

The square root of A is "defined" as:

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

and $log(A)$ is defined as

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

Note that one would hope that operator functions would respect various properties of the functions, such as inverses. To formally prove such a thing, let us consider the diagonal representation $A = \sum_a a\,|\lambda = a\rangle\langle\lambda = a|$ and the induced definition of $f(A) = \sum_a f(a)\,|\lambda = a\rangle\langle\lambda = a|$. Note that the $f(a)$ are eigenvalues of $f(A)$, with corresponding eigenvectors $|\lambda = a\rangle$, since the fact that the $|\lambda = a\rangle$ are orthonormal gives $\left(\sum_a f(a)\,|\lambda = a\rangle\langle\lambda = a|\right)|\lambda = a'\rangle = f(a')\,|\lambda = a'\rangle$. So $\sum_a f(a)\,|\lambda = a\rangle\langle\lambda = a|$ is a diagonal representation of $f(a)$. Now $f^-1(f(A)) = f^{-1}(\sum_a f(a)\,|\lambda = a\rangle\langle\lambda = a|) = \sum_a f^{-1}(f(a))\,|\lambda = a\rangle\langle\lambda = a| = \sum_a a\,|\lambda = a\rangle\langle\lambda = a| = A$.

2.35) (Exponentiation of Pauli Matrices) let $\vec{v}$ be any real three-dimensional unit vector and θ a real number. Prove that

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

where $\vec{v}\cdot\vec{\sigma} \equiv \sum_{i=1}^{3} v_i\sigma_i$. This exercise is generalized in problem 2.1 on page 117.
Soln: To find the eigenvalues of $\vec{v}\cdot\vec{\sigma}$, we first express in matrix form:

$$\begin{aligned}\vec{v}\cdot\vec{\sigma} &= \sum_{i=1}^{3} v_i\sigma_i(= v_1\sigma_x + v_2\sigma_y + v_3\sigma_z = v_1X + v_2Y + v_3Z) && \text{(definition)} \\ &= v_1\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + v_2\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} + v_3\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} && \text{(substitute)} \\ &= \begin{bmatrix} v_3 & v_1 - iv_2 \\ v_1 + iv_2 & -v_3 \end{bmatrix} && \text{(collect terms)} \\ 0 = \det(\vec{v}\cdot\vec{\sigma} - \lambda I) &= (v_3 - \lambda)(-v_3 - \lambda) - (v_1 - iv_2)(v_1 + iv_2) && \text{(characteristic equation)} \\ &= \lambda^2 - (v_1^2 + v_2^2 + v_3^2) && \text{(expand and collect } v\text{'s)} \\ &= \lambda^2 - 1 && (v \text{ is a unit vector})\end{aligned}$$

Solving yields that the eigenvalues of $\vec{v}\cdot\vec{\sigma}$ are $\lambda = \pm 1$. Let $|\lambda_{\pm 1}\rangle$ be eigenvectors with eigenvalues ± 1. Since $\vec{v}$ is a real valued vector, it is easily seen that $\vec{v}\cdot\vec{\sigma}$ is Hermitian, and so is diagonalizable, and we may take the $|\lambda_{\pm 1}\rangle$ to be orthonormal by Exercise 2.22. In particular, this gives

$$\begin{aligned}|\lambda_1\rangle\langle\lambda_1| + |\lambda_{-1}\rangle\langle\lambda_{-1}| &= I && \text{(completeness)} \\ |\lambda_1\rangle\langle\lambda_1| - |\lambda_{-1}\rangle\langle\lambda_{-1}| &= \vec{v}\cdot\vec{\sigma} && \text{(diagonalization)}\end{aligned}$$

Now

$$\begin{aligned}\exp(i\theta\vec{v}\cdot\vec{\sigma}) &= e^{i\theta}\,|\lambda_1\rangle\langle\lambda_1| + e^{-i\theta}\,|\lambda_{-1}\rangle\langle\lambda_{-1}| && \text{(definition)} \\ &= (\cos\theta + i\sin\theta)\,|\lambda_1\rangle\langle\lambda_1| + (\cos\theta - i\sin\theta)\,|\lambda_{-1}\rangle\langle\lambda_{-1}| && \text{(Euler's formula)} \\ &= \cos\theta(|\lambda_1\rangle\langle\lambda_1| + |\lambda_{-1}\rangle\langle\lambda_{-1}|) + i\sin\theta(|\lambda_1\rangle\langle\lambda_1| - |\lambda_{-1}\rangle\langle\lambda_{-1}|) && \text{(group cos and sin terms)} \\ &= \cos(\theta)I + i\sin(\theta)\vec{v}\cdot\vec{\sigma}. && \text{(from above)}\end{aligned}$$

2.36) Show that the Pauli matrices except for I have trace zero.
Soln:

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

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

$$\operatorname{tr}(AB) = \operatorname{tr}(BA).$$

$$\begin{aligned}
\operatorname{tr}(AB) &= \sum_i \langle i|AB|i\rangle && \text{(using matrix representation of } AB\text{)}\\
&= \sum_i \langle i|AIB|i\rangle && \text{(insert } I\text{)}\\
&= \sum_{i,j} \langle i|A|j\rangle \langle j|B|i\rangle && \text{(completeness: } I = \textstyle\sum_j |j\rangle\langle j|\text{)}\\
&= \sum_{i,j} \langle j|B|i\rangle \langle i|A|j\rangle && \text{(commutativity in } \mathbb{C}\text{)}\\
&= \sum_j \langle j|BA|j\rangle && \text{(completeness: } I = \textstyle\sum_i |i\rangle\langle i|\text{)}\\
&= \operatorname{tr}(BA) && \text{(using matrix representation of } BA\text{)}
\end{aligned}$$

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

$$\operatorname{tr}(A+B) = \operatorname{tr}(A) + \operatorname{tr}(B)$$

and if z is an arbitrary complex number show that

$$\operatorname{tr}(zA) = z\operatorname{tr}(A).$$

Soln:

$$\begin{aligned}
\operatorname{tr}(A+B) &= \sum_i \langle i|A+B|i\rangle && \text{(using matrix representation of } A+B\text{)}\\
&= \sum_i (\langle i|A|i\rangle + \langle i|B|i\rangle) && \text{(llinearity of } \langle\cdot|\cdot|\cdot\rangle\text{)}\\
&= \sum_i \langle i|A|i\rangle + \sum_i \langle i|B|i\rangle && \text{(separate terms)}\\
&= \operatorname{tr}(A) + \operatorname{tr}(B). && \text{(using matrix representation of } A \text{ and } B\text{)}
\end{aligned}$$

$$
\begin{aligned}
\operatorname{tr}(zA) &= \sum_i \langle i|zA|i\rangle && \text{(matrix representation)}\\
&= \sum_i z\,\langle i|A|i\rangle && \text{(linearity)}\\
&= z\sum_i \langle i|A|i\rangle && \text{(linearity of sum)}\\
&= z\operatorname{tr}(A). && \text{(matrix representation)}
\end{aligned}
$$

2.39) (The Hilbert-Schmidt inner product on operators) The set L_V of linear operators on a Hilbert space V is a vector space. An important additional result is that the vector space L_V can be given a natural inner product structure, turning it into a Hilbert space.
(1) Show that the function $(\cdot,\cdot)$ on $L_V \times L_V$ defined by

$$(A, B) \equiv \operatorname{tr}(A^\dagger B)$$

is an inner product function. This inner product is known as the *Hilbert-Schmidt* or *trace* inner product.
(2) If V has d dimensions, show that L_V has dimension d^2.
(3) Find an orthonormal basis of Hermitian matrices for the Hilbert space L_V.
Soln: (1) We check the three properties of inner products on page 65:
(i) linearity in the second argument:

$$
\begin{aligned}
\left(A, \sum_i \lambda_i B_i\right) &= \operatorname{tr}\left(A^\dagger\left(\sum_i \lambda_i B_i\right)\right) && \text{(definition of $(\cdot,\cdot)$)}\\
&= \operatorname{tr}\left(\sum_i \lambda_i A^\dagger B_i\right) && \text{(linearity in L_V)}\\
&= \sum_i \lambda_i \operatorname{tr}(A^\dagger B_i) && \text{(linearity of tr from Exercise 2.38)}\\
&= \sum_i \lambda_i (A, B_i) \quad \square && \text{(definition of $(\cdot,\cdot)$)}
\end{aligned}
$$

(ii) conjugate symmetry:

$$
\begin{aligned}
(A,B)^* &= \left(\operatorname{tr}(A^\dagger B)\right)^* && \text{(definition)}\\
&= \left(\sum_{i,j} \langle i|A^\dagger|j\rangle\,\langle j|B|i\rangle\right)^* && \text{(matrix representation, insert I, apply completeness)}\\
&= \sum_{i,j} \langle j|B|i\rangle^*\,\langle i|A^\dagger|j\rangle^* && \text{(distributivity and multiplicativity of *, commutativity in $\mathbb{C}$)}\\
&= \sum_{i,j} |j\rangle^{\dagger *}\,B^*\,|i\rangle^*\,|i\rangle^{\dagger *}\,A^{\dagger *}\,|j\rangle^* && \text{(distribute conjugation and make †s explicit)}\\
&= \sum_{i,j} |j\rangle^T\,B^{\dagger T}\,|i\rangle^{\dagger T}\,|i\rangle^T\,A^T\,|j\rangle^{\dagger T} && \text{($^{\dagger *} = {}^T$, $^* = {}^{\dagger T}$)}\\
&= \sum_{i,j} (|i\rangle^\dagger\,B^\dagger\,|j\rangle)^T(|j\rangle^\dagger\,A\,|i\rangle)^T && \text{(^{T} is cyclic)}\\
&= \sum_{i,j} \langle i|B^\dagger|j\rangle\,\langle j|A|i\rangle && \text{(definition of $\langle\cdot|\cdot|\cdot\rangle$, transpose in $\mathbb{C}$)}\\
&= \sum_i \langle i|B^\dagger A|i\rangle && \text{(completeness)}\\
&= \operatorname{tr}(B^\dagger A) = (B,A) && \text{(matrix representation, definition of $(\cdot,\cdot)$)}
\end{aligned}
$$

(iii) positivity:

$$\begin{aligned}(A,A) &= \mathrm{tr}(A^\dagger A) && \text{(definition)}\\ &= \sum_i \langle i|A^\dagger A|i\rangle && \text{(matrix representation)}\\ &\geq 0 && (A^\dagger \text{ A is positive by Exercise 2.25)}\end{aligned}$$

We are left only show that if $(A,A)=0$, then $A=0$. Note that $(A,A)=0$ implies that $\langle i|A^\dagger A|i\rangle = 0$ for all i, where here, the $|i\rangle$ are an orthonormal *basis* with respect to which the matrix representation of A is constructed. Note that $A\,|i\rangle$ is the i-th column of A, as constructed on page 64. Also $\|A\,|i\rangle\|^2 = \langle i|A^\dagger A|i\rangle = 0$, so the i-th column of A is 0, for all i. This gives that $(A,A)=0$ iff $A=0$, completing the proof of positivity.

(2) To show that L_V has dimension d^2, note that the elements of L_V are linear operators from V to V and thus have a matrix representation by a $d\times d$ matrix. The vector space of $d\times d$ matrices over $\mathbb{C}$ is clearly at most d^2-dimensional. An easily constructed basis is the set of matrices populated with all 0s except for a single 1, where the position of the 1's ranges over all d^2 positions. This set is clearly linearly independent and spans L_V, from which dimension d^2 follows.. More generally, if $|k\rangle_k = 0^{d-1}$ is any orthonormal basis for V, then $\{|k\rangle\langle\ell|\}_{k,\ell}$ is a basis. The first example is produced used the standard (computational) basis.
(3) The obvious choices of basis matrices discussed above are orthonormal, but not Hermitian. Note that for $d=2$, the Pauli matrices are orthonormal with respect to the Hilbert-Schmidt inner product, and are Hermitian. Other bases can be constructed using a "discrete Weyl system". See Example 1.6 on page 11 of the lecture notes titled "Mathematical Introduction to Quantum Information Processing", from Professor Michael Wolf of the Technical University of Munich (`https://www-m5.ma.tum.de/foswiki/pub/M5/Allgemeines/MA5057_2019S/QIPlecture.pdf`), and a convenient diagram included in "Holevo Capacity of Discrete Weyl Channels", by Rehman et. al (`https://www.nature.com/articles/s41598-018-35777-7`). Unfortunately, neither of those constructions are Hermitian. To construct an orthonormal basis of Hermitian matrices, we use a related construction. First, let $|k\rangle_{k=0}^{d-1}$ be an orthonormal basis for V. The matrices we construct, $U_{k,\ell}$, will be doubly indexed by elements of the basis. When $|k\rangle_{k=0}^{d-1}$ is the standard basis, each will have two non-zero entries, unless $k=\ell$, in which case there will be a single non-zero entry. Define

$$\alpha_{k,\ell} := \begin{cases} \frac{1}{\sqrt{2}} & k<\ell\\ \frac{1}{2} & k=\ell\\ \frac{i}{\sqrt{2}} & k>\ell\end{cases}$$

Now let $U_{k,\ell} = \alpha_{k,\ell}\,|k\rangle\langle\ell| + \alpha^*_{k,\ell}\,|\ell\rangle\langle k|$. When $k=\ell$, $|k\rangle\langle\ell|$ and $|\ell\rangle\langle k|$ coincide, placing $\alpha_{k,\ell}+\alpha^*_{k,\ell}=1$ somewhere along the diagonal. More generally, for fixed k,ℓ

$$\begin{aligned}U^\dagger_{k,\ell} &= \Big(\alpha_{k,\ell}\,|k\rangle\langle\ell| + \alpha^*_{k,\ell}\,|\ell\rangle\langle k|\Big)^\dagger && \text{(definition)}\\ &= \alpha^*_{k,\ell}\,\langle\ell|^\dagger\,|k\rangle^\dagger + \alpha_{k,\ell}\,\langle\ell|^\dagger\,|k\rangle^\dagger && (\text{properties of }{}^\dagger)\\ &= \alpha^*_{k,\ell}\,|\ell\rangle\langle k| + \alpha_{k,\ell}\,|k\rangle\langle\ell| && (\text{simplify }{}^\dagger)\\ &= U_{k,\ell},\end{aligned}$$

so each $U_{k,\ell}$ is Hermitian. To show they are orthonormal, consider the inner-product of two arbitrary matrices $U_{k,\ell}$ and $U_{n.m}$.

$$\begin{aligned}(U_{k,\ell},U_{n,m}) &= \mathrm{tr}(U^\dagger_{k,\ell}U_{n,m}) && (\text{definition of }(\cdot,\cdot))\\ &= \mathrm{tr}(U_{k,\ell}U_{n,m}) && (\text{Hermiticity of }U_{k,\ell})\\ &= \mathrm{tr}\bigg(\Big(\alpha_{k,\ell}\,|k\rangle\langle\ell| + \alpha^*_{k,\ell}\,|\ell\rangle\langle k|\Big)\Big(\alpha_{n,m}\,|n\rangle\langle m| + \alpha^*_{n,m}\,|m\rangle\langle n|\Big)\bigg) && (\text{definition of }U_{i,j})\end{aligned}$$

$$\begin{aligned}
&= \operatorname{tr}\begin{pmatrix} & \alpha_{k,\ell}\alpha_{n,m}\left|k\right\rangle\left\langle \ell|n\right\rangle\left\langle m\right| & + & \alpha_{k,\ell}\alpha^*_{n,m}\left|k\right\rangle\left\langle \ell|m\right\rangle\left\langle n\right| \\ + & \alpha^*_{k,\ell}\alpha_{n,m}\left|\ell\right\rangle\left\langle k|n\right\rangle\left\langle m\right| & + & \alpha^*_{k,\ell}\alpha^*_{n,m}\left|\ell\right\rangle\left\langle k|m\right\rangle\left\langle n\right| \end{pmatrix} && \text{(F.O.I.L.)} \\
&= \begin{matrix} & \alpha_{k,\ell}\alpha_{n,m}\delta_{\ell,n}\operatorname{tr}(\left|k\right\rangle\left\langle m\right|) & + & \alpha_{k,\ell}\alpha^*_{n,m}\delta_{\ell,m}\operatorname{tr}(\left|k\right\rangle\left\langle n\right|) \\ + & \alpha^*_{k,\ell}\alpha_{n,m}\delta_{k,n}\operatorname{tr}(\left|\ell\right\rangle\left\langle m\right|) & + & \alpha^*_{k,\ell}\alpha^*_{n,m}\delta_{k,m}\operatorname{tr}(\left|\ell\right\rangle\left\langle n\right|) \end{matrix} && \text{(linearity, orthonormality)} \\
&= \begin{matrix} & \alpha_{k,\ell}\alpha_{n,m}\delta_{\ell,n}\operatorname{tr}(\left\langle m|k\right\rangle) & + & \alpha_{k,\ell}\alpha^*_{n,m}\delta_{\ell,m}\operatorname{tr}(\left\langle n|k\right\rangle) \\ + & \alpha^*_{k,\ell}\alpha_{n,m}\delta_{k,n}\operatorname{tr}(\left\langle m|\ell\right\rangle) & + & \alpha^*_{k,\ell}\alpha^*_{n,m}\delta_{k,m}\operatorname{tr}(\left\langle n|\ell\right\rangle) \end{matrix} && \text{(cyclicity of tr)} \\
&= \begin{matrix} & \alpha_{k,\ell}\alpha_{n,m}\delta_{\ell,n}\delta_{k,m} & + & \alpha_{k,\ell}\alpha^*_{n,m}\delta_{\ell,m}\delta_{k,n} \\ + & \alpha^*_{k,\ell}\alpha_{n,m}\delta_{k,n}\delta_{\ell,m} & + & \alpha^*_{k,\ell}\alpha^*_{n,m}\delta_{k,m}\delta_{\ell,n} \end{matrix} && \text{(orthonormality)} \\
&= \begin{matrix} & (\alpha_{k,\ell}\alpha_{n,m} + (\alpha_{k,\ell}\alpha_{n,m})^*)\delta_{k,m}\delta_{\ell,n} \\ + & (\alpha_{k,\ell}\alpha^*_{n,m} + (\alpha_{k,\ell}\alpha^*_{n,m})^*)\delta_{k,n}\delta_{\ell,m} \end{matrix} && \text{(group like δs, property of *)} \\
&= \begin{matrix} & 2\,\mathfrak{Re}(\alpha_{k,\ell}\alpha_{n,m})\delta_{k,m}\delta_{\ell,n} \\ + & 2\,\mathfrak{Re}(\alpha_{k,\ell}\alpha^*_{n,m})\delta_{k,n}\delta_{\ell,m} \end{matrix} && \text{(property of *)}
\end{aligned}$$

When $k \neq n$ or $\ell \neq m$ only the first term contributes, so $(U_{k,\ell}, U_{n,m}) = 2\,\mathfrak{Re}(\alpha_{k,\ell}\alpha_{n,m})\delta_{k,m}\delta_{\ell,n}$. This can only be nonzero if $m = k$ and $n = \ell$, in which case $(U_{k,\ell}, U_{n,m}) = (U_{k,\ell}, U_{\ell,k}) = 2\,\mathfrak{Re}(\alpha_{k,\ell}\alpha_{\ell,k}) = 2\,\mathfrak{Re}(\frac{i}{2}) = 0$, since $k \neq n = \ell$ implies one of the $\alpha_{k,\ell}$ and $\alpha_{\ell,k}$ is $\frac{1}{\sqrt{2}}$, and the other $\frac{i}{\sqrt{2}}$. Hence, the $U_{k,\ell}$ are orthogonal. If $k = n$ and $\ell = m$, then $(U_{k,\ell}, U_{k,\ell}) = 2\,\mathfrak{Re}(\alpha_{k,\ell}\alpha_{k,\ell})\delta^2_{k,\ell} + 2\,\mathfrak{Re}(\alpha_{k,\ell}\alpha^*_{k,\ell})$. There are two cases. If $k = \ell$, then $\alpha_{k,\ell} = \alpha^*_{k,\ell} = \frac{1}{2}$, so $(U_{k,\ell}, U_{k,\ell}) = 2(\frac{1}{4}) + 2(\frac{1}{4}) = 1$. When $k \neq \ell$, $(U_{k,\ell}, U_{k,\ell}) = 2\,\mathfrak{Re}(\alpha_{k,\ell}\alpha^*_{k,\ell}) = 2\,\mathfrak{Re}(\|\alpha_{k,\ell}\|^2) = 2(\frac{1}{2}) = 1$. So in all cases $(U_{k,\ell}, U_{k,\ell}) = 1$ and thus the $U_{k,\ell}$ are a set of d^2 Hermitian matrices which are orthonormal. Orthonormality implies linear independence, in which case the subspace spanned by the $U_{k,\ell}$ has dimension d^2 and so must be L_V itself, making $U_{k,\ell}$ a basis.

2.40) (Commutation relations for the Pauli matrices) Verify the commutation relations

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

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

$$[\sigma_j, \sigma_k] = 2i\sum_{\ell=1}^{3}\epsilon_{jk\ell}\sigma_\ell$$

.

Soln:

$$\begin{aligned}
[\sigma_1, \sigma_2] = [X, Y] &= XY - YX \\
&= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} - \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\
&= \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} - \begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix} \\
&= \begin{bmatrix} 2i & 0 \\ 0 & -2i \end{bmatrix} \\
&= 2iZ = 2i\epsilon_{12}\sigma_3
\end{aligned}$$

$$\begin{aligned}
[\sigma_2, \sigma_3] = [Y, Z] &= \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 2i \\ 2i & 0 \end{bmatrix} \\
&= 2iX = 2i\epsilon_{231}\sigma_1
\end{aligned}$$

$$\begin{aligned}
[\sigma_3, \sigma_1] = [Z, X] &= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 2 \\ -2 & 0 \end{bmatrix} \\
&= 2i \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \\
&= 2iY = 2i\epsilon_{312}\sigma_2
\end{aligned}$$

2.41) (Anti-commutation relations for the Pauli Matrices) Verify that the anticommutation relations

$$\sigma_i, \sigma_j = 0$$

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

$$\sigma_i^2 = I.$$

Soln:

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

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

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

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

2.42) Verify that

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

Soln:

$$\frac{[A,B] + \{A,B\}}{2} = \frac{AB - BA + AB + BA}{2} = AB \qquad \text{(definition of } [\cdot,\cdot] \text{ and } \{\cdot,\cdot\}\text{)}$$

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

$$\sigma_j \sigma_k = \delta_{jk} I + i \sum_{\ell=1}^{3} \epsilon_{jk\ell} \sigma_\ell.$$

From Exercises 2.41 (eq (2.75), (2.76)), $\{\sigma_j, \sigma_k\} = 2\delta_{jk} I$.

$$
\begin{aligned}
\sigma_j \sigma_k &= \frac{[\sigma_j, \sigma_k] + \{\sigma_j, \sigma_k\}}{2} && \text{(Exercise 2.42 (eq 2.77))} \\
&= \frac{2i \sum_{l=1}^{3} \epsilon_{jkl} \sigma_l + 2\delta_{jk} I}{2} && \text{(Exercise 2.40 (eq 2.74) and above)} \\
&= \delta_{jk} I + i \sum_{\ell=1}^{3} \epsilon_{jk\ell} \sigma_\ell && \text{(cancel 2s)}
\end{aligned}
$$

2.44) Suppose $[A,B] = 0$, $\{A,B\} = 0$, and A is invertible. Show that B must be 0.
Soln: By assumption, $[A,B] = AB - BA = 0$ implies $AB = BA$, now $\{A,B\} = AB + BA = 2AB = 0$, so $AB = 0$. Since A is invertible, multiplying by A^{-1} from the left gives

$$
\begin{aligned}
A^{-1}AB &= 0 && \text{(A is invertible, multiply both sides by } A^{-1}\text{)} \\
IB &= 0 && (A^{-1}A = I) \\
B &= 0.
\end{aligned}
$$

2.45) Show that $[A,B]^\dagger = [B^\dagger, A^\dagger]$.
Soln:

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

2.46) Show that $[A, B] = -[B, A]$.
Soln:

$$\begin{aligned} [A, B] &= AB - BA && \text{(definition of } [\cdot,\cdot]) \\ &= -(BA - AB) && \text{(reverse signs)} \\ &= -[B, A] && \text{(definition of } [\cdot,\cdot]) \end{aligned}$$

2.47) Suppose A and B are Hermitian. Show the $i[A, B]$ is Hermitian.
Soln:

$$\begin{aligned} (i[A, B])^\dagger &= -i[A, B]^\dagger && \text{(distribute } ^\dagger) \\ &= -i\left[B^\dagger, A^\dagger\right] && \text{(Exercise 2.45)} \\ &= -i[B, A] && (A \text{ and } B \text{ are Hermitian}) \\ &= i[A, B] && \text{(Exercise 2.46)} \end{aligned}$$

2.48) What is the polar decomposition of positive matrix P? Of a unitary matrix U? Of a Hermitian matrix H?
Soln:

(Positive) Since P is positive, it is Hermitian by Exercise 2.24, thus normal, so the spectral decomposition theorem gives that it is diagonalizable. Then $P = \sum_i \lambda_i |i\rangle\langle i|$, where $\lambda_i \geq 0$ by positivity. Note that this implies that $\sqrt{\lambda_i^2} = \lambda_i$.

$$\begin{aligned} J &= \sqrt{P^\dagger P} && \text{(uniqueness of } J) \\ &= \sqrt{P^2} && \text{(Hermiticity)} \\ &= \sum_i \sqrt{\lambda_i^2}\, |i\rangle\langle i| && \text{(spectral decomposition, definition of } \sqrt{\cdot}) \\ &= \sum_i \lambda_i |i\rangle\langle i| && (\sqrt{\cdot} \text{ in } \mathbb{R}) \\ &= P. && \text{(spectral decomposition)} \end{aligned}$$

Therefore, for any positive operator P, the polar decomposition of P is $P = UP$. *If P were positive *definite**, it would be easy to show that P is invertible. If $a = \sum_j a_j |j\rangle$, then if $Pa = 0$, $Pa = P\left(\sum_j a_j |j\rangle\right) = \left(\sum_i \lambda_i |i\rangle\langle i|\right)\left(\sum_j a_j |j\rangle\right) = \sum_i \lambda_i a_i |i\rangle = 0$. Since $|i\rangle$ is a basis, we must have $a_i\lambda_i = 0$ for all i, but the λ_i cannot be 0 since, being a basis vector, $|i\rangle \neq 0$ implies $\langle i|P|i\rangle = \lambda_i > 0$ by positive definiteness. Now all $a_i = 0$. Having 0-dimensional null-space, P must be invertible, in which case $U = I$ by uniqueness of U, since I satisfies $P = IP$. Then $P = P$ *is* the polar decomposition of 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, the polar decomposition of U is $U = U$.

Alternatively, since J is unique, note that U satisfies $U = UJ$, where $J = I$, so $U = U$ is again the polar decomposition of U. This is essentially the same argument, as above, where we use the stated uniqueness of J instead of the unique formula for it.

(Hermitian) Suppose $H = UJ$. By Hermiticity

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

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

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

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

unless H is positive.

2.49) Express the polar decomposition of a normal matrix in the outer product representation.
Soln: Let A be a normal matrix, which by the spectral decomposition theorem we may write $A = \sum_i \lambda_i |i\rangle\langle i|$ for an orthonormal basis $|i\rangle$. As in the proof of the polar decomposition theorem, define $|e_i\rangle = A|i\rangle$ for those $|i\rangle$ for which $\lambda_i \neq 0$, and extend via Graham-Schmidt to find $|e_i\rangle$ for those $|i\rangle$ for which $\lambda_i = 0$.

$$\begin{aligned}
J &= \sqrt{A^\dagger A} = \sum_i |\lambda_i| |i\rangle\langle i|. && \text{(uniqueness)}\\
U &= \sum_i |e_i\rangle\langle i|. && \text{(as constructed in the proof)}\\
A &= UJ && \text{(polar decomposition)}\\
&= \left(\sum_i |e_j\rangle \langle j|\right)\left(\sum_j |\lambda_j| |j\rangle\langle j|\right) && \text{(defined above)}\\
&= \sum_i |\lambda_i| |e_i\rangle\langle i|. && \text{(orthonormality)}
\end{aligned}$$

2.50) Find the left and right polar decomposition of the matrix

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

Soln: We have $A^\dagger A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$. To construct J, we must find a spectral decomposition of $A^\dagger A$. The characteristic equation of $A^\dagger A$ is $\det(A^\dagger A - \lambda I) = (2-\lambda)(1-\lambda) - 1 = \lambda^2 - 3\lambda + 1 = 0$. By the quadratic formula, the eigenvalues of $A^\dagger A$ are $\lambda_\pm = \frac{3\pm\sqrt{5}}{2}$. The assosiated orthonormal eigenvectors are $|\lambda_\pm\rangle = \frac{5\pm\sqrt{5}}{10}\begin{bmatrix} \frac{1\pm\sqrt{5}}{2} \\ 1 \end{bmatrix}$. We have $|\lambda_\pm\rangle\langle\lambda_\pm| = \begin{bmatrix} 1 \pm \frac{2\sqrt{5}}{5} & \frac{1}{2} \pm \frac{3\sqrt{5}}{10} \\ \frac{1}{2} \pm \frac{3\sqrt{5}}{10} & \frac{1}{2} \pm \frac{\sqrt{5}}{10} \end{bmatrix}$. By the spectral decomposition theorem, $A^\dagger A = \lambda_+ |\lambda_+\rangle\langle\lambda_+| + \lambda_- |\lambda_-\rangle\langle\lambda_-|$, and

$$\begin{aligned}
J = \sqrt{A^\dagger A} &= \sqrt{\lambda_+} |\lambda_+\rangle\langle\lambda_+| + \sqrt{\lambda_-} |\lambda_-\rangle\langle\lambda_-| && \text{(definition of } \sqrt{\cdot}\text{)}\\
&= \sum_\pm \sqrt{\frac{3\pm\sqrt{5}}{2}} \begin{bmatrix} 1 \pm \frac{2\sqrt{5}}{5} & \frac{1}{2} \pm \frac{3\sqrt{5}}{10} \\ \frac{1}{2} \pm \frac{3\sqrt{5}}{10} & \frac{1}{2} \pm \frac{\sqrt{5}}{10} \end{bmatrix}
\end{aligned}$$

$$J^{-1} = \frac{1}{\sqrt{\lambda_+}} |\lambda_+\rangle\langle\lambda_+| + \frac{1}{\sqrt{\lambda_-}} |\lambda_-\rangle\langle\lambda_-|. \qquad \text{(Applying the }^{-1}\text{ function to } J\text{)}$$

$$U = AJ^{-1} \qquad (A \text{ is invertible, } U \text{ is unique})$$

It is not worth simplifying J, J^{-1}, or U, nor is it worth finding the right polar decomposition.

2.51) Verify that the Hadamard gate H is unitary.
Soln:

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

2.52) Very that $H^2 = I$.
Soln: It was shown above that $H^\dagger = H$, so $H^2 = H^\dagger H = I$.

2.53) What are the eigenvalues and eigenvectors of H?
Soln: Using the characteristic equation

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

the eigenvalues are $\lambda_\pm = \pm 1$. The associated orthonormal eigenvectors can be calculated to be $|\lambda_\pm\rangle = \frac{1}{\sqrt{4\mp 2\sqrt{2}}}\begin{bmatrix}1\\ -1\pm\sqrt{2}\end{bmatrix}$.

2.54) Suppose A and B are commuting Hermitian operators. Prove that $\exp(A)\exp(B) = \exp(A+B)$.
Soln: Since $[A,B] = 0$, A and B are simultaneously diagonalizable. Let $|i\rangle$ be an orthonormal basis such that $A = \sum_i a_i |i\rangle\langle i|$, $B = \sum_i b_i |i\rangle\langle i|$. Note that $A+B$ is also simultaneously diagonalizable, since $A + B = \sum_i (a_i + b_i)|i\rangle\langle i|$.

$$\begin{aligned}\exp(A)\exp(B) &= \left(\sum_i \exp(a_i)|i\rangle\langle i|\right)\left(\sum_i \exp(b_i)|i\rangle\langle i|\right) && \text{(from above)}\\ &= \sum_{i,j}\exp(a_i + b_j)|i\rangle\langle i|j\rangle\langle j| && \text{(group sum)}\\ &= \sum_{i,j}\exp(a_i+b_j)|i\rangle\langle j|\,\delta_{i,j} && \text{($|i\rangle$ is orthonormal)}\\ &= \sum_i \exp(a_i+b_i)|i\rangle\langle i| && \text{(group non-zero terms)}\\ &= \exp(A+B). && \text{(definition of $\exp(\cdot)$)}\end{aligned}$$

2.55) Prove that $U(t_1, t_2)$ defined in Equation (2.91) is unitary.
Soln: Some definitions: H is the Hamiltonian of some closed system. It is Hermitian, and hence has a spectral decomposition: $H = \sum E\,|E\rangle\langle E|$. Note that $H^\dagger = \left(\sum E\,|E\rangle\langle E|\right)^\dagger = \sum E^*\left(|E\rangle\langle E|\right)^\dagger = \sum E^*\,|E\rangle\langle E|$ by Exercise 2.13. U is defined as $U(t_1,t_2) \equiv \exp\left[\frac{-iH(t_2-t_1)}{\hbar}\right]$. To prove U is unitary, we show $UU^\dagger = I$.

Now

$$\begin{aligned}
(U(t_1,t_2))^\dagger &= \left(\exp\left[\frac{-iH(t_2-t_1)}{\hbar}\right]\right)^\dagger && \text{(definition of } U\text{)}\\
&= \left(\sum_E \exp\left(\frac{-iE(t_2-t_2)}{\hbar}\right)|E\rangle\langle E|\right)^\dagger && \text{(definition of } \exp(A)\text{)}\\
&= \sum_E \exp\left(\frac{-iE(t_2-t_1)}{\hbar}\right)^* |E\rangle\langle E|^\dagger && \text{(linearity of }^\dagger\text{)}\\
&= \sum_E \exp\left(\frac{iE(t_2-t_1)}{\hbar}\right)|E\rangle\langle E| && \text{(complex conjugation, exercise 2.13)}\\
&= \exp\left(\frac{iH(t_2-t_1)}{\hbar}\right) && \text{(definition of } \exp(A)\text{)}
\end{aligned}$$

We have

$$\begin{aligned}
U(t_2-t_1)(U(t_2-t_1))^\dagger &= \exp\left(-\frac{iH(t_2-t_1)}{\hbar}\right)\exp\left(\frac{iH(t_2-t_1)}{\hbar}\right) && \text{(from above)}\\
&= \sum_{E,E'}\left(\exp\left(\frac{-iE(t_2-t_1)}{\hbar}\right)|E\rangle\langle E|\right)\left(\exp\left(\frac{iE'(t_2-t_1)}{\hbar}\right)|E'\rangle\langle E'|\right)\\
& && \text{(}E'\text{ used to distinguish from }E\text{)}\\
&= \sum_{E,E'}\exp\left(-\frac{i(E-E')(t_2-t_1)}{\hbar}\right)|E\rangle\,\langle E'|\,\delta_{E,E'} && \text{(orthonormality)}\\
&= \sum_E \exp(0)\,|E\rangle\langle E| && \text{(group nonzero terms)}\\
&= \sum_E |E\rangle\langle E| && (e^0=1)\\
&= I && \text{(completeness)}
\end{aligned}$$

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

2.56) Use the spectral decomposition to show that $K \equiv -i\log(U)$ is Hermitian for any unitary U and thus $U = \exp(iK)$ for some Hermitian K.
Soln: Since U is unitary, it has a spectral decomposition with respect to some orthonormal basis, say $|\lambda\rangle$. For each eigenvalue λ, note that exerces 2.18 gives that $\|\lambda\| = 1$, so express $\lambda = e^{i\theta}$ for some real valued argument θ. Then $\log(U) = \sum \log(\lambda)\,|\lambda\rangle\langle\lambda| = \sum i\theta\,|\lambda\rangle\langle\lambda|$. Now $K \equiv -i\log(U) = \sum\theta\,|\lambda\rangle\langle\lambda|$, and $K^\dagger = (\sum\theta\,|\lambda\rangle\langle\lambda|)^\dagger = \sum\theta^*(|\lambda\rangle\langle\lambda|)^\dagger = \sum\theta\,|\lambda\rangle\langle\lambda| = K$, since θ is real and exercise 2.13 gives that $(|\lambda\rangle\langle\lambda|)^\dagger = |\lambda\rangle\langle\lambda|$. So, K is Hermitian, and by applying the fact that exp and log are inverse complex valued functions, at least for the λ, we can conclude that $\exp(iK) = \exp(i\cdot(-i\log(U))) = \exp(\log(U)) = U$.

2.57) (Cascaded measurements are single measurements) Suppose $\{L_\ell\}$ and $\{M_m\}$ are two sets of measurement operators. Show that a measurement defined by the measurement operators $\{L_\ell\}$ followed by a measurement defined by the measurement operators $\{M_m\}$ is physically equivalent to a single measurement defined by measurement operators $\{N_{\ell m}\}$ with the representation $N_{\ell m} \equiv M_m L_\ell$.
Soln: Let the state of a physical system be ψ. Note that by definition the state after applying measurement operators $\{L_\ell\}$ is $|\phi\rangle \equiv \frac{L_\ell|\psi\rangle}{\sqrt{\langle\psi|L_\ell^\dagger L_\ell|\psi\rangle}}$ for some ℓ. The state after applying measurement operators

$\{M_m\}$ is $\xi \equiv \frac{M_m|\phi\rangle}{\sqrt{\langle\phi|M_m^\dagger M_m|\phi\rangle}}$ for some m. Now

$$\begin{aligned}
\xi &= \frac{M_m\,|\phi\rangle}{\sqrt{\langle\phi|M_m^\dagger M_m|\phi\rangle}} && \text{(from above)}\\
&= \frac{M_m \frac{L_\ell|\psi\rangle}{\sqrt{\langle\psi|L_\ell^\dagger L_\ell|\psi\rangle}}}{\sqrt{\left\langle \frac{L_\ell|\psi\rangle}{\sqrt{\langle\psi|L_\ell^\dagger L_\ell|\psi\rangle}} \middle| M_m^\dagger M_m \middle| \frac{L_\ell|\psi\rangle}{\sqrt{\langle\psi|L_\ell^\dagger L_\ell|\psi\rangle}} \right\rangle}} && \text{(substitute for } \phi\text{)}\\
&= \frac{\frac{M_m L_\ell|\psi\rangle}{\sqrt{\langle\psi|L_\ell^\dagger L_\ell|\psi\rangle}}}{\sqrt{\frac{\left\langle L_\ell|\psi\rangle \middle| M_m^\dagger M_m \middle| L_\ell\psi\right\rangle}{\langle\psi|L_\ell^\dagger L_\ell|\psi\rangle}}} && \text{(group internal scalar values)}\\
&= \frac{\frac{\sqrt{\langle\psi|L_\ell^\dagger L_\ell|\psi\rangle}}{1}\cdot\frac{M_m L_\ell|\psi\rangle}{\sqrt{\langle\psi|L_\ell^\dagger L|\psi\rangle}}}{\sqrt{\langle\psi|L_\ell^\dagger M_m^\dagger M_m L_\ell|\psi\rangle}} && \text{(move the scalar to numerator)}\\
&= \frac{(M_m L_\ell)\psi}{\sqrt{\langle\psi|(M_m L_\ell)^\dagger (M_m L_\ell)|\psi\rangle}} && \text{(cancel scalars, group operators)}\\
&= \frac{N_{\ell m}\,|\psi\rangle}{\sqrt{\langle\psi|N_{\ell m}^\dagger N_{\ell m}|\psi\rangle}} && \text{(definition of } N_{\ell m}\text{)}
\end{aligned}$$

This gives that the state of the system after applying measurement operators $\{L_\ell\}$ followed by measurement operators $\{M_m\}$ can be expressed in terms of applying measurement operators $N_{\ell m} = M_m L_\ell$. Note, though, that the state of the system after applying $N_{\ell m}$ is exactly the state of the system after applying M_m, for any ℓ. In practice, the intermediate measurement result ℓ would be unknown, and to find the probability that the system is in state m after application of the measurement operators $\{N_{\ell m}\}$, one would have to sum over the possible intermediate measurement results.

2.58) Suppose we prepare a quantum system in an eigenstate $|\psi\rangle$ of some observable M, with corresponding eigenvalue m. What is the average observed value of M, and the standard deviation?
Soln: Express $M = \sum_\mu \mu P_\mu$, where the $|\mu\rangle$ are an orthonormal set of eigenvectors each with eigenvalue μ. We may assume $\||\psi\rangle\| = 1$, so that $|\psi\rangle = |\mu\rangle$ for some $|\mu\rangle$ (and $m = \mu$ for some μ).

$$\begin{aligned}
\mathbf{E}(M) = \langle M\rangle &= \langle\psi|M|\psi\rangle && (\text{ definition of } \mathbf{E}(M))\\
&= \left\langle \psi \middle| \sum_\mu \mu P_\mu \middle| \psi \right\rangle && (\text{ spectral decomposition of } M)\\
&= \sum_\mu \mu \,\langle\psi|P_\mu|\psi\rangle && \text{(linearity)}\\
&= \sum_\mu \mu \,\langle\psi|\mu\rangle\,\langle\mu|\psi\rangle && (P_\mu \text{ is a projector})\\
&= \sum_\mu \mu\delta_{\psi,\mu}^2 && \text{(orthonormality)}\\
&= m && \text{(collect nonzero terms)}
\end{aligned}$$

Similarly, $\langle M^2\rangle = \langle\psi|M^2|\psi\rangle = m^2$. Now $\Delta(M) = \sqrt{\langle M^2\rangle - \langle M\rangle^2} = \sqrt{m^2 - (m)^2} = 0$.

2.59) Suppose we have (a) qubit in the state $|0\rangle$, and we measure the observable X. What is the average

value of X? What is the standard deviation?
Soln: There are two ways to proceed. First, we can apply X, noting that $X\left|0\right\rangle = \left|1\right\rangle$ and $X\left|1\right\rangle = \left|0\right\rangle$.

$$\begin{aligned}
\langle X\rangle &= \langle 0|X|0\rangle = \langle 0|1\rangle = 0 \\
\langle X^2\rangle &= \langle 0|X^2|0\rangle = \langle 0|X|1\rangle = \langle 0|0\rangle = 1 \\
\Delta(X) &= \sqrt{\langle X^2\rangle - \langle X\rangle^2} = \sqrt{1-0^2} = 1
\end{aligned}$$

Alternatively, write $X = (|+\rangle\langle+|) - (|-\rangle\langle-|)$, and note that $X^2 = (|+\rangle\langle+|) + (|-\rangle\langle-|)$ Now

$$\begin{aligned}
\langle X\rangle &= \langle 0|X|0\rangle = \langle 0|+\rangle\left\langle +|0\right\rangle - \langle 0|-\rangle\left\langle -|0\right\rangle = \frac{1}{2} - \frac{1}{2} = 0 \\
\langle X^2\rangle &= \langle 0|X^2|0\rangle = \langle 0|+\rangle\left\langle +|0\right\rangle + \langle 0|-\rangle\left\langle -|0\right\rangle = \frac{1}{2} + \frac{1}{2} = 1 \\
\Delta(X) &= \sqrt{\langle X^2\rangle - \langle X\rangle^2} = \sqrt{1-0^2} = 1
\end{aligned}$$

2.60) Show that $\vec{v}\cdot\vec{\sigma}$ has eigenvalues ±1, and that the projectors onto the corresponding eigenspaces are given by $P_\pm = (I \pm \vec{v}\cdot\vec{\sigma})/2$.
Soln: We calculate eigenvalues by expressing $\vec{v}\cdot\vec{\sigma}$ explicitly in terms of v_1, v_2, and v_3.

$$\begin{aligned}
\vec{v}\cdot\vec{\sigma} &= \sum_{i=1}^{3} v_i\sigma_i && \text{(definition)} \\
&= v_1\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix} + v_2\begin{bmatrix}0 & -i\\ i & 0\end{bmatrix} + v_3\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix} && (\sigma_1 = X, \sigma_2 = Y, \sigma_3 = Z) \\
&= \begin{bmatrix} v_3 & v_1 - iv_2 \\ v_1 + iv_2 & -v_3\end{bmatrix} && \text{(add corresponding entries)}
\end{aligned}$$

$$\begin{aligned}
\det(\vec{v}\cdot\vec{\sigma} - \lambda I) &= (v_3-\lambda)(-v_3-\lambda) - (v_1 - iv_2)(v_1+iv_2) && \text{(characteristic equation)} \\
&= \lambda^2 - (v_1^2 + v_2^2 + v_3^2) && \text{(simplify)} \\
&= \lambda^2 - 1 && (\vec{v}\text{ is a unit vector})
\end{aligned}$$

So the eigenvalues of $\vec{v}\cdot\vec{\sigma}$ are $\lambda = \pm 1$.

To calculate the projectors onto the corresponding eigenspaces, note that $|\lambda_1\rangle = \begin{bmatrix} 1 \\ \frac{1-v_3}{v_1-iv_2}\end{bmatrix}$ can easily be seen to be an eigenvector with eigenvalue $\lambda = 1$. To normalize, we'll use the fact that $1 - v_3^2 = v_1^2 + v_2^2$, since $\vec{v}$ is a unit vector. Factoring both sides and dividing yields $\frac{1\pm_1 v_3}{v_1 \pm_2 iv_2} = \frac{v_1 \mp_2 iv_2}{1\mp_1 v_3}$, that is, such rational functions can be flipped by negating both binary operations.

$$\begin{aligned}
\||\lambda_1\rangle\|^2 &= 1 + \left\|\frac{1-v_3}{v_1 - iv_2}\right\|^2 && (\text{definition of } \|\cdot\|^2) \\
&= 1 + \left(\frac{v_1+iv_2}{1+v_3}\right)\left(\frac{v_1-iv_2}{1+v_3}\right) && (\text{flip, } \|c\|^2 = c\cdot c^*) \\
&= 1 + \frac{v_1^2+v_2^2}{(1+v_3)^2} && \text{(expand)} \\
&= \frac{1 + 2v_3 + v_1^2 + v_2^2 + v_3^2}{(1+v_3)^2} && \text{(common denominator)} \\
&= \frac{2(1+v_3)}{(1+v_3)^2} && (\vec{v}\text{ is a unit vector}) \\
&= \frac{2}{1+v_3} && \text{(cancel)}
\end{aligned}$$

So $|\lambda_1\rangle = \sqrt{\frac{1+v_3}{2}}\begin{bmatrix} 1 \\ \frac{1-v_3}{v_1-iv_2}\end{bmatrix}$ is a normalized eigenvector with eigenvalue 1. Similarly, $|\lambda_{-1}\rangle = \sqrt{\frac{1-v_3}{2}}\begin{bmatrix} 1 \\ -\frac{1+v_3}{v_1-iv_2}\end{bmatrix}$ is a normalized eigenvector with eigenvalue -1. For convenience, we write $|\lambda_{\pm 1}\rangle = \sqrt{\frac{1\pm v_3}{2}}\begin{bmatrix} 1 \\ \frac{v_3\mp 1}{v_1-iv_2}\end{bmatrix}$. Calculating the projectors:

$$
\begin{aligned}
|\lambda_{\pm 1}\rangle\langle\lambda_{\pm 1}| &= \frac{1\pm v_3}{2}\begin{bmatrix} 1 \\ \frac{v_3\mp 1}{v_1-iv_2}\end{bmatrix}\begin{bmatrix} 1 & \left(\frac{v_3\mp 1}{v_1-iv_2}\right)^*\end{bmatrix} && \text{(definition)} \\
&= \frac{1\pm v_3}{2}\begin{bmatrix} 1 \\ \frac{v_1+iv_2}{v_3\pm 1}\end{bmatrix}\begin{bmatrix} 1 & \left(\frac{v_1+iv_2}{v_3\pm 1}\right)^*\end{bmatrix} && \text{(flip)} \\
&= \frac{1\pm v_3}{2}\begin{bmatrix} 1 \\ \frac{v_1+iv_2}{v_3\pm 1}\end{bmatrix}\begin{bmatrix} 1 & \frac{v_1-iv_2}{v_3\pm 1}\end{bmatrix} && (v_1, v_2 \text{ are real, conjugate}) \\
&= \frac{1\pm v_3}{2}\begin{bmatrix} 1 & \frac{v_1-iv_2}{1\pm v_3} \\ \frac{v_1+iv_2}{1\pm v_3} & \frac{v_1^2+v_2^2}{(1\pm v_3)^2}\end{bmatrix} && \text{(multiply)} \\
&= \frac{1}{2}\begin{bmatrix} 1\pm v_3 & v_1-iv_2 \\ v_1+iv_2 & \frac{1-v_3^2}{1\pm v_3}\end{bmatrix} && (v \text{ is a unit vector, cancel external numerator}) \\
&= \frac{1}{2}\begin{bmatrix} 1\pm v_3 & v_1-iv_2 \\ v_1+iv_2 & 1\mp v_3^2\end{bmatrix} && \text{(cancel internal denominator)} \\
&= \frac{1}{2}\left(I \pm \begin{bmatrix} v_3 & v_1-iv_2 \\ v_1+iv_2 & -v_3\end{bmatrix}\right) && \text{(separate)} \\
&= \frac{1}{2}(I\pm\vec{v}\cdot\vec{\sigma}) && \text{(definition)}
\end{aligned}
$$

The first author points out that when $v_1 - iv_2 = 0$, or equivalently when $v_3 \pm 1 = 0$, various expressions above are indeterminant. The first author attempts to circumvent this by working more generally below, however, the second author is suspicious that the argument is circular. Some of the details are instructive however, so it is left below. To deal with the degenerate cases, note that, since $\vec{v}$ is a real-valued vector, $v_1 - iv_2 = 0$ implies $v_1 = 0, v_2 = 0$, and $v_3 = \pm 1$, in which case $\vec{v}\cdot\vec{\sigma} = \pm Z$. The normalized eigenvectors are $|\lambda_1\rangle = \pm|0\rangle$ and $|\lambda_{-1}\rangle = \pm|1\rangle$, where here the ± indicate the sign of v_3 instead of the sign of the eigenvalue. Now

$$
\begin{aligned}
|\lambda_1\rangle\langle\lambda_1| = (\pm 1)^2\,|0\rangle\langle 0| = \begin{bmatrix} 1 & 0 \\ 0 & 0\end{bmatrix} = \frac{1}{2}(I+Z) = \frac{1}{2}(I+\vec{v}\cdot\vec{\sigma}) \\
|\lambda_{-1}\rangle\langle\lambda_{-1}| = (\pm 1)^2\,|1\rangle\langle 1| = \begin{bmatrix} 0 & 0 \\ 0 & 1\end{bmatrix} = \frac{1}{2}(I-Z) = \frac{1}{2}(I-\vec{v}\cdot\vec{\sigma}),
\end{aligned}
$$

completing the proof. The first author's attempted resolution follows:

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.

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

$$\begin{aligned}
(\vec{v}\cdot\vec{\sigma})^2 &= \sum_{j,k=1}^{3} (v_j\sigma_j)(v_k\sigma_k) \\
&= \sum_{j,k=1}^{3} v_j v_k \sigma_j \sigma_k \\
&= \sum_{j,k=1}^{3} v_j v_k \left(\delta_{jk} I + i\sum_{l=1}^{3} \epsilon_{jkl}\sigma_l\right) \quad (\because \text{Exercise 2.43, eqn (2.78) page 78}) \\
&= \sum_{j,k=1}^{3} v_j v_k \delta_{jk} I + i \sum_{j,k,l=1}^{3} \epsilon_{jkl} v_j v_k \sigma_l \\
&= \sum_{j=1}^{3} v_j^2 I \\
&= I \quad \left(\because \sum_j v_j^2 = 1\right)
\end{aligned}$$

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

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

$$\begin{aligned}
\langle\psi|(P_\pm - |\lambda_\pm\rangle\langle\lambda_\pm|)|\psi\rangle &= \langle\psi|P_\pm|\psi\rangle - \langle\psi|\lambda_\pm\rangle\,\langle\lambda_\pm|\psi\rangle\,. \\
\langle\psi|P_\pm|\psi\rangle &= \langle\psi|\frac{1}{2}(I \pm \vec{v}\cdot\vec{\sigma})|\psi\rangle & &\left(\begin{matrix}\text{implicit assumption the above equals 0} \\ \text{and that } |\lambda_{\pm 1}\rangle\langle\lambda_{\pm 1}| = \tfrac{1}{2}(I \pm \vec{v}\cdot\sigma)\end{matrix}\right) \\
&= \frac{1}{2} \pm \frac{1}{2}\,\langle\psi|\vec{v}\cdot\vec{\sigma})|\psi\rangle \\
&= \frac{1}{2} \pm \frac{1}{2}(|\alpha|^2 - |\beta|^2) \\
&= \frac{1}{2} \pm \frac{1}{2}(2|\alpha|^2 - 1) \quad (\because |\alpha|^2 + |\beta|^2 = 1) \\
\langle\psi|\lambda_1\rangle\,\langle\lambda_1|\psi\rangle &= |\alpha|^2 \\
\langle\psi|\lambda_{-1}\rangle\,\langle\lambda_{-1}|\psi\rangle &= |\beta|^2 = 1 - |\alpha|^2
\end{aligned}$$

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

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

$$\begin{aligned}
p(1) = \langle 0|P_1|0\rangle &= \left\langle 0\middle|\frac{1}{2}(I + \vec{v}\cdot\vec{\sigma})\middle|0\right\rangle & &\text{(definition)} \\
&= \langle 0|\lambda_1\rangle\,\langle\lambda_1|0\rangle & &\text{(use eigenvector directly)} \\
&= \frac{1+v_3}{2}\begin{bmatrix}1 & 0\end{bmatrix}\begin{bmatrix}1 \\ \frac{1-v_3}{v_1-iv_2}\end{bmatrix}\begin{bmatrix}1 & \frac{1-v_3}{v_1+iv_2}\end{bmatrix}\begin{bmatrix}1 \\ 0\end{bmatrix} & &\text{(substitute)} \\
&= \frac{1+v_3}{2}\begin{bmatrix}1 & 0\end{bmatrix}\begin{bmatrix}1 & \frac{v_1-iv_2}{1+v_3} \\ \frac{v_1+iv_2}{1+v_3} & \frac{1-v_3}{1+v_3}\end{bmatrix}\begin{bmatrix}1 \\ 0\end{bmatrix} & &\text{(multiply)} \\
&= \frac{1}{2}(1+v_3) & &\text{(extract 0,0 entry)}
\end{aligned}$$

The post-measurement state is

$$\begin{aligned}
\frac{|\lambda_1\rangle\,\langle\lambda_1|0\rangle}{\sqrt{p(1)}} &= \left(\frac{\sqrt{\frac{1+v_3}{2}}}{\sqrt{\frac{1+v_3}{2}}}\begin{bmatrix}1 & \left(\frac{1-v_3}{v_1-iv_2}\right)^*\end{bmatrix}\begin{bmatrix}1 \\ 0\end{bmatrix}\right)|\lambda_1\rangle & &\left(\begin{matrix}\text{numerator from normalization} \\ \text{denominator from above}\end{matrix}\right) \\
&= |\lambda_1\rangle & &\text{(simplify, multiply)}
\end{aligned}$$

2.62) Show that any measurement where the measurement operators and the POVM elements coincide is a projective measurement.
Soln: We can no longer avoid the converse of Exercise 2.16. Before we prove it, we quibble about semantics. The assumption in the exercise is that the POVM elements "coincide" with the measurement operators. This can be taken to mean one of two things. Let $\{M_m\}$ be the set of measurement operators. One interpretation is that $M_m = E_m = M_m^\dagger M_m$ for all m. We'll call this direct coincidence. Another interpretation is that they coincide as sets, i.e., that $\{M_m\} = \{E_m\}$, without requiring $M_m = E_m$ for each m. A mathematician may argue that what is to be assumed is set-wise coincidence, but we'll show that this is not the case. We'll argue for the assumption of direct coincidence by contradiction. If it were the case that POVM measurements which satisfied the setwise coincidence assumption but not the direct coincidence assumption were projective measurements, then there would exist $M_m = E_\mu = \overline{M_\mu^\dagger} M_\mu$, where $M_m \neq M_\mu$. However, being a projective measurement, M_μ must be Hermitian, in which case $M_m = M_\mu^2 = M_\mu$, by Exercise 2.16 (not the converse). This is a contradiction, but it is important to note

what we can conclude from it. Our assumption was that setwise coincidence of measurement operators and POVM measurements was enough to prove the measurement projective. Having contradicted this, we conclude that the exercise must assume direct coincidence instead. We have not proven that setwise coincidence without direct coincidence is impossible.

Theorem. 2.0.2. *Let P be a normal linear operator. If $P^2 = P$, then P is a projector, that is $P = \sum_i |i\rangle\langle i|$ for some orthonormal basis $|i\rangle$.*

Proof. Having assumed normality (the statement is not true in general if we do not), we may apply the Structural Decomposition Theorem to write $P = \sum_i \lambda_i |i\rangle\langle i|$. Note, we may assume $\lambda_i \neq 0$. Being idempotent, $\sum_i \lambda_i |i\rangle\langle i| = P = P^2 = \left(\sum_i \lambda_i |i\rangle\langle i|\right)^2 = \sum_i \lambda_i^2 |i\rangle\langle i|$ by orthonormality. The $|i\rangle\langle i|$ are linearly independent (see Exercise 2.10), from which we conclude that all λ_i satisfy $\lambda_i = \lambda_i^2$, from which $\lambda_i(\lambda_i - 1) = 0$, or that $\lambda_i = 1$, since $\lambda_i \neq 0$. Now $P = \sum_i |i\rangle\langle i|$, as required.
Note: Some definitions of projectors define them in terms of relations between their kernels and images as opposed to the formulaic definition given in Equation 2.35. They are equivalent (at least for normal matrices). With the set theoretic definition, Exercise 2.16 and it's converse Theorem 2.0.2 follow by definition. □

On to the exercise: Suppose M_m is a measurement operator such that $E_m = M_m^\dagger M_m = M_m$. Note that $M_m = M_m^\dagger M_m$ is positive by exercise 2.25, thus Hermitian, so $M_m = E_m = M_m^\dagger M_m = M_m^2$. Being Hermitian, M_m is normal, in which case Theorem 2.0.2 above applies, giving that M_m is a projector. Thus the measurement is a projective measurement.

2.63) Suppose a measurement is described by measurement operators M_m. Show that there exist unitary operators U_m such that $M_m = U_m\sqrt{E_m}$, where E_m is the POVM associated to the measurement.
Soln: By the singular value decomposition (Corrolary 2.4, eq 2.80, p 79), there exists unitary U, V, and real-valued diagonal D such that $M_m = UDV$. Now

$$\begin{aligned}
\sqrt{E_m} &= \sqrt{M_m^\dagger M_m} && \text{(definition)}\\
&= \sqrt{V^\dagger D^\dagger U^\dagger U D V} && \text{(properties of }^\dagger)\\
&= \sqrt{V^\dagger D^2 V} && (U \text{ unitary, } D \text{ real-valued diagonal})\\
&= V^\dagger D V && (V \text{ unitary}\to (V^\dagger D V)^2 = V^\dagger D^2 V)\\
&= V^\dagger (U^\dagger U) D V && (U \text{ unitary, } U^\dagger U = I)\\
UV\sqrt{E_m} &= M_m && (U, V \text{ unitary, solve for } M_m)
\end{aligned}$$

Define $U_m \equiv UV$ so that $M_m = U_m\sqrt{E_m}$, completing the solution.

2.64) Suppose Bob is given a quantum state chosen from a set $|\psi_1\rangle, \ldots, |\psi_m\rangle$ of linearly independent states. Construct a POVM $\{E_1, \ldots, E_{m+1}\}$ such that if outcome E_i occurs, $1 \leq i \leq m$, then Bob knows with certainty that he was given the state $|\psi_i\rangle$.
Soln: Being linearly independent, S forms a basis for the subspace it spans. Consider the subspaces spanned by $S_i' = \{|\psi_j\rangle\}_{i\neq j}$. Let $|\psi_i'\rangle$ be a non-zero unit vector in the orthogonal complement. Note that if $i \neq j$, then $\langle\psi_i|\psi_j'\rangle = 0$, since ψ_j' is in the orthogonal complement of a subspace containing ψ_i. Also note that $\langle\psi_i|\psi_i'\rangle \neq 0$, since if it were, then ψ_i would be in the subspace spanned by S_i', violating linear independence. Combining, we write $\langle\psi_i|\psi_j'\rangle = \delta_{i,j} \cdot p_i$, for some non-zero p_i. We may scale the ψ_i' so that $p_i = 1$. For, $1 \leq i \leq m$, define $E_i = |\psi_i'\rangle\langle\psi_i'|$, then, to cover the define $E_{m+1} = I - \sum_m E_i$. Now, for

$1 \le i, j \le m$,

$$\begin{aligned}
p_i(j) &\equiv \langle\psi_i|E_j^\dagger E_j|\psi_i\rangle && \text{(definition)}\\
&= \langle\psi_i|(|\psi_j'\rangle\langle\psi_j'|)^\dagger\,|\psi_j'\rangle\langle\psi_j'||\psi_i\rangle && \text{(definition of } E_j\text{)}\\
&= \langle\psi_i|\psi_j'\rangle\,\langle\psi_j'|\psi_j'\rangle\,\langle\psi_j'|\psi_i\rangle && \text{(Exercise 2.13: } (|\phi\rangle\langle\phi|)^\dagger = |\phi\rangle\langle\phi|\text{)}\\
&= \delta_{i,j}\cdot 1\cdot\delta_{i,j} && \text{(construction)}\\
&= \delta_{i,j} && \text{(simplify)}
\end{aligned}$$

That is, $E_1, \ldots, E_{m+1}$ is a POVM such that state $|\psi_i\rangle$ and outcome E_i correspond exactly, for $1 \le i \le m$. Outcome E_{m+1} will result from measuring any state outside of the span of S, that is, in their orthogonal complement, with probability 1 as well The first author included links to several relevant journal articles. The second author has not evaluated them, but they are include below.

- 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) Express the states $(|0\rangle + |1\rangle)/\sqrt{2}$ and $(|0\rangle - |1\rangle)/\sqrt{2}$ in a basis in which there are *not* the same up to a relative phase shift.
Soln: Note that $(|0\rangle + |1\rangle)/\sqrt{2} = |+\rangle$ and $(|0\rangle - |1\rangle)/\sqrt{2} = |-\rangle$ are an orthonormal basis. Since the amplitude of $|+\rangle$ in the expression $(|0\rangle + |1\rangle)/\sqrt{2} = |+\rangle$ is 1, and that in $(|0\rangle - |1\rangle)/\sqrt{2} = |-\rangle$ is 0, there is no relative phase θ such that such that $e^{i\theta}\cdot 1 = 0$, so $|+\rangle$ and $|-\rangle$ do not differ by a relative phase in the basis they comprise.
2.66) Show that the average value of the observable X_1Z_2 for a two qubit system measured in the state $(|00\rangle + |11\rangle)/\sqrt{2}$ is zero.
Soln: Let $|\Phi^+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$.

$$\begin{aligned}
\mathbf{E}(X_1Z_2) = \langle X_1Z_2\rangle &= \langle\Phi^+|X_1Z_2|\Phi^+\rangle\\
&= \frac{\langle 00| + \langle 11|}{\sqrt{2}}\cdot\frac{X_1Z_2(|00\rangle + |11\rangle)}{\sqrt{2}}\\
&= \frac{\langle 00| + \langle 11|}{\sqrt{2}}\cdot\frac{|10\rangle - |01\rangle}{\sqrt{2}}\\
&= \frac{\langle 00|10\rangle - \langle 00|01\rangle + \langle 11|10\rangle - \langle 11|01\rangle}{2}\\
&= \frac{0 - 0 + 0 - 0}{2}\\
&= 0
\end{aligned}$$

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

$$\langle w_1|U^\dagger U|w_2\rangle = \langle w_1|w_2\rangle\,.$$

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

to denote the extension.
Soln: Let $|w_i\rangle$ be an orthonormal basis for W. Since U preserves inner products in W, $|u_i\rangle \equiv U\,|w_i\rangle$ is an orthonormal basis for image(U), hence $\langle u_i|u_j\rangle = \delta_{i,j}$. Consider the orthogonal complement, $W^{\perp}$. By definition $V = W \oplus W^{\perp}$. Let $|w_j'\rangle$ and $|u_j'\rangle$ be orthonormal bases for $W^{\perp}$ and $(\text{image}(U))^{\perp}$. Note that, as provided, U is not defined on the $|w_j'\rangle$, so we may not state that ~~$|u_j'\rangle = U\,|w_j'\rangle$~~, where here we strike the statement not because it is necessarily false, but because it is not assumed. We can, however, state that $\langle u_i'|u_j'\rangle = \delta i, j$, since the $|u_j\rangle$ are orthonormal, and that $\langle u_i|u_j'\rangle = 0$, since the $|u_j'\rangle$ are in a space orthogonal to the $|u_i\rangle$. Define $U' : V \to V$ as $U' = \sum_i |u_i\rangle\langle w_i| + \sum_j |u_j'\rangle\langle w_j'|$. First, we prove unitarity:

$$\begin{aligned}
(U')^{\dagger}U' &= \left(\sum_i |w_i\rangle\langle u_i| + \sum_j |w_j'\rangle\langle u_j'|\right)\left(\sum_k |u_k\rangle\langle w_k| + \sum_\ell |u_\ell'\rangle\langle w_\ell'|\right) && \text{(definition, linearity, Exercise 2.13)}\\
&= \sum_{i,k} |w_i\rangle\,\langle u_i|u_k\rangle\,\langle w_k| + \sum_{i,\ell} |w_i\rangle\,\langle u_i|u_\ell'\rangle\,\langle w_\ell'| + \sum_{j,k} |w_j'\rangle\,\langle u_j'|u_k\rangle\,\langle w_k'| + \sum_{j,\ell} |w_j'\rangle\,\langle u_j'|u_\ell'\rangle\,\langle w_\ell'|\\
& && \text{(F.O.I.L., linearity)}\\
&= \sum_{i,k} \delta_{i,k}\,|w_i\rangle\langle w_k| + \sum_{j,\ell} \delta_{j,\ell}\,|w_j'\rangle\langle w_\ell'| && \text{(orthonormality, orthogonality)}\\
&= \sum_i |w_i\rangle\langle w_i| + \sum_j |w_j'\rangle\langle w_j'| && \text{(collect non-zero term)}\\
&= I && \text{(completeness)}
\end{aligned}$$

where the last equality holds because $\{|w_i\rangle\}\bigcup\{|w_j'\rangle\}$ forms a basis for the entire space V. Similarly $U'(U')^{\dagger} = I$, so U' is unitary. It is left only to show that U' is an extension of U, that is $U'\,|w\rangle = U\,|w\rangle$ for all w in W.

$$\begin{aligned}
U'\,|w\rangle &= \left(\sum_i |u_i\rangle\langle w_i| + \sum_j |u_j'\rangle\langle w_j'|\right)|w\rangle && \text{(definition of } U')\\
&= \left(\sum_i |u_i\rangle\,\langle w_i|\right)|w\rangle + \sum_j |u_j'\rangle\,\langle w_j'|w\rangle && \text{(linearity)}\\
&= \left(\sum_i |u_i\rangle\,\langle w_i|\right)|w\rangle && (|w\rangle \in W,\ |w_j'\rangle \in W^{\perp})\\
&= \left(\sum_i U\,|w_i\rangle\,\langle w_i|\right)|w\rangle && \text{(definition of } u_i)\\
&= U\left(\sum_i |w_i\rangle\langle w_i|\right)|w\rangle && \text{(linearity)}\\
&= U\,|w\rangle\,. && \text{(completeness)}
\end{aligned}$$

where the last inequality holds only when applied to vectors in the subspace W, as is being done.

2.68) Prove that $(|\Psi^+\rangle =)\,|\psi\rangle \equiv (|00\rangle + |11\rangle)/\sqrt{2} \neq |a\rangle\,|b\rangle$ for all single qubit states $|a\rangle$ and $|b\rangle$.
Soln: Suppose $|a\rangle = a_0\,|0\rangle + a_1\,|1\rangle$ and $|b\rangle = b_0\,|0\rangle + b_1\,|1\rangle$. Then

$$|a\rangle\,|b\rangle = |a\rangle \otimes |b\rangle = a_0b_0\,|00\rangle + a_0b_1\,|01\rangle + a_1b_0\,|10\rangle + a_1b_1\,|11\rangle\,.$$

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. Since $a_0b_1 = 0$, either $a_0 = 0$ or $b_1 = 0$, however, $a_0 = 0$ contradicts $a_0b_0 = 1$, and $b_1 = 0$ contradicts $a_1b_1 = 1$. Thus $|\psi\rangle \neq |a\rangle\,|b\rangle$.

2.69) Verify that the Bell basis forms an orthonormal basis for the two qubit state space. Define Bell states and the Bell Matrix as follows.

$$|\Phi^+\rangle = |\psi_0\rangle \equiv \frac{|00\rangle + |11\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\0\\0\\1\end{bmatrix} \qquad |\Phi^-\rangle = |\psi_1\rangle \equiv \frac{|00\rangle - |11\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\0\\0\\-1\end{bmatrix}$$

$$|\Psi^+\rangle = |\psi_2\rangle \equiv \frac{|01\rangle + |10\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}}\begin{bmatrix}0\\1\\1\\0\end{bmatrix} \qquad |\Psi^-\rangle = |\psi_3\rangle \equiv \frac{|01\rangle - |10\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}}\begin{bmatrix}0\\1\\-1\\0\end{bmatrix}$$

$$B = \begin{bmatrix}|\Phi^+\rangle & |\Phi^-\rangle & |\Psi^+\rangle & |\Psi^-\rangle\end{bmatrix}$$

Note that $\langle\psi_i|\psi_j\rangle$ is the i, j-entry in $B^\dagger B$, so orthonormality will follow if $B^\dagger B = I$.

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

Linear independence follows from orthogonality. Being 4 linearly independent vectors in a 4-dimensional vector space, the Bell states form a basis.

2.70) Suppose E is any positive operator acting on Alice's qubit. Show that $\langle\psi|E \otimes I|\psi\rangle$ *takes the same value* when $|\psi\rangle$ is any of the four Bell states. Suppose some malevolent third party ('Eve') intercepts Alice's qubit on the way to Bob in the superdense coding protocol. Can Eve infer anything about which of the four possible bit strings $00, 01, 10, 11$ Alice is trying to send? If so, how, or if not, why not?
Soln: For any of the Bell states we get $\langle\psi_i|E \otimes I|\psi_i\rangle = \frac{1}{2}(\langle 0|E|0\rangle + \langle 1|E|1\rangle)$. We exhibit the calculation for $|\psi_0\rangle$. Others are similar.

$$\begin{aligned}\langle\psi_0|E \otimes I|\psi_0\rangle &= \frac{1}{2}\Big((\langle 00| + \langle 11|) \cdot (E \otimes I)(|00\rangle + |11\rangle)\Big) && \text{(definition)}\\ &= \frac{1}{2}\Big((\langle 00| + \langle 11|) \cdot \Big(((E\,|0\rangle) \otimes |0\rangle + (E\,|1\rangle) \otimes |1\rangle\Big)\Big) && (E \text{ and } I \text{ act independently})\\ &= \frac{1}{2}\Big(\langle 0|E|0\rangle \cdot \langle 0|0\rangle + \langle 0|E|1\rangle \cdot \langle 0|1\rangle + \langle 1|E|0\rangle \cdot \langle 1|0\rangle + \langle 1|E|1\rangle \cdot \langle 1|1\rangle\Big) && \text{(F.O.I.L.)}\\ &= \frac{1}{2}(\langle 0|E|0\rangle + \langle 1|E|1\rangle) && \text{(collect non-zero terms)}\end{aligned}$$

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, the result above applies, in which case the $p_i(m)$ take on the same values for all $|\psi_i\rangle$, that is, for all i and m, $p_i(m) = 1/4$. In other words, no matter the outcome m, the probability that ψ was in any of the Bell states is uniform. So Eve can't distinguish Bell states given only access to a single qubit. [Note, the exercise above only proves Eve can't distinguish Bell states given only access to the first qubit, but the only difference from Bob's perspective is that Eve's Bell basis uses $-|\psi_3\rangle = -|\Psi^-\rangle$. This negative does not change the result of any of the calculations, so Eve can't distinguish the Bell states given access to Bob's qubit either.]

2.71) (Criterion to decide if a state is mixed or pure) Let ρ be a density operator. Show that $\text{tr}(\rho^2) \leq 1$, with equality if and only if ρ is a pure state.
Soln: By definition, there exists an ensemble of pure states $\{p_i, |\psi_i\rangle\}$, where $0 \leq p_i \leq 1$ and $\sum_i p_i = 1$ such that $\rho = \sum_i p_i\,|\psi_i\rangle\langle\psi_i|$. Note that on this range $0 \leq p_i^2 \leq p_i$. Also, $\rho = \sum_i \lambda_i\,|i\rangle\langle i|$ for some orthonormal

basis $|i\rangle\langle i|$, by the spectral decomposition theorem. Express $|\psi_i\rangle = \sum_j \alpha_j |i\rangle$ ... I believe the first author's proof below assumes that $|\psi_i\rangle = |i\rangle$, i.e. that the $|\psi_i\rangle$ are orthonormal, but this is not guaranteed, and in fact, very much not assumed.

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

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

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) (Bloch sphere for mixed states) The Bloch sphere picture for pure states of a single qubit was introduced in Section 1.2. This description has an important generalization to mixed states as follows.
(1) Show that an arbitrary density matrix for a mixed state qubit may be written as

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

where $\vec{r}$ is a real three-dimensional vector such that $\|\vec{r}\| \leq 1$. This vector is known as the *Bloch vector* for the state ρ.
(2) What is the Bloch vector representation for the state $\rho = I/2$?
(3) Show that a state ρ is pure if and only if $\|\vec{r}\| = 1$.
(4) Show that for pure states the description of the Bloch vector we have given coincides with that in Section 1.2.

Soln: Note, even though the topic of this problem includes mixed states, the representation we are constructing is a representation of a single qubit. That qubit could be entangled with others, but these other qubits are not explicitly represented in the Bloch sphere representation of the qubit of interest. The level of entanglement of the qubit of interest with other qubits is represented, though.
(1) Let ρ be an arbitrary density matrix for a single complex-dimensional state. ρ is Hermitian, so we may set $\rho = \begin{bmatrix} a & b \\ b^* & d \end{bmatrix}$, where $a, d \in \mathbb{R}$ and $b \in \mathbb{C}$. Because ρ is density matrix, $\mathrm{tr}(\rho) = a + d = 1$. Define $r_1 = \mathfrak{Re}(b)/2$, $r_2 = -\mathfrak{Im}(b)/2$, $r_3 = a - d$, and finally $\vec{r} = (r_1, r_2, r_3)$. Expressing a, b, and d in terms of $\vec{r}$, we have $a = \frac{1+r_3}{2}$, $b = \frac{r_1 - ir_2}{2}$, and $d = \frac{1-r_3}{2}$. Now

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

Thus an arbitrary density matrix ρ can be written as $\rho = \frac{1}{2}(I + \vec{r}\cdot\vec{\sigma})$ for some real-valued three-dimensional vector $\vec{r}$. It remains to show that $\|\vec{r}\| \leq 1$. To do so, note that Theorem 2.5 gives that ρ is positive. Being positive, the eigenvalues of ρ must be non-negaive. So, let's find the eigenvalues:

$$\begin{aligned}
\det(\rho - \lambda I) &= (a-\lambda)(d-\lambda) - \|b\|^2 && \text{(characteristic equation)}\\
&= \lambda^2 - (a+d)\lambda + ad - \|b\|^2 = 0 && \text{(simplify)}\\
&= \lambda^2 - \lambda + \left(\frac{1-r_3^2}{4} - \frac{r_1^2+r_2^2}{4}\right) && (\mathrm{tr}(\rho)=1\text{, express in terms of } \vec{r})\\
&= \lambda^2 - \lambda + \left(\frac{1-\|r\|^2}{4}\right) && (\text{definition of } \|\cdot\| \text{ in } \mathbb{R}^3)\\
\lambda &= \frac{1 \pm \sqrt{1-(1-\|r\|^2)}}{2} && \text{(quadratic formula)}\\
&= \frac{1 \pm \|\vec{r}\|^2}{2} && \text{(simplify)}
\end{aligned}$$

Now $\frac{1-\|\vec{r}\|^2}{2} \geq 0 \to \|\vec{r}\| \leq 1$.

(2) If $\rho = I/2$, then $a = d = 1/2$ and $b = 0$. So $\vec{v} = (0,0,0)$, and $\rho = I/2$ corresponds to the origin of Bloch sphere.
(3) By exercise 2.71, ρ is a pure state if and only if $\mathrm{tr}(\rho^2) = 1$. Let's calculate $\mathrm{tr}(\rho^2)$.

$$\begin{aligned}
\rho^2 &= \frac{1}{2}(I + \vec{r}\cdot\vec{\sigma})\ \frac{1}{2}(I + \vec{r}\cdot\vec{\sigma}) && \text{(part (1) above)}\\
&= \frac{1}{4}\left[I + 2\vec{r}\cdot\vec{\sigma} + \|r\|^2\left(\frac{\vec{r}\cdot\vec{\sigma}}{\|r\|}\right)^2\right] && \text{(F.O.I.L., normalize)}\\
&= \frac{1}{4}\left(I + 2\vec{r}\cdot\vec{\sigma} + \|\vec{r}\|^2 I\right) && \left(\begin{array}{r}\text{see the first author's attempted resolution}\\ \text{of the special case of Exercise 2.60}\end{array}\right)\\
\mathrm{tr}(\rho^2) &= \frac{1}{4}\left[\mathrm{tr}(I) + 2\,\mathrm{tr}(\vec{r}\cdot\vec{\sigma}) + \|\vec{r}\|^2\,\mathrm{tr}(I)\right] && (\text{linearity of } \mathrm{tr}(\cdot))\\
&= \frac{2 + 2\,\|\vec{r}\|^2}{4}. && (\mathrm{tr}(I) = 2 \text{ in } \mathbb{C}^2,\ \mathrm{tr}(\vec{r}\cdot\vec{\sigma}) = r_3 - r_3 = 0)
\end{aligned}$$

Now ρ is a pure state if and only if $\mathrm{tr}(\rho^2) = \frac{2+2\|\vec{r}\|^2}{4} = 1$, which occurs if and only if $\|\vec{r}\| = 1$.
(4) TODO

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

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

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

Theorem 2.6)

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

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

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

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

$$\sum_i \left|\tilde{\psi}_i\right\rangle\!\left\langle\tilde{\psi}_i\right| = \left[\left|\tilde{\psi}_1\right\rangle \cdots \left|\tilde{\psi}_k\right\rangle\right] \begin{bmatrix} \left\langle\tilde{\psi}_1\right| \\ \vdots \\ \left\langle\tilde{\psi}_k\right| \end{bmatrix} \tag{2.2}$$

$$= \left[\left|\tilde{\varphi}_1\right\rangle \cdots \left|\tilde{\varphi}_k\right\rangle\right] U^T U^* \begin{bmatrix} \left\langle\tilde{\varphi}_1\right| \\ \vdots \\ \left\langle\tilde{\varphi}_k\right| \end{bmatrix} \tag{2.3}$$

$$= \left[\left|\tilde{\varphi}_1\right\rangle \cdots \left|\tilde{\varphi}_k\right\rangle\right] \begin{bmatrix} \left\langle\tilde{\varphi}_1\right| \\ \vdots \\ \left\langle\tilde{\varphi}_k\right| \end{bmatrix} \tag{2.4}$$

$$= \sum_j \left|\tilde{\varphi}_j\right\rangle\!\left\langle\tilde{\varphi}_j\right|. \tag{2.5}$$

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 $\rho = \sum_{k=1}^{l} p_k \left|k\right\rangle\!\left\langle k\right| = \sum_{k=1}^{l} \left|\tilde{k}\right\rangle\!\left\langle\tilde{k}\right|$, where $\left|\tilde{k}\right\rangle = \sqrt{\lambda_k} \left|k\right\rangle$.

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

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

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

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

[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) Suppose a composite of system A and B is in the state $|a\rangle\,|b\rangle$, where $|a\rangle$ is a pure state of system A, and *ketb* is a pure state of system B. Show that the reduced density operator of a system A alone is a pure state.
Soln:

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

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

2.77)

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

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.

(⇒) Proved by exercise 2.74.

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

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

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

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

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

$$\rho_A = \lambda_0 \left|\lambda_0\right\rangle\!\left\langle\lambda_0\right| + \lambda_1 \left|\lambda_1\right\rangle\!\left\langle\lambda_1\right|.$$

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

Then

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

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

2.80)

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

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

2.81)

Let the Schmidt decomposition of $\left|AR_1\right\rangle$ be $\left|AR_1\right\rangle = \sum_i \sqrt{p_i}\left|\psi_i^A\right\rangle\left|\psi_i^R\right\rangle$ and let $\left|AR_2\right\rangle = \sum_i \sqrt{q_i}\left|\phi_i^A\right\rangle\left|\phi_i^R\right\rangle$.
Suppose ρ^A has orthonormal decomposition $\rho^A = \sum_i p_i \left|i\right\rangle\!\left\langle i\right|$.
Since $\left|AR_1\right\rangle$ and $\left|AR_2\right\rangle$ are purifications of the ρ^A, we have

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

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

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

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

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

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

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

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,

$$\operatorname{tr}_R(|\psi_i\rangle |i\rangle) = |\psi_i\rangle .$$

(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

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

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

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

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

About $|i\rangle$,

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

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

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

Problem 2.1)

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

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

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

Thus

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

Problem 2.2) Unsolved

Problem 2.3) Unsolved

Chapter 3

Introduction to computer science

3.1) (Non-computable processes in Nature) How might we recognize that a process in Nature computes a function non computable by a Turing machine?
Soln: There are several well known non-Turing-computable functions which if identified to be computable by a process in nature would provide examples. For instance, the Halting problem: `https://en.wikipedia.org/wiki/Halting_problem`. More specifically, since Turing machines map non-negative integers to non-negative integers, their input and output spaces are countable (`https://en.wikipedia.org/wiki/Countable_set`). If any process in nature was found to compute a function taking input or providing output from an uncountable space, this could not be computed using a Turing machine. Note, Turing machines could compute the function within any desired level of approximation, but could not compute the function exactly.

3.2) (Turing numbers) Show that single-tape Turing machines can each be given a number from a list 1,2,3,. . . in such a way that the number uniquely specifies the corresponding machine. We call this number the *Turing number* of the corresponding machine. (*Hint*: Every positiive integer has a unique prime factorizatization $p_1^{a_1} p_2^{a_2} \dots p_k^{a_k}$, where p_i are distinct prime numbers and $a_1, \dots, a_k$ are non-negative integers.)
Soln: Per the hint, we show that a Turing machine can be encoded uniquely be a finite ordered list of integer values $[a_1, a_2, \dots, a_k]$. Unique prime factorization can be used to encode the Turing machine as the non-negative integer $\prod_i p_i^{a_i}$, where p_i is the i-th prime, starting with $p_1 = 2, p_2 = 3, \dots$. A non-negative integer corresponding to a Turing machine can then be decoded to reproduce the unique Turing machine whose encoding gives rise to it via the exponents in it's unique prime factorization. What follows is likely overly detailed for some. The basic idea is that each part of the the Turing machine can be encoded in a finite sequence of non-negative integers and decoded from that sequence. Concatenating those sequences (carefully) is then enough to specify the Turing machine. Note, it will not be the case that all non-negative integers correspond to valid Turing machines, but this is not required. We'll extend this encoding to encode Turing machines in operation, and explain how operation of a Turing machine can be simulated by multiplication of it's Turing number by a rational number determined conditionally by the Turing number itself. [Note: it is unclear how useful this extension will be. The idea is relatively simple, but its formal specification is intricate and very much not necessary to understand. Feel free to skip it.]

To produce an encoding of a Turing machine, we encode each of it's elements separately. We start with the finite state control. The finite state control consists of a finite set of $m+2$ states $Q = \{q_s, q_1, \dots, q_m, q_h\}$. Individually, it doesn't matter what form the q_i take, only that they are distinguishable. The integers $0, 1, \dots, m, m+1$ are distinguishable, so all that is required to encode a finite state machine is a single integer. So, setting $a_1 = m$ allows a_1 to track the size of the finite state machine and is enough to encode it.

To encode the tape, we let a_2 be the size of the alpthabet Γ: $a_2 = |\Gamma|$, where here Γ includes the starting character $\triangleright$, corresponding to tape value 0, and blank character b corresponding to tape value

$a_2 - 1$. The other states may be assumed to be non-negative integers $1, \ldots, a_2 - 2$. To encode the entirety of the tape, note that only a finite number of squares are non-blank. Let β be the largest index of a non-blank tape square. Then set $a_3 = \beta$, and for each tape square with index i, for $1 \le i \le \beta$, set a_{3+i} equal to the non-negative integer value assigned to the alphabet character occupying tape square i. All tape squares with index more than β are blank and need not be encoded. Note that by construction $a_4 = 0$ for all Turing machines, since tape square 1 always contains $\triangleright$, which was assigned value 0.

Next, we encode the program. The program contains a finite ordered list of program lines, say π of them. Set $a_{3+\beta+1} = \pi$. For $1 \le i \le \pi$, we encode program line i with a second prime factorization. Program line i consists of 5 elements: $\langle q_i, x_i, q_i', x_i', s_i \rangle$. Here, q_i and q_i' are states in Q which can be indexed with non-negative integers, say $\ell_{i,1}$ and $\ell_{i,3}$, with $0 \le \ell_{i,1}, \ell_{i,3} \le m + 1$. x_i and x_i' are characters in the alphabet Γ which can be indexed with non-negative integers, say $\ell_{i,2}$ and $\ell_{i,4}$, with $0 \le \ell_{i,2}, \ell_{i,4} < |\Gamma| (= a_2)$. s_i is an integer value that is either -1, 0, or 1. Setting $\ell_{i,5} = s_i$ directly leaves open the possibility that $\ell_{i,5} = -1$, which in turn will yield non-integer encodings of the program line. There are several ways to circumvent this, the likely easiest of which is to set $\ell_{i,5} = s_i + 1$. However, the author prefers setting $\ell_{i,5} = s_i \ \% \ 3$, the remainder of s_i when divided by 3 (its residue modulo 3). This allows $s_i = 0$ and $s_i = 1$ to be encoded as $\ell_{i,5} = 0$ and $\ell_{i,5} = 1$, which are natural Boolean indicators that the tape-head should advance to the right, but requires $s_i = -1$ be encoded as $\ell_{i,5} = 2$, indicating that the tape-head should move to the left. Now, to encode program line i, for $1 \le i \le \pi$, set $a_{3+\beta+1+i} = 2^{\ell_{i,1}} \cdot 3^{\ell_{i,2}} \cdot 5^{\ell_{i,3}} \cdot 7^{\ell_{i,4}} \cdot 11^{\ell_{i,5}}$.

Now, for a Turing machine M, assigning Turing number $\tau(M) = \prod_{i=1}^{3+\beta+1+\pi} p_i^{a_i}$ produces an integer encoding. To show that it is unique, we reverse the encoding process and argue that all pieces of the Turing machine can be recovered uniquely from this integer value. Let an encoding of a Turing machine, $\tau(M)$, be given and begin with its unique prime factorization $\tau(M) = \prod_{i=1}^{\omega(\tau(M))} p_i^{a_i}$, where here ω is a function that returns the largest index of a prime that divides input integer. [Note, $a_4 = 0$ will mean that this isn't the number of distinct prime factors]. Immediately, we recover the size of the finite state machine, *i.e.* m. It contains a_1 states indexed by integers, along with the special starting and halting state q_s and q_h. Next, the size of the alphabet Γ is given by a_2, where here Γ includes the starting character $\triangleright$ and the blank character b. Next, the encoding of the tape starts with $a_3 = \beta$ which indicates the maximum index of a non-blank tape square. β encodings of tape squares follow, starting with $a_4 = 0$, indicating that tape square 1 contains the starting character $\triangleright$, which was assigned character value 0. If $a_4 \ne 0$, the integer provided could not be an encoding of a Turing machine, violating the assumption that $\tau(M)$ was such an integer. a_{3+i} encodes the value stored on the tape at index i, for $i \le i \le \beta$, where $a_{3+i} = a_2 - 1$ indicates the tape square i is blank. All tape squares with index i, for $i > \beta$, are assumed to be blank. It is left only to decode the program. We start with its length $\pi = a_{3+\beta+1}$. π encodings of individual program lines should follow, each of which should be of the form $a_{3+\beta+1+i} = 2^{\ell_{i,1}} \cdot 3^{\ell_{i,2}} \cdot 5^{\ell_{i,3}} \cdot 7^{\ell_{i,4}} \cdot 11^{\ell_{i,5}}$, from which we can recover $q_i = \ell_{i,1}$, $x_i = \ell_{i,2}$, $q_i' = \ell_{i,3}$, $x_i' = \ell_{i,4}$, and $s_i = \ell_{i,5} \ \widetilde{\%} \ 3$, where here $\widetilde{\%} \ 3$ is modular reduction on to the set of residues $-1, 0,$ and 1, instead of the standard set of residues $0, 1, 2$. Note that $r \ \widetilde{\%} \ 3 = ((r+1) \ \% \ 3) - 1$. It is easy to see that the program line encoded by $a_{3+\beta_1+i}$ is uniquely determined, as is the initial state of the tape from $a_3, \ldots, a_{3+\beta}$. The alphabet Γ is uniquely determined by a_2, and the finite state machine is uniquely determined by its size, given by a_1. So, the entirety of the Turing machine M can be uniquely recovered from it's Turing number $\tau(M)$, so $\tau(M)$ is unique.

(Extension): Note that, as defined, our encoding uniquely encodes Turing machines in their initial state q_s, with read-write tape-head positioned on tape square 1 holding the starting character $\triangleright$. The encoding scheme could be extended to encode Turing machines in operation by adding an encoding of the current state in the finite state control and current position of the read-write tape-head which will require only two additional prime factors and exponents. For compatibility, to encode the current state of the finite state machine, we use $a_{3+\beta+1+\pi+1}$, where $a_{3+\beta+1+\pi+1} = 0$ indicates the state machine is in starting state q_s, and $a_{3+\beta+1+\pi+1} = m + 1$ indicates the state machine is in state q_h and has halted. To encode the position of the read-write tape-head we require one more additional prime factor and exponent. For compatibility, we use $a_{3+\beta+1+\pi+2}$. Full compatibility of encoding will require $a_{3+\beta+1+\pi+2} = 0$ to indicate that the machine is not yet operating and that the read-write tape-head has not yet been positioned on a tape square, neither tape square 1 holding $\triangleright$, as encoded by $a_4 = 0$, or another subsequent tape square holding any other value. Having $a_{3+\beta+1+\pi+2} > 0$ indicates that the Turing machine M is in operation in state specified by $a_{3+\beta+1+\pi+1}$, which we'll call σ, and read-write tape-head on the tape square specified by $a_{3+\beta+1+\pi+2}$, which we'll call σ. Note then that the tape-square pointed to by the read-write tape-head would contain the value specified by a_{3+a_σ}, which we'll call ν.

Now, to simulate execution of the Turing machine, note that in each step the program list is searched for a pattern matching it's current state and the character in the tape square being read by the read-write tape-head, that is, for $\langle \sigma, \nu, \cdot, \cdot, \cdot \rangle$. This is equivalent to searching $a_{4+\beta+1}, \dots, a_{4+\beta+\pi}$ for an integer divisible by $2^\sigma \cdot 3^\nu$, but no more powers of 2 or 3. Once a matching $a_{4+\beta+i}$ is found, the multiplicities of $5, 7$, and 11 in its factorization will give values for $\ell_{i,3}, \ell_{i,4}$, and $\ell_{i,5}$. The finite state machine can then be updated by multiplying by $p_{3+\beta+1+\pi+1}^{\ell_{i,3}-\sigma}$. The contents of the tape can be updated by multiplying by $p_{3+a_\sigma}^{\ell_{i,4}-\nu}$. The read-write tape-head can be moved by multiplying by $p_{3+\beta+1+\pi+2}^{\ell_{i,5} \, \widetilde{\%} \, 3}$. Doing so will change the Turing number of the machine in operation M to the number encoding M after a single step.

3.3) (Turing machine to reverse a bit string) Describe a Turing machine which takes a binary number x as input, and outputs the bits of x in reverse order. (*Hint*: In this and the next exercise it may help to use a mutli-tape Turing machine and/or symbols other than $\triangleright$, 0, 1, and the blanks.)
Soln: By "takes a binary number x as input", what is meant is the non-blank portion of the tape contains the value x, encoded somehow. In general, the tap can hold more than just function input, but for this problem that won't be necessary. We'll use a two-tape machine, with both tapes containing symbols from the alphabet $\triangleright, 0, 1, b$. Tape 1 will contain $\triangleright$, followed by the input value x in binary, followed by blanks indicated with bs. The second tape will contain $\triangleright$ and blanks. The Turing machine will populate the second tape with the reversed binary value of x, followed by blanks. It will not clear the first tape (although that could be done without too much trouble). Before we define the program, we specify that the finite state machine will contain 4 states, the starting state q_s, the halted state q_h, a search state s, and a write state w. Now, consider the program

$$P = \begin{cases} 1: & \langle \; q_s, & \triangleright, & \triangleright, & s, & \triangleright, & \triangleright, & +1, & 0 \; \rangle \\ 2: & \langle \; s, & 0, & \triangleright, & s, & 0, & \triangleright, & +1, & 0 \; \rangle \\ 3: & \langle \; s, & 1, & \triangleright, & s, & 1, & \triangleright, & +1, & 0 \; \rangle \\ 4: & \langle \; s, & b, & \triangleright, & w, & b, & \triangleright, & -1, & +1 \; \rangle \\ 5: & \langle \; w, & 0, & b, & w, & 0, & 0, & -1, & +1 \; \rangle \\ 6: & \langle \; w, & 1, & b, & w, & 1, & 1, & -1, & +1 \; \rangle \\ 7: & \langle \; w, & \triangleright, & b, & q_h, & \triangleright, & b, & 0, & 0 \; \rangle \end{cases}$$

Execution of the Turing machine begins by executing line 1 of P, which moves tape-head 1 forward, leaves tape-head 2 in place, and sets the state of the finite state machine to the search state. While in the search state, the program operates by executing lines 2 and 3, advancing tape-head 1 leaving the content of tape 1 unchanged, until it reaches a blank indicating that the end of the input has been reached, finally matching line 4. Once the blank on tape 1 is reached, line 4 changes the finite state machine to the write state, shifts tape-head 1 to the last bit of input, and advancing tape-head 2 to the first position in tape 2. Until the start of tape 1 is encountered, the program operates by executing lines 5 and 6, each of which copies the character on tape 1 pointed to be tape-head 1 onto tape 2 in the position pointed to by tape-head 2. It then moves the tape-heads in opposite directions so that tape-head 1 points to the preceeded bit of input and tape-head 2 points to the next bit of output. When the start of tape 1 is encountered, line 7 explicitly halts the program. Explicitly halting is not necessary in this case. Note, tape 1 could be cleared by replacing x_1' in lines 5 and 6 with bs. Then, the output could be moved to tape 1 while simultaneously clearing tape 2 by replacing line 7 with:

$$\begin{array}{rlrrrrrrrrl}
7: & \langle & w, & \triangleright, & b, & w, & \triangleright, & b, & 0, & -1 & \rangle \\
8: & \langle & w, & \triangleright, & 0, & w, & \triangleright, & 0, & 0, & -1 & \rangle \\
9: & \langle & w, & \triangleright, & 1, & w, & \triangleright, & 1, & 0, & -1 & \rangle \\
10: & \langle & w, & \triangleright, & \triangleright, & w, & \triangleright, & \triangleright, & +1, & +1 & \rangle \\
11: & \langle & w, & b, & 1, & w, & 1, & b, & +1, & +1 & \rangle \\
12: & \langle & w, & b, & 0, & w, & 0, & b, & +1, & +1 & \rangle \\
13: & \langle & w, & b, & b, & q_h, & b, & b, & 0, & 0 & \rangle
\end{array}$$

Here, line 7 reverses the direction of tape-head 2. Lines 8 and 9 allow it to retreat to the start of tape 2 in line 10, at which point tape-head 1 and 2 are advanced in tandem and the blank in tape 1 is swapped with the bit in tape 2, one bit at a time, by executing lines 11 and 12. Once tape-head 2 is at the end of the reversed bit-string, line 13 is reached, explicitly halting the program.

3.4) (Turing machine to add modulo 2) Describe a Turing machine to add two binary numbers x and y modulo 2. The numbers are input on the Turing machine tape in binary, in the form x, followed by a single blank, followed by y. If one number is not as long as the other then you may assume that it has been padded with leading 0's to make the two numbers the same length.

Soln: The specification that x and y can be padded so that they have the same length clouds the interpretation of the exercise. If adding x and y modulo 2 means finding the parity of $x + y$, the more natural interpretation to a mathematician, then a rather natural machine achieves this without padding. Alternatively, adding x and y modulo 2 could mean $x \wedge y$. Here, padding would be convenient. We'll start with the first interpretation, the parity of $x + y$, that is $x + y \pmod 2$. Here, only the last bits of x and y matter.

We define a single-tape machine, using the standard alphabet $\Gamma = \{\triangleright, 0, 1, b\}$. For convenience, define a set of states $S = \{s, 0, 1, h\}$ and let the finite state control consist of (a subset of) $S \bigoplus S$, the Cartesian product of S with itself, with $q_s \equiv (s, s)$ and $q_h \equiv (h, h)$. This will allow the state to represent a bit from x and a bit from y simultaneously. For convenience, let a $*$ in a state contained within a program line be a wildcard, where a $*$ will only occur in an output state if one also occurred in the input state in the same coordinate, in which case the coordinate state is left unchanged in the output state. For added convenience, let the # wildcard represent a 0 or 1 character read from the tape, and a corresponding 0 or 1 state in a coordinate of the output state. The # character will never be used as an output character to be written to the tape. The program below could be specified without wildcards by replicating the program lines containing them, producing a fully specified program line for each value the wildcards could represent. If multiple wildcards occur in a line, the result would be distinct fully specified program lines, one for each pair of state, character pair in $S \bigoplus \{0, 1\}$.

$$P = \left\{ \begin{array}{rlcccccl}
1: & \langle & (s,s), & \triangleright, & (s,s), & \triangleright, & +1 & \rangle \\
2: & \langle & (*,s), & \#, & (\#,s), & b, & +1 & \rangle \\
3: & \langle & (*,s), & b, & (*,b), & b, & +1 & \rangle \\
4: & \langle & (*,b), & \#, & (*,\#), & b, & +1 & \rangle \\
5: & \langle & (*,0), & \#, & (*,\#), & b, & +1 & \rangle \\
6: & \langle & (*,1), & \#, & (*,\#), & b, & +1 & \rangle \\
7: & \langle & (*,0), & b, & (*,0), & b, & -1 & \rangle \\
8: & \langle & (*,1), & b, & (*,1), & b, & -1 & \rangle \\
9: & \langle & (0,0), & \triangleright, & (0,h), & \triangleright, & +1 & \rangle \\
10: & \langle & (0,1), & \triangleright, & (1,h), & \triangleright, & +1 & \rangle \\
11: & \langle & (1,0), & \triangleright, & (1,h), & \triangleright, & +1 & \rangle \\
12: & \langle & (1,1), & \triangleright, & (0,h), & \triangleright, & +1 & \rangle \\
13: & \langle & (0,h), & b, & (h,h), & 0, & 0 & \rangle \\
14: & \langle & (1,h), & b, & (h,h), & 1, & 0 & \rangle
\end{array} \right.$$

Line 1 initializes the program, advancing the tape head to the first bit of x. Line 2 iteratively stores the last bit of x read from the tape in the first coordinate of the state and erases the bit from the tape, advancing the tape head to the next bit of x. Line 3 identifies the end of x and advances to the first bit of y. Lines 4-6 iteratively store the last bit of y read from the tape in the second coordinate of the state and erases the bit from the tape, advancing the tape head to the next bit of y. Lines 7 and 8 identify the end of y, then iteratively returns the tape head to the start of the tape, since the tape has been erased entirely except for the starting $\triangleright$ character, which triggers one of lines 9-12 when read. While the tape head is returning to the start of the tape, and in particular, when it reaches it, the current state contains the last bit of x and the last bit of y by construction. In lines 9-12, these bits are XORed, the tape head is advanced to the first position (which holds the blank character), and the output state is set equal to the XOR paired with h in the second coordinate indicating halting is in progress. The XOR is the parity of $x + y$ that needs to be output onto the tape. Once the second state-coordinate indicates halting, the XOR is written to the tape in the current/first character, replacing the blank that was there, and the output state is explicitly set to $(h, h) \equiv q_h$, the halting state of the Cartesian product finite state control.

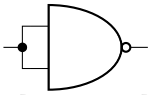
To compute $x \wedge y$ we use a two-tape machine. The finite state control will consist of the states $\{q_s, s, m, r, w, q_h\}$, which we'll refer to as begin, search, move, return, write, and halt. The standard alphabet $\Gamma = \{\triangleright, 0, 1, b\}$ will be sufficient. Once again, we use # wildcards, where now a subscript will indicate which tape the character was read from, and $\#_1 \wedge \#_2$ is the XOR of the binary integers represented by the characters $\#_1$ and $\#_2$ read from the respective tapes. Consider the program:

$$P = \begin{cases} \begin{array}{rcccccccccl} 1: & \langle & q_s, & \triangleright, & \triangleright, & s, & \triangleright, & \triangleright, & +1, & 0 & \rangle \\ 2: & \langle & s, & \#_1, & \triangleright, & s, & \#_1, & \triangleright, & +1, & 0 & \rangle \\ 3: & \langle & s, & b, & \triangleright, & m, & b, & \triangleright, & +1, & +1 & \rangle \\ 4: & \langle & m, & \#_1, & b, & m, & b, & \#_1, & +1, & +1 & \rangle \\ 5: & \langle & m, & b, & b, & r, & b, & b, & -1, & -1 & \rangle \\ 6: & \langle & r, & b, & \#_2, & r, & b, & \#2, & -1, & -1 & \rangle \\ 7: & \langle & r, & b, & \triangleright, & r, & b, & \triangleright, & -1, & 0 & \rangle \\ 8: & \langle & r, & \#_1, & \triangleright, & r, & \#_1, & \triangleright, & -1, & 0 & \rangle \\ 9: & \langle & r, & \triangleright, & \triangleright, & w, & \triangleright, & \triangleright, & +1, & +1 & \rangle \\ 10: & \langle & w, & \#_1, & \#_2, & w, & \#_1 \wedge \#_2, & b, & +1, & +1 & \rangle \\ 11: & \langle & w, & b, & b, & q_h, & b, & b, & 0, & 0 & \rangle \end{array} \end{cases}$$

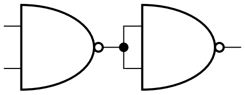
Line 1 starts the program, transitioning the machine to the search state, advancing the first tape to the first bit of x. Here, "search" is searching for the start of y. While in the search state, line 2 causes tape head 2 to stay in place, advances tape head 1, and leaves the contents of x unchanged on tape 1. Once tape head 1 reaches the end of x, line 3 transitions the finite state control to the move state, and advances both tape heads. In the move state, the bits of y are iteratively moved from tape 1 to tape 2, erasing tape 1, by executing line 4. Since x and y have the same length, this aligns x and y in their respective tapes, with corresponding bits in tape squares with equal indices. When tape head 1 reaches the end of y, line 5 sets the state to return, and retreats both tapes. Line 6 is then executed iteratively, retreating tape head 1 across the erased portion of tape 1 which previously held y, and tape head 2 across the portion of tape 2 that *now* holds y, leaving y in place. Since x and y have the same length, we reach the blank which previously separated x and y on tape 1 and the $\triangleright$ on tape 2 at the same time. This executes line 7, which transitions the machine to only retreating tape head 1, while leaving tape head 2 in place on the starting $\triangleright$ character. Then line 8 is executed iteratively, continuing to retreat tape head 1 to its start. When both tape heads have returned to their starts, we execute line 9, transitioning into the write state and move both tape heads forward: tape head 1 to the first bit of x and tape head 2 to the first bit of y. While in the write state,executing line 10 writes the bitwise XOR of x and y onto tape 1 and erases tape 2, advancing both tape heads in turn. Since x and y have the same length, we are guaranteed to reach the blank characters at the ends of x and y at the same time, causing line 11 to be executed, explicitly halting the program, leaving tape 1 containing $x \wedge y$ and tape 2 blank.

3.8) (Universality of `NAND`) Show that the `NAND` gate can be used to simulate `AND`, `XOR`, and `NOT` gates, provided wires, ancilla bits and `FANOUT` are available

NOT:



AND:



(Landau Big-O notation) A brief note about wording: Section 3.2.1 says that $f(n)$ *is* $O(g(n))$, but this leads to a temptation to write equations like $f(n) = O(g(n))$. Such equations aren't reflexive though; writing$O(g(n)) = f(n)$ doesn't make sense. We'll say that$f(n)$ is *in* $O(g(n))$, where $O(g(n))$ is taken to be a *class* of functions, as originally presented. So, *in* refers to class membership and may be written $f(n) \in O(g(n))$. Also, when a parameter is used inside the O, Ω, or Θ, it is unnecessary to include the parameter in statements about function's membership in such a class. For example, if $f(n) \equiv n$, then $f \in \Theta(n)$ is unambiguous.

3.9) Prove that $f(n)$ is $O(g(n))$ if and only if $g(n)$ is $\Omega(f(n))$. Deduce that $f(n)$ is $\Theta(g(n))$ if and only if $g(n)$ is $\Theta(f(n))$.
Soln: Let us assume that $f(n) \in O(g(n))$. By definition there exists c and n_0 such that for all $n > n_0$, $f(n) \leq c \cdot g(n)$. Dividing by c and setting $c' = \frac{1}{c}$ yields that there exists c' and (the same) n_0 such that for all $n > n_0$, $c' \cdot f(n) \leq g(n)$. This is exactly the defining property of the statement: $g(n) \in \Omega(f(n))$, so $f(n) \in O(g(n))$ implies $g(n) \in \Omega(f(n))$. to prove the converse, assume that $g(n) \in \Omega(f(n))$, divide by c' to re-recover c, and recognize the definition of $f(n) \in O(g(n))$.

Now, to show that $f(n) \in \Theta(g(n))$ if and only if $g(n) \in \Theta(f(n))$, note that by definition$f(n) \in \Theta(g(n))$ means that $f(n) \in O(g(n))$ and $f(n) \in \Omega(g(n))$, which by the first part of the exercise means that $g(n) \in \Omega(f(n))$ and $g(n) \in O(f(n))$, which is exactly the definition of $g(n) \in \Omega(f(n))$.

3.10) Suppose $g(n)$ is a polynomial of degree k. Show that $g(n)$ is $O(n^\ell)$ for any $\ell \geq k$.
Soln: Let k be fixed and $\ell \geq k$ be given. Let $g(n) \equiv \sum_{i=0}^{k} c_i n^i$, define $c \equiv 2 \cdot \sum_{j=0}^{k} |c_j|$ and let $n > 2 \equiv n_0$.

$$\begin{aligned}
g(n) &= \sum_{i=0}^{k} c_i n^i && \text{(definition)} \\
&\leq \sum_{i=0}^{k} |c_i| n^i && (c_i \leq |c_i|) \\
&\leq \sum_{i=0}^{k} \left(\sum_{j=0}^{k} |c_j| \right) n^i && (|c_i| \leq |c_0| + \ldots + |c_i| + \ldots + |c_k|) \\
&= \left(\frac{c}{2}\right) \sum_{i=0}^{k} n^i && \text{(definition of } c\text{)} \\
&= \left(\frac{c}{2}\right) \left(n^k + \sum_{i=0}^{k-1} n^i \right) && \text{(separate leading term)} \\
&= \left(\frac{c}{2}\right) \left(n^k + \frac{n^k - 1}{n - 1} \right) && \text{(finitie geometric series)} \\
&< \left(\frac{c}{2}\right) \left(n^k + \frac{n^k}{1} \right) && \text{(denominator} > 1\text{, numerator smaller)} \\
&= c \cdot n^k && \text{(simplify)} \\
&\leq c \cdot n^\ell && (\ n > 1)
\end{aligned}$$

By definition, $g \in O(n^\ell)$.

3.11) Show that $\log n$ is $O(n^k)$ for any $k > 0$.
Soln: We prove the exercise using the *natural* log. Logarithms in other bases differ from the natural log by constant multiplicative factors which can be absorbed into the constant c, so the result will hold for logarithms in any base. Let $k > 0$ be given and define $\kappa = \left\lceil \frac{1}{k} \right\rceil$. Note that κ is an integer with value at least 1. Define $c = (\kappa!)^{1/\kappa}$, $n_0 = 1$, and let $n > n_0$. Now, consider the Taylor series expansion of $e^x = \sum_{i=0}^{\infty} \frac{x^i}{i!}$, which converges and is valid for all $x \in \mathbb{R}$. In particular, this series is valid for $x = (\kappa! n)^{1/\kappa}$.

$$\begin{aligned}
e^{(\kappa! n)^{1/\kappa}} &= \sum_{i=0}^{\infty} \frac{(\kappa! n)^{i/\kappa}}{i!} && \text{(Taylor series)}\\
&> \frac{(\kappa! n)^{\kappa/\kappa}}{\kappa!} && \text{(all terms positive, take only the } \kappa \text{ term)}\\
&= n && \text{(simplify)}\\
\log n &< (\kappa!)^{1/\kappa} n^{1/\kappa} && \left(\begin{array}{r}\text{take logarithm, increasing functions preserve}\\ \text{inequalities, but we've switch sides}\end{array}\right)\\
&= c n^{1/\kappa} && \text{(definition of } c\text{)}\\
&< c n^k && (n > 1,\ \kappa = \lceil 1/k \rceil \geq 1/k \implies k \geq 1/\kappa)
\end{aligned}$$

So $\log \in O(n^k)$.

3.12) ($n^{\log n}$ is super-polynomial) Show that n^k is $O(n^{\log n})$ for any k, but that $n^{\log n}$ is never $O(n^k)$.
Soln: To avoid assumptions about behavior of constant exponents and asymptotic functions (as opposed to constant multiples), in this problem we allow the logarithm base, say a, be arbitrary, but we must assume $a > 1$. For $a \leq 1$, the result is not true. Let k be given, define $c \equiv 1$, $n_0 \equiv \max(1, a^k)$, and let $n > n_0$.

$$\begin{aligned}
c \cdot n^{\log_a n} &= n^{\log_a n} && \text{(definition of } c\text{)}\\
&\geq n^{\log_a a^k} && \left(\begin{array}{r} n > 1, a > 1 \Rightarrow n^t \text{ and } \log_a t \text{ increasing}\\ n > a^k \text{ by construction}\end{array}\right)\\
&= n^k && \text{(simplify)}
\end{aligned}$$

So $n^k \in O(n^{\log_a n})$. To show that $n^{\log_a n} \notin O(n_k)$, we again let k be given, $c > 0$ be any fixed constant, and let n_0 be a fixed positive integer. We show there exists $n > n_0$ such that $n^{\log_a n} > c \cdot n^k$. Consider, for example, $n = log_{n_0} a + k$ ... to be continued

3.13)
3.14)

3.15) (Lower bound for compare-and-swap based sorts) Suppose an n element list is sorted by applying some sequence of compare-and-swap operations to the list. There are $n!$ possible initial orderings of the list. Show that after k of the compare-and-swap operations have been applied, at most 2^k of the possible initial orderings will have been sorted into the correct order. Conclude that $\Omega(n \log n)$ compare-and-swap operations are required to sort all posible initial orderings into the correct order.
Soln: In each compare-and-swap operation there are two choices, either swap or do not swap according to the result of the comparison. After each comparison-based decision, no matter how the next pair of entries to compare is decided, there are still only two options, swap or do not swap. The resulting logical control flow can be modeled with a binary tree with depth at most k. Each leaf of this binary

tree corresponds to a sequence of compare-and-swap operations, where we've decided to swap after some comparisons and not after others. Swaps apply transpositions to the original permuted list. At a leaf, the original list will be sorted if and only if the product of transpositions applied is equal to the inverse of the permutation representing the original order of the list. Importantly, this can only be true for a single ordering of the original list. So, each leaf corresponds to having sorted a single ordering of the list, so after k compare-and-swap operations at most 2^k initial orderings can be sorted.

In order for a compare-and-swap based sorting algorithm to sort all possible orderings of an n-long list, we need there to be at least $n!$ leaves in the binary tree. As there are at most 2^k leaves, we need k to be such that $2^k \geq n!$. It can easily be show that $n! > \left(\frac{n}{2}\right)^{\frac{n}{2}}$, so we need $2^k \geq \left(\frac{n}{2}\right)^{\frac{n}{2}}$. Applying $\log_2$ to both sides yields $k \geq \left(\frac{n}{2}\right)(\log_2 n - 1)$. This is easily seen to imply that $k \in \Omega(n \log_2 n)$.

Chapter 4

Quantum circuits

4.1) In Exercise 2.11, which you should do now if you haven't already done it, you computed the eigenvectors of the Pauli matrices. Find the points on the Block sphere which correspond to the normalized eigenvectors of the different Pauli matrices.
Soln: The normalized eigenvectors of X are $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, and $|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$. Let's consider $|+\rangle$. On the Bloch sphere, $|+\rangle$ has $\theta = 2\arccos(\frac{1}{\sqrt{2}}) = \frac{\pi}{2}$, and $\phi = \frac{\ln\left(\frac{1}{\sin(\frac{\pi}{2})}\right)}{i} = \frac{\ln(1)}{i} = 0$. This corresponds to the Bloch vector $|+\rangle \simeq (1,0,0)$. For $|-\rangle$ we again have $\theta = \frac{\pi}{2}$, but now $\phi = \frac{\ln\left(\frac{-1}{\sin(\frac{\pi}{2})}\right)}{i} = \pi$, so $|-\rangle$ corresponds to a Bloch vector $|-\rangle \simeq (-1,0,0)$.

The normalized eigenvectors of Y are $\frac{1}{2}(|0\rangle \pm i\,|1\rangle) = |\psi_{y^\pm}\rangle$. For $|\psi_{y^+}\rangle$, $\theta = 2\arccos(\frac{1}{\sqrt{2}}) = \frac{\pi}{2}$, and $\phi = \frac{\ln\left(\frac{i}{\sin(\frac{\pi}{2})}\right)}{i} = \frac{\pi}{2}$, so $|\psi_{y^+}\rangle \simeq (0,1,0)$. Similarly, for $|\psi_{y^-}\rangle$ we have $\phi = -\frac{\pi}{2}$, in which case $|\psi_{y^-}\rangle \simeq (0,-1,0)$.

The normalized eigenvectors of Z are $|0\rangle$ and $|1\rangle$. For $|0\rangle$, $\theta = 0$ and ϕ is indeterminate. Still $|0\rangle \simeq (0,0,1)$. For $|1\rangle$, $\theta = \pi$ and again ϕ is indeterminate, but $|1\rangle \simeq (0,0,-1)$.

4.2)
4.3) Show that, up to a global phase, the $\pi/8$ gate satifies $T = R_z(\pi/4)$.
Soln: What is meant by $T = R_z(\pi/4)$ is that the result of applying both operations to the same input vector results in outputs that are scalar multiples. That is, $T = e^{i\phi}R_z(\pi/4)$, for some θ. It is given that $T = \exp(i\pi/8)\begin{bmatrix}\exp(-i\pi/8) & 0 \\ 0 & \exp(i\pi/8)\end{bmatrix} = \exp(i\pi/8)R_z(\pi/4)$, so $\phi = \pi/8$ suffices.

4.4) Express the Hadamard gate H as a product of R_x and R_z rotations and $e^{i\phi}$ for some ϕ.
Soln: First, note that $H = \frac{X+Z}{\sqrt{2}}$, and by equations (4.4) and (4.6),

$$R_x(-\pi/2) = e^{\frac{i\pi X}{4}} = \cos(-\pi/4)I - i\sin(-\pi/4)X = \frac{I+iX}{\sqrt{2}} \text{ and}$$
$$R_z(-\pi/2) = e^{\frac{i\pi Z}{4}} = \cos(-\pi/4)I - i\sin(-\pi/4)Z = \frac{I+iZ}{\sqrt{2}}$$

Now:

$$\begin{aligned}
R_x(-\pi/2)R_z(-\pi/2)R_x(-\pi/2) &= \frac{(I+iX)}{\sqrt{2}}\frac{(I+iZ)}{\sqrt{2}}\frac{(I+iX)}{\sqrt{2}} \\
&= \frac{1}{2\sqrt{2}}(I+2iX+iZ-ZX-XZ-X^2-iXZX) \\
&= \frac{i}{\sqrt{2}}(X+Z) && (X^2=I,\, XZ=-ZX,\, XZX=-Z) \\
&= iH.
\end{aligned}$$

So, $H = e^{-i\pi/2}R_x(-\pi/2)R_z(-\pi/2)R_x(-\pi/2)$. $\phi = \pi/2$ suffices.

4.5) Prove that $(\hat{n}\cdot\vec{\sigma})^2 = I$, and use this to verify Equation (4.8).
Soln: $\hat{n} = (n_x, n_y, n_z)$, and $\vec{\sigma} = (X, Y, Z)$, so

$$\begin{aligned}
(\hat{n}\cdot\vec{\sigma})^2 &= (n_xX+n_yY+n_zZ)^2 \\
&= n_x^2X^2 + n_xn_yXY + n_xn_zXZ + n_yn_xYX + n_y^2Y^2 + n_yn_zYZ + n_zn_xZX + n_zn_yZY + n_z^2Z^2 \\
&= (n_x^2+n_y^2+n_z^2)I && (XY=-YX,\, XZ=-XZ,\, YZ=-ZY,\, X^2=Y^2=Z^2=I) \\
&= \|\hat{n}\|^2 I \\
&= I && (\hat{n} \text{ is a unit vector})
\end{aligned}$$

4.6)
4.7) Show that $XYX = -Y$ and use this to prove that $XR_y(\theta)X = R_y(-\theta)$.
Soln: Since $YX = -XY$, we have $XYX = X(-XY) = -X^2Y = -Y$. Now,

$$\begin{aligned}
XR_y(\theta)X &= X\Big(\cos(\theta/2)I - i\sin(\theta/2)Y\Big)X \\
&= \cos(\theta/2)X^2 - i\sin(\theta/2)XYX \\
&= \cos(\theta/2) + i\sin(\theta/2)Y && (X^2=I,\, XYX=-Y) \\
&= \cos(-\theta/2) - i\sin(-\theta/2)Y && (\cos(\phi)=\cos(-\phi),\, \sin(\phi)=-\sin(-\phi)) \\
&= R_y(-\theta). && (\text{definition of } R_y)
\end{aligned}$$

4.8) An arbitrary single qubit unitary operator can be written in the form

$$U = \exp(i\alpha)R_{\hat{n}}(\theta)$$

for some real numbers α and θ, and a real three-dimensional unit vector $\hat{n}$.

1. Prove this fact.
2. Find values for α, θ, and $\hat{n}$ giving the Hadamard gate H.
3. Find values for α, θ, and $\hat{n}$ giving the phase gate

$$S = \begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix}.$$

Soln: We skip proving fact #1. Note that $H = (X+Z)/\sqrt{2}$, and setting $\theta = \pi/2$, $\alpha = \pi/2$, and $\hat{n} = (1,0,1)/\sqrt{2}$ gives $\hat{n}\cdot\vec{\sigma} = (X+Z)/\sqrt{2} = H$ and

$$\begin{aligned}
\exp(i\alpha)R_{\hat{n}}(\theta) &= e^{i\pi/2}\Big(\cos(\pi/2)I - i\sin(\pi/2)H\Big) \\
&= -i^2H = H && (e^{i\pi/2}=i, -i^2=1)
\end{aligned}$$

For S, set $\theta = \pi/4$, $\alpha = \pi/4$, and $\hat{n} = (0,0,1)$ so that $\hat{n}\cdot\vec{\sigma} = Z$ and

$$\begin{aligned}\exp(i\alpha)R_{\hat{n}}(\theta) &= e^{i\pi/4}\Big(\cos(\pi/4)I - i\sin(\pi/4)Z\Big)\\ &= \frac{1+i}{\sqrt{2}}\left(\frac{I}{\sqrt{2}} - \frac{iZ}{\sqrt{2}}\right)\\ &= \frac{1+i}{2}\begin{bmatrix}1-i & 0\\ 0 & 1+i\end{bmatrix} = \begin{bmatrix}1 & 0\\ 0 & i\end{bmatrix} = S\end{aligned}$$

4.9)
4.10)
4.11)
4.12)
4.13) (Circuit identities) It is useful to be able to simplify circuits by inspection, using well-known identities. Prove the following three identities:

$$HXH = Z; \quad HYH = -Y; \quad HZH = X$$

Soln: Note that $X^2 = Z^2 = I, H = (X+Z)/\sqrt{2}$, and that $XZ = -ZX$, so

$$\begin{aligned}HXH &= \frac{1}{2}(X+Z)X(X+Z)\\ &= \frac{1}{2}(X^3 + X^2Z + ZX^2 + ZXZ)\\ &= \frac{1}{2}(X+Z+Z-X) = Z && (X^3 = X;\ ZXZ = -Z^2X = -X)\end{aligned}$$

Since H is self-inverse, if follows easily that $HZH = X$. It can be shown via direct calculation that $Y = iXZ = -iZX$, so $HYH = iHXZH = iH(HZH)(HXH)H = iZX = -Y$.

4.14) Use the previous exercies to show that $HTH = R_x(\pi/4)$, up to a global phase.
Soln: It can be verfied via linear algebra and simple complex arithmetic that

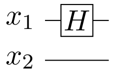
$$T = e^{-i\pi/8}\Big(\cos(\pi/8)I - i\sin(\pi/8)Z\Big).$$

Now

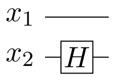
$$\begin{aligned}HTH &= \frac{e^{-i\pi/8}}{2}(X+Z)\Big(\cos(\pi/8)I - i\sin(\pi/8)Z\Big)(X+Z)\\ &= \frac{e^{-i\pi/8}}{2}\Big(\cos(\pi/8)X^2 + \cos(\pi/8)XZ - i\sin(\pi/8)XZX - i\sin(\pi/8)XZ^2\\ &\qquad + \cos(\pi/8)ZX + \cos(\pi/8)Z^2 - i\sin(\pi/8)Z^2X - i\sin(\pi/8)Z^3\Big)\\ &= \frac{e^{-i\pi/8}}{2}\Big(\cos(\pi/8)(X^2+Z^2) - i\sin(\pi/8)(2X + XZX + Z)\Big) && (XZ = -ZX,\ Z^2 = I)\\ &= e^{-i\pi/8}\Big(\cos(\pi/8)I - i\sin(\pi/8)X\Big) && (XZX = -Z)\\ &= e^{-i\pi/8}R_x(\pi/4)\end{aligned}$$

4.15)

4.16) (Matrix representation of multi-qubit gates) What is the 4×4 unitary matrix for the circuit



in the computation basis? What is the unitary matrix for the circuit

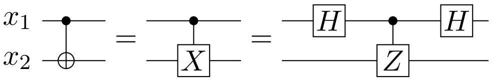


in the computational basis?
Soln: Note: we've changed the qubit labels so that reading states top to bottom corresponds to reading them left to right in the concatenated computation basis representation. The unitary matrices are:
$\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 0 & 1 & 0\\ 0 & 1 & 0 & 1\\ 1 & 0 & -1 & 0\\ 0 & 1 & 0 & -1\end{bmatrix}$ and $\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1 & 0 & 0\\ 1 & -1 & 0 & 0\\ 0 & 0 & 1 & 1\\ 0 & 0 & 1 & -1\end{bmatrix}$.

4.17) (Building `CNOT` from controlled-Z gates) Construct a `CNOT` gate from one controled-Z gate, that is, the gate whose action in the computational basis is specified by the unitary matrix

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

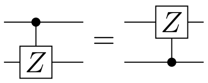
and two Hadamard gates, specifying the control and target qubits.
Soln: The Hadamard gate maps the eigenvectors of the X operator to those of the Z operator, so applying a Hadamard to the control before and after executing a controlled-Z should effect a `CNOT`.



In matrix form, this corresponds to

$$\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & 1 & 0\end{bmatrix} = \frac{1}{2}\begin{bmatrix}1 & 0 & 1 & 0\\ 0 & 1 & 0 & 1\\ 1 & 0 & -1 & 0\\ 0 & 1 & 0 & -1\end{bmatrix}\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & -1\end{bmatrix}\begin{bmatrix}1 & 0 & 1 & 0\\ 0 & 1 & 0 & 1\\ 1 & 0 & -1 & 0\\ 0 & 1 & 0 & -1\end{bmatrix}.$$

4.18) Show that

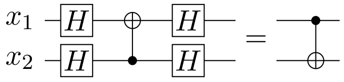


Soln: In the computational basis the controlled-Z changes the state if and only if the control and target qubits are both $|1\rangle$. In this criteria the control and target are interchangeable, so these gates perform the same action.

4.19) (`CNOT` action on density matrices) The `CNOT` gate is a simple permutation whose action on a density matrix ρ is to rearrange the elements in the matrix. Write out this action explicitly in the computational basis.
Soln:

$$|x_2x_1\rangle = \alpha\,|00\rangle + \beta\,|01\rangle + \gamma\,|10\rangle + \delta\,|11\rangle \xrightarrow{CX} \alpha\,|00\rangle + \beta\,|01\rangle + \underline{\delta}\,|10\rangle + \underline{\gamma}\,|11\rangle\,.$$

4.20) (`CNOT` basis transformation) Unlike ideal classical gates, ideal quantum gates do not have (as electrical engineers say) 'high-impedence' inputs. In fact, the role of 'control' and 'target' are arbitrary – they depend on what basis you think of a device as operating in. We have described how the `CNOT` behaves with respect to the computational basis, and in this description the state of the control qubit is not changed. However, if we work in a different basis then the control qubit *does* change: we will show that its phase is flipped depending on the state of the 'target' qubit! Show that



Introducing basis states $|\pm\rangle \equiv (|0\rangle \pm |1\rangle)/\sqrt{2}$, use this circuit identity to show that the effect of a `CNOT` with the first qubit as control and the second qubit as target is as follows:

$$\begin{aligned}
&|x_1\rangle\,|x_2\rangle\\
&\;|+\rangle\,|+\rangle \to |+\rangle\,|+\rangle\\
&\;|-\rangle\,|+\rangle \to |-\rangle\,|+\rangle\\
&\;|+\rangle\,|-\rangle \to |-\rangle\,|-\rangle\\
&\;|-\rangle\,|-\rangle \to |+\rangle\,|-\rangle\,.
\end{aligned}$$

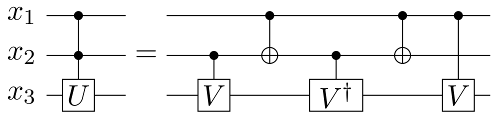
Thus, with respect to the this new basis, the state of the target qubit is not changed, while the state of the control qubit is flipped if the target starts as $|-\rangle$, otherwise it is left alone. That is, in this basis, the target and control have essentially interchanged roles!
Soln: Note, we've switched the roles of qubits in the diagram (and labeled them) so that reading states top to bottom in the diagram corresponds to reading them left to right in the concatenated basis representation. To show the circuit identity, we use matrix multiplication:

$$\frac{1}{4}\underset{H(x_1,\_)}{\begin{bmatrix}1&1&0&0\\1&-1&0&0\\0&0&1&1\\0&0&1&-1\end{bmatrix}}\underset{H(\_,x_2)}{\begin{bmatrix}1&0&1&0\\0&1&0&1\\1&0&-1&0\\0&1&0&-1\end{bmatrix}}\underset{\texttt{CNOT}(x_2,x_1)}{\begin{bmatrix}1&0&0&0\\0&0&0&1\\0&0&1&0\\0&1&0&0\end{bmatrix}}\underset{H(x_1,\_)}{\begin{bmatrix}1&1&0&0\\1&-1&0&0\\0&0&1&1\\0&0&1&-1\end{bmatrix}}\underset{H(\_,x_2)}{\begin{bmatrix}1&0&1&0\\0&1&0&1\\1&0&-1&0\\0&1&0&-1\end{bmatrix}}=\underset{\texttt{CNOT}(x_1,x_2)}{\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&0&1\\0&0&1&0\end{bmatrix}}.$$

The matrix in the middle on the left can be verified to be the representation of the action of `CNOT` in the computational basis, with x_2 as control, and x_1 as target. In functional notation below, we'll let $\texttt{CNOT}(x_1, x_2)$ denote the action of a `CNOT` controlled by $|x_1\rangle$ and targetting $|x_2\rangle$, with $\texttt{CNOT}'(x_1, x_2) \equiv \texttt{CNOT}(x_2, x_1)$ switching target and control, and let $H(x_1, x_2)$ denote the application of Hadamard gates to both qubits. The circuit identity is that $H(\texttt{CNOT}'(H(x_1, x_2)) = \texttt{CNOT}(x_1, x_2)$. Now:

$$\begin{aligned}
&(|x_1\rangle\,,|x_2\rangle)\\
&\texttt{CNOT}(|+\rangle\,,|+\rangle) = H(\texttt{CNOT}'(H(|+\rangle\,,|+\rangle))) = H(\texttt{CNOT}'(|0\rangle\,,|0\rangle)) = H(|0\rangle\,,|0\rangle) = |+\rangle\,|+\rangle\\
&\texttt{CNOT}(|-\rangle\,,|+\rangle) = H(\texttt{CNOT}'(H(|-\rangle\,,|+\rangle))) = H(\texttt{CNOT}'(|1\rangle\,,|0\rangle)) = H(|1\rangle\,,|0\rangle) = |-\rangle\,|+\rangle\\
&\texttt{CNOT}(|+\rangle\,,|-\rangle) = H(\texttt{CNOT}'(H(|+\rangle\,,|-\rangle))) = H(\texttt{CNOT}'(|0\rangle\,,|1\rangle)) = H(|1\rangle\,,|1\rangle) = |-\rangle\,|-\rangle\\
&\texttt{CNOT}(|-\rangle\,,|-\rangle) = H(\texttt{CNOT}'(H(|-\rangle\,,|-\rangle))) = H(\texttt{CNOT}'(|1\rangle\,,|1\rangle)) = H(|0\rangle\,,|1\rangle) = |+\rangle\,|-\rangle
\end{aligned}$$

4.21) Suppose U is a single qubit unitary operator, and V is a unitary operator chosen so that $V^2 = U$. Verify that Figure 4.8 implements the $C^2(U)$ operation.



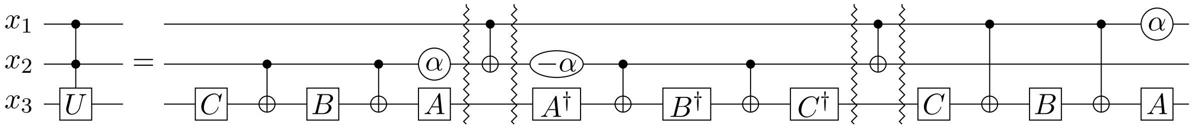
Soln: We verify by tracking the applications of V and $V^\dagger$ to $|x_3\rangle$ for each computational basis state representing $|x_1\rangle|x_2\rangle$. The first V is applied when $|x_2\rangle = |1\rangle$. The first CNOT calculates the parity of $|x_1\rangle|x_2\rangle$ so that the middle $V^\dagger$ is applied for when $|x_1\rangle|x_2\rangle = |0\rangle|1\rangle$ or $|1\rangle|0\rangle$. The second CNOT uncomputes the parity so that in the end $|x_1\rangle|x_2\rangle$ is unchanged. The final V is then applied if $|x_1\rangle = |1\rangle$. In the end, the output of the circuit is

$$\begin{aligned}
&|x_1\rangle|x_2\rangle \\
&|0\rangle|0\rangle|x_3\rangle \to |0\rangle|0\rangle \qquad |x_3\rangle \\
&|0\rangle|1\rangle|x_3\rangle \to |0\rangle|1\rangle VV^\dagger|x_3\rangle = |0\rangle|1\rangle \quad |x_3\rangle && (V \text{ is unitary, } VV^\dagger = I) \\
&|1\rangle|0\rangle|x_3\rangle \to |1\rangle|0\rangle V^\dagger V|x_3\rangle = |1\rangle|0\rangle \quad |x_3\rangle && (V \text{ is unitary, } V^\dagger V = I) \\
&|1\rangle|1\rangle|x_3\rangle \to |1\rangle|1\rangle \quad V^2|x_3\rangle = |1\rangle|1\rangle U|x_3\rangle && (V^2 = U)
\end{aligned}$$

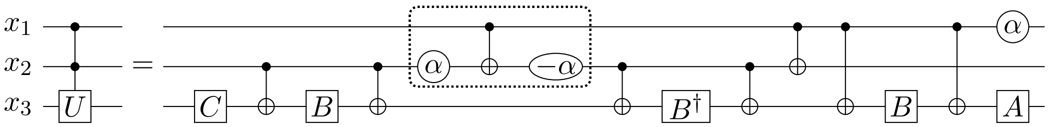
4.22) Prove that a $C^2(U)$ gate (for any single qubit unitary U) can be constructed using at most eight one-qubit gates, and six controlled-NOTs

Soln: Note: this solution borrows heavily from DaftWullie's answer to the quantumcomputing.stackexchange question here: https://quantumcomputing.stackexchange.com/questions/7082/how-to-reduce-circuit-elements-of-a-decomposed-c2u-operation.

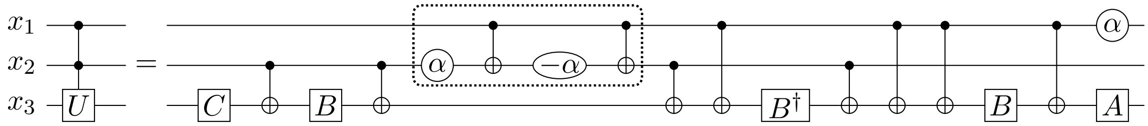
Let V be a single-qubit unitary operator such that $V^2 = U$, and By Corollary 4.2, let A, B, and C be single-qubit unitary operators and α an overall phase factor such that $ABC = I$ and $V = e^{i\alpha}AXBXC$. Combining figures 4.6 and 4.8 gives



where the $\pm\alpha$ gates correspond to the action of $\begin{bmatrix} 1 & 0 \\ 0 & e^{\pm i\alpha} \end{bmatrix}$. The $AA^\dagger$ and the $C^\dagger C$ cancel, each resulting in the identity, so we arrive at:



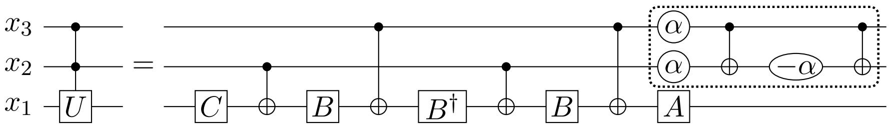
where we've drawn attention to the grouped α, $-\alpha$, and CNOT because we will investigate them as a group. First the group must gain another CNOT. The CNOT we'll gain is the 6th, but to gain it we need to temporarily introduce more CNOTs. Neither having been targeted by a CNOT prior, the result of the third CNOT is to calculate the "parity" of $|x_1\rangle$ and $|x_2\rangle$ in-place in $|x_2\rangle$ (phase of $|x_2\rangle$ ignored). The 4th and 5th CNOTs are then controlled by this parity. We can move the 6-th CNOT to before the 4th, which will uncompute the parity (but not the phase), but we'll need to reconfigure CNOT's 4 and 5 so that the result is once again a NOT applied to $|x_3\rangle$ controlled by the parity of $|x_1\rangle$ and $|x_2\rangle$, not just a NOT controlled by $|x_2\rangle$. To do this, note that applying CNOTs to $|x_3\rangle$ controlled by $|x_1\rangle$, then another controlled by $|x_2\rangle$ results in a NOT applied to $|x_3\rangle$ if and only if $|x_1\rangle|x_2\rangle = |0\rangle|1\rangle$ or $|1\rangle|0\rangle$, *i.e.* a CNOT controlled by the parity of $|x_1\rangle|x_2\rangle$. So, we arrive at



Now, the operation of the grouped gates on $|x_1\rangle\,|x_2\rangle$ can be expressed as

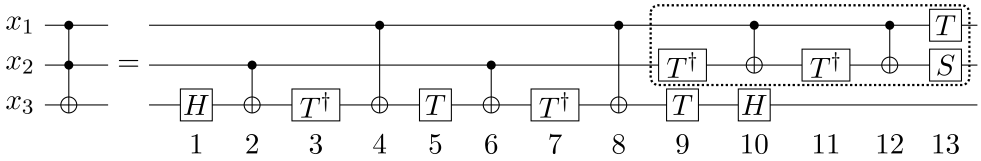
$$\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&0&1\\0&0&1&0\end{bmatrix}\begin{bmatrix}1&0&0&0\\0&e^{-i\alpha}&0&0\\0&0&1&0\\0&0&0&e^{-i\alpha}\end{bmatrix}\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&0&1\\0&0&1&0\end{bmatrix}\begin{bmatrix}1&0&0&0\\0&e^{i\alpha}&0&0\\0&0&1&0\\0&0&0&e^{i\alpha}\end{bmatrix}=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&e^{-i\alpha}&0\\0&0&0&e^{i\alpha}\end{bmatrix}$$

Being diagonal, the group can be transposed with any gate not targeting $|x_1\rangle$ or $|x_2\rangle$. Doing so will only change when the relative phase corrections are performed in the circuit, not the final result. We'll move it to the end to group the action only on $|x_1\rangle$ and $|x_2\rangle$. While we're at it, we'll move the final α to the start of the group, since it is diagonal and can be transposed with any gate not targeting $|x_1\rangle$. Doing this allows the two `CNOT`(x_2, x_3)s previously beside the group to cancel, along with the consecutive `CNOT`(x_1, x_3)s between the $B^\dagger$ and B, giving the final circuit below with 8 single-qubit gates (including the α's and $-\alpha$) and 6 `CNOT`s.



4.23)

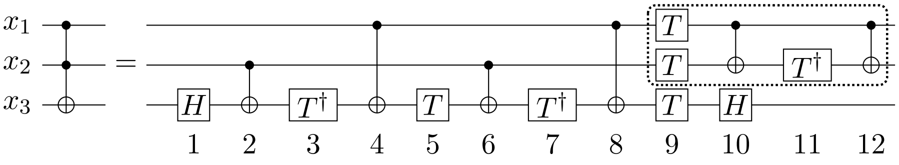
4.24) Verify that Figure 4.9 implements the Toffoli gate.



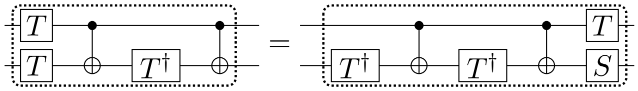
Soln: Define $V = \frac{1+i}{2}(I - iX), A = HT, B = T^\dagger, C = H$, and $\alpha = \pi/4$. Note that this choice of V is similar to that on page 182, but not exactly the same. Still, we have $V^2 = (\frac{1+i}{2})^2(I - iX)^2 = \frac{i}{2}(I - 2iX - X^2) = -i^2X = X$. Also, $ABC = HTT^\dagger H = I$, and

$$e^{i\alpha}AXBXC = \frac{1+i}{\sqrt{2}}HTXT^\dagger XH = \frac{1+i}{(\sqrt{2})^3}\begin{bmatrix}1&1\\1&-1\end{bmatrix}\begin{bmatrix}1&0\\0&e^{i\pi/4}\end{bmatrix}\begin{bmatrix}0&1\\1&0\end{bmatrix}\begin{bmatrix}1&0\\0&e^{-i\pi/4}\end{bmatrix}\begin{bmatrix}0&1\\1&0\end{bmatrix}\begin{bmatrix}1&1\\1&-1\end{bmatrix}$$
$$= \frac{1+i}{2}\begin{bmatrix}1&-i\\-i&1\end{bmatrix} = V$$

We conclude that the construction from Problem 4.22 applies. With $\alpha = \pi/4$, note that $\textcircled{-\alpha} = \boxed{T^\dagger}$ and $\textcircled{\alpha} = \boxed{T}$. So:



All that remains is to show that the grouped subcircuits are equivalent, that is, to show that



We do this by matrix mutiplication. On the left, we have:

$$\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&0&1\\0&0&1&0\end{bmatrix}\begin{bmatrix}1&0&0&0\\0&e^{-i\pi/4}&0&0\\0&0&1&0\\0&0&0&e^{-i\pi/4}\end{bmatrix}\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&0&1\\0&0&1&0\end{bmatrix}\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&e^{i\pi/4}&0\\0&0&0&e^{i\pi/4}\end{bmatrix}\begin{bmatrix}1&0&0&0\\0&e^{i\pi/4}&0&0\\0&0&1&0\\0&0&0&e^{i\pi/4}\end{bmatrix}$$

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

On the right:

$$\begin{bmatrix}1&0&0&0\\0&i&0&0\\0&0&1&0\\0&0&0&i\end{bmatrix}\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&e^{i\pi/4}&0\\0&0&0&e^{i\pi/4}\end{bmatrix}\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&0&1\\0&0&1&0\end{bmatrix}\begin{bmatrix}1&0&0&0\\0&e^{-i\pi/4}&0&0\\0&0&1&0\\0&0&0&e^{-i\pi/4}\end{bmatrix}\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&0&1\\0&0&1&0\end{bmatrix} \backslash$$

$$\cdot\begin{bmatrix}1&0&0&0\\0&e^{-i\pi/4}&0&0\\0&0&1&0\\0&0&0&e^{-i\pi/4}\end{bmatrix}=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&i\end{bmatrix}$$

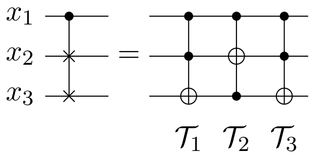
So, the subcircuits are equivalent and Figure 4.9 implements the Toffoli gate. Note that both subcircuits execute a controlled-S gate on $|x_2\rangle$, controlled by $|x_1\rangle$, The subcircuit in the construction provided by the exercise is the result of a direct application of Figure 6, where $A = S, B = C = T^\dagger$, and $\alpha = \pi/4$, since $ABC = S(T^\dagger)^2 = SS^\dagger = I$, and $e^{i\alpha}AXBXC = e^{i\pi/4}SXT^\dagger XT^\dagger = S$ can be easily verified via matrix-multiplication

4.25) Recall that the Fredkin (controlled-swap) gate performs the transform

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

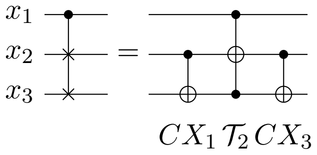
1. Give a quantum circuit which uses three Toffoli gates to construct the Fredkin gate (*Hint:* think of the swap gate construction - you can control each gate one at a time)
2. Show that the first and last Toffoli gates can be replaced by CNOT gates.
3. Now replace the middle Toffoli gate with the circuit in Figure 4.8 to obtain a Fredkin gate construction using only six two-qubit gates
4. Can you come up with an even simpler construction, with only five two-qubit gates?

Soln: For part 1, we'll show the following circuit identity by tracing the result of each computational basis state through each of the three Toffolis, indexed 1-3.



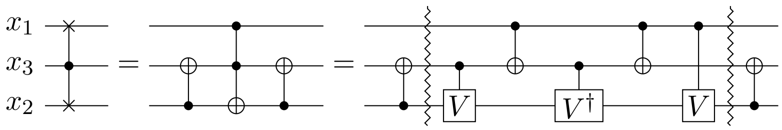
$\lvert x_1x_2x_3\rangle$	$\mathcal{T}_1$	$\mathcal{T}_2$	$\mathcal{T}_3$
$\lvert 000\rangle$	$\lvert 000\rangle$	$\lvert 000\rangle$	$\lvert 000\rangle$
$\lvert 001\rangle$	$\lvert 001\rangle$	$\lvert 001\rangle$	$\lvert 001\rangle$
$\lvert 010\rangle$	$\lvert 010\rangle$	$\lvert 010\rangle$	$\lvert 010\rangle$
$\lvert 011\rangle$	$\lvert 011\rangle$	$\lvert 011\rangle$	$\lvert 011\rangle$
$\lvert 100\rangle$	$\lvert 100\rangle$	$\lvert 100\rangle$	$\lvert 100\rangle$
$\lvert 101\rangle$	$\lvert 101\rangle$	$\lvert 111\rangle$	$\lvert 110\rangle$
$\lvert 110\rangle$	$\lvert 111\rangle$	$\lvert 101\rangle$	$\lvert 101\rangle$
$\lvert 111\rangle$	$\lvert 110\rangle$	$\lvert 110\rangle$	$\lvert 111\rangle$

For part 2, we replace $\mathcal{T}_1$ and $\mathcal{T}_3$ with $\texttt{CNOT}(x_2, x_3)$.

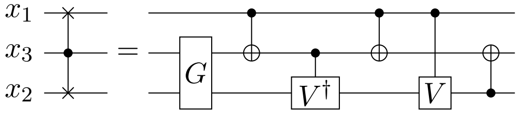


$\lvert x_1x_2x_3\rangle$	CX_1	$\mathcal{T}_2$	CX_3
$\lvert 000\rangle$	$\lvert 000\rangle$	$\lvert 000\rangle$	$\lvert 000\rangle$
$\lvert 001\rangle$	$\lvert 001\rangle$	$\lvert 001\rangle$	$\lvert 001\rangle$
$\lvert 010\rangle$	$\lvert 011\rangle$	$\lvert 011\rangle$	$\lvert 010\rangle$
$\lvert 011\rangle$	$\lvert 010\rangle$	$\lvert 010\rangle$	$\lvert 011\rangle$
$\lvert 100\rangle$	$\lvert 100\rangle$	$\lvert 100\rangle$	$\lvert 100\rangle$
$\lvert 101\rangle$	$\lvert 101\rangle$	$\lvert 111\rangle$	$\lvert 110\rangle$
$\lvert 110\rangle$	$\lvert 111\rangle$	$\lvert 101\rangle$	$\lvert 101\rangle$
$\lvert 111\rangle$	$\lvert 110\rangle$	$\lvert 110\rangle$	$\lvert 111\rangle$

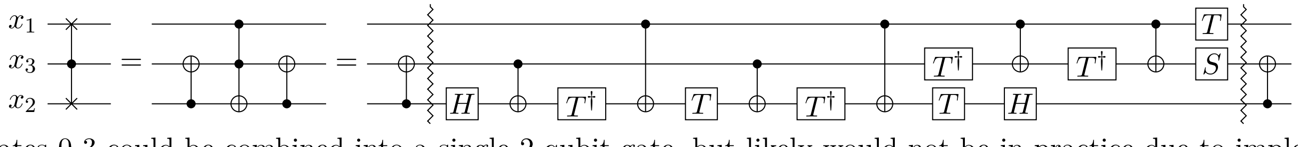
For part 3, we rearrange x_2 and x_3. To apply Figure 4.8 we let $U = X$, so that if $V = \frac{1-i}{2}(I + iX)$, then $V^2 = X = U$, and we obtain



This circuit contains 7 2-qubit gates, but consecutive applications of any 2-qubit gates to the same pair of qubits can be combined. The resulting 2-qubit gate can likely not be thought of a simple unitary applied to one of the qubits, controlled by the other though. Still, combining the first two gates into a 2-qubit gate, say G, gives

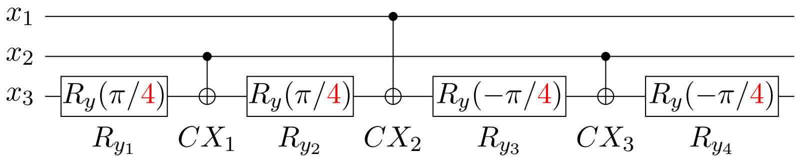


We answer part 4 in the negative in that, in the opinion of the second author, it is not worth removing another 2-qubit gate from the construction. Arbitrary 2-qubit gates are likely to be extremely hard to implement in practice. The controlled V and $V^\dagger$ are likely to be hard as well, so in practice, instead of using the circuit in Figure 4.8, that in Figure 4.9 would likely be used, resulting in a circuit more along the lines of:



Gates 0-3 could be combined into a single 2-qubit gate, but likely would not be in practice due to implementation difficulty.

4.26) Show that the circuit:



differs from a Toffoli gate only by relative phases. That is, the circuit takes $|c_1, c_2, t\rangle$ to $e^{i\theta(c_1,c_2,t)} \cdot |c_1, c_2, t \oplus c_1 \cdot c_2\rangle$, where $e^{i\theta(c_1,c_2,t)}$ is some relative phase factor. Such gates can sometimes be useful in experimental implementations, where it may be much easier to implement a gate that is the same as the Toffoli up to relative phases than it is to do the Toffoli directly.

Soln: Note the correction changing the rotations to $\pi/4$ instead of π. Using rotations of π, the circuit is equivalent to a single `CNOT`(x_1, x_3). Now, none of the gates alter the states of $|x_1\rangle$ or $|x_2\rangle$, so we can analyze this circuit by focusing only on it's effect on $|x_3\rangle$ while conditioning on the static state of $|x_1x_2\rangle$. First, we'll expand $R_y(\pm\pi/4)$. Note that $\cos(\pm\pi/8) = \frac{\sqrt{2+\sqrt{2}}}{2}$, and $\sin(\pm\pi/8) = \pm\frac{\sqrt{2-\sqrt{2}}}{2}$. Now:

$$R_y(\pm\pi/4) = \cos(\pm\pi/8)I - i\sin(\pm\pi/8)Y = \begin{bmatrix}\cos(\pm\pi/8) & -\sin(\pm\pi/8)\\ \sin(\pm\pi/8) & \cos(\pm\pi/8)\end{bmatrix} = \begin{bmatrix}\frac{\sqrt{2+\sqrt{2}}}{2} & \mp\frac{\sqrt{2-\sqrt{2}}}{2}\\ \pm\frac{\sqrt{2-\sqrt{2}}}{2} & \frac{\sqrt{2+\sqrt{2}}}{2}\end{bmatrix}.$$

When $|x_1x_2\rangle = |00\rangle$, the operations applied to $|x_3\rangle$ are $R_y(\pi/4)^2R_y(-\pi/4)^2$. Note that $R_y(\theta)R_y(-\theta) = I$, so in this case the circuit does not change the state of $|x_3\rangle$. When $|x_1x_2\rangle = |01\rangle$, the operations applied to $|x_3\rangle$ are $R_y(\pi/8)XR_y(\pi/8)R_y(-\pi/8)XR_y(-\pi/8) = R_y(\pi/8)X^2R_y*(-\pi/8) = R_y(\pi/8)R_y*(-\pi/8) = I$, so again, $|x_3\rangle$ is unchanged. When $|x_1x_2\rangle = |10\rangle$, the operations are:

$$\begin{aligned}R_y(\pi/4)^2XR_y(-\pi/4)^2 &= R_y(\pi/2)XR_y(-\pi/2)\\ &= \begin{bmatrix}\frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2}\end{bmatrix}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2}\\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2}\end{bmatrix}\\ &= \begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix}\end{aligned}$$

Finally, when $|x_1x_2\rangle = |11\rangle$, the operations are:

$$R_y(\pi/4)XR_y(\pi/4)XR_y(-\pi/4)XR_y(-\pi/4) =$$

$$\begin{bmatrix}\frac{\sqrt{2+\sqrt{2}}}{2} & -\frac{\sqrt{2-\sqrt{2}}}{2}\\ \frac{\sqrt{2-\sqrt{2}}}{2} & \frac{\sqrt{2+\sqrt{2}}}{2}\end{bmatrix}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}\frac{\sqrt{2+\sqrt{2}}}{2} & -\frac{\sqrt{2-\sqrt{2}}}{2}\\ \frac{\sqrt{2-\sqrt{2}}}{2} & \frac{\sqrt{2+\sqrt{2}}}{2}\end{bmatrix}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}\frac{\sqrt{2+\sqrt{2}}}{2} & \frac{\sqrt{2-\sqrt{2}}}{2}\\ -\frac{\sqrt{2-\sqrt{2}}}{2} & \frac{\sqrt{2+\sqrt{2}}}{2}\end{bmatrix}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}\frac{\sqrt{2+\sqrt{2}}}{2} & \frac{\sqrt{2-\sqrt{2}}}{2}\\ -\frac{\sqrt{2-\sqrt{2}}}{2} & \frac{\sqrt{2+\sqrt{2}}}{2}\end{bmatrix}$$

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

So, the circuit performs the transformation

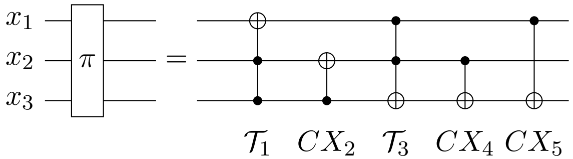
$$\begin{bmatrix}1 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1\\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0\end{bmatrix}$$

This transformation matches a Toffoli gate, except that it introduces a relative phase of π to the result of the computational basis state $|100\rangle$.

4.27) Using just CNOTs and Toffoli gates, construct a quantum circuit to perform the tansformation

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

This kind of partial cyclic permutation operation will be useful later, in Chapter 7.
Soln: CNOTs and Toffoli gates permute the computational basis states as does the target transformation. But for lack of intuition, a python script was written to exhaustively search products of the permutations they represent, looking for the target transformation as the result. A product of 5 gates was found involving 2 Toffolis and 3 CNOTs:



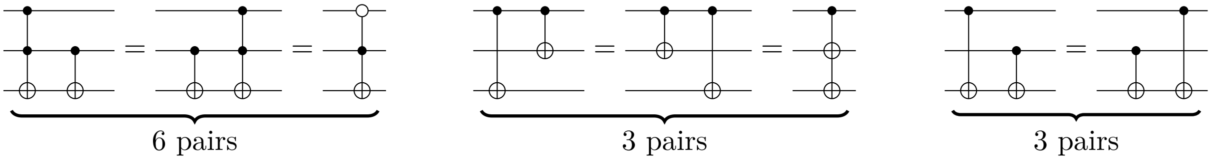
$\|x_1x_2x_3\rangle$	$\mathcal{T}_1$	CX_2	$\mathcal{T}_3$	CX_4	CX_5
$\|000\rangle$	$\|000\rangle$	$\|000\rangle$	$\|000\rangle$	$\|000\rangle$	$\|000\rangle$
$\|001\rangle$	$\|001\rangle$	$\|011\rangle$	$\|011\rangle$	$\|010\rangle$	$\|010\rangle$
$\|010\rangle$	$\|010\rangle$	$\|010\rangle$	$\|010\rangle$	$\|011\rangle$	$\|011\rangle$
$\|011\rangle$	$\|111\rangle$	$\|101\rangle$	$\|101\rangle$	$\|101\rangle$	$\|100\rangle$
$\|100\rangle$	$\|100\rangle$	$\|100\rangle$	$\|100\rangle$	$\|100\rangle$	$\|101\rangle$
$\|101\rangle$	$\|101\rangle$	$\|111\rangle$	$\|110\rangle$	$\|111\rangle$	$\|110\rangle$
$\|110\rangle$	$\|110\rangle$	$\|110\rangle$	$\|111\rangle$	$\|110\rangle$	$\|111\rangle$
$\|111\rangle$	$\|011\rangle$	$\|001\rangle$	$\|001\rangle$	$\|001\rangle$	$\|001\rangle$

The last 3 gates effectively execute a NOT on $|x_3\rangle$ if $|x_1\rangle = |1\rangle$ *and/or* $|x_2\rangle = |1\rangle$. They can be performed in any order, leading to 6 equivalent circuits realizing π. These 6 are the only "minimal realizations", that is, the only circuits realizing π using the least Toffolis and/or CNOTs possible.

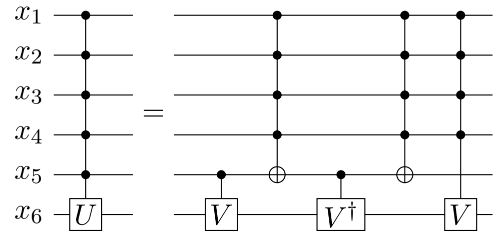
In an effort to determine which permutations could be implemented via Toffolis and CNOTs, the script was used to tabulate the number of permutations minimally representable as products of these gates of various lengths. Because the usable gates are all controlled and no controlled gate alters the state $|000\rangle$, all products of such gates fix $|000\rangle$. There are $7! = 5040$ permutations of the remaining computation basis states, so there are potentially 5040 permutations of the computational basis that could be represented. Suprisingly, at least to the second author, all of them are representable. The following table lists counts of the minimum number of required gates:

gates	count
0	1
1	9
2	60
3	261
4	845
5	1784
6	1688
7	386
8	6

The unique permutation representable as a product of 0 `CNOT`s and/or Toffolis is, of course, the identity. The 9 representable as products of single gates are the gates themselves. There are $9^2 = 81$ possible products of two gates, 9 of which consist of a repeated gate and so produce the identity. There are 3 types of symmetries that cause 12 of the 81 pairs to produce the same non-identity result, so the 81 pairs ultimately only produce 60 dinstinct permutations as products of 2 `CNOT`s and/or Toffolis. The symmetries are depicted below.



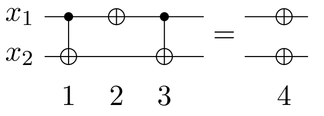
4.28) For $U = V^2$ with V unitary, construct a $C^5(U)$ gate analogous to that in Figure 4.10, but using no work qubits. You may use controlled-V and controlled-$V^\dagger$ gates.
Soln: NOTE: this task is only feasible if "controlled-V and controlled-$V^\dagger$ gates" is interpreted to mean $C(V)$, $C(V^\dagger)$, and $C^4(V)$, along with $C^4(X)$ gates. The $C^4(V)$ gate will implicitly require the existence of a unitary gate W such that $W^2 = V$ and the use of $C(W)$, $C(W^\dagger)$, $C^3(W)$ and $C^3(X)$ gates, which in turn require the existence of a unitary gate P such that $P^2 = W$ and the use of $C(P)$, $C(P^\dagger)$, $C^2(P)$ and $C^2(X)$ (Toffoli) gates. Finally, the $C^2(P)$ gate requires the existence of unitary Q such that $Q^2 = P$ and the use of controlled-Q, controlled-$Q^\dagger$ along with $C(X)$ gates. We'll construct $C^n(X)$ gates from Toffolis and 2-qubit gate in Exercise 4.29. In Exercise 4.24 we constructed Toffoli gates from 2-qubit gates, but still, in order to execute a $C^n(U)$ using only 1- and 2-qubit gates, we'd need a 1-qubit gate performing a unitary operator whose 2^n-th power is U. This gate may be guaranteed to exist theoretically, possibly under suitable hypotheses, but is almost certainly extremely difficult to implement physically. So, a construction of such a circuit is of questionable value. Instead, we offer a circuit using $C(V), C(V^\dagger)$ and $C^4(V)$. The circuits using $C^3(W)$, $C^2(P)$, and $C(Q)$ can be constructed recursively in turn. For a relatively rigorous mathematical proof of the infeasibility of the exercise without the use of $C^4(V)$-gates, see Wilfred Lee's answer to this cs.stackexchange question https://cs.stackexchange.com/questions/80538/is-it-possible-to-construct-a-c5u-with-v2-u-and-no-work-qubits-nielsen-ch



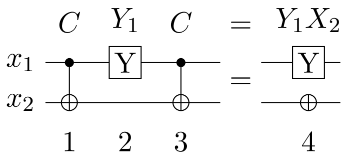
The verification of this circuit is identical to that of the circuit in figure 8 in Exercise 4.21, where instead of checking each computational basis state of $|x_1\rangle |x_2\rangle$, we check the states of $|x_1x_2x_3x_4\rangle |x_5\rangle$, where the possible states of $|x_5\rangle$ are the usual $|0\rangle$ and $|1\rangle$, but the states of $|x_1x_2x_3x_4\rangle$ are $|1111\rangle$ and $\neg |1111\rangle$, *i.e.* anything other than $|1111\rangle$.

4.29) Find a circuit containing $O(n^2)$ Toffoli, `CNOT` and single qubit gates which implements a $C^n(X)$ gate (for $n > 3$), using no work qubits.
Soln: This exercise is also impossible as specified and interpreted naturally.
4.30)
4.31) (More circuit identities) Let subscripts denote which qubit an operator acts on, and let C be a `CNOT` with qubit 1 the control and qubit 2 the target qubit. Prove the following identities:

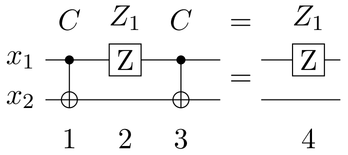
$$C \; X_1 \; C \; = X_1X_2$$



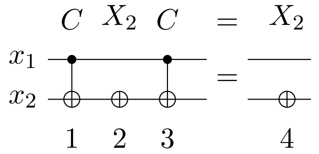
$\|x_1x_2\rangle$	1	2	3	4
$\|00\rangle$	$\|00\rangle$	$\|10\rangle$	$\|11\rangle$	$\|11\rangle$
$\|01\rangle$	$\|01\rangle$	$\|11\rangle$	$\|10\rangle$	$\|10\rangle$
$\|10\rangle$	$\|11\rangle$	$\|01\rangle$	$\|01\rangle$	$\|01\rangle$
$\|11\rangle$	$\|10\rangle$	$\|00\rangle$	$\|00\rangle$	$\|00\rangle$



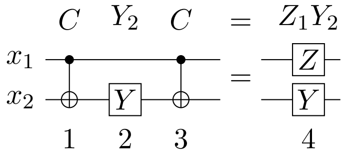
$\|x_1x_2\rangle$	1	2	3	4
$\|00\rangle$	$\|00\rangle$	$i\|10\rangle$	$i\|11\rangle$	$i\|11\rangle$
$\|01\rangle$	$\|01\rangle$	$i\|11\rangle$	$i\|10\rangle$	$i\|10\rangle$
$\|10\rangle$	$\|11\rangle$	$-i\|01\rangle$	$-i\|01\rangle$	$-i\|01\rangle$
$\|11\rangle$	$\|10\rangle$	$-i\|00\rangle$	$-i\|00\rangle$	$-i\|00\rangle$



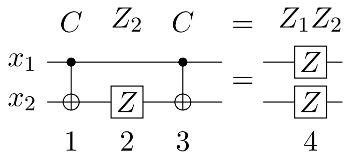
$\|x_1x_2\rangle$	1	2	3	4
$\|00\rangle$	$\|00\rangle$	$\|00\rangle$	$\|00\rangle$	$\|00\rangle$
$\|01\rangle$	$\|01\rangle$	$\|01\rangle$	$\|01\rangle$	$\|01\rangle$
$\|10\rangle$	$\|11\rangle$	$-\|11\rangle$	$-\|10\rangle$	$-\|10\rangle$
$\|11\rangle$	$\|10\rangle$	$-\|10\rangle$	$-\|11\rangle$	$-\|11\rangle$



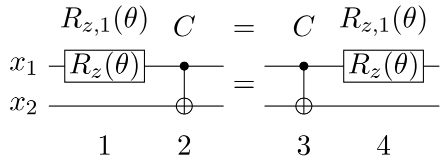
$\|x_1x_2\rangle$	1	2	3	4
$\|00\rangle$	$\|00\rangle$	$\|01\rangle$	$\|01\rangle$	$\|01\rangle$
$\|01\rangle$	$\|01\rangle$	$\|00\rangle$	$\|00\rangle$	$\|00\rangle$
$\|10\rangle$	$\|11\rangle$	$\|10\rangle$	$\|11\rangle$	$\|11\rangle$
$\|11\rangle$	$\|10\rangle$	$\|11\rangle$	$\|10\rangle$	$\|10\rangle$



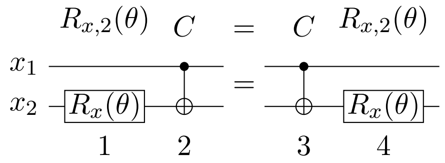
$\|x_1x_2\rangle$	1	2	3	4
$\|00\rangle$	$\|00\rangle$	$i\|01\rangle$	$i\|01\rangle$	$i\|01\rangle$
$\|01\rangle$	$\|01\rangle$	$-i\|00\rangle$	$-i\|00\rangle$	$-i\|00\rangle$
$\|10\rangle$	$\|11\rangle$	$-i\|10\rangle$	$-i\|11\rangle$	$-i\|11\rangle$
$\|11\rangle$	$\|10\rangle$	$i\|11\rangle$	$i\|10\rangle$	$i\|10\rangle$



$\|x_1x_2\rangle$	1	2	3	4
$\|00\rangle$	$\|00\rangle$	$\|00\rangle$	$\|00\rangle$	$\|00\rangle$
$\|01\rangle$	$\|01\rangle$	$-\|01\rangle$	$-\|01\rangle$	$-\|01\rangle$
$\|10\rangle$	$\|11\rangle$	$-\|11\rangle$	$-\|10\rangle$	$-\|10\rangle$
$\|11\rangle$	$\|10\rangle$	$\|10\rangle$	$\|11\rangle$	$\|11\rangle$



$\|x_1x_2\rangle$	1	2	3	4
$\|00\rangle$	$e^{-i\theta/2}\|00\rangle$	$e^{-i\theta/2}\|00\rangle$	$\|00\rangle$	$e^{-i\theta/2}\|00\rangle$
$\|01\rangle$	$e^{-i\theta/2}\|01\rangle$	$e^{-i\theta/2}\|01\rangle$	$\|01\rangle$	$e^{-i\theta/2}\|01\rangle$
$\|10\rangle$	$e^{i\theta/2}\|10\rangle$	$e^{i\theta/2}\|11\rangle$	$\|11\rangle$	$e^{i\theta/2}\|11\rangle$
$\|11\rangle$	$e^{i\theta/2}\|11\rangle$	$e^{i\theta/2}\|10\rangle$	$\|10\rangle$	$e^{i\theta/2}\|10\rangle$



$\|x_1x_2\rangle$	1	2	3	4
$\|00\rangle$	$\cos(\theta/2)\|00\rangle - i\sin(\theta/2)\|01\rangle$	$\cos(\theta/2)\|00\rangle - i\sin(\theta/2)\|01\rangle$	$\|00\rangle$	$\cos(\theta/2)\|00\rangle - i\sin(\theta/2)\|01\rangle$
$\|01\rangle$	$-i\sin(\theta/2)\|00\rangle + \cos(\theta/2)\|01\rangle$	$-i\sin(\theta/2)\|00\rangle + \cos(\theta/2)\|01\rangle$	$\|01\rangle$	$-i\sin(\theta/2)\|00\rangle + \cos(\theta/2)\|01\rangle$
$\|10\rangle$	$\cos(\theta/2)\|10\rangle - i\sin(\theta/2)\|11\rangle$	$-i\sin(\theta/2)\|10\rangle + \cos(\theta/2)\|11\rangle$	$\|11\rangle$	$-i\sin(\theta/2)\|10\rangle + \cos(\theta/2)\|11\rangle$
$\|11\rangle$	$-i\sin(\theta/2)\|10\rangle + \cos(\theta/2)\|11\rangle$	$\cos(\theta/2)\|10\rangle - i\sin(\theta/2)\|11\rangle$	$\|10\rangle$	$\cos(\theta/2)\|10\rangle - i\sin(\theta/2)\|11\rangle$

Chapter 5

The quantum Fourier transform and its applications

5.1) Give a direct proof that the linear transformation defined by Equation (5.2) is unitary.

$$|j\rangle \longrightarrow \sum_{k=0}^{N-1} e^{2\pi ijk/N}|k\rangle$$

Soln: First note that $e^{2\pi i/N}$ is an N-th root of unity, which we'll denote ω. The quantum Fourier transform (QFT) transforms $|j\rangle \to \sum_{k=0}^{N-1}\omega^{jk}|k\rangle$. In matrix form:

$$QFT = \frac{1}{\sqrt{N}}\begin{bmatrix} 1 & 1 & 1 & 1 & \dots & 1 & 1 \\ 1 & \omega^{1\cdot 1} & \omega^{2\cdot 1} & \omega^{3\cdot 1} & \dots & \omega^{(N-2)\cdot 1} & \omega^{(N-1)\cdot 1} \\ 1 & \omega^{1\cdot 2} & \omega^{2\cdot 2} & \omega^{3\cdot 2} & \dots & \omega^{(N-2)\cdot 2} & \omega^{(N-1)\cdot 2} \\ 1 & \omega^{1\cdot 3} & \omega^{2\cdot 3} & \omega^{3\cdot 3} & \dots & \omega^{(N-2)\cdot 3} & \omega^{(N-1)\cdot 3} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & \omega^{1\cdot(N-2)} & \omega^{2\cdot(N-2)} & \omega^{3\cdot(N-2)} & \dots & \omega^{(N-2)\cdot(N-2)} & \omega^{(N-1)\cdot(N-2)} \\ 1 & \omega^{1\cdot(N-1)} & \omega^{2\cdot(N-1)} & \omega^{3\cdot(N-1)} & \dots & \omega^{(N-2)\cdot(N-1)} & \omega^{(N-1)\cdot(N-1)} \end{bmatrix}$$

Noting that $\omega^* = \omega^{-1}$, we have

$$QFT^\dagger = \frac{1}{\sqrt{N}}\begin{bmatrix} 1 & 1 & 1 & 1 & \dots & 1 & 1 \\ 1 & \omega^{-1\cdot 1} & \omega^{-1\cdot 2} & \omega^{-1\cdot 3} & \dots & \omega^{-1\cdot(N-2)} & \omega^{-1\cdot(N-1)} \\ 1 & \omega^{-2\cdot 1} & \omega^{-2\cdot 2} & \omega^{-2\cdot 3} & \dots & \omega^{-2\cdot(N-2)} & \omega^{-2\cdot(N-1)} \\ 1 & \omega^{-3\cdot 1} & \omega^{-3\cdot 2} & \omega^{-3\cdot 3} & \dots & \omega^{-3\cdot(N-2)} & \omega^{-3\cdot(N-1)} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & \omega^{-(N-2)\cdot 1} & \omega^{-(N-2)\cdot 2} & \omega^{-(N-2)\cdot 3} & \dots & \omega^{-(N-2)\cdot(N-2)} & \omega^{-(N-2)\cdot(N-1)} \\ 1 & \omega^{-(N-1)\cdot 1} & \omega^{-(N-1)\cdot 2} & \omega^{-(N-1)\cdot 3} & \dots & \omega^{-(N-1)\cdot(N-2)} & \omega^{-(N-1)\cdot(N-1)} \end{bmatrix}$$

Now $QFT * QFT^\dagger = \frac{1}{N}\left[\sum_{\ell=0}^{N-1}\omega^{\ell j}\omega^{-k\ell}\right]_{j,k} = \frac{1}{N}\left[\sum_{\ell=0}^{N-1}\omega^{\ell(j-k)}\right]_{j,k}$. When $j = k$, *i.e.* on the diagonal, all exponents in the summation formula for the j,k entry become 0, producing a sum of N 1s, canceling the $\frac{1}{N}$ scalar and giving that the diagonal entries of $QFT * QFT^\dagger$ are all 1s. For off-diagonal entries, *i.e.* when $j \neq k$, let $a = j - k$. Recognizing a finite geometric series and that ω is an N-th root of unity, we have:

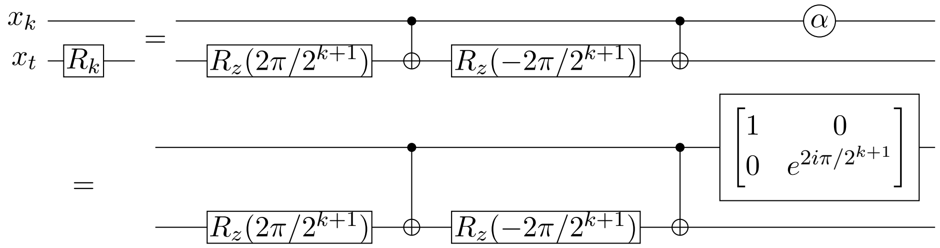
$$\sum_{\ell=0}^{N-1}\omega^{\ell(j-k)} = \sum_{\ell=0}^{N-1}\omega^{a\cdot\ell} = 1 + \omega^a + \omega^{2a} + \ldots + \omega^{(N-1)a} = \frac{1-\omega^{N\cdot a}}{1-\omega^a} = \frac{1-1^a}{1-\omega^a} = 0$$

Note that the denominator is non-zero, since $a \neq 0$. Ultimately, $QFT * QFT^\dagger = I$, and QFT is unitary.

5.2) Explicitly compute the Fourier transform of the n qubit state $|00\dots0\rangle$.
Soln: $QFT(|00\dots0\rangle) = \frac{1}{2^{n-1}}\sum_{k=0}^{\sqrt{2^n}} |k\rangle$.
5.3)
5.4) Give a decomposition of the controlled-R_k gate into single qubit and `CNOT` gates.
Soln: By Theorem 4.1 we may write $R_k = e^{i\alpha}R_z(\beta)R_y(\gamma)R_z(\delta)$ for some α, β, γ and δ. In this case, $\alpha = 2\pi/2^{k+1}$, $\beta = 0$, $\gamma = 0$, and $\delta = 2\pi/2^k$ suffice, since

$$e^{2i\pi/2^{k+1}}R_z(0)R_y(0)R_z(2\pi/2^k) = e^{2i\pi/2^{k+1}}I^2\begin{bmatrix} e^{-2i\pi/2^{k+1}} & 0 \\ 0 & e^{2i\pi/2^{k+1}}\end{bmatrix} \qquad = \begin{bmatrix} 1 & 0 \\ 0 & e^{2i\pi/2^k}\end{bmatrix} = R_k$$

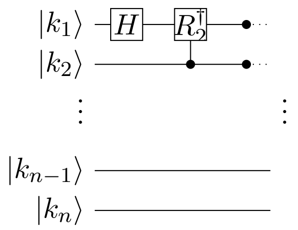
Following the proof of Corollary 4.2, set $A = R_z(0)R_y(0) = I, B = R_y(0)R_z(-2\pi/2^{k+1}) = R_z(-2\pi/2^{k+1}), C = R_z(2\pi/2^{k+1})$ so that $ABC = I$ and $U = e^{i\alpha}AXBXC$. Applying the construction in Figure 4.6 gives:



5.5) Give a quantum circuit to perform the inverse quantum Fourier transform.
Soln: The inverse quantum Fourier transform is the $QFT^\dagger$ transformation shown in Exercise 5.1. For a fixed computational basis state $|k\rangle$,

$$\begin{aligned} QFT^\dagger(|k\rangle) &= \frac{1}{\sqrt{N}}\sum_{j=0}^{N-1} e^{-2\pi ijk/2^n}|j\rangle \\ &= \frac{\Big(|0\rangle + e^{-2\pi i0.k_n}|1\rangle\Big)\Big(|0\rangle + e^{-2\pi i0.k_{n-1}k_n}|1\rangle\Big)\cdots\Big(|0\rangle + e^{-2\pi i0.k_1k_2\dots k_{n-1}k_n}|1\rangle\Big)}{\sqrt{N}}\end{aligned}$$

Using $R_\ell^\dagger = \begin{bmatrix} 1 & 0 \\ 0 & e^{-2\pi i/2^\ell}\end{bmatrix}$ instead of R_ℓ in Figure 5.1 will implement $QFT^\dagger$. It's verification is effectively identical to that of Figure 5.1 on page 219. For completeness, the circuit is diagrammed here:



5.6) (Approximate quantum Fourier transform) The quantum circuit construction of the quantum Fourier transform apparently requires gates of exponential precision in the number of qubits used. However, such precision is never required in any quantum circuit of polynomial size. For example, let U be the ideal quantum Fourier transform on n qubits, and V be the transform which results if the controlled R_k gates are performed to a precision of $\Delta = 1/p(n)$ for some polynomial $p(n)$. Show that the error $E(U, V) = \max_{|\psi\rangle}\|(U - V)|\psi\rangle\|$ scales as $\Theta(n^2/p(n))$, and thus polynomial precision in each gate is sufficient to guarantee polynomial accuracy in the output state.

Soln: Let R_j^i be the controlled R_j gate, controlled by the qubit with index j in Figure 5.1 (1-up), targeting the qubit with index i. Let $R_j'^i$ be the same gate performed with precision to a precision of Δ. Note that the superscripts indicate the target and are not exponents.

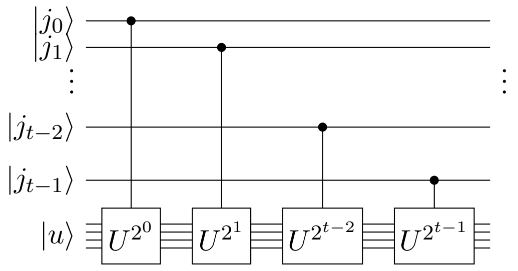
$$\begin{aligned} E(U,V) &\equiv \max_{|\psi\rangle} \|(U-V)\,|\psi\rangle\| && \text{(definition)} \\ &= \max_{|\psi\rangle} \Big\|(H^1 R_2^1 R_3^1 \cdots R_n^1 H^2 R_2^2 \cdots R_{n-1}^2 \cdots H^{n-1} R_2^{n-1} H^n \\ &\qquad - H^1 R_2'^1 R_3'^1 \cdots R_n'^1 H^2 R_2'^2 \cdots R_{n-1}'^2 \cdots H^{n-1} R_2'^{n-1} H^n)\,|\psi\rangle\Big\| \end{aligned}$$

The trick seems to be to apply a (supposedly true) lemma that $E(AB, A'B') \leq E(A, A') + E(B, B')$. This doesn't seem to be entirely obvious, but if allowed to apply it:

$$\begin{aligned} &\leq E(H^1, H^1) + E(R_2^1, R_2'^1) + \ldots + E(R_n^1, R_n'^1) + E(H^2, H^2) \\ &\quad + E(R_2^2, R_2'^2) + \ldots + E(R_{n-1}^2, R_{n-1}'^2) + \ldots \\ &\quad + E(H^{n-1}, H^{n-1}) + E(R_2^{n-1}, R_2'^{n-1}) + E(H^n, H^n) \\ &\leq 0 + \Delta + \ldots + \Delta + 0 + \Delta + \ldots + \Delta + \ldots + 0 + \Delta + 0 \\ &= (n-1)\Delta + (n-2)\Delta + \ldots + \Delta \\ &= \Delta \sum_{i=1}^{n-1} i \\ &= \left(\frac{n(n-1)}{2}\right)\delta \leq \frac{n*(n-1)}{2p(n)} = O(n^2/p(n)) \end{aligned}$$

This only gives an upper bound. A lower bound is less important and would seem to require a lower bound in the lemma, which I'm suspicious of. It would also require the assumption that the controlled R_k gates are performed with causal error Δ. In practice, this is unlikely to be the case. In application, the only causal errors are likely to be the result of simply not performing the controlled-R_k gates at all, for k above some fixed threshold. Doing so would introduce causal errors, but those errors would depend on k, not n. The threshold for k may depend on n, and the *total* error depends on the number of controlled-R_k gates not performed which is determined by n, but individually the error in not performing a single controlled-R_k would depend only on k.

5.7) Additional insight into the circuit in Figure 5.2 amy be obtained by showing, as you should now do, that the effect of the sequence of controlled-U operations like that in Figure 5.2 is to take the state $|j\rangle\,|u\rangle$ to $|j\rangle\,U^j\,|u\rangle$. (Note that this does not depend on $|u\rangle$ being an eigenstate of U.
Soln: Note that $|j\rangle$ is implicitly assumed to be a computational basis state, so let's write $j = j_{t-1}j_{t-2}\ldots j_1 j_0$ in binary. Preferring to write quantum registers in big-endian: $|j\rangle = |j_0 j_1 \ldots j_{t-2} j_{t-1}\rangle$. The sequence of controlled-U gates is thus:



with each $|j_s\rangle$ being either $|0\rangle$ or $|1\rangle$. The controlled U^{2^s} gate is applied to $|u\rangle$ if and only if $|j_s\rangle = |1\rangle$

which can be represented by multiplication of $|u\rangle$ by $U^{j_s 2^s}$.

$$\begin{aligned}
|j\rangle|u\rangle &= |j_0 j_1 \ldots j_{t-2} j_{t-1}\rangle |u\rangle \\
&\to |j_0 j_1 \ldots j_{t-2} j_{t-1}\rangle U^{j_{t-1}2^{t-1}} U^{j_{t-2}2^{t-2}} \cdots U^{j_1 2^1} U^{j_0 2^0} |u\rangle \\
&= |j_0 j_1 \ldots j_{t-2} j_{t-1}\rangle U^{j_0 j_1 \ldots j_{t-2} j_{t-1}} |u\rangle && \text{(collect powers of } U \text{ in binary)} \\
&= |j\rangle U^j |u\rangle && \text{(recognize binary representation of } j\text{)}
\end{aligned}$$

5.8) Suppose the phase estimation algorithm takes the state $|0\rangle|u\rangle$ to the state $|\widetilde{\varphi_u}\rangle|u\rangle$, so that given then input $|0\rangle\left(\sum_u c_u |u\rangle\right)$, the algorithm outputs $\sum_u c_u |\widetilde{\varphi_u}\rangle|u\rangle$. Show that is t is chosen according to (5.35), then the probability for measuring φ_u accurate to n bits at the conclusion of the phase estimation algorithm is at least $|c_u|^2(1-\epsilon)$.

Soln: Note that in this problem we assume $\{|u\rangle\}$ forms an eigenbasis, but that the bound on probability is requested for measuring φ_u for some fixed eigenvalue, say v. Measurement of the second register collapses the state to $|\widetilde{\varphi}_v\rangle|v\rangle$ with probabily $|c_v|^2$, from which we can determine φ_v accurate to n bits with probability $1-\epsilon$. Multiplying probabilities gives that performing phase estimation on $|0\rangle\left(\sum_u c_u |u\rangle\right)$ gives a probability of measuring φ_v accurate to n bits of $|c_v|^2(1-\epsilon)$.

5.9)

5.10) Show that the (multiplicative) order of $(x=)5 \bmod (N=)21$ is 6.

Soln: It suffices to show that $5^t \not\equiv 1 \pmod{21}$ for $1 \le t < 5$, and that $5^6 \equiv 1 \pmod{21}$

t	$5^t \pmod{21}$
0	1
1	5
2	4
3	20 ($\equiv -1$)
4	16 ($\equiv -5$)
5	17 ($\equiv -4$)
6	1 ($\equiv -20$)

5.11) Show that the (multiplicative) order of $x \pmod N$ satisfies $r \le N$.

Soln: Note that $x \neq 0$, since $0^r \equiv 0 \pmod N$ for all r, and the order of 1 is 1 for all N. For $N \ge 2$, see Euler's totient theorem which states that when $\gcd(x,N)=1$, $x^{\phi(N)} \equiv 1 \pmod N$, where $\phi(N)$ is the totient of N, *i.e.* the number of integers $1 \le n \le N$ such that $\gcd(n,N)=1$. For $N \ge 2$, $\gcd(N,N) \neq 1$, so $\phi(N) \le N-1$. Equality holds (for ϕ and the order of x) if and only if N is prime. In the end, for $N \ge 2$, the order of any integer x comprime to the modulus N is at most $\phi(N) \le N-1$. Note, with a little additional consideration it can be shown that the order of x is, in fact, a divisor $\phi(N)$ by using Lagrange's Theorem on subgroup orders applied to the cyclic group generated by x in the multiplicative group mod N.

5.12) Show that U is unitary (*Hint:* x is co-prime to N, and therefore has an inverse modulo N)

Soln: U performs modular multiplication by x. When x is coprime to N, modular multiplication by x executes a permutation of the computational basis states, that is, U is a permutation matrix. Inverses of permutation matrices are easily seen to be their transposes. So, $U^\dagger U = U^T U = U^{-1} U = I$, and U is unitary.

5.13) Prove (5.44):

$$\frac{1}{\sqrt{r}} \sum_{s=0}^{r-1} |u_s\rangle = |1\rangle .$$

(*Hint:* $\sum_{s=0}^{r-1} \exp(-2\pi i s k/r) = r\delta_{k0}$.) In fact, prove that

$$\frac{1}{\sqrt{r}} \sum_{s=0}^{r-1} e^{2\pi i s k/r} |u_s\rangle = |x^k \pmod N\rangle .$$

Soln: Note that in this exercise k is fixed, whereas in the definition of $|u_s\rangle$ in equation (5.37), k is an index that is iterated over. Below, we'll iterate over k' instead of k. In the hint, k is arbitrary. We'll apply it to $(k-k')$ at a point at which k' will also be fixed, yielding $\sum_{s=0}^{r-1} \exp(-2\pi i s(k-k')/r) = r\delta_{(k-k')0} = r\delta_{kk'}$.

$$\begin{aligned}
\frac{1}{\sqrt{r}} \sum_{s=0}^{r-1} e^{2\pi i s k/r} |u_s\rangle &= \frac{1}{\sqrt{r}} \sum_{s=0}^{r-1} e^{2\pi i s k/r} \left(\frac{1}{\sqrt{r}} \sum_{k'=0}^{r-1} e^{-2\pi i s k'/r} |x^{k'} \pmod N\rangle \right) \\
&= \frac{1}{r} \sum_{s=0}^{r-1} \sum_{k'=0}^{r-1} e^{2\pi i s(k-k')/r} |x^{k'} \pmod N\rangle && \text{(distribute)} \\
&= \frac{1}{r} \sum_{k'=0}^{r-1} \left(\sum_{s=0}^{r-1} e^{2\pi i s(k-k')/r} \right) |x^{k'} \pmod N\rangle && \text{(change order)} \\
&= \frac{1}{r} \sum_{k'=0}^{r-1} r\delta_{kk'} |x^{k'} \pmod N\rangle && \text{(apply hint)} \\
&= |x^k \pmod N\rangle && \text{(collect non-zero terms)}
\end{aligned}$$

5.14) The quantum state produced in the order-finding algorithm, before the inverse Fourier transform, is

$$|\psi\rangle = \sum_{j=0}^{2^t-1} |j\rangle U^j |1\rangle = \sum_{j=0}^{2^t-1} |j\rangle |x^j \pmod N\rangle ,$$

if we initialize the second register as $|1\rangle$. Show that the same state is obtained if we replace U^j with a *different* unitary transform V, which computes

$$V|j\rangle|k\rangle = |j\rangle|k + x^j \pmod N\rangle ,$$

and start the second register in the state $|0\rangle$. Also show how to construct V in $O(L^3)$ gates.
Soln: Initializing the second register in the state $|k\rangle = |0\rangle$ means that applying V produces

$$V\left(\sum_{j=0}^{2^t-1} |j\rangle|0\rangle\right) = \sum_{j=0}^{2^t-1} V|j\rangle|0\rangle = \sum_{j=0}^{2^t-1} |j\rangle|0 + x^j \pmod N\rangle = \sum_{j=0}^{2^t-1} |j\rangle|x^j \pmod N\rangle .$$

This might be missing the point, but if explicitly required, V can be implemented by first applying U to $|j\rangle|1\rangle|k\rangle$ to produce the state $|j\rangle|x^j \pmod N\rangle|0\rangle$, where the $|k\rangle$ is the state of a third register whose length is the same as the second. Then, adding the second register to the third $\pmod N$ produces $|j\rangle|x^j \pmod N\rangle|x^j \pmod N\rangle$. Modular addition can be performed using $O(L)$ gates, so the complexity of V can be seen to be $O(L^3 + L) = O(L^3)$.

5.15) Show that the least common multiple of positive integers x and y is $xy/\gcd(x,y)$, and thus may be computed in $O(L^2)$ operations if x and y are L bit numbers.
Soln: This is trivial ... still, note that $\gcd(x,y)|x$ and $\gcd(x,y)|y$, so $x = k\gcd(x,y)$ and $y = \ell\gcd(x,y)$ for some k and ℓ. Now $xy/\gcd(x,y) = x\ell = ky$, so $xy/\gcd(x,y)$ is a common multiple of x and y. To show it is the least common multiple, it is most instructive to rely on the unique prime factorization formula, but instead, note that by the Euclidean algorithm (or Bezout's Lemma), there exists integers a and b such that $\gcd(x,y) = ax + by$ (one of a and b is negative). Let m be any common multiple, meaning

$m = \kappa x = \lambda y$. Now $m \gcd(x, y) = max + mby = \lambda yax + \kappa xby = xy(\lambda a + \kappa b)$. Now xy is divisible by $\gcd(x, y)$ by definition, so we can write $m = \frac{xy}{\gcd(x,y)}(\lambda a + \kappa b)$, and $\frac{xy}{\gcd(x,y)}$ divides m. So, m cannot be the least common multiple, unless m is itself $\frac{xy}{\gcd(x,y)}$.

The complexity of classical (schoolbook) multiplication and division are both $O(L^2)$. The complexity of calculating the gcd is at most $O(L^2)$ so that in total the least common multiple can be calculated in $O(L^2)$ operations.

5.16) For all $x \geq 2$ prove that $\int_x^{x+1} 1/y^2 dy \geq 2/3x^2$. Show that

$$\sum_q \frac{1}{q^2} \leq \frac{3}{2}\int_2^\infty \frac{1}{y^2}dy = 3/4,$$

and thus that (5.58) holds.

Soln: $\int_x^{x+1} 1/y^2 dy = \frac{-1}{y}\Big|_x^{x+1} = \frac{1}{x} - \frac{1}{x+1} = \frac{1}{x(x+1)}$. for $x \geq 2$, note that both x and $(x-2)$ are positive, so that

$$\begin{aligned}
0 &\leq x(x-2) && \text{(by assumption)}\\
&= x^2 - 2x\\
2x^2 + 2x &\leq 3x^2 && \text{(add } 2x^2 + 2x \text{ to both sides)}\\
2/3x^2 &\leq 1/(x^2 + x) && \text{(multiply by } 1/(3x^2)(x^2 + x)\text{, note: its positive)}\\
&= \int_x^{x+1} 1/y^2 dy
\end{aligned}$$

So $\int_x^{x+1} 1/y^2 dy \geq 2/3x^2$, or more applicably, $\frac{1}{x^2} \leq \frac{3}{2}\int_x^{x+1}\frac{1}{y^2}dy$. For the second part:

$$\sum_{q \text{ prime}} \frac{1}{q^2} < \sum_{x=2}^{\infty}\frac{1}{x^2} \leq \sum_{x=2}^{\infty}\frac{3}{2}\int_x^{x+1}\frac{1}{y^2}dy = \int_2^\infty \frac{1}{y^2}dy = \frac{-3}{2y}\Big|_2^\infty = \frac{3}{4}.$$

Using this bound for the sum of prime square reciprocals in equation 5.57 gives the lower bound on probability in equation 5.58. This bound is quite loose though. The prime square reciprocal sum has been studies extensively. Summing over all integers, not just primes, we can apply the solution (due to Euler) to the Basel problem:

$$\sum_{q \text{ prime}} \frac{1}{q^2} \leq -1 + \sum_{x=1}^{\infty}\frac{1}{x^2} = -1 + \frac{\pi^2}{6} \approx 0.6450$$

. Euler also evaluated the sum over primes specifically to high precision. See page 480 of `http://www.17centurymaths.com/contents/euler/introductiontoanalysisvolone/ch15vol1.pdf`

$$\sum_{q \text{ prime}} \frac{1}{q^2} \approx 0.45225.$$

5.17) Suppose N is L bits long. The aim of this exercise is to find an efficient classical algorithm to determine whether $N = a^b$ for some integers $a \geq 2$ and $b \geq 2$. This may be done as follows:

(1) Show that b, if it exists, satisfies $b \leq L$.

(2) Show that it takes at most $O(L^2)$ operations to compute $\log_2 N, x = y/b$ for $b \leq L$, and the two integers u_1 and u_2 nearest to 2^x.

(3) Show that it takes at most $O(L^2)$ operations to compute u_1^b and u_2^b (use repeated squaring) and check to see if either is equal to N.

(4) Combine the previous results to give an $O(L^3)$ operation algorithm to dtermine whether $N = a^b$ for integers a and b.

Soln:

(1) In fact, $b \leq L - 1$. By assumption $a^b = N$, so $b = \log_a N = \log_2 N / \log_2 a \leq \log_2 N \leq \lceil \log_2 N \rceil \equiv L$. Being integers, $b \leq L - 1$ unless equality holds throughout. Equality can hold only if $a = 2$ and $\log_2 N$ is an integer, *i.e* $N = 2^b$, but in this case $L = b + 1$, so once again $b \leq L - 1$. In (2) and (3) below, we'll assume b is fixed. The result will be that algorithm generated actually has complexity $O(L^4)$. If for fixed a the value of b can be determined without iteration, such as during the sequence of square and (maybe) multiply operations constructing u_1^b and u_2^b described in (3), then this might reduce the complexity to $O(L^3)$.

(2) Note, we don't need $\log_2 N$ as a decimal number, we actually need $\lceil \log_2 N \rceil$. To calculate this, start with 1 and iteratively double by adding the result to itself until it is at least N. Doing so requires L iterations, each of which consists of a single standard schoolbook addition and an integer comparison, both of which can be done in $O(L)$ steps, so that in total calculating $\lceil \log_2 N \rceil$ requires $L * O(L) = O(L^2)$ steps. [Note: one additional comparison might be required to initialize (or terminate) the algorithm, but this will not change the asymptotics.] I don't know what y is, but it must be a fixed L-bit value ($y \leq N - 1$). Schoolbook division of one L-bit integer by another is well-known to have complexity $O(L^2)$. The result, x, will be another L-bit integer at most $N - 1$. We construct 2^x by summing various powers of 2. Calculating 2^{2^i} can be done iteratively with the same sequence of $O(L)$ doublings described above. Then 2^x can be assembled by adding the at most $O(L)$ powers of 2 corresponding to the 1's in the binary representation of x. Computing the binary representation of x (in reverse order) can be done with an iterative sequence of $O(L)$ subtractions of the largest powers of 2 smaller than the current value (which can be determined with at most $O(L^2)$ comparisons). The $O(L)$ subtractions, each requiring $O(L)$ operations, take at total of $O(L^2)$ operations. u_1 and u_2 can be computed by adding and subtracting 1, taking $O(L)$ steps each. So, all requested operations can be performed in $O(L^2)$ steps.

(3) Calculating u_1^b and u_2^b can be done in $O(L^2)$ steps similar to the calculation of 2^x above, but instead of adding powers of 2 specified by the binary representation of x, we square, or square and multiply according the binary representation of b. The binary representation of b can be calculated in $O(L^2)$ steps, as described for x above. Then, start with $z = u_i$ ($i = 1, 2$). For each bit except the most significant (in order from least significant to second-most significant), square z and if the bit of b is a 1, multiply by u_i. Using comparisons (with complexity $O(L)$) we can decide to stop when the result is at least N and test for equality if we've completed the exponentiation. Squaring doubles the bitlength of the result, so lets assume we perform ℓ squarings. Note that $\ell \leq \log_2 L$. Then u_i must have had at most $L/2^\ell$ bits. After the j-th square and (maybe) multiply operation, the result has at most $L/2^{\ell-j}$ bits. In the next step, the square and multiply each require $O((L/2^{\ell-j})^2)$ steps. Over all ℓ iterations, $\sum_{j=0}^{\ell-1} O((L/2^{\ell-j})^2) = O(L^2 \cdot \sum_{j=1}^{\ell} 1/2^{2j}) = O(\frac{1}{3}L^2) = O(L^2)$ operations are required to perform the square and (maybe) multiplies (where the additional 1/3 constant is the value of the infinite geometric series as opposed to the finite). Adding the comparisons with N contributes only $O(\ell L) = O(L \log_2 L) = O(L^2)$ (where here, equality indicates set-containment, not complexity class equality). So, in the end, for fixed b, calculating u_i^b and testing whether either is equal to N requires only $O(L^2)$ operations.

(4) The final step in the $O(L^4)$ algorithm to determine if $N = a^b$ for some integers a and b is to iterate over a and b. *** TODO: Fix ... I can't iterate over a. Doing so introduces a factor of N into the complexity. For now, I'm moving on.

5.19) (Factoring 91) Suppose we wish to factor N=91. Confirm steps 1 and 2 are passed. For step 3, suppose we choose $x = 4$, which is co-prime to 91. Compute the order r of x with respect to N, and show that $x^{r/2} \pmod{91} = 64 \neq -1 \pmod{91}$, so the algorithm succeeds, giving $\gcd(64 - 1, 91) = 7$.
Soln: in the 5-step description of the factoring by order-finding algorithm on pages 233,234, step 1 consists of returning the factor 2 if N is even. It is not, so step 1 is passed (as in the algorithm does not yet terminate). Step 2 is to determine if $N = a^b$ for some integers a and $b \geq 2$. For completeness, the table below lists all potential values of a^b that could equal 91, none of which actually do.

$a\backslash b$	2	3	4	5	6
2	4	8	16	32	64
3	9	27	81		
4	16	64			
5	25				
6	36				
7	49				
8	64				
9	81				

Now, note that $\gcd(4, 91) = 1$, so step 3 is passed. To compute the (multiplicative) order of $x = 4$ $(\text{mod } N = 91)$:

t	$4^t \pmod{91}$
0	1
1	4
2	16
3	64
4	74
5	23
6	1

So the order of $4 \pmod{91}$ is $r = 6$. Being even, we check $4^{r/2} \equiv 64 \not\equiv -1 \pmod{91}$. The nearest integers to $4^{r/2}$ are 63 and 65. We'll start with 63. Noting that $\gcd(63, 91) = 7$, the algorithm would return 7. Had we tested $\gcd(65, 91)$ first, we could have found the other factor, 13. Noting that $91 = 7 \times 13$, the algorithm has succeeded and returned one (or both) factors.

5.19) Show that $N = 15$ is the smallest number for which the order-finding subroutine is required, that is, it is the smallest composite number that is not even or a power of some smaller integer.
Soln:

N	not applicable because
2	even,prime
3	prime
4	even, $4 = 2^2$
5	prime
6	even
7	prime
8	even, $8 = 2^3$
9	$9 = 3^2$
10	even
11	prime
12	even
13	prime
14	even

5.20) Suppose $f(x+r) = f(x)$, and $0 \leq x < N$, for n an integer multiple of r. Compute

$$\hat{f}(\ell) \equiv \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} e^{-2\pi i \ell x/N} f(x),$$

and relate the result to (5.63). You will need to use the fact that

$$\sum_{k \in \{0, r, 2r, \ldots, N-r\}} e^{2\pi i k\ell/N} = \begin{cases} \sqrt{N/r} & \text{if } \ell \text{ is an integer multiple of } N/r \\ 0 & \text{otherwise.} \end{cases}$$

Soln:

$$\begin{aligned}
\hat{f}(\ell) &\equiv \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} e^{-2\pi i \ell x/N} f(x) && \text{(definition)} \\
&= \frac{1}{\sqrt{N}} \sum_{k \in \{0, r, 2r, \ldots, N-r\}} \sum_{j=0}^{r-1} e^{-2\pi i \ell (k+j)/N} f(k+j) && \text{(separate summation)} \\
&= \frac{1}{\sqrt{N}} \sum_{k \in \{0, r, 2r, \ldots, N-r\}} e^{-\pi i \ell k/N} \sum_{j=0}^{r-1} e^{-2\pi i \ell j/N} f(k+j) && \text{(group contributions of } k \text{ to scalar)} \\
&= \frac{1}{\sqrt{N}} \sum_{k \in \{0, r, 2r, \ldots, N-r\}} e^{-\pi i \ell k/N} \sum_{j=0}^{r-1} e^{-2\pi i \ell j/N} f(j) && (f \text{ has period } r) \\
&= \begin{cases} \frac{1}{\sqrt{r}} \sum_{j=0}^{r-1} e^{-2\pi i \ell j/N} f(j) & \text{if } \ell \text{ is an integer multiple of } N/r \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}$$

Note, in the last equality, we apply the fact given in the exercise. The fractional powers of $e^{2\pi i}$ as provided in the statement of the exercise are positive, but those in the solution are negative. Negating fractional powers of $e^{2\pi i}$ has the effect of complex conjugation, so the resulting value would also need to be conjugated, but the resulting values are real numbers, so complex conjugation does not change their value.

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) What is the trace distance between the probability distribution $(1,0)$ and the probability distribution $(1/2,1/2)$? Between $(1/2,1/3,1/6)$ and $(3/4,1/8,1/8)$?
Soln:

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

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

9.2) Show that the trace distance between probability distributions $(p,1-p)$ and $(q,1-q)$ is $|p-q|$.
Soln:

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

9.3) What is the fidelity of the probability distributions $(1,0)$ and $(1/2,1/2)$? Of $(1/2,1/3,1/6)$ and $(3/4,1/8,1/8)$?
Soln:

$$F((1,0),(1/2,1/2)) = \sqrt{1\cdot 1/2}+\sqrt{0\cdot 1/2} = \frac{1}{\sqrt{2}} \simeq 0.707$$

$$\begin{aligned} F\left((1/2,1/3,1/6),(3/4,1/8,1/8)\right) &= \sqrt{1/2\cdot 3/4}+\sqrt{1/3\cdot 1/8}+\sqrt{1/6\cdot 1/8} \\ &= \frac{4\sqrt{6}+\sqrt{3}}{12} \simeq 0.961 \end{aligned}$$

9.4) Prove (9.3):

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

Soln: Let the index set be U, and define $S^p = \{x | x \in U, p_x > q_x\}$, $S^q = \{x | x \in U, p_x < q_x\}$, and $S^0 = \{x | x \in U, p_x = q_x\}$. That is, S^p is the subset of the index set in which the outcomes are more probable under p_x than under q_x, and vice versa in S^q, where S^0 is the subset of the outcomes equally likely under both p_x and q_x. To prove that $D(p_x,q_x) = \max_S |p(S)-q(S)|$, we'll prove that both are equal to $p(S^p)-q(S^p)\,(=-p(S^q)+q(S^q))$.

To see that $\max_S |p(S)-q(S)| = p(S^p)-q(S^p)$, it is enough to argue that S^p maximizes $|p(S)-q(S)|$. S^p contains all elements of U that contribute positively to $p(S)-q(S)$. So, the inclusion of any other element of U would decrease the quantity $p(S)-q(S)$. So, $p(S)-q(S)$ is maximized by S^p. Similarly, $p(S)-q(S)$ is minimized by S^q. All subsets of U produce values of $p(S)-q(S)$ in between. To determine the maximum absolute value, we'll argue that $p(S^q)-q(S^q) = -(p(S^p)-q(S^p))$, that is, these two extremes produce values equal in magnitude, but opposite in sign. To see this, note that $p(U)-q(U) \equiv \sum_{x\in U} p(x) - \sum_{x\in U} q(x) = 1-1 = 0$, since p and q are probability distributions and U is the entire index set on which they are defined. Alternatively,

$$\begin{aligned} p(U)-q(U) &= p(S^p\cup S^0\cup S^q) - q(S^p\cup S^0\cup S^q) && \text{(set equality)} \\ &= p(S^p)+p(S^0)+p(S^q)-q(S^p)-q(S^0)-q(S^q) && \text{(separate implicit sums)} \\ &= p(S^p)-q(S^p)+p(S^0)-q(S^0)+p(S^q)-q(S^q) && \text{(rearrange)} \\ &= p(S^p)-q(S^p)+p(S^q)-q(S^q) && (p_x = q_x \text{ for } x\in S^0) \end{aligned}$$

So $0 = p(U)-q(U) = p(S^p)-q(S^p)+p(S^q)-q(S^q)$ and $p(S^q)-q(S^q) = -(p(S^p)-q(S^p))$. So, $p(S)-q(S)$ achieves extremes of equal magnitude but opposite sign on S^p and S^q, so $\max_S |p(S)-q(S)| = p(S^p)-q(S^p)\,(=-p(S^q)+q(S^q))$.

Now we must show that $D(p_x,q_x) = p(S^p)-q(S^p)$:

$$\begin{aligned} D(p_x,q_x) &\equiv \frac{1}{2}\sum_{x\in U}|p_x-q_x| \\ &= \frac{1}{2}\left(\sum_{x\in S^p}|p_x-q_x| + \sum_{x\in S^0}|p_x-q_x| + \sum_{x\in S^q}|p_x-q_x|\right) && \text{(separate explicit sum)} \\ &= \frac{1}{2}\left(\sum_{x\in S^p}|p_x-q_x| + \sum_{x\in S^0}0 + \sum_{x\in S^q}|p_x-q_x|\right) && (p_x-q_x = 0 \text{ for } x\in S^0) \\ &= \frac{1}{2}\left(\sum_{x\in S^p}p_x-q_x + \sum_{x\in S^q}-(p_x-q_x)\right) && (p_x-q_x \text{ has consistent sign within } S^p \text{ and } S^q) \\ &= \frac{1}{2}(p(S^p)-q(S^p)-(p(S^q)-q(S^q))) && \text{(definition of } p,q \text{ applied to subsets)} \\ &= p(S^p)-q(S^p) && (p(S^q)-q(S^q) = -(p(S^p)-q(S^p))) \end{aligned}$$

So, $D(p_x,q_x) = p(S^p)-q(S^p) = \max_S |p(S)-q(S)|$.

9.5) Show that the absolute value signs may be removed from Equation (9.3), that is,

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

Soln: This has effectively already been proven in the solution to exercise 9.4 by the observation that $p(S) - q(S)$ achieves extremes equal in magnitude, but opposite in sign, on S^p and S^q. So,

$$\max_S (p(S) - q(S)) = p(S^p) - q(S^p) = \max_S |p(S) - q(S)|.$$

9.6) What is the trace distance between the density operators

$$\frac{3}{4}|0\rangle\langle 0| + \frac{1}{4}|1\rangle\langle 1|\,; \frac{2}{3}|0\rangle\langle 0| + \frac{1}{3}|1\rangle\langle 1|?$$

Between:

$$\frac{3}{4}|0\rangle\langle 0| + \frac{1}{4}|1\rangle\langle 1|\,; \frac{2}{3}|+\rangle\langle +| + \frac{1}{3}|-\rangle\langle -|?$$

(Recall that $|\pm\rangle \equiv (|0\rangle + |1\rangle)/\sqrt{2}$)

Soln: Define $\rho = \frac{3}{4}|0\rangle\langle 0| + \frac{1}{4}|1\rangle\langle 1|$, $\sigma = \frac{2}{3}|1\rangle\langle 1| + \frac{1}{3}|1\rangle\langle 1|$. Both being diagonal in the same basis, we may apply equation (9.14):

$$\begin{aligned} D(\rho, \sigma) &= \frac{1}{2}\operatorname{tr}|\rho - \sigma| && \text{(definition)} \\ &= D((3/4, 1/4), (2/3, 1/3)) && \text{(equation (9.14))} \\ &= \frac{1}{2}\left(\left| \frac{3}{4} - \frac{2}{3} \right| + \left| \frac{1}{4} - \frac{1}{3} \right| \right) && \text{(definition of trace-distance of probability distributions)} \\ &= \frac{1}{2}\left(\frac{1}{12} + \frac{1}{12} \right) \\ &= \frac{1}{12} \end{aligned}$$

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

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

so, $\sigma = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{6}|0\rangle\langle 1| + \frac{1}{6}|1\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1|$.

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

$$\begin{aligned} (\rho - \sigma)^\dagger(\rho - \sigma) &= (\rho - \sigma)^2 && (\rho - \sigma \text{ is symmetric, real-valued}) \\ &= \begin{bmatrix} \frac{1}{4} & \frac{-1}{6} \\ \frac{-1}{6} & \frac{-1}{4} \end{bmatrix} \begin{bmatrix} \frac{1}{4} & \frac{-1}{6} \\ \frac{-1}{6} & \frac{-1}{4} \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{4^2} + \frac{1}{6^2} & \frac{-1}{4\cdot 6} + \frac{1}{4\cdot 6} \\ \frac{-1}{4\cdot 6} + \frac{1}{4\cdot 6} & \frac{1}{6^2} + \frac{1}{4^2} \end{bmatrix} \\ &= \left(\frac{1}{4^2} + \frac{1}{6^2} \right) I \end{aligned}$$

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

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 $\sigma = \sum_i p_i \sigma_i$.

$$
\begin{aligned}
D\left(\sum_i p_i \rho_i, \sigma\right) &= D\left(\sum_i p_i \rho_i, \sum_i p_i \sigma_i\right) && (9.1) \\
&\leq \sum_i p_i D(\rho_i, \sigma_i) \quad (\because \text{eqn}(9.50)) && (9.2) \\
&= \sum_i p_i D(\rho_i, \sigma). \quad (\because \text{assumption}). && (9.3)
\end{aligned}
$$

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

$$A = \sum_i a_i \left|i\right\rangle\!\left\langle i\right|.$$

Let f be a function. Then

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

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

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

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. $\operatorname{tr}(A^\dagger B) = \left\langle m\right|A\otimes B\left|m\right\rangle$ should be $\operatorname{tr}(A^T B) = \left\langle m\right|A\otimes B\left|m\right\rangle$. See errata list.

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

$$\operatorname{tr}(A) = \left\langle m\right|(I\otimes A)\left|m\right\rangle, \quad (I\otimes A)\left|m\right\rangle = (A^T\otimes I)\left|m\right\rangle$$

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)

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

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

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 $\mathcal{E}(\rho_j) = \rho_j$.

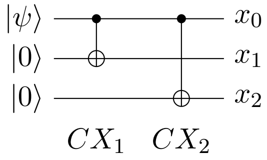
Problem 1)
Problem 2)

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 10

Quantum error-correction

10.1) Verify that the encoding circuit in Figure 10.2 works as claimed:



Soln: Let $|\psi\rangle = a\,|0\rangle + b\,|1\rangle$. Our goal is to show that application of the diagrammed circuit maps $|\psi\rangle\,|00\rangle = a\,|000\rangle + b\,|100\rangle$ to $a\,|000\rangle + b\,|111\rangle = a\,|0_L\rangle + b\,|1_L\rangle$. Note that $|\psi\rangle\,|00\rangle = \begin{bmatrix} a & 0 & 0 & 0 & b & 0 & 0 & 0 \end{bmatrix}^T$, and

$$CX_1 = \begin{bmatrix} 1&0&0&0&0&0&0&0\\ 0&1&0&0&0&0&0&0\\ 0&0&1&0&0&0&0&0\\ 0&0&0&1&0&0&0&0\\ 0&0&0&0&0&0&1&0\\ 0&0&0&0&0&0&0&1\\ 0&0&0&0&1&0&0&0\\ 0&0&0&0&0&1&0&0 \end{bmatrix},\ CX_2 = \begin{bmatrix} 1&0&0&0&0&0&0&0\\ 0&1&0&0&0&0&0&0\\ 0&0&1&0&0&0&0&0\\ 0&0&0&1&0&0&0&0\\ 0&0&0&0&0&1&0&0\\ 0&0&0&0&1&0&0&0\\ 0&0&0&0&0&0&0&1\\ 0&0&0&0&0&0&1&0 \end{bmatrix}.$$

So

$$\begin{aligned} CX_2 CX_1\,|\psi\rangle\,|00\rangle &= \begin{bmatrix} 1&0&0&0&0&0&0&0\\ 0&1&0&0&0&0&0&0\\ 0&0&1&0&0&0&0&0\\ 0&0&0&1&0&0&0&0\\ 0&0&0&0&0&1&0&0\\ 0&0&0&0&1&0&0&0\\ 0&0&0&0&0&0&0&1\\ 0&0&0&0&0&0&1&0 \end{bmatrix} \begin{bmatrix} 1&0&0&0&0&0&0&0\\ 0&1&0&0&0&0&0&0\\ 0&0&1&0&0&0&0&0\\ 0&0&0&1&0&0&0&0\\ 0&0&0&0&0&0&1&0\\ 0&0&0&0&0&0&0&1\\ 0&0&0&0&1&0&0&0\\ 0&0&0&0&0&1&0&0 \end{bmatrix} \begin{bmatrix} a\\0\\0\\0\\b\\0\\0\\0 \end{bmatrix} \\ &= \begin{bmatrix} a & 0 & 0 & 0 & 0 & 0 & 0 & b \end{bmatrix}^T \\ &= a\,|000\rangle + b\,|111\rangle \end{aligned}$$

as desired.

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

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

11.4)
11.5)

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

From the non-negativity of the relative entropy,

$$\begin{aligned}H(X) + H(Y) - H(X,Y) &\geq 0 && (11.21)\\ \therefore H(X) + H(Y) &\geq H(X,Y). && (11.22)\end{aligned}$$

11.6)

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

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)

Chapter 12

Quantum information theory

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

$$|ABE\rangle = U\,|\beta_{00}^{\otimes n}\rangle\,|0\rangle_E \tag{12.1}$$

$$\rho^E = tr_{AB}(|ABE\rangle\,\langle ABE|) \tag{12.2}$$

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^E_k and p_k are the corresponding Eve's states and probabilities. Note,

$$\rho_E = \sum_k p_k \rho^E_k. \tag{12.3}$$

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,

$$H(K:E) \leq S(\rho^E) - \sum_k p_k \rho^E_k \leq S(\rho^E) = S(\rho). \tag{12.4}$$

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 $P(E \mid k) = P(E \mid f(x_0) = \cdots = f(x_{k-1}))$, $P(k \mid E) = P(f(x_0) = \cdots = f(x_{k-1}) \mid E)$, and $P(k) = P(f(x_0) = \cdots = f(x_{k-1}))$, 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
$n = 4$	$\frac{3}{4} \simeq 0.7500$	$\frac{1}{1} \simeq 1.0000$	$\rightarrow$						
$n = 6$	$\frac{5}{7} \simeq 0.7143$	$\frac{10}{11} \simeq 0.9091$	$\frac{1}{1} \simeq 1.0000$	$\rightarrow$					
$n = 8$	$\frac{7}{10} \simeq 0.7000$	$\frac{7}{8} \simeq 0.8750$	$\frac{35}{36} \simeq 0.9722$	$\frac{1}{1} \simeq 1.0000$	$\rightarrow$				
$n = 10$	$\frac{9}{13} \simeq 0.6923$	$\frac{6}{7} \simeq 0.8571$	$\frac{21}{22} \simeq 0.9545$	$\frac{126}{127} \simeq 0.9921$	$\frac{1}{1} \simeq 1.0000$	$\rightarrow$			
$n = 12$	$\frac{11}{16} \simeq 0.6875$	$\frac{11}{13} \simeq 0.8462$	$\frac{33}{35} \simeq 0.9429$	$\frac{66}{67} \simeq 0.9851$	$\frac{462}{463} \simeq 0.9978$	$\frac{1}{1} \simeq 1.0000$	$\rightarrow$		
$n = 14$	$\frac{13}{19} \simeq 0.6842$	$\frac{26}{31} \simeq 0.8387$	$\frac{143}{153} \simeq 0.9346$	$\frac{143}{146} \simeq 0.9795$	$\frac{429}{431} \simeq 0.9954$	$\frac{1716}{1717} \simeq 0.9994$	$\frac{1}{1} \simeq 1.0000$	$\rightarrow$	
$n = 16$	$\frac{15}{22} \simeq 0.6818$	$\frac{5}{6} \simeq 0.8333$	$\frac{13}{14} \simeq 0.9286$	$\frac{39}{40} \simeq 0.9750$	$\frac{143}{144} \simeq 0.9931$	$\frac{715}{716} \simeq 0.9986$	$\frac{6435}{6436} \simeq 0.9998$	$\frac{1}{1} \simeq 1.0000$	$\rightarrow$
$n = 18$	$\frac{17}{25} \simeq 0.6800$	$\frac{34}{41} \simeq 0.8293$	$\frac{85}{92} \simeq 0.9239$	$\frac{34}{35} \simeq 0.9714$	$\frac{221}{223} \simeq 0.9910$	$\frac{442}{443} \simeq 0.9977$	$\frac{2431}{2432} \simeq 0.9996$	$\frac{24310}{24311} \simeq 1.0000$	$\frac{1}{1} \simeq 1.0000$
$n = 20$	$\frac{19}{28} \simeq 0.6786$	$\frac{19}{23} \simeq 0.8261$	$\frac{323}{351} \simeq 0.9202$	$\frac{646}{667} \simeq 0.9685$	$\frac{646}{653} \simeq 0.9893$	$\frac{323}{324} \simeq 0.9969$	$\frac{4199}{4202} \simeq 0.9993$	$\frac{8398}{8399} \simeq 0.9999$	$\frac{92378}{92379} \simeq 1.0000$
$n = 22$	$\frac{21}{31} \simeq 0.6774$	$\frac{14}{17} \simeq 0.8235$	$\frac{133}{145} \simeq 0.9172$	$\frac{57}{59} \simeq 0.9661$	$\frac{323}{327} \simeq 0.9878$	$\frac{1292}{1297} \simeq 0.9961$	$\frac{969}{970} \simeq 0.9990$	$\frac{4522}{4523} \simeq 0.9998$	$\frac{29393}{29394} \simeq 1.0000$
$n = 24$	$\frac{23}{34} \simeq 0.6765$	$\frac{23}{28} \simeq 0.8214$	$\frac{161}{176} \simeq 0.9148$	$\frac{161}{167} \simeq 0.9641$	$\frac{437}{443} \simeq 0.9865$	$\frac{437}{439} \simeq 0.9954$	$\frac{7429}{7439} \simeq 0.9987$	$\frac{14858}{14863} \simeq 0.9997$	$\frac{14858}{14859} \simeq 0.9999$
$n = 26$	$\frac{25}{37} \simeq 0.6757$	$\frac{50}{61} \simeq 0.8197$	$\frac{115}{126} \simeq 0.9127$	$\frac{230}{239} \simeq 0.9623$	$\frac{805}{817} \simeq 0.9853$	$\frac{575}{578} \simeq 0.9948$	$\frac{10925}{10943} \simeq 0.9984$	$\frac{2185}{2186} \simeq 0.9995$	$\frac{37145}{37149} \simeq 0.9999$
$n = 28$	$\frac{27}{40} \simeq 0.6750$	$\frac{9}{11} \simeq 0.8182$	$\frac{225}{247} \simeq 0.9109$	$\frac{270}{281} \simeq 0.9609$	$\frac{690}{701} \simeq 0.9843$	$\frac{345}{347} \simeq 0.9942$	$\frac{1035}{1037} \simeq 0.9981$	$\frac{1725}{1726} \simeq 0.9994$	$\frac{6555}{6556} \simeq 0.9998$
$n = 30$	$\frac{29}{43} \simeq 0.6744$	$\frac{58}{71} \simeq 0.8169$	$\frac{261}{287} \simeq 0.9094$	$\frac{261}{272} \simeq 0.9596$	$\frac{1305}{1327} \simeq 0.9834$	$\frac{1740}{1751} \simeq 0.9937$	$\frac{10005}{10027} \simeq 0.9978$	$\frac{10005}{10012} \simeq 0.9993$	$\frac{10005}{10007} \simeq 0.9998$
$n = 32$	$\frac{31}{46} \simeq 0.6739$	$\frac{31}{38} \simeq 0.8158$	$\frac{899}{990} \simeq 0.9081$	$\frac{899}{938} \simeq 0.9584$	$\frac{8091}{8234} \simeq 0.9826$	$\frac{8091}{8146} \simeq 0.9932$	$\frac{4495}{4506} \simeq 0.9976$	$\frac{13485}{13496} \simeq 0.9992$	$\frac{310155}{310232} \simeq 0.9998$
$n = 34$	$\frac{33}{49} \simeq 0.6735$	$\frac{22}{27} \simeq 0.8148$	$\frac{341}{376} \simeq 0.9069$	$\frac{2046}{2137} \simeq 0.9574$	$\frac{9889}{10071} \simeq 0.9819$	$\frac{1798}{1811} \simeq 0.9928$	$\frac{24273}{24338} \simeq 0.9973$	$\frac{5394}{5399} \simeq 0.9991$	$\frac{13485}{13489} \simeq 0.9997$
$n = 36$	$\frac{35}{52} \simeq 0.6731$	$\frac{35}{43} \simeq 0.8140$	$\frac{77}{85} \simeq 0.9059$	$\frac{22}{23} \simeq 0.9565$	$\frac{682}{695} \simeq 0.9813$	$\frac{1705}{1718} \simeq 0.9924$	$\frac{4495}{4508} \simeq 0.9971$	$\frac{12586}{12599} \simeq 0.9990$	$\frac{37758}{37771} \simeq 0.9997$
$n = 38$	$\frac{37}{55} \simeq 0.6727$	$\frac{74}{91} \simeq 0.8132$	$\frac{1295}{1431} \simeq 0.9050$	$\frac{259}{271} \simeq 0.9557$	$\frac{407}{415} \simeq 0.9807$	$\frac{1628}{1641} \simeq 0.9921$	$\frac{12617}{12656} \simeq 0.9969$	$\frac{11470}{11483} \simeq 0.9989$	$\frac{33263}{33276} \simeq 0.9996$
$n = 40$	$\frac{39}{58} \simeq 0.6724$	$\frac{13}{16} \simeq 0.8125$	$\frac{481}{532} \simeq 0.9041$	$\frac{1443}{1511} \simeq 0.9550$	$\frac{3367}{3435} \simeq 0.9802$	$\frac{481}{485} \simeq 0.9918$	$\frac{1221}{1225} \simeq 0.9967$	$\frac{814}{815} \simeq 0.9988$	$\frac{2294}{2295} \simeq 0.9996$
$n = 42$	$\frac{41}{61} \simeq 0.6721$	$\frac{82}{101} \simeq 0.8119$	$\frac{533}{590} \simeq 0.9034$	$\frac{1066}{1117} \simeq 0.9543$	$\frac{19721}{20129} \simeq 0.9797$	$\frac{19721}{19891} \simeq 0.9915$	$\frac{19721}{19789} \simeq 0.9966$	$\frac{1517}{1519} \simeq 0.9987$	$\frac{16687}{16695} \simeq 0.9995$
$n = 44$	$\frac{43}{64} \simeq 0.6719$	$\frac{43}{53} \simeq 0.8113$	$\frac{1763}{1953} \simeq 0.9027$	$\frac{3526}{3697} \simeq 0.9537$	$\frac{45838}{46807} \simeq 0.9793$	$\frac{22919}{23123} \simeq 0.9912$	$\frac{848003}{851063} \simeq 0.9964$	$\frac{848003}{849193} \simeq 0.9986$	$\frac{65231}{65265} \simeq 0.9995$
$n = 46$	$\frac{45}{67} \simeq 0.6716$	$\frac{30}{37} \simeq 0.8108$	$\frac{129}{143} \simeq 0.9021$	$\frac{387}{406} \simeq 0.9532$	$\frac{1763}{1801} \simeq 0.9789$	$\frac{35260}{35583} \simeq 0.9909$	$\frac{343785}{345077} \simeq 0.9963$	$\frac{22919}{22953} \simeq 0.9985$	$\frac{848003}{848479} \simeq 0.9994$
$n = 48$	$\frac{47}{70} \simeq 0.6714$	$\frac{47}{58} \simeq 0.8103$	$\frac{705}{782} \simeq 0.9015$	$\frac{141}{148} \simeq 0.9527$	$\frac{6063}{6196} \simeq 0.9785$	$\frac{2021}{2040} \simeq 0.9907$	$\frac{82861}{83184} \simeq 0.9961$	$\frac{414305}{414951} \simeq 0.9984$	$\frac{1077193}{1077839} \simeq 0.9994$
$n = 50$	$\frac{49}{73} \simeq 0.6712$	$\frac{98}{121} \simeq 0.8099$	$\frac{2303}{2556} \simeq 0.9010$	$\frac{658}{691} \simeq 0.9522$	$\frac{987}{1009} \simeq 0.9782$	$\frac{1974}{1993} \simeq 0.9905$	$\frac{14147}{14204} \simeq 0.9960$	$\frac{198058}{198381} \simeq 0.9984$	$\frac{4060189}{4062773} \simeq 0.9994$
$n = 52$	$\frac{51}{76} \simeq 0.6711$	$\frac{17}{21} \simeq 0.8095$	$\frac{833}{925} \simeq 0.9005$	$\frac{4998}{5251} \simeq 0.9518$	$\frac{11186}{11439} \simeq 0.9779$	$\frac{5593}{5648} \simeq 0.9903$	$\frac{50337}{50546} \simeq 0.9959$	$\frac{11186}{11205} \simeq 0.9983$	$\frac{28294}{28313} \simeq 0.9993$
$n = 54$	$\frac{53}{79} \simeq 0.6709$	$\frac{106}{131} \simeq 0.8092$	$\frac{901}{1001} \simeq 0.9001$	$\frac{901}{947} \simeq 0.9514$	$\frac{44149}{45161} \simeq 0.9776$	$\frac{25228}{25481} \simeq 0.9901$	$\frac{296429}{297694} \simeq 0.9958$	$\frac{592858}{593903} \simeq 0.9982$	$\frac{296429}{296638} \simeq 0.9993$
$n = 56$	$\frac{55}{82} \simeq 0.6707$	$\frac{55}{68} \simeq 0.8088$	$\frac{583}{648} \simeq 0.8997$	$\frac{583}{613} \simeq 0.9511$	$\frac{9911}{10141} \simeq 0.9773$	$\frac{4505}{4551} \simeq 0.9899$	$\frac{31535}{31673} \simeq 0.9956$	$\frac{12614}{12637} \simeq 0.9982$	$\frac{592858}{593295} \simeq 0.9993$
$n = 58$	$\frac{57}{85} \simeq 0.6706$	$\frac{38}{47} \simeq 0.8085$	$\frac{1045}{1162} \simeq 0.8993$	$\frac{1254}{1319} \simeq 0.9507$	$\frac{11077}{11337} \simeq 0.9771$	$\frac{11077}{11192} \simeq 0.9897$	$\frac{51357}{51587} \simeq 0.9955$	$\frac{85595}{85756} \simeq 0.9981$	$\frac{119833}{119925} \simeq 0.9992$
$n = 60$	$\frac{59}{88} \simeq 0.6705$	$\frac{59}{73} \simeq 0.8082$	$\frac{1121}{1247} \simeq 0.8990$	$\frac{2242}{2359} \simeq 0.9504$	$\frac{24662}{25247} \simeq 0.9768$	$\frac{12331}{12461} \simeq 0.9896$	$\frac{653543}{656533} \simeq 0.9954$	$\frac{59413}{59528} \simeq 0.9981$	$\frac{1010021}{1010826} \simeq 0.9992$
$n = 62$	$\frac{61}{91} \simeq 0.6703$	$\frac{122}{151} \simeq 0.8079$	$\frac{3599}{4005} \simeq 0.8986$	$\frac{3599}{3788} \simeq 0.9501$	$\frac{68381}{70019} \simeq 0.9766$	$\frac{273524}{276449} \simeq 0.9894$	$\frac{752191}{755701} \simeq 0.9954$	$\frac{752191}{753686} \simeq 0.9980$	$\frac{3624193}{3627183} \simeq 0.9992$
$n = 64$	$\frac{63}{94} \simeq 0.6702$	$\frac{21}{26} \simeq 0.8077$	$\frac{1281}{1426} \simeq 0.8983$	$\frac{549}{578} \simeq 0.9498$	$\frac{3599}{3686} \simeq 0.9764$	$\frac{3599}{3638} \simeq 0.9893$	$\frac{68381}{68706} \simeq 0.9953$	$\frac{478667}{479642} \simeq 0.9980$	$\frac{5265337}{5269822} \simeq 0.9991$