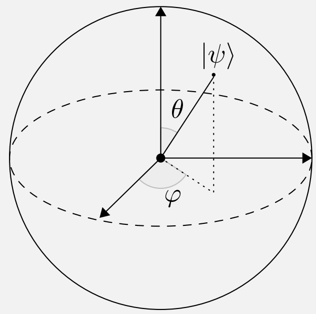
Solutions to Quantum Computation and Quantum Information (Vol. 1)



Arnab Adhikary & Rio Weil

This document was typeset on October 13, 2025

Introduction:

This document is the first volume to a collection of comprehensive solutions to the exercises and problems in Nielsen and Chuang's "Quantum Computation and Quantum Information". Each solution has the involved concepts (and hence rough pre-requisite knowledge) necessary for the problem in addition to the solution. Some exercises may contain additional remarks about implications. Starred exercises were considered to be more difficult (difficulty is assumed for the problems at the end of the chapter).

The first volume contains solutions to all exercises/problems in Chapters 2 (Introduction to quantum mechanics), 4 (Quantum circuits), 7 (Quantum computers: physical realization), 8 (Quantum noise and quantum operations), 9 (Distance measures for quantum information), and 10 (Quantum error-correction).

The most up-to-date version of this document, as well as a feedback form, can be found at `https://nielsenandchuangsolutions.github.io`.

Contents

2 Introduction to quantum mechanics

Exercise 2.1: Linear dependence: example

Show that $(1, -1)$, $(1, 2)$ and $(2, 1)$ are linearly dependent.

Solution

Concepts Involved: Linear Independence/Dependence

We observe that:

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

showing that the three vectors are linearly dependent by definition. $\square$

Remark: Alternatively, we can apply theorem that states that for any vector space V with $\dim V = n$, any list of $m > n$ vectors in V will be linearly dependent (here, $V = \mathbb{R}^2, n = 2, m = 3$).

Exercise 2.2: Matrix representations: example

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

Solution

Concepts Involved: Linear Algebra, Matrix Representation of Operators.

Identifying $|0\rangle \cong \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $|1\rangle \cong \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, we have:

$$A = \begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{bmatrix}$$

Using the given relations, we have:

$$A|0\rangle = 0|0\rangle + 1|1\rangle \implies \begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 0 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 1 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \implies a_{00} = 0, a_{10} = 1$$

$$A|1\rangle = 1|0\rangle + 0|1\rangle \implies \begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 0 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \implies a_{01} = 1, a_{11} = 0$$

Therefore with respect to the input basis $\{|0\rangle, |1\rangle\}$ and output basis $\{|0\rangle, |1\rangle\}$, A has matrix representation:

$$A \cong \begin{array}{c c} & \begin{array}{cc} \langle 0| & \langle 1| \end{array} \\ \begin{array}{c} |0\rangle \\ |1\rangle \end{array} & \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \end{array}$$

Suppose we instead choose the input and output basis to be $\left\{ |+\rangle = \frac{|0\rangle+|1\rangle}{\sqrt{2}}, |-\rangle = \frac{|0\rangle-|1\rangle}{\sqrt{2}} \right\}$. Identifying $|+\rangle \cong \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $|-\rangle \cong \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, we have:

$$A = \begin{bmatrix} a_{++} & a_{+-} \\ a_{-+} & a_{--} \end{bmatrix}$$

Using the linearity of A, we have:

$$A\,|+\rangle = \frac{1}{\sqrt{2}} A(|0\rangle + |1\rangle) = \frac{1}{\sqrt{2}}(A\,|0\rangle + A\,|1\rangle) = \frac{1}{\sqrt{2}}(|1\rangle + |0\rangle) = |+\rangle$$

and:

$$A\,|-\rangle = \frac{1}{\sqrt{2}} A(\,|0\rangle - \,|1\rangle) = \frac{1}{\sqrt{2}}(A\,|0\rangle - A\,|1\rangle) = \frac{1}{\sqrt{2}}(\,|1\rangle - \,|0\rangle) = -\,|-\rangle\,,$$

which can be used to determine the matrix elements:

$$A\,|+\rangle = 1\,|+\rangle + 0\,|-\rangle \implies \begin{bmatrix} a_{++} & a_{+-} \\ a_{-+} & a_{--} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 0 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \implies a_{++} = 1, a_{+-} = 0$$

$$A\,|-\rangle = 0\,|+\rangle - 1\,|-\rangle \implies \begin{bmatrix} a_{++} & a_{+-} \\ a_{-+} & a_{--} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0 \begin{bmatrix} 1 \\ 0 \end{bmatrix} - 1 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \implies a_{-+} = 0, a_{--} = -1$$

Therefore with respect to the input basis $\{\,|+\rangle\,, \,|-\rangle\}$ and output basis $\{\,|+\rangle\,, \,|-\rangle\}$, A has matrix representation:

$$A \cong \begin{array}{c c} & \begin{array}{cc} \langle +| & \langle -| \end{array} \\ \begin{array}{c} |+\rangle \\ |-\rangle \end{array} & \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \end{array}$$

□

Remark: If we choose the input and output bases to be different, we can even represent the A operator as an identity matrix. Specifically, if the input basis to be chosen to be $\{\,|0\rangle\,, \,|1\rangle\}$ and output basis as $\{\,|1\rangle\,, \,|0\rangle\}$, the matrix representation of A looks like:

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

Exercise 2.3: Matrix representation for operator products

Suppose A is a linear operator from vector space V to vector space W, and B is a linear operator from vector space W to vector space X. Let $|v_i\rangle, |w_j\rangle, |x_k\rangle$ 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.

Solution

Concepts Involved: Matrix Representation of Operators

Taking the matrix of representations of A and B to the appropriate bases $|v_i\rangle, |w_j\rangle, |x_k\rangle$ of V, W and X, we have:

$$A|v_j\rangle = \sum_i A_{ij}|w_i\rangle, \quad B|w_i\rangle = \sum_k B_{ki}|x_k\rangle$$

Hence, looking at $BA : V \mapsto X$, we have:

$$\begin{aligned} BA|v_j\rangle &= B(A|v_j\rangle) \\ &= B\left(\sum_i A_{ij}|w_i\rangle\right) \\ &= \sum_i A_{ij}B|w_i\rangle \\ &= \sum_i A_{ij}\left(\sum_k B_{ki}|x_k\rangle\right) \\ &= \sum_k\sum_i B_{ki}A_{ij}|x_k\rangle \\ &= \sum_k (BA)_{kj}|x_k\rangle \end{aligned}$$

which shows that the matrix representation of BA is indeed the matrix product of the representations of B and A. $\square$

Exercise 2.4: Matrix representation for identity

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

Solution

Concepts Involved: Matrix Representation of Operators

Let V be a vector space and $|v_i\rangle$ be a basis of V. Let $A : V \mapsto V$ be a linear operator, and let its matrix representation taken to be respect to $|v_i\rangle$ as the input and output basis. We then have for each

$i \in \{1, \dots, n\}$:

$$A|v_i\rangle = 1|v_i\rangle + \sum_{j \neq i} 0|v_j\rangle = \sum_j \delta_{ij}|v_j\rangle$$

From which we obtain that A has the matrix representation:

$$A \cong \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & & & \ddots & \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

□

Exercise 2.5

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

Solution

Concepts Involved: Inner Products
Recall that on $\mathbb{C}^n$, $(\cdot, \cdot)$ was defined as:

$$((y_1, \dots, y_n), (z_1, \dots, z_n)) \equiv \sum_i y_i^* z_i = [y_1^* \dots y_n^*] \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}.$$

Furthermore, recall the three conditions for the function $(\cdot, \cdot) : V \times V \mapsto \mathbb{C}$ to be considered an inner product:

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

(2) $(|v\rangle, |w\rangle) = (|w\rangle, |v\rangle)^*$.

(3) $(|v\rangle, |v\rangle) \geq 0$ with equality if and only if $|v\rangle = \mathbf{0}$.

We check that $(\cdot, \cdot) : \mathbb{C}^n \times \mathbb{C}^n \mapsto \mathbb{C}$ satisfies the three conditions:

(1) We see that:

$$\begin{aligned} ((y_1, \dots, y_n), \sum_k \lambda_k (z_1, \dots, z_n)_k) &= \sum_i y_i^* \sum_k \lambda_k z_{i_k} \\ &= \sum_k \lambda_k \sum_i y_i^* z_{i_k} \\ &= \sum_k \lambda_k ((y_1, \dots, y_n), (z_1, \dots, z_n)_k) \end{aligned}$$

(2) We have:

$$((y_1, \dots, y_n), (z_1, \dots, z_n)) = \sum_i y_i^* z_i$$
$$= \sum_i (y_i z_i^*)^*$$
$$= \left(\sum_i z_i^* y_i\right)^*$$
$$= \left(((z_1, \dots, z_n), (y_1, \dots, y_n))\right)^*$$

(3) We observe for $\mathbf{0} = (0, \dots 0)$:

$$(\mathbf{0}, \mathbf{0}) = \sum_i 0 \cdot 0 = 0$$

For $\mathbf{y} = (y_1, \dots, y_n) \neq 0$ we have at least one y_i (say, y_j) is nonzero, and hence:

$$((y_1, \dots, y_n), (y_1, \dots, y_n)) = \sum_i y_i^2 \geq y_j^2 > 0$$

which proves the claim.

□

Exercise 2.6

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

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

Solution

Concepts Involved: Inner Products

Applying properties (2) (conjugate symmetry), (1) (linearity in second argument), and (2) (again) in succession, we have:

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

□

Exercise 2.7

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

Solution

Concepts Involved: Inner Products, Orthogonality, Normalization

Recall that two vectors $|v\rangle, |w\rangle$ are orthogonal if $\langle v|w\rangle = 0$, and the norm of $|v\rangle$ is given by $\| |v\rangle \| = \sqrt{\langle v|v\rangle}$.

First we show the two vectors are orthogonal:

$$\langle w|v\rangle = 1 \cdot 1 + 1 \cdot (-1) = 0$$

The norms of $\left|w\right\rangle, \left|v\right\rangle$ are given by:

$$\| |w\rangle \| = \sqrt{\langle w|w\rangle} = \sqrt{1^2 + 1^2} = \sqrt{2},$$
$$\| |c\rangle \| = \sqrt{\langle v|v\rangle} = \sqrt{1^2 + (-1)^2} = \sqrt{2}$$

So the normalized forms of the vectors are:

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

□

Exercise 2.8

Verify that the Gram-Schmidt procedure produces and orthonormal basis for V.

Solution

Concepts Involved: Linear Independence, Bases, Inner Products, Orthogonality, Normalization, Gram-Schmidt Procedure, Induction

Recall that given $|w_1\rangle, \dots, |w_d\rangle$ as a basis set for a vector space V, the Gram-Schmidt procedure constructs a basis set $|v_1\rangle, \dots, |v_d\rangle$ by defining $|v_1\rangle \equiv |w_1\rangle / \left\| |w_1\rangle \right\|$ and then defining $|v_{k+1}\rangle$ inductively for $1 \leq k \leq d-1$ as:

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

It is evident that each of the $|v_j\rangle$ have unit norm as they are defined in normalized form. It therefore suffices to show that each of the $|v_1\rangle, \dots, |v_d\rangle$ are orthogonal to each other, and that this set of vectors forms a basis of V. We proceed by induction. For $k = 1$, we have:

$$|v_2\rangle = \frac{|w_2\rangle - \langle v_1|w_2\rangle |v_1\rangle}{\left\| |w_2\rangle + \langle v_1|w_2\rangle |v_1\rangle \right\|}$$

Therefore:

$$\langle v_1|v_2\rangle = \frac{\langle v_1|w_2\rangle - \langle v_1|w_2\rangle \langle v_1|v_1\rangle}{\left\| |w_2\rangle + \langle v_1|w_2\rangle |v_1\rangle \right\|} = \frac{\langle v_1|w_2\rangle - \langle v_1|w_2\rangle}{\left\| |w_2\rangle + \langle v_1|w_2\rangle |v_1\rangle \right\|} = 0$$

so the two vectors are orthogonal. Furthermore, they are linearly independent; if they were linearly dependent, we could write $|v_1\rangle = \lambda |v_2\rangle$ for some $\lambda \in \mathbb{C}$, but then multiplying both sides by $\langle v_1|$ we get:

$$\langle v_1|v_1\rangle = \lambda \langle v_1|v_2\rangle \implies 1 = 0$$

which is a contradiction. This concludes the base case. For the inductive step, let $k \geq 1$ and suppose that $|v_1\rangle, \dots, |v_k\rangle$ are orthogonal and linearly independent. We then have:

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

Then for any $j \in \{1, \dots k\}$, we have:

$$\langle v_j|v_{k+1}\rangle = \frac{\langle v_j|w_{k+1}\rangle - \sum_{i=1}^{k} \langle v_i|w_{k+1}\rangle \langle v_j|v_j\rangle |v_i\rangle}{\left\| |w_{k+1}\rangle - \sum_{i=1}^{k} \langle v_i|w_{k+1}\rangle |v_i\rangle \right\|} = \frac{\langle v_j|w_{k+1}\rangle - \langle v_j|w_{k+1}\rangle}{\left\| |w_{k+1}\rangle - \sum_{i=1}^{k} \langle v_i|w_{k+1}\rangle |v_i\rangle \right\|} = 0$$

where in the second equality we use the fact that $\left\langle v_j \middle| v_i \right\rangle = \delta_{ij}$ for $i, j \in \{1, \dots k\}$ by the inductive hypothesis. We therefore find that $|v_{k+1}\rangle$ is orthogonal to all of $|v_1\rangle, \dots, |v_k\rangle$. Furthermore, $|v_1\rangle, \dots, |v_k\rangle, |v_{k+1}\rangle$ is lienarly independent. Suppose for the sake of contradiction that this was false. Then, there would exist $\lambda_1, \dots \lambda_k$ not all nonzero such that:

$$\lambda_1 |v_1\rangle + \dots + \lambda_k |v_k\rangle = |v_{k+1}\rangle$$

but then multiplying both sides by $\langle v_{k+1}|$ we have:

$$\lambda_1 \langle v_{k+1}|v_1\rangle + \ldots + \lambda_k \langle v_{k+1}|v_k\rangle = \langle v_{k+1}|v_{k+1}\rangle \implies 0 = 1$$

by orthonormality. This gives a contradiction, and hence $|v_1\rangle, \ldots, |v_k\rangle, |v_{k+1}\rangle$ are linearly independent, finishing the inductive step. Therefore, $|v_1\rangle, \ldots, |v_d\rangle$ is an orthonormal list of vectors which is linearly independent. Since $|w_1\rangle, \ldots, |w_d\rangle$ is a basis for V, then V has dimension d. Hence, $|v_1\rangle, \ldots, |v_d\rangle$ being a linearly independent list of d vectors in V is a basis of V. We conclude that it is an orthonormal basis of V, as claimed. □

Exercise 2.9: Pauli operators and the outer product

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

Solution

Concepts Involved: Matrix Representation of Operators, Outer Products
Recall that if A has matrix representation:

$$A \cong \begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{bmatrix}$$

with respect to $|0\rangle, |1\rangle$ as the input/output bases, then we can express A in outer product notation as:

$$A = a_{00} |0\rangle\langle 0| + a_{01} |0\rangle\langle 1| + a_{10} |1\rangle\langle 0| + a_{11} |1\rangle\langle 1|$$

Furthermore, recall the representation of the Pauli matrices with respect to the orthonormal basis $|0\rangle, |1\rangle$:

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

We immediately see that:

$$\begin{aligned} I &= |0\rangle\langle 0| + |1\rangle\langle 1| \\ X &= |0\rangle\langle 1| + |1\rangle\langle 0| \\ Y &= -i\,|0\rangle\langle 1| + i\,|1\rangle\langle 0| \\ Z &= |0\rangle\langle 0| - |1\rangle\langle 1| \end{aligned}$$

□

Exercise 2.10

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

Solution

Concepts Involved: Matrix Representation of Operators, Outer Products

The matrix representation of $|v_j\rangle\langle v_j|$ with respect to the $|v_i\rangle$ basis is a matrix with 1 in the jth column and row (i.e. the (j,j)th entry in the matrix) and 0 everywhere else. □

Exercise 2.11

Find the eigenvectors, eigenvalues, and diagonal representations of the Pauli matrices X, Y and Z.

Solution

Concepts Involved: Eigenvalues, Eigenvectors, Diagonalization, Pauli Operators

Given an operator A on a vector space V, recall that an eigenvector $|v\rangle$ of A and its corresponding eigenvalue λ are defined by:

$$A|v\rangle = \lambda|v\rangle$$

Furthermore, recall the diagonal representation of A is given by

$$A = \sum_i \lambda_i \, |i\rangle\langle i|$$

Where $|i\rangle$ form an orthonormal set of eigenvectors for A, and λ_i are the corresponding eigenvalues.

We start with X. Solving for the eigenvalues, we have:

$$\det(X - I\lambda) = 0 \implies \det \begin{bmatrix} -\lambda & 1 \\ 1 & -\lambda \end{bmatrix} = 0 \implies \lambda^2 - 1 = 0$$

From which we obtain $\lambda_1 = 1, \lambda_2 = -1$. Solving for the eigenvectors, we then have:

$$(X - I\lambda_1)|v_1\rangle = \mathbf{0} \implies \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} v_{11} \\ v_{12} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \implies v_{11} = 1, v_{12} = 1$$
$$(X - I\lambda_2)|v_2\rangle = \mathbf{0} \implies \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} v_{21} \\ v_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \implies v_{21} = 1, v_{22} = -1$$

Hence we find that $|v_1\rangle = |0\rangle + |1\rangle$, $|v_2\rangle = |0\rangle - |1\rangle$. Normalizing these eigenvectors (Also see Exercise 2.7), we divide by $\big\||v_1\rangle\big\| = \big\||v_2\rangle\big\| = \sqrt{2}$, giving us:

$$|v_1\rangle = |+\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}, \quad |v_2\rangle = |-\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}.$$

The diagonal representation of X is then given by:

$$X = \lambda_1 \, |v_1\rangle\langle v_1| + \lambda_2 \, |v_2\rangle\langle v_2| = |+\rangle\langle +| - |-\rangle\langle -|$$

We do the same for Y. Solving for the eigenvalues:

$$\det(A - I\lambda) = 0 \implies \det \begin{bmatrix} -\lambda & -i \\ i & -\lambda \end{bmatrix} = 0 \implies \lambda^2 - 1 = 0$$

From which we obtain $\lambda_1 = 1, \lambda_2 = -1$. Solving for the eigenvectors, we then have:

$$(Y - I\lambda_1)|v_1\rangle = \mathbf{0} \implies \begin{bmatrix} -1 & -i \\ i & -1 \end{bmatrix} \begin{bmatrix} v_{11} \\ v_{12} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \implies v_{11} = 1, v_{12} = i$$

$$(Y - I\lambda_2)|v_2\rangle = \mathbf{0} \implies \begin{bmatrix} 1 & -i \\ i & 1 \end{bmatrix} \begin{bmatrix} v_{21} \\ v_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \implies v_{21} = 1, v_{22} = -i$$

We therefore have $|v_1\rangle = |0\rangle + i|1\rangle, |v_2\rangle = |0\rangle - i|1\rangle$. Normalizing by dividing by $\||v_1\rangle\| = \||v_2\rangle\|$, we obtain that:

$$|v_1\rangle = |y_+\rangle = \frac{|0\rangle + i|1\rangle}{\sqrt{2}}, \quad |v_2\rangle = |y_-\rangle = \frac{|0\rangle - i|1\rangle}{\sqrt{2}}.$$

The diagonal representation of Y is then given by:

$$Y = |y_+\rangle\langle y_+| - |y_-\rangle\langle y_-|$$

For Z, the process is again the same. We give the results and omit the details:

$$\lambda_1 = 1, |v_1\rangle = |0\rangle \quad \lambda_2 = -1, |v_2\rangle = |1\rangle$$

$$Z = |0\rangle\langle 0| - |1\rangle\langle 1|$$

□

Exercise 2.12

Prove that the matrix

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

is not diagonalizable.

Solution

Concepts Involved: Eigenvalues, Eigenvectors, Diagonalization

---

Solving for the eigenvalues of the matrix, we have:

$$\det \begin{bmatrix} 1-\lambda & 0 \\ 1 & 1-\lambda \end{bmatrix} = 0 \implies (1-\lambda)^2 = 0 \implies \lambda_1, \lambda_2 = 1$$

But since the eigenvalue 1 is degenerate, the matrix only has one eigenvector; it therefore cannot be diagonalized. $\square$

Exercise 2.13

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

Solution

Concepts Involved: Adjoints

We observe that:

$$\begin{aligned}((\,|w\rangle\langle v|)^\dagger|x\rangle, |y\rangle) = (|x\rangle, (\,|w\rangle\langle v|)|y\rangle) = (|x\rangle, \langle v|y\rangle\,|w\rangle) &= \langle x|\,\langle v|y\rangle\,|w\rangle \\ &= \langle x|w\rangle\,\langle v|y\rangle \\ &= \langle x|w\rangle\,(|v\rangle, |y\rangle) \\ &= (\langle x|w\rangle^*\,|v\rangle, |y\rangle) \\ &= (\langle w|x\rangle\,|v\rangle, |y\rangle) \\ &= ((\,|v\rangle\langle w|)|x\rangle, |y\rangle)\end{aligned}$$

Where in the third-to last equality we use the conjugate linearity in the first argument (see Exercise 2.6) and in the second-to last equality we use that $\langle a|b\rangle^* = \langle b|a\rangle$. Comparing the first and last expressions, we conclude that $(\,|w\rangle\langle v|)^\dagger = |v\rangle\langle w|$. $\square$

Exercise 2.14: Anti-linearity of the adjoint

Show that the adjoint operator is anti-linear,

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

Solution

Concepts Involved: Adjoints

We observe that:

$$\begin{aligned}\left(\left(\sum_i a_i A_i\right)^\dagger |a\rangle, |b\rangle\right) = \left(|a\rangle, \sum_i a_i A_i|b\rangle\right) &= \sum_i a_i\,(|a\rangle, A_i|b\rangle) \\ &= \sum_i a_i\left(A_i^\dagger|a\rangle, |b\rangle\right) \\ &= \left(\sum_i a_i^* A_i^\dagger|a\rangle, |b\rangle\right)\end{aligned}$$

where we invoke the definition of the adjoint in the first and third equalities, the linearity in the second argument in the second equality, and the conjugate linearity in the first argument in the last equality. The claim is proven by comparing the first and last expressions. □

Exercise 2.15

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

Solution

Concepts Involved: Adjoints

Applying the definition of the Adjoint twice (and using the conjugate symmetry of the inner product) we have

$$((A^\dagger)^\dagger|a\rangle, |b\rangle) = (|a\rangle, A^\dagger|b\rangle) = (A^\dagger|b\rangle, |a\rangle)^* = (|b\rangle, A|a\rangle)^* = \left((A|a\rangle, |b\rangle)^*\right)^* = (A|a\rangle, |b\rangle).$$

The claim follows by comparison of the first and last expressions. □

Exercise 2.16

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

Solution

Concepts Involved: Projectors

Let $|1\rangle, \dots, |k\rangle$ be an orthonormal basis for the subspace W of V. Then, using the definition of the projector onto W, we have:

$$P^2 = P \cdot P = \left(\sum_{i=1}^k |i\rangle\langle i|\right)\left(\sum_{i'=1}^k |i'\rangle\langle i'|\right) = \sum_{i=1}^k\sum_{i'=1}^k |i\rangle\langle i|i'\rangle\langle i'| = \sum_{i=1}^k\sum_{i'=1}^k |i\rangle\delta_{ii'}\langle i'| = \sum_{i=1}^k |i\rangle\langle i| = P$$

where in the fourth/fifth equality we use the orthonormality of the basis to collapse the double sum. □

Exercise 2.17

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

Solution

Concepts Involved: Hermitian Operators, Normal Operators, Spectral Decomposition

$\boxed{\implies}$ Let A be a Normal and Hermitian matrix. Then, it has a diagonal representation $A = \sum_i \lambda_i |i\rangle\langle i|$ where $|i\rangle$ is an orthonormal basis for V and each $|i\rangle$ is an eigenvector of A with eigenvalue λ_i. By the

Hermicity of A, we have $A = A^\dagger$. Therefore, we have:

$$A^\dagger = \left(\sum_i \lambda_i |i\rangle\langle i|\right)^\dagger = \sum_i \lambda_i^* |i\rangle\langle i| = A = \sum_i \lambda_i |i\rangle\langle i|$$

where we use the results of Exercises 2.13 and 2.14 in the second equality. Comparing the third and last expressions, we have $\lambda_i = \lambda_i *$ and hence the eigenvalues are real.

$\boxed{\Longleftarrow}$ Let A be a Normal matrix with real eigenvalues. Then, A has diagonal representation $A = \sum_i \lambda_i |i\rangle\langle i|$ where λ_i are all real. We therefore have:

$$A^\dagger = \left(\sum_i \lambda_i |i\rangle\langle i|\right)^\dagger = \lambda_i^* |i\rangle\langle i| = \sum_i \lambda_i |i\rangle\langle i| = A$$

where in the third equality we use that $\lambda_i^* = \lambda_i$. We conclude that A is Hermitian. □

Exercise 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.

Solution

Concepts Involved: Unitary Operators, Spectral Decomposition

Let U be a unitary matrix. It is then normal as $U^\dagger U = U^\dagger U = I$. It therefore has spectral decomposition $U = \sum_i \lambda_i |i\rangle\langle i|$ where $|i\rangle$ is an orthonormal basis of V, and $|i\rangle$ are the eigenvectors of U with eigenvalues λ_i. We then have:

$$\begin{aligned}
UU^\dagger = I &\implies \left(\sum_i \lambda_i |i\rangle\langle i|\right)\left(\sum_{i'} \lambda_{i'} |i'\rangle\langle i'|\right)^\dagger = I \\
&\implies \left(\sum_i \lambda_i |i\rangle\langle i|\right)\left(\sum_{i'} \lambda_{i'}^* |i'\rangle\langle i'|\right) = I \\
&\implies \sum_i \sum_{i'} \lambda_i \lambda_{i'} |i\rangle \langle i|i'\rangle \langle i'| = I \\
&\implies \sum_i \sum_{i'} \lambda_i \lambda_{i'}^* |i\rangle \delta_{ii'} \langle i'| = I \\
&\implies \sum_i \lambda_i \lambda_{i'} |i\rangle\langle i| = I \\
&\implies \sum_i |\lambda_i|^2 |i\rangle\langle i| = \sum_i 1 |i\rangle\langle i|
\end{aligned}$$

From which we obtain that $|\lambda_i|^2 = 1$, and hence $|\lambda_i| = 1$, proving the claim. □

Exercise 2.19: Pauli matrices: Hermitian and unitary

Show that the Pauli matrices are Hermitian and unitary.

Solution

Concepts Involved: Hermitian Operators, Unitary Operators, Pauli Operators

We check I, X, Y, Z in turn.

$$I^\dagger = \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^T\right)^* = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^* = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

$$I^\dagger I = II = I$$

$$X^\dagger = \left(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}^T\right)^* = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}^* = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = X$$

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

$$Y^\dagger = \left(\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}^T\right)^* = \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix}^* = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} = Y$$

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

$$Z^\dagger = \left(\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}^T\right)^* = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}^* = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = Z$$

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

□

Exercise 2.20: Basis changes

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

Solution

Concepts Involved: Matrix Representations of Operators, Completeness Relation

Using the completeness relation twice, we get:

$$\begin{aligned}
A'_{ij} = \langle v_i|A|v_j\rangle = \langle v_i|IAI|v_j\rangle = \langle v_i| \left(\sum_{i'} |w_{i'}\rangle\langle w_{i'}|\right) A \left(\sum_{j'} |w_{j'}\rangle\langle w_{j'}|\right) |v_j\rangle \\
= \sum_{i'}\sum_{j'}\langle v_i|w_{i'}\rangle\langle w_{i'}|A|w_{j'}\rangle\langle w_{j'}|v_j\rangle \\
= \sum_{i'}\sum_{j'}\langle v_i|w_{i'}\rangle A''_{ij}\langle w_{j'}|v_j\rangle
\end{aligned}$$

□

Exercise 2.21

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

Solution

Concepts Involved: Hermitian Operators, Spectral Decomposition

Note that the converse of Theorem 2.1 does not hold if we replace "normal" with "Hermitian". Diagonalizability does not imply Hermicity, with a concrete example being $S = |0\rangle\langle 0| + i\,|1\rangle\langle 1|$. So, we just prove the forwards direction.

We proceed by induction on the dimension d of V. The $d = 1$ case is trivial as M is already diagonal in any representation in this case. Let λ be an eigenvalue of M, P the projector onto the λ subspace, and Q the projector onto the orthogonal complement. Then $M = (P+Q)M(P+Q) = PMP + QMP + PMQ + QMQ$. Obviously $PMP = \lambda P$. Furthermore, $QMP = 0$, as M takes the subspace P into itself. We claim that $PMQ = 0$ also. To see this, we recognize that $(PMQ)^\dagger = Q^\dagger M^\dagger P^\dagger = QMP = 0.$ and hence $PMQ = 0$. Thus $M = PMP + QMQ$. QMQ is normal, as $(QMQ)^\dagger = Q^\dagger M^\dagger Q^\dagger = QMQ$ (and Hermiticity implies that the operator is normal). By induction, QMQ is diagonal with respect to some orthonormal basis for the subspace Q, and PMP is already diagonal with respect to some orthonormal basis for P. It follows that $M = PMP + QMQ$ is diagonal with respect to some orthonormal basis for the total vector space. □

Exercise 2.22

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

Solution

Concepts Involved: Eigenvalues, Eigenvectors, Hermitian Operators

Let A be a Hermitian operator, and let $|v_1\rangle, |v_2\rangle$ be two eigenvectors of A with corresponding eigenvalues λ_1, λ_2 such that $\lambda_1 \neq \lambda_2$. We then have:

$$\langle v_1|A|v_2\rangle = \langle v_1|\lambda_2|v_2\rangle = \lambda_2\langle v_1|v_2\rangle$$
$$\langle v_1|A|v_2\rangle = \langle v_1|A^\dagger|v_2\rangle = \langle v_1|\lambda_1|v_2\rangle = \lambda_1\langle v_1|v_2\rangle$$

where we use the Hermiticity of A in the second line. Substracting the first line from the second, we have:

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

Since $\lambda_1 \neq \lambda_2$ by assumption, the only way this equality is satisfied is if $\langle v_1|v_2\rangle = 0$. Hence, $|v_1\rangle, |v_2\rangle$ are orthogonal. □

Exercise 2.23

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

Solution

Concepts Involved: Linear Algebra, Eigenvalues, Eigenvectors, Projectors.

Let P be a projector, and $|v\rangle$ be an eigenvector of P with corresponding eigenvalue λ. From Exercise 2.16, we have $P^2 = P$, and using this fact, we observe:

$$P|v\rangle = \lambda|v\rangle$$
$$P|v\rangle = P^2|v\rangle = PP|v\rangle = P\lambda|v\rangle = \lambda P|v\rangle = \lambda^2|v\rangle.$$

Subtracting the first line from the second, we get:

$$\mathbf{0} = (\lambda^2 - \lambda)|v\rangle = \lambda(\lambda - 1)|v\rangle.$$

Since $|v\rangle$ is not the zero vector, we therefore obtain that either $\lambda = 0$ or $\lambda = 1$. □

Exercise 2.24: Hermiticity of positive operator

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

Solution

Concepts Involved: Hermitian Operators, Positive Operators

Let A be an operator. We first make the observation that we can write A as:

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

So let $B = \frac{A+A^\dagger}{2}$ and $C = \frac{A-A^\dagger}{2i}$. B and C are Hermitian, as:

$$B^\dagger = \left(\frac{A + A^\dagger}{2}\right)^\dagger = \frac{A^\dagger + (A^\dagger)^\dagger}{2} = \frac{A^\dagger + A}{2} = B$$
$$C^\dagger = \left(\frac{A - A^\dagger}{2i}\right)^\dagger = \frac{A^\dagger - (A^\dagger)^\dagger}{-2i} = \frac{A - A^\dagger}{2i} = C$$

so we have hence proven that we can write $A = B + iC$ for hermitian B, C for any operator A. Now, assume that A is positive. We then have for any vector $|v\rangle$:

$$\langle v|A|v\rangle \geq 0.$$

Using the identity derived above, we have:

$$\langle v|B|v\rangle + i\langle v|C|v\rangle \geq 0.$$

The positivity forces $C = 0$. Therefore, $A = B$ and hence A is Hermitian. □

Exercise 2.25

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

Solution

Concepts Involved: Adjoints, Positive Operators

Let A be an operator. Let $|v\rangle$ be an arbitrary vector, and then we then have:

$$\Big(|v\rangle, A^\dagger A|v\rangle\Big) = \Big((A^\dagger)^\dagger|v\rangle, A|v\rangle\Big) = \big(A|v\rangle, A|v\rangle\big).$$

By the property of inner products, the expression must be greater than zero. □

Exercise 2.26

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

Solution

Concepts Involved: Tensor Products, Kronecker Products

Using the definition of the tensor product, we have:

$$|\psi\rangle^{\otimes 2} = \frac{|0\rangle|0\rangle + |0\rangle|1\rangle + |1\rangle|0\rangle + |1\rangle|1\rangle}{2} \cong \begin{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \\ \frac{1}{\sqrt{2}} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$$

$$\begin{aligned} |\psi\rangle^{\otimes 3} &= \frac{|0\rangle|0\rangle|0\rangle + |0\rangle|0\rangle|1\rangle + |0\rangle|1\rangle|0\rangle + |0\rangle|1\rangle|1\rangle + |1\rangle|0\rangle|0\rangle + |1\rangle|0\rangle|1\rangle + |1\rangle|1\rangle|0\rangle + |1\rangle|1\rangle|1\rangle}{2\sqrt{2}} \\ &= |\psi\rangle \otimes |\psi\rangle^{\otimes 2} \cong \begin{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} \\ \frac{1}{\sqrt{2}} \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} \end{bmatrix} = \begin{bmatrix} \frac{1}{2\sqrt{2}} \\ \frac{1}{2\sqrt{2}} \\ \frac{1}{2\sqrt{2}} \\ \frac{1}{2\sqrt{2}} \\ \frac{1}{2\sqrt{2}} \\ \frac{1}{2\sqrt{2}} \\ \frac{1}{2\sqrt{2}} \\ \frac{1}{2\sqrt{2}} \end{bmatrix} \end{aligned}$$

□

Exercise 2.27

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

Solution

Concepts Involved: Tensor Products, Kronecker Products, Pauli Operators

---

Using the Kronecker product, we have:

(a)

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

(b)

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

(c)

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

Comparing (b) and (c), we conclude that the tensor product is not commutative. □

Exercise 2.28

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

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

Solution

Concepts Involved: Adjoints, Tensor Products, Kronecker Products

Using the Kronecker product representaton of $a \otimes B$, we have:

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

$$(A \otimes B)^T = \begin{bmatrix} A_{11}B & A_{12}B & \dots & A_{1n}B \\ A_{21}B & A_{22}B & \dots & A_{2n}B \\ \vdots & \vdots & \vdots & \vdots \\ A_{m1}B & A_{m2}B & \dots & A_{mn}B \end{bmatrix}^T = \begin{bmatrix} A_{11}B^T & A_{21}B^T & \dots & A_{n1}B^T \\ A_{12}B^T & A_{22}B & \dots & A_{n2}B^T \\ \vdots & \vdots & \vdots & \vdots \\ A_{1m}B^T & A_{2m}B^T & \dots & A_{nm}B^T \end{bmatrix} = A^T \otimes B^T.$$

The relation for the distributivity of the hermitian conjugate over the tensor product then follows from the former two relations:

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

$\square$

Exercise 2.29

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

Solution

Concepts Involved: Unitary Operators, Tensor Products

Suppose A, B are unitary. Then, $A^\dagger A = I$ and $B^\dagger B = I$. Using the result of the Exercise 2.28, we then have:

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

$\square$

Exercise 2.30

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

Solution

Concepts Involved: Hermitian Operators, Tensor Products

Suppose A, B are Hermitian. Then, $A^\dagger = A$ and $B^\dagger = B$. Then, using the result of Exercise 2.28, we have:

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

$\square$

Exercise 2.31

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

Solution

Concepts Involved: Positive Operators

Suppose A, B are positive operators. We then have $\langle v|A|v\rangle \geq 0$ and $\langle w|B|w\rangle \geq 0$. Therefore, for any $|v\rangle \otimes |w\rangle$:

$$\big(|v\rangle \otimes |w\rangle, A \otimes B(|v\rangle \otimes |w\rangle)\big) = \langle v|A|v\rangle\langle w|B|w\rangle \geq 0$$

$\square$

Exercise 2.32

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

Solution

Concepts Involved: Projectors

Let P_1, P_2 be projectors. We then have $P_1^2 = P_1$ and $P_2^2 = P_2$ by Exercise 2.16. Therefore:

$$(P_1 \otimes P_2)^2 = (P_1 \otimes P_2)(P_1 \otimes P_2) = P_1^2 \otimes P_2 = P_1 \otimes P_2$$

so $P_1 \otimes P_2$ is a projector. $\square$

Exercise 2.33

The Hadamard operator on one qubit may be written as

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

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 $H^{\otimes 2}$

Solution

Concepts Involved: Linear algebra, Matrix Representation of Operators, Outer Products.

Looking at the form of the Hadamard operator on one qubit, we observe that:

$$H = \frac{1}{\sqrt{2}} \left[\, |0\rangle\langle 0| + |0\rangle\langle 1| + |1\rangle\langle 0| - |1\rangle\langle 1| \right]$$

Hence:

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

Where x, y run over 0 and 1. Taking the n-fold tensor product of this expression, we get:

$$\begin{aligned} H^{\otimes} &= \frac{1}{\sqrt{2}} \sum_{x,y} (-1)^{x \cdot y} |x\rangle\langle y| \otimes \frac{1}{\sqrt{2}} \sum_{x,y} (-1)^{x \cdot y} |x\rangle\langle y| \otimes \ldots \otimes \frac{1}{\sqrt{2}} \sum_{x,y} (-1)^{x \cdot y} |x\rangle\langle y| \\ &= \frac{1}{\sqrt{2^n}} \sum_{\mathbf{x},\mathbf{y}} (-1)^{\mathbf{x} \cdot \mathbf{y}} |\mathbf{x}\rangle\langle \mathbf{y}| \end{aligned}$$

Where $\mathbf{x}, \mathbf{y}$ are length n-binary strings. This proves the claim.

Now explicitly writing $H^{\otimes 2}$, we have:

$$H^{\otimes 2} = \frac{1}{\sqrt{2^2}} \sum_{\mathbf{x},\mathbf{y}} (-1)^{(\mathbf{x}\cdot\mathbf{y})} |\mathbf{x}\rangle\langle\mathbf{y}|$$
$$\cong \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

Note that here, $\mathbf{x}, \mathbf{y}$ are binary length 2 strings. The sum goes through all pairwise combinations of $\mathbf{x}, \mathbf{y} \in \{00, 01, 10, 11\}$. □

Remark: Sylvester's Construction gives an interesting recursive construction of Hadamard matrices. See `https://en.wikipedia.org/wiki/Hadamard_matrix`. Discussion on interesting (related) open problem concerning the maximal determinant of matrices consisting of entries of 1 and -1 can be found here `https://en.wikipedia.org/wiki/Hadamard%27s_maximal_determinant_problem`.

Exercise 2.34

Find the square root and logarithm of the matrix

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

Solution

Concepts Involved: Spectral Decomposition, Operator Functions

---

We begin by diagonalizing the matrix (which we call A) as to be able to apply the definition of operator functions. By inspection, A is Hermitian as it is equal to its conjugate transpose, so the spectral decomposition exists. Solving for the eigenvalues, we consider the characterstic equation:

$$\det(A - \lambda I) = 0 \implies \det \begin{bmatrix} 4-\lambda & 3 \\ 3 & 4-\lambda \end{bmatrix} = 0 \implies (4-\lambda)^2 - 9 = 0 \implies \lambda^2 - 8\lambda + 7 = 0$$

Using the quadratic equation, we get $\lambda_1 = 1, \lambda_2 = 7$. Using this to find the eigenvectors of the matrix, we have:

$$\begin{bmatrix} 4-1 & 3 \\ 3 & 4-1 \end{bmatrix} \begin{bmatrix} v_{11} \\ v_{12} \end{bmatrix} = \mathbf{0} \implies v_{11} = 1, v_{12} = -1$$

$$\begin{bmatrix} 4-7 & 3 \\ 3 & 4-7 \end{bmatrix} \begin{bmatrix} v_{21} \\ v_{22} \end{bmatrix} = \mathbf{0} \implies v_{21} = 1, v_{22} = 1$$

Hence our normalized eigenvectors are:

$$|v_1\rangle = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \end{bmatrix}, \quad |v_2\rangle = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

Therefore the spectral composition of the matrix is given by:

$$A = 1\,|v_1\rangle\langle v_1| + 7\,|v_2\rangle\langle v_2|$$

Calculating the square root of A, we then have:

$$\sqrt{A} = \sqrt{1}\,|v_1\rangle\langle v_1| + \sqrt{7}\,|v_2\rangle\langle v_2| = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} + \frac{\sqrt{7}}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1+\sqrt{7} & -1+\sqrt{7} \\ -1+\sqrt{7} & 1+\sqrt{7} \end{bmatrix}.$$

Calculating the logarithm of A, we have:

$$\log(A) = \log(1)\,|v_1\rangle\langle v_1| + \log(7)\,|v_2\rangle\langle v_2| = \frac{\log(7)}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

□

Exercise 2.35: Exponential of Pauli matrices

Let $\mathbf{v}$ be any real, three-dimensional unit vector and θ a real number. Prove that

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

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

Solution

Concepts Involved: Spectral Decomposition, Operator Functions, Pauki Operators
Recall that $\sigma_1 \equiv X, \sigma_2 \equiv Y$, and $\sigma_3 \equiv Z$.

First, we compute $\mathbf{v}\cdot\boldsymbol{\sigma}$ in matrix form:

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

In order to compute the complex exponential of this matrix, we will want to find its spectral decomposition.

Using the characterstic equation to find the eigenvalues, we have:

$$\begin{aligned}\det(\mathbf{v}\cdot\boldsymbol{\sigma} - I\lambda) = 0 &\implies \det\begin{bmatrix} v_3-\lambda & v_1-iv_2 \\ v_1+iv_2 & -v_3-\lambda\end{bmatrix} = 0 \\ &\implies (v_3-\lambda)(-v_3-\lambda) - (v_1-iv_2)(v_1+iv_2) = 0 \\ &\implies \lambda^2 - v_3^2 - v_1^2 - v_2^2 = \lambda^2 - (v_1^2+v_2^2+v_3^2) = 0 \\ &\implies \lambda^2 - 1 = 0 \\ &\implies \lambda_1 = 1, \lambda_2 = -1\end{aligned}$$

where in the second-to-last implication we use the fact that $\mathbf{v}$ is a unit vector. Letting $|v_1\rangle, |v_2\rangle$ be the associated eigenvectors, $\mathbf{v}\cdot\boldsymbol{\sigma}$ has spectral decomposition:

$$\mathbf{v}\cdot\boldsymbol{\sigma} = |v_1\rangle\langle v_1| - |v_2\rangle\langle v_2|$$

Applying the complex exponentiation operator, we then have:

$$\exp(i\theta\mathbf{v}\cdot\boldsymbol{\sigma}) = \exp(i\theta)\,|v_1\rangle\langle v_1| + \exp(-i\theta)\,|v_2\rangle\langle v_2|\,.$$

Using Euler's formula, we then have:

$$\begin{aligned}\exp(i\theta\mathbf{v}\cdot\boldsymbol{\sigma}) &= (\cos\theta + i\sin\theta)\,|v_1\rangle\langle v_1| + (\cos\theta - i\sin\theta)\,|v_2\rangle\langle v_2| \\ &= \cos(\theta)\left(|v_1\rangle\langle v_1| + |v_2\rangle\langle v_2|\right) + i\sin(\theta)\left(|v_1\rangle\langle v_1| - |v_2\rangle\langle v_2|\right) \\ &= \cos(\theta)I + i\sin(\theta)\mathbf{v}\cdot\boldsymbol{\sigma}.\end{aligned}$$

Where in the last line we use the completeness relation and the spectral decomposition. □

Exercise 2.36

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

Solution

Concepts Involved: Trace, Pauli Operators

We have:

$$\begin{aligned}\mathrm{tr}(X) &= \mathrm{tr}\begin{bmatrix}0 & 1 \\ 1 & 0\end{bmatrix} = 0 + 0 = 0 \\ \mathrm{tr}(Y) &= \mathrm{tr}\begin{bmatrix}0 & -i \\ i & 0\end{bmatrix} = 0 + 0 = 0 \\ \mathrm{tr}(Z) &= \mathrm{tr}\begin{bmatrix}1 & 0 \\ 0 & -1\end{bmatrix} = 1 - 1 = 0\end{aligned}$$

□

Exercise 2.37: Cyclic property of the trace

If A and B are two linear operators show that

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

Solution

Concepts Involved: Trace

Let A, B be linear operators. Then, $C = AB$ has matrix representation with entries $C_{ij} = \sum_k A_{ik}B_{kj}$ and $D = BA$ has matrix representation with entries $D_{ij} = \sum_k B_{ik}A_{kj}$. We then have:

$$\operatorname{tr}(AB) = \operatorname{tr}(C) = \sum_i C_{ii} = \sum_i \sum_k A_{ik}B_{ki} = \sum_k \sum_i B_{ki}A_{ik} = \sum_k D_{kk} = \operatorname{tr}(D) = \operatorname{tr}(BA)$$

□

Exercise 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

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

Solution

Concepts Involved: Trace

From the definition of trace, we have:

$$\operatorname{tr}(A + B) = \sum_i (A + B)_{ii} = \sum_i A_{ii} + B_{ii} = \sum_i A_{ii} + \sum_i B_{ii} = \operatorname{tr}(A) + \operatorname{tr}(B)$$

$$\operatorname{tr}(zA) = \sum_i (zA)_{ii} = z\sum_i A_{ii} = z\operatorname{tr}(A)$$

□

Exercise 2.39: The Hilbert-Schmidt inner product on operators

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

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

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

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

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

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

Solution

Concepts Involved: Trace, Inner Products, Hermitian Operators, Bases

(1) We show that $(\cdot, \cdot)$ satisfies the three properties of an inner product. Showing that it is linear in the second argument, we have:

$$\left(A, \sum_i \lambda_i B_i\right) = \mathrm{tr}\left(A \sum_i \lambda_i B_i\right) = \sum_i \lambda_i \,\mathrm{tr}(AB_i) = \sum_i \lambda_i (A, B_i)$$

where in the second to last equality we use the result of Exercise 2.38. To see that it is conjugate-symmetric, we have:

$$(A, B) = \mathrm{tr}\Big(A^\dagger B\Big) = \mathrm{tr}\Big((B^\dagger A)^\dagger\Big) = \mathrm{tr}\Big(B^\dagger A\Big)^* = (B, A)^*$$

Finally, to show positive definitness, we have:

$$(A, A) = \mathrm{tr}\Big(A^\dagger A\Big) = \sum_i \sum_k A^\dagger_{ik} A_{ki} = \sum_i \sum_k A^*_{ki} A_{ki} = \sum_i \sum_k |A_{ki}|^2 \geq 0$$

so we conclude that $(\cdot, \cdot)$ is an inner product function.

(2) Suppose V has d dimensions. Then, the elements of L_V which consist of linear operators $A : V \mapsto V$ have representations as $d \times d$ matrices. There are d^2 such linearly independent matrices (take the matrices with 1 in one of the d^2 entries and 0 elsewhere), and we conclude that L_V has d^2 linearly independent vectors and hence dimension d^2.

(3) As discussed in the previous part of the question, one possible basis for this vector space would be $|v_i\rangle\langle v_j|$ where $|v_k\rangle$ form an orthonormal basis of V with $i, j \in \{1, \ldots d\}$. These of course are just matrices with 1 in one entry and 0 elsewhere. It is easy to see that this is a basis as for any

$A \in L_V$ we can write $A = \sum_{ij} \lambda_{ij} |v_i\rangle\langle v_j|$. We can verify that these are orthonormal; suppose $|v_{i_1}\rangle\langle v_{j_1}| \neq |v_{i_2}\rangle\langle v_{j_2}|$. Then, we have:

$$\begin{aligned}\big(\,|v_{i_1}\rangle\langle v_{j_1}|\,,\,|v_{i_2}\rangle\langle v_{j_2}|\big) &= \mathrm{tr}\Big((\,|v_{i_1}\rangle\langle v_{j_1}|)^\dagger\,|v_{i_2}\rangle\langle v_{j_2}|\Big)\\ &= \mathrm{tr}\big(|v_{j_1}\rangle\langle v_{i_1}|v_{i_2}\rangle\langle v_{j_2}|\big)\end{aligned}$$

If $|v_{i_1}\rangle \neq |v_{i_2}\rangle$, then the above expression reduces to $\mathrm{tr}(0) = 0$. If $|v_{i_1}\rangle = |v_{i_2}\rangle$, then it follows that $|v_{j_1}\rangle \neq |v_{j_2}\rangle$ (else this would contradict $|v_{i_1}\rangle\langle v_{j_1}| \neq |v_{i_2}\rangle\langle v_{j_2}|$) and in this case we have:

$$\begin{aligned}\big(\,|v_{i_1}\rangle\langle v_{j_1}|\,,\,|v_{i_2}\rangle\langle v_{j_2}|\big) &= \mathrm{tr}\big(|v_{j_1}\rangle\langle v_{i_1}|v_{i_2}\rangle\langle v_{j_2}|\big)\\ &= \mathrm{tr}\big(|v_{j_1}\rangle\langle v_{j_2}|\big)\\ &= 0\end{aligned}$$

So we therefore have the inner product of two non-identical elements in the basis is zero. Furthermore, we have:

$$\big(\,|v_{i_1}\rangle\langle v_{j_1}|\,,\,|v_{i_1}\rangle\langle v_{j_1}|\big) = \mathrm{tr}\big(\,|v_{i_1}\rangle\langle v_{j_1}|\;|v_{i_1}\rangle\langle v_{j_1}|\big) = \mathrm{tr}\big(|v_{i_1}\rangle\langle v_{i_1}|\big) = 1$$

so we confirm that this basis is orthonormal. However, evidently this basis is *not* Hermitian as if $i \neq j$, then $(\,|v_i\rangle\langle v_j|)^\dagger = |v_j\rangle\langle v_i| \neq |v_i\rangle\langle v_j|$. To fix this, we can modify our basis slightly. We keep the diagonal entries as is (as these are indeed Hermitian!) but for the off-diagonals, we replace every pair of basis vectors $|v_i\rangle\langle v_j|$, $|v_j\rangle\langle v_i|$ (for $i > j$) with:

$$\frac{|v_i\rangle\langle v_j| + |v_j\rangle\langle v_i|}{\sqrt{2}}, \quad i\frac{|v_i\rangle\langle v_j| - |v_j\rangle\langle v_i|}{\sqrt{2}}.$$

A quick verification shows that these are indeed Hermitian:

$$\begin{aligned}\left(\frac{|v_i\rangle\langle v_j| + |v_j\rangle\langle v_i|}{\sqrt{2}}\right)^\dagger &= \frac{(\,|v_i\rangle\langle v_j|)^\dagger + (\,|v_j\rangle\langle v_i|)^\dagger}{\sqrt{2}} = \frac{|v_i\rangle\langle v_j| + |v_j\rangle\langle v_i|}{\sqrt{2}}\\ \left(i\frac{|v_i\rangle\langle v_j| - |v_j\rangle\langle v_i|}{\sqrt{2}}\right)^\dagger &= -i\frac{(\,|v_i\rangle\langle v_j|)^\dagger - (\,|v_j\rangle\langle v_i|)^\dagger}{\sqrt{2}} = i\frac{|v_i\rangle\langle v_j| - |v_j\rangle\langle v_i|}{\sqrt{2}}\end{aligned}$$

It now suffices to show that these new vectors (plus the diagonals) form a basis and are orthonormal. To see that these form a basis, observe that:

$$\begin{aligned}\frac{1}{\sqrt{2}}\frac{|v_i\rangle\langle v_j| + |v_j\rangle\langle v_i|}{\sqrt{2}} - \frac{i}{\sqrt{2}}\left(i\frac{|v_i\rangle\langle v_j| - |v_j\rangle\langle v_i|}{\sqrt{2}}\right) &= |v_i\rangle\langle v_j|\\ \frac{1}{\sqrt{2}}\frac{|v_i\rangle\langle v_j| + |v_j\rangle\langle v_i|}{\sqrt{2}} + \frac{i}{\sqrt{2}}\left(i\frac{|v_i\rangle\langle v_j| - |v_j\rangle\langle v_i|}{\sqrt{2}}\right) &= |v_j\rangle\langle v_i|\end{aligned}$$

and since we know that $|v_i\rangle\langle v_j|$ for all $i, j \in \{1, \ldots d\}$ form a basis, this newly defined set of vectors must be a basis as well. Furthermore, since the new basis vectors are constructed from orthogonal $|v_i\rangle\langle v_j|$, the newly defined vectors will be orthogonal to each other if $i_1, j_1 \neq i_2, j_2$. The only things left to check is that for any choice of i, j that:

$$\frac{|v_i\rangle\langle v_j| + |v_j\rangle\langle v_i|}{\sqrt{2}} \text{ and } i\frac{|v_i\rangle\langle v_j| - |v_j\rangle\langle v_i|}{\sqrt{2}}.$$

are orthogonal, and that these vectors are normalized. Checking the orthogonality, we have:

$$\begin{aligned}
\left(\frac{|v_i\rangle\langle v_j| + |v_j\rangle\langle v_i|}{\sqrt{2}}, i\frac{|v_i\rangle\langle v_j| - |v_j\rangle\langle v_i|}{\sqrt{2}}\right) &= \operatorname{tr}\left(\left(\frac{|v_i\rangle\langle v_j| + |v_j\rangle\langle v_i|}{\sqrt{2}}\right)\left(i\frac{|v_i\rangle\langle v_j| - |v_j\rangle\langle v_i|}{\sqrt{2}}\right)\right) \\
&= \frac{i}{2}\operatorname{tr}\big(|v_j\rangle\langle v_j| - |v_i\rangle\langle v_i|\big) \\
&= 0.
\end{aligned}$$

And checking the normalization, we have:

$$\begin{aligned}
\left(\frac{|v_i\rangle\langle v_j| + |v_j\rangle\langle v_i|}{\sqrt{2}}, \frac{|v_i\rangle\langle v_j| + |v_j\rangle\langle v_i|}{\sqrt{2}}\right) &= \operatorname{tr}\left(\left(\frac{|v_i\rangle\langle v_j| + |v_j\rangle\langle v_i|}{\sqrt{2}}\right)\left(\frac{|v_i\rangle\langle v_j| + |v_j\rangle\langle v_i|}{\sqrt{2}}\right)\right) \\
&= \frac{1}{2}\operatorname{tr}\big(|v_i\rangle\langle v_i| + |v_j\rangle\langle v_j|\big) \\
&= 1
\end{aligned}$$

$$\begin{aligned}
\left(i\frac{|v_i\rangle\langle v_j| - |v_j\rangle\langle v_i|}{\sqrt{2}}, i\frac{|v_i\rangle\langle v_j| - |v_j\rangle\langle v_i|}{\sqrt{2}}\right) &= \operatorname{tr}\left(\left(i\frac{|v_i\rangle\langle v_j| - |v_j\rangle\langle v_i|}{\sqrt{2}}\right)\left(i\frac{|v_i\rangle\langle v_j| - |v_j\rangle\langle v_i|}{\sqrt{2}}\right)\right) \\
&= -\frac{1}{2}\operatorname{tr}\big(-|v_i\rangle\langle v_i| - |v_j\rangle\langle v_j|\big) \\
&= 1
\end{aligned}$$

□

Exercise 2.40: Commutation relations for the Pauli matrices

Verify the commutation relations

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

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

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

Solution

Concepts Involved: Commutators, Pauli Operators

We verify the proposed relations via computation in the computational basis:

$$[X,Y] = XY - YX = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} - \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} - \begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix} = 2iZ$$
$$[Y,Z] = YZ - ZY = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} - \begin{bmatrix} 0 & -i \\ -i & 0 \end{bmatrix} = 2iX$$
$$[Z,X] = ZX - XZ = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} - \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = 2iY$$

□

Exercise 2.41: Anti-commutation relations for the Pauli matrices

Verify the anticommutation relations

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

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

$$\sigma_i^2 = I$$

Solution

Concepts Involved: Linear Algebra, Anticommutators, Pauli Operators

---

We again verify the proposed relations via computation in the computational basis:

$$\{X,Y\} = XY + YX = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} + \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} + \begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$
$$\{Y,Z\} = YZ + ZY = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} + \begin{bmatrix} 0 & -i \\ -i & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$
$$\{Z,X\} = ZX + XZ = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

This proves the first claim as $\{A, B\} = AB + BA = BA + AB = \{B, A\}$ and the other 3 relations are equivalent to the ones already proven. Verifying the second claim, we have:

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

□

Remark: Note that we can write this result consicely as $\{\sigma_j, \sigma_k\} = 2\delta_{ij} I$

Exercise 2.42

Verify that

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

Solution

Concepts Involved: Commutators, Anticommutators

By algebraic manipulation we obtain:

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

□

Exercise 2.43

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

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

Solution

Concepts Involved: Commutators, Anticommutators, Pauli Operators

Applying the results of Exercises 2.40, 2.41, and 2.42, we have:

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

□

Exercise 2.44

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

Solution

Concepts Involved: Commutators, Anticommutators

By assumption, we have:

$$\begin{aligned}[A, B] &= AB - BA = 0 \\ \{A, B\} &= AB + BA = 0.\end{aligned}$$

Adding the first line to the second we have:

$$2AB = 0 \implies AB = 0.$$

A^{-1} exists by the invertibility of A, so multiplying by A^{-1} on the left we have:

$$A^{-1}AB = A^{-1}0 \implies IB = 0 \implies B = 0.$$

□

Exercise 2.45

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

Solution

Concepts Involved: Commutators, Adjoints

Using the properties of the adjoint, we have:

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

□

Exercise 2.46

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

Solution

Concepts Involved: Commutators

By the definition of the commutator:

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

□

Exercise 2.47

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

Solution

Concepts Involved: Commutators, Hermitian Operators

Suppose A, B are Hermitian. Using the results of Exercises 2.45 and 2.46, we have:

$$(i[A, B])^\dagger = -i([A, B])^\dagger = -i[B^\dagger, A^\dagger] = i[A^\dagger, B^\dagger] = i[A, B].$$

□

Exercise 2.48

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

Solution

Concepts Involved: Polar Decomposition, Positive Operators, Unitary Operators, Hermitian Operators

If P is a positive matrix, then no calculation is required; $P = IP = PI$ is the polar decomposition (as I is unitary and P is positive). If U is a unitary matrix, then $J = \sqrt{U^\dagger U} = \sqrt{I} = I$ and $K = \sqrt{UU^\dagger} = \sqrt{I} = I$ so the polar decomposition is $U = UI = IU$ (where U is unitary and I is positive). If H is hermitian, we then have:

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

and $K = \sqrt{HH^\dagger} = \sum_i |\lambda_i| |i\rangle\langle i|$ in the same way. Hence the polar decomposition is $H = U \sum_i |\lambda_i| |i\rangle\langle i| = \sum_i |\lambda_i| |i\rangle\langle i| U$. □

Exercise 2.49

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

Solution

Concepts Involved: Polar Decomposition, Outer Products

Let A be a normal matrix. Then, A has spectral decomposition $A = \sum_i \lambda_i |i\rangle\langle i|$. Therefore, we have:

$$A^\dagger A = AA^\dagger = \sum_i \sum_{i'} \lambda_i \lambda_{i'}^* |i\rangle\langle i| |i'\rangle\langle i'| = \sum_i \sum_{i'} \lambda_i \lambda_{i'}^* |i\rangle\langle i'| \delta_{ii'} = \sum_i |\lambda_i|^2 |i\rangle\langle i|$$

We then have:

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

and $K = \sum_i |\lambda_i| |i\rangle\langle i|$ identically. Furthermore, U is unitary, so it also has a spectral decomposition of $\sum_j \mu_j |j\rangle\langle j|$. Hence we have the polar decomposition in the outer product representation as:

$$\begin{aligned}
A &= UJ = KU \\
A &= U \sum_i |\lambda_i| |i\rangle\langle i| \sum_j = \sum_i |\lambda_i| |i\rangle\langle i| \sum_j U \\
A &= \sum_j \sum_i \mu_j |\lambda_i| |j\rangle\langle j|i\rangle\langle i| = \sum_i \sum_j |\lambda_i| \mu_j |i\rangle\langle i|j\rangle\langle j|
\end{aligned}$$

□

Exercise 2.50

Find the left and right polar decompositions of the matrix

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

Solution

Concepts Involved: Polar Decomposition

Let $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$. We start with the left polar decomposition, and hence find $J = \sqrt{A^\dagger A}$. In order to do this, we find the spectral decompositions of $A^\dagger A$ and $AA^\dagger$.

$$\det\Big(A^\dagger A - I\lambda\Big) = 0 \implies \det\left(\begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix}\begin{bmatrix}1 & 0\\ 1 & 1\end{bmatrix} - \begin{bmatrix}\lambda & 0\\ 0 & \lambda\end{bmatrix}\right) = 0 \implies \det\begin{bmatrix}2-\lambda & 1\\ 1 & 1-\lambda\end{bmatrix} = 0$$
$$\implies \lambda^2 - 3\lambda + 1 = 0$$
$$\implies \lambda_1 = \frac{3+\sqrt{5}}{2}, \lambda_2 = \frac{3-\sqrt{5}}{2}$$

Solving for the eigenvectors, we have:

$$\begin{bmatrix}2-\frac{3+\sqrt{5}}{2} & 1\\ 1 & 1-\frac{3+\sqrt{5}}{2}\end{bmatrix}|v_1\rangle = \mathbf{0} \implies |v_1\rangle = \begin{bmatrix}1+\sqrt{5}\\ 2\end{bmatrix}$$
$$\begin{bmatrix}2-\frac{3-\sqrt{5}}{2} & 1\\ 1 & 1-\frac{3-\sqrt{5}}{2}\end{bmatrix}|v_2\rangle = \mathbf{0} \implies |v_2\rangle = \begin{bmatrix}1-\sqrt{5}\\ 2\end{bmatrix}$$

Normalizing, we get:

$$|v_1\rangle = \frac{1}{\sqrt{10+2\sqrt{5}}}\begin{bmatrix}1+\sqrt{5}\\ 2\end{bmatrix}, \quad |v_2\rangle = \frac{1}{\sqrt{10-2\sqrt{5}}}\begin{bmatrix}1-\sqrt{5}\\ 2\end{bmatrix}$$

The spectral decomposition of $A^\dagger A$ is therefore:

$$A^\dagger A = \lambda_1 |v_1\rangle\langle v_1| + \lambda_2 |v_2\rangle\langle v_2|$$

Calculating J, we therefore have:

$$J = \sqrt{A^\dagger A} = \sqrt{\lambda_1}\,|v_1\rangle\langle v_1| + \sqrt{\lambda_2}\,|v_2\rangle\langle v_2| = \frac{1}{\sqrt{5}}\begin{bmatrix}3 & 1\\ 1 & 2\end{bmatrix}$$

The last equality is not completely trivial, but the algebra is tedious so we invite the reader to use a symbolic calculator, as we have. We make the observation that:

$$A = UJ \implies U = AJ^{-1}$$

So calculating J^{-1}, we have:

$$J^{-1} = \frac{1}{\sqrt{\lambda_1}}|v_1\rangle\langle v_1| + \frac{1}{\sqrt{\lambda_2}}|v_2\rangle\langle v_2| = \frac{1}{\sqrt{5}}\begin{bmatrix}2 & -1\\ -1 & 3\end{bmatrix}$$

Where we again have used the help of a symbolic calculator. Calculating U, we then have:

$$U = AJ^{-1} = \begin{bmatrix}1 & 0\\ 1 & 1\end{bmatrix}\frac{1}{\sqrt{5}}\begin{bmatrix}2 & -1\\ -1 & 3\end{bmatrix} = \frac{1}{\sqrt{5}}\begin{bmatrix}2 & -1\\ 1 & 2\end{bmatrix}$$

Hence the left polar decomposition of A is given by:

$$A = UJ = \left(\frac{1}{\sqrt{5}}\begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}\right)\left(\frac{1}{\sqrt{5}}\begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}\right)$$

We next solve for the right polar decomposition. We could repeat the procedure of solving for the spectral decomposition of $AA^\dagger$, but we take a shortcut; since we know the K that satisfies:

$$A = KU$$

will be unique, and U is unitary, we can simply multiply both sides of the above equation on the right by $U^{-1} = U^\dagger$ to obtain K. Hence:

$$K = AU^\dagger = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\frac{1}{\sqrt{5}}\begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix} = \frac{1}{\sqrt{5}}\begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}.$$

Therefore the right polar decomposition of A is given by:

$$A = KU = \left(\frac{1}{\sqrt{5}}\begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}\right)\left(\frac{1}{\sqrt{5}}\begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}\right)$$

□

Exercise 2.51

Verify that the Hadamard gate H is unitary.

Solution

Concepts Involved: Unitary Operators

We observe that:

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

showing that H is indeed unitary. □

Remark: The above calculation shows that H is also Hermitian and Idempotent.

Exercise 2.52

Verify that $H^2 = I$.

Solution

Concepts Involved: Unitary Operators

See the calculation and remark in the previous exercise. □

Exercise 2.53

What are the eigenvalues and eigenvectors of H?

Solution

Concepts Involved: Eigenvalues, Eigenvectors

Using the characteristic equation to find the eigenvalues, we have:

$$\det(H - I\lambda) = 0 \implies \det\begin{bmatrix} \frac{1}{\sqrt{2}} - \lambda & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} - \lambda \end{bmatrix} = 0 \implies \lambda^2 - 1 = 0$$
$$\implies \lambda_1 = 1, \lambda_2 = -1$$

Finding the eigenvectors, we then have:

$$\begin{bmatrix} \frac{1}{\sqrt{2}} - 1 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} - 1 \end{bmatrix} |v_1\rangle = \mathbf{0} \implies |v_1\rangle = \begin{bmatrix} 1 + \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$
$$\begin{bmatrix} \frac{1}{\sqrt{2}} + 1 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} + 1 \end{bmatrix} |v_2\rangle = \mathbf{0} \implies |v_2\rangle = \begin{bmatrix} -1 + \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

Normalizing, we have:

$$|v_1\rangle = \frac{1}{\sqrt{2+\sqrt{2}}}\begin{bmatrix} 1 + \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, |v_2\rangle = \frac{1}{\sqrt{2-\sqrt{2}}}\begin{bmatrix} -1 + \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

□

Exercise 2.54

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

Solution

Concepts Involved: Operator Functions, Simultaneous Diagonalization

Since A, B commute, they can be simulatneously diagonalized; that is, there exists some orthonormal basis $|i\rangle$ of V such that $A = \sum_i a_i |i\rangle\langle i|$ and $B = \sum_i b_i |i\rangle\langle i|$. Hence, using the definition of operator functions, we have:

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

□

Exercise 2.55

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

Solution

Concepts Involved: Unitary Operators, Spectral Decomposition, Operator Functions

Since the Hamiltonian H is Hermitian, it is normal and hence has spectral decomposition:

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

where all E are real by the Hermicity of H, and $|E\rangle$ is an orthonormal basis of the Hilbert space. We then have:

$$\begin{aligned}U(t_1,t_2) \equiv \exp\left[\frac{-iH(t_2-t_1)}{\hbar}\right] &= \exp\left[\frac{-i\sum_E E\,|E\rangle\langle E|\,(t_2-t_1)}{\hbar}\right]\\ &= \sum_E \exp\left(\frac{-iE(t_2-t_1)}{\hbar}\right)|E\rangle\langle E|\end{aligned}$$

Hence calculating $U^\dagger(t_1, t_2)$ we have:

$$U^\dagger(t_1,t_2) = \left(\sum_E \exp\left(\frac{-iE(t_2-t_1)}{\hbar}\right)|E\rangle\langle E|\right)^\dagger = \sum_E \left(\exp\left(\frac{-iE(t_2-t_1)}{\hbar}\right)\right)^* (|E\rangle\langle E|)^\dagger$$
$$= \sum_E \exp\left(\frac{iE(t_2-t_1)}{\hbar}\right)|E\rangle\langle E|$$

Therefore computing $U^\dagger(t_1, t_2)U(t_2, t_1)$ we have:

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

where in the second equality we use the fact that the eigenstates are orthogonal. We conclude that U is unitary. □

Exercise 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.

Solution

Concepts Involved: Hermitian Operators, Unitary Operators, Spectral Decomposition, Operator Functions

---

Suppose U is unitary. Then, U is normal and hence has spectral decomposition:

$$U = \sum_j \lambda_j |j\rangle\langle j|$$

where $|j\rangle$ are eigenvectors of U with eigenvalues λ_j, and $|j\rangle$ forms an orthonormal basis of the Hilbert space. By Exercise 2.18, all eigenvalues of unitary operators have eigenvalues of modulus 1, so we can let $\lambda_j = \exp(i\theta_j)$ where $\theta_j \in \mathbb{R}$ and hence write the above as:

$$U = \sum_j \exp(i\theta_j)|j\rangle\langle j|$$

We then have:

$$K \equiv -i\log(U) = -i\log\left(\sum_j \exp(i\theta_j)\,|j\rangle\langle j|\right) = \sum_j -i\log\Big(\exp(i\theta_j)\Big)\,|j\rangle\langle j| = \sum_j -i(i\theta_j)\,|j\rangle\langle j|$$
$$= \sum_j \theta_j\,|j\rangle\langle j|$$

We then observe that:

$$K^\dagger = \left(\sum_j \theta_j\,|j\rangle\langle j|\right)^\dagger = \sum_j \theta_j\,|j\rangle\langle j|$$

as the θ_js are real and $(|j\rangle\langle j|)^\dagger = |j\rangle\langle j|$. Hence K is Hermitian. Then, multiplying both sides in $K = -i\log(U)$ by i and exponentiating both sides, we obtain the desired relation. □

Exercise 2.57: Cascaded measurements are single measurements

Suppose $\{L_l\}$ and $\{M_m\}$ are two sets of measurement operators. Show that a measurement defined by the measurement operators $\{L_l\}$ followed by a measurement defined by the measurement operators $\{M_m\}$ is physically equivalent to a single measurement defined by measurement operators $\{N_{lm}\}$ with the representation $N_{lm} \equiv M_m L_l$.

Solution

Concepts Involved: Quantum Measurement

Suppose we have (normalized) initial quantum state $|\psi_0\rangle$. Then, the state after measurement of L_l is given by definition to be:

$$|\psi_0\rangle \mapsto |\psi_1\rangle = \frac{L_l|\psi_0\rangle}{\sqrt{\langle\psi_0|L_l^\dagger L_l|\psi_0\rangle}}.$$

The state after measurement of M_m on $|\psi_1\rangle$ is then given to be:

$$|\psi_1\rangle \mapsto |\psi_2\rangle = \frac{M_m|\psi_1\rangle}{\sqrt{\langle\psi_1|M_m^\dagger M_m|\psi_1\rangle}} = \frac{M_m\left(\frac{L_l|\psi_0\rangle}{\sqrt{\langle\psi_0|L_l^\dagger L_l|\psi_0\rangle}}\right)}{\sqrt{\left(\frac{L_l^\dagger\langle\psi_0|}{\sqrt{\langle\psi_0|L_l^\dagger L_l|\psi_0\rangle}}\right)M_m^\dagger M_m\left(\frac{L_l|\psi_0\rangle}{\sqrt{\langle\psi_0|L_l^\dagger L_l|\psi_0\rangle}}\right)}}$$
$$= \frac{M_m L_l|\psi_0\rangle}{\sqrt{\langle\psi_0|L_l^\dagger L_l|\psi_0\rangle}}\frac{\sqrt{\langle\psi_0|L_l^\dagger L_l|\psi_0\rangle}}{\sqrt{\langle\psi_0|L_l^\dagger M_m^\dagger M_m L_l^\dagger|\psi_0\rangle}}$$
$$= \frac{M_m L_l|\psi_0\rangle}{\sqrt{\langle\psi_0|L_l^\dagger M_m^\dagger M_m L_l^\dagger|\psi_0\rangle}}.$$

Conversely, the state of $|\psi_0\rangle$ after measurement of $N_{lm} = M_m L_l$ is given by:

$$|\psi_0\rangle \mapsto |\psi_3\rangle = \frac{M_m L_l |\psi_0\rangle}{\sqrt{\langle\psi_0|L_l^\dagger M_m^\dagger M_m L_l|\psi_0\rangle}}.$$

We see that $|\psi_2\rangle = |\psi_3\rangle$ (that is, the cascaded measurment produces the same result as the single measurement), proving the claim. □

Exercise 2.58

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

Solution

Concepts Involved: Quantum Measurement, Expectation, Standard Deviation

---

By the definition of expectation, we have:

$$\langle M\rangle_{|\psi\rangle} = \langle\psi|M|\psi\rangle = \langle\psi|m|\psi\rangle = m\langle\psi|\psi\rangle = m$$

Where in the second equality we use that $|\psi\rangle$ is an eigenstate of M with eigenvalue m, and in the last equality we use that $|\psi\rangle$ is a normalized quantum state. Next, calculating $\left\langle M^2\right\rangle_{|\psi\rangle}$, we have:

$$\left\langle M^2\right\rangle_{|\psi\rangle} = \langle\psi|M^2|\psi\rangle = \langle\psi|MM|\psi\rangle = \langle\psi|M^\dagger M|\psi\rangle = \langle\psi|m^* m|\psi\rangle = \langle\psi|m^2|\psi\rangle = m^2\langle\psi|\psi\rangle = m^2.$$

Note that we have used the fact that M is Hermitian (it is an observable) to use that $M^\dagger = M$ and $m^* = m$ as all eigenvalues of Hermitian operators are real. Now calculating the standard deviation, we have:

$$\Delta(M) = \sqrt{\langle M^2\rangle - \langle M\rangle^2} = \sqrt{m^2 - (m)^2} = 0$$

□

Exercise 2.59

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

Solution

Concepts Involved: Quantum Measurement, Projective Measurement, Expectation, Standard Deviation, Pauli Operators

---

By the definition of expectation, we have:

$$\langle X\rangle_{|0\rangle} = \langle 0|X|0\rangle = \langle 0|1\rangle = 0$$

Next calculating $\langle X^2 \rangle_{|0\rangle}$, we have:

$$\left\langle X^2 \right\rangle_{|0\rangle} = \langle 0|XX|0\rangle = \langle 1|1\rangle = 1$$

Hence the standard deviation of X is given by:

$$\Delta(X) = \sqrt{\langle X^2\rangle - \langle X\rangle^2} = \sqrt{1-0} = 1$$

□

Exercise 2.60

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

Solution

Concepts Involved: Eigenvalues, Projectors, Pauli Operators

Let $|v\rangle$ be a unit vector. We already showed in Exercise 2.35 that $\mathbf{v}\cdot\boldsymbol{\sigma}$ has eigenvalues $\lambda_+ = 1, \lambda_- = -1$. We next prove a general statement; namely, that for a observable on a 2-dimensional Hilbert space with eigenvalues $\lambda_\pm = \pm 1$ has projectors

$$P_\pm = \frac{I \pm O}{2}$$

To see this is the case, let $P_+ = |o_+\rangle\langle o_+|$, $P_- = |o_-\rangle\langle o_-|$, $I = |o_+\rangle\langle o_+| + |o_-\rangle\langle o_-|$, and $O = |o_+\rangle\langle o_+| - |o_-\rangle\langle o_-|$. We then have:

$$\frac{I+O}{2} = \frac{|o_+\rangle\langle o_+| - |o_-\rangle\langle o_-|}{2} = |o_+\rangle\langle o_+| = P_+$$
$$\frac{I-O}{2} = \frac{|o_-\rangle\langle o_-| + |o_-\rangle\langle o_-| - |o_+\rangle\langle o_+| + |o_-\rangle\langle o_-|}{2} = |o_-\rangle\langle o_-| = P_-$$

Hence the general statement is proven. Applying this to $O = \mathbf{v}\cdot\boldsymbol{\sigma}$ (which is indeed Hermitian and hence an observable as each of X, Y, Z are Hermitian), we get that:

$$P_\pm = \frac{I \pm \mathbf{v}\cdot\boldsymbol{\sigma}}{2}$$

as claimed. □

Exercise 2.61

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

Solution

Concepts Involved: Quantum Measurement, Projective Measurement, Pauli Operators

The probability of obtaining the result $+1$ is given by:

$$p(+) = \langle 0|P_+|0\rangle = \langle 0|\frac{I + \mathbf{v}\cdot\boldsymbol{\sigma}}{2}|0\rangle$$

We recall from from Exercise 2.35 that:

$$\mathbf{v}\cdot\boldsymbol{\sigma} = \begin{bmatrix} v_3 & v_1 - iv_2 \\ v_1 + iv_2 & -v_3 \end{bmatrix} = v_3 |0\rangle\langle 0| + (v_1 - iv_2) |0\rangle\langle 1| + (v_1 + iv_2) |1\rangle\langle 0| - v_3 |1\rangle\langle 1| \,.$$

Hence computing $p(+)$, we get:

$$\begin{aligned} p(+) &= \langle 0| \left(\frac{1}{2}|0\rangle + \frac{1}{2}\left(v_3|0\rangle + (v_1 + iv_2)|1\rangle\right)\right) \\ &= \langle 0| \left(\frac{1+v_3}{2}|0\rangle + \frac{v_1 + iv_2}{2}|1\rangle\right) \\ &= \frac{1+v_3}{2}\langle 0|0\rangle + \frac{v_1+iv_2}{2}\langle 0|1\rangle = \frac{1+v_3}{2} \end{aligned}$$

so the probability of measuring the $+1$ outcome is $\frac{1+v_3}{2}$. The state after the measurement of the $+1$ outcome is given by:

$$|0\rangle \mapsto \frac{P_+|0\rangle}{\sqrt{p(+)}} = \frac{\frac{1+v_3}{2}|0\rangle + \frac{v_1+iv_2}{2}|1\rangle}{\sqrt{\frac{1+v_3}{2}}} = \frac{1}{\sqrt{2(1+v_3)}}\left((1+v_3)|0\rangle + (v_1+iv_2)|1\rangle\right)$$

□

Exercise 2.62

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

Solution

Concepts Involved: Quantum Measurement, Projective Measurement, POVM Measurement

Suppose we have the measurement operators M_m are equal to the POVM elements E_m. In this case, we have:

$$M_m = E_m \equiv M_m^\dagger M_m$$

$M_m^\dagger M_m$ is positive by Exercise 2.25, so it follows that M_m is positive and hence Hermitian by Exercise 2.24. Hence, $M_m^\dagger = M_m$, and therefore:

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

From which we conclude that M_m are projective measurement operators. □

Exercise 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.

Solution

Concepts Involved: Quantum Measurement, POVM Measurement, Polar Decomposition

Since M_m is a linear operator, by the left polar decomposition there exists unitary U such that:

$$M_m = U\sqrt{M_m^\dagger M_m} = U\sqrt{E_m},$$

where in the last equality we use that $M_m^\dagger M_m = E_m$. □

Exercise 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 $\{E_1, E_2, \ldots, E_{m+1}\}$ such that if outcome E_i occurs, $1 \leq i \leq m$, then Bob knows with certainty that he was given the state $|\psi_i\rangle$. (The POVM must be such that $\langle\psi_i|E_i|\psi_i\rangle > 0$ for each i.)

Solution

Concepts Involved: POVM Measurement, Orthogonality

Let $\mathcal{H}$ be the Hilbert space where the given states lie, and let V be the m-dimensional subspace spanned by $|\psi_1\rangle, \ldots, |\psi_m\rangle$. For each $i \in \{1, \ldots, m\}$, let W_i be the subspace of V spanned by $\{|\psi_j\rangle : j \neq i\}$. Let $W_i^\perp$ be the orthogonal complement of W_i which consists of all states in $\mathcal{H}$ orthogonal to all states in W_i. We then have any vector in V can be written as the sum of a vector in W_i and $W_i^\perp \cap V$ (see for example Theorem 6.47 in Axler's *Linear Algebra Done Right*). Therefore, for any $|\psi_i\rangle$ we can write:

$$|\psi_i\rangle = |w_i\rangle + |p_i\rangle$$

Where $|w_i\rangle \in W_i$ and $|p_i\rangle \in W_i^\perp \cap V$. Define $E_i = \frac{|p_i\rangle\langle p_i|}{m}$. By construction, we have for any $|\psi\rangle \in \mathcal{H}$:

$$\langle\psi|E_i|\psi\rangle = \frac{\left|\langle\psi|p_i\rangle\right|^2}{m} \geq 0$$

so the E_is are positive are required. Furthermore, defining $E_{i+1} = I - \sum_{i=1}^m E_i$ we again see that for any $|\psi\rangle \in \mathcal{H}$:

$$\langle\psi|E_{i+1}|\psi\rangle = \langle\psi|I|\psi\rangle - \sum_{i=1}^m \langle\psi|E_i|\psi\rangle = 1 - \sum_{i=1}^m \langle\psi|E_i|\psi\rangle \geq 1 - \sum_{i=1}^m \frac{1}{m} = 0$$

so E_{i+1} is also positive as required. Finally, to see that the $E_1, \ldots E_m$ have the desired properties, observe by construction that since $|p_i\rangle \in W_i^\perp \cap V$, it follows that $\langle\psi_j|p_i\rangle = 0$ for any $j \neq i$ (as the $|p_i\rangle$ will be orthogonal to all the vectors in $\{|\psi_j\rangle : j \neq i\}$ by construction). Calculating $\langle\psi_i|E_i|\psi_i\rangle$, we observe that:

$$\langle\psi_i|E_i|\psi_i\rangle = (\langle w_i| + \langle p_i|)\,\frac{|p_i\rangle\langle p_i|}{m}\,(|w_i\rangle + |p_i\rangle) = \frac{|\langle p_i|p_i\rangle|^2}{m} = \frac{1}{m} \geq 0$$

so if Bob measures E_i, he can be certain that he was given the state $|\psi_i\rangle$. □

Exercise 2.65

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

Solution

Concepts Involved: Linear Algebra, Phase

Let us define our basis to be $|+\rangle := (|0\rangle + |1\rangle)/\sqrt{2}$ and $|-\rangle := (|0\rangle - |1\rangle)/\sqrt{2}$. Our two states are then just the basis vectors of this basis $(|+\rangle, |-\rangle)$ and are not the same up to relative phase shift. □

Exercise 2.66

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

Solution

Concepts Involved: Quantum Measurement, Expectation, Composite Systems, Pauli Operators

Computing the expectation value of X_1Z_2, we get:

$$\begin{aligned}
\langle X_1Z_2\rangle &= \left(\frac{\langle 00| + \langle 11|}{\sqrt{2}}\right) X_1Z_2 \left(\frac{|00\rangle + |11\rangle}{\sqrt{2}}\right) \\
&= \left(\frac{\langle 00| + \langle 11|}{\sqrt{2}}\right)\left(\frac{X_1Z_2|00\rangle + X_1Z_2|11\rangle}{\sqrt{2}}\right) \\
&= \left(\frac{\langle 00| + \langle 11|}{\sqrt{2}}\right)\left(\frac{|10\rangle - |01\rangle}{\sqrt{2}}\right) \\
&= \frac{1}{2}\left(\langle 00|10\rangle - \langle 00|01\rangle + \langle 11|10\rangle - \langle 11|01\rangle\right) \\
&= \frac{1}{2}(0 + 0 + 0 + 0) \\
&= 0.
\end{aligned}$$

□

Exercise 2.67

Suppose V is a Hilbert space with a subspace W. Suppose $U : W \mapsto V$ 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 $U' : V \mapsto V$ 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.

Solution

Concepts Involved: Inner Products, Unitary Operators

By assumption we have U is unitary on W as $\langle w_1|U^\dagger U|w_2\rangle = \langle w_1|w_2\rangle$ and hence $U^\dagger U = I_W$. Hence, it has spectral decomposition:

$$U = \sum_j \lambda_j |j\rangle\langle j|$$

where $\{|j\rangle\}$ is an orthonormal basis of the subspace W. Then, let $\{|j\rangle\} \cup \{|i\rangle\}$ be an orthnormal basis of the full space V. We then define:

$$U' = \sum_j \lambda_j |j\rangle\langle j| + \sum_i |i\rangle\langle i| = U + \sum_i |i\rangle\langle i|$$

We can then see that for any $|w\rangle \in W$ that:

$$U'|w\rangle = \left(U + \sum_i |i\rangle\langle i|\right)|w\rangle = U|w\rangle + \sum_j |i\rangle\langle i|w\rangle = U|w\rangle$$

where in the last line we use that $\langle i|w\rangle = 0$ as $|i\rangle$ are not in the subspace W. Finally, verifying the unitarity of U' we have:

$$\begin{aligned}
U'^\dagger U' &= \left(\sum_j \lambda_j^* |j\rangle\langle j| + \sum_i |i\rangle\langle i|\right)\left(\sum_{j'} \lambda_j |j\rangle\langle j| + \sum_{i'} |i\rangle\langle i|\right) \\
&= \sum_j\sum_{j'} |j\rangle\langle j|j'\rangle\langle j'| + \sum_j\sum_{i'} |j\rangle\langle j|i'\rangle\langle i'| + \sum_i\sum_{j'} |i\rangle\langle i|j'\rangle\langle j'| + \sum_i\sum_{i'} |i\rangle\langle i|i'\rangle\langle i'| \\
&= \sum_j\sum_{j'} \langle j|j'\rangle\delta_{jj'} + \sum_i\sum_{i'} \langle i|i'\rangle\delta_{ii'} \\
&= \sum_j |j\rangle\langle j| + \sum_i |i\rangle\langle i| \\
&= I
\end{aligned}$$

□

Exercise 2.68

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

Solution

Concepts Involved: Composite Systems, Entanglement
Recall that:

$$|\psi\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}$$

Suppose for the take of contradiction that $|\psi\rangle = |a\rangle|b\rangle$ for some single qubit states $|a\rangle$ and $|b\rangle$. Then, we have $|a\rangle = \alpha|0\rangle + \beta|1\rangle$ and $|b\rangle = \gamma|0\rangle + \delta|1\rangle$ for some $\alpha, \beta, \gamma, \delta \in \mathbb{C}$ such that $|\alpha|^2 + |\beta|^2 = 1$ and $|\gamma|^2 + |\delta|^2 = 1$. We then have:

$$|a\rangle|b\rangle = (\alpha|0\rangle + \beta|1\rangle)(\gamma|0\rangle + \delta|1\rangle) = \alpha\gamma|00\rangle + \alpha\delta|01\rangle + \beta\gamma|10\rangle + \beta\delta|11\rangle.$$

Where we have used the linearity of the tensor product (though we supress the $\otimes$ symbols in the above expression). We then have:

$$|\psi\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}} = \alpha\gamma|00\rangle + \alpha\delta|01\rangle + \beta\gamma|10\rangle + \beta\delta|11\rangle$$

which forces $\alpha\delta = 0$ and $\beta\gamma = 0$. However, we then have at least one of $\alpha\gamma$ or $\beta\delta$ is also zero, and we thus reach a contradiction. □

Exercise 2.69

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

Solution

Concepts Involved: Orthogonality, Bases, Composite Systems
Recall that the bell basis is given by:

$$|B_{00}\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}, \quad |B_{01}\rangle\frac{|00\rangle - |11\rangle}{\sqrt{2}}, \quad |B_{10}\rangle = \frac{|01\rangle + |10\rangle}{\sqrt{2}}, \quad |B_{11}\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}}$$

We first verify orthonormality. We observe that:

$$
\begin{aligned}
\langle B_{00}|B_{00}\rangle &= \left(\frac{\langle 00| + \langle 11|}{\sqrt{2}}\right)\left(\frac{|00\rangle + |11\rangle}{\sqrt{2}}\right) = \frac{1}{2}\left(\langle 00|00\rangle + \langle 00|11\rangle + \langle 11|00\rangle + \langle 11|11\rangle\right) = 1 \\
\langle B_{00}|B_{01}\rangle &= \left(\frac{\langle 00| + \langle 11|}{\sqrt{2}}\right)\left(\frac{|00\rangle - |11\rangle}{\sqrt{2}}\right) = \frac{1}{2}\left(\langle 00|00\rangle - \langle 00|11\rangle + \langle 11|00\rangle - \langle 11|11\rangle\right) = 0 \\
\langle B_{00}|B_{10}\rangle &= \left(\frac{\langle 00| + \langle 11|}{\sqrt{2}}\right)\left(\frac{|01\rangle + |10\rangle}{\sqrt{2}}\right) = \frac{1}{2}\left(\langle 00|01\rangle + \langle 00|10\rangle + \langle 11|01\rangle + \langle 11|10\rangle\right) = 0 \\
\langle B_{00}|B_{11}\rangle &= \left(\frac{\langle 00| + \langle 11|}{\sqrt{2}}\right)\left(\frac{|01\rangle - |10\rangle}{\sqrt{2}}\right) = \frac{1}{2}\left(\langle 00|01\rangle - \langle 00|10\rangle + \langle 11|01\rangle - \langle 11|10\rangle\right) = 0 \\
\langle B_{01}|B_{01}\rangle &= \left(\frac{\langle 00| - \langle 11|}{\sqrt{2}}\right)\left(\frac{|00\rangle - |11\rangle}{\sqrt{2}}\right) = \frac{1}{2}\left(\langle 00|00\rangle - \langle 00|11\rangle - \langle 11|00\rangle + \langle 11|11\rangle\right) = 1 \\
\langle B_{01}|B_{10}\rangle &= \left(\frac{\langle 00| - \langle 11|}{\sqrt{2}}\right)\left(\frac{|01\rangle + |10\rangle}{\sqrt{2}}\right) = \frac{1}{2}\left(\langle 00|01\rangle + \langle 00|10\rangle - \langle 11|01\rangle - \langle 11|10\rangle\right) = 0 \\
\langle B_{01}|B_{11}\rangle &= \left(\frac{\langle 00| - \langle 11|}{\sqrt{2}}\right)\left(\frac{|01\rangle - |10\rangle}{\sqrt{2}}\right) = \frac{1}{2}\left(\langle 00|01\rangle - \langle 00|10\rangle - \langle 11|01\rangle + \langle 11|10\rangle\right) = 0 \\
\langle B_{10}|B_{10}\rangle &= \left(\frac{\langle 01| + \langle 10|}{\sqrt{2}}\right)\left(\frac{|01\rangle + |10\rangle}{\sqrt{2}}\right) = \frac{1}{2}\left(\langle 01|01\rangle + \langle 01|10\rangle + \langle 10|01\rangle + \langle 10|10\rangle\right) = 1 \\
\langle B_{10}|B_{11}\rangle &= \left(\frac{\langle 01| + \langle 10|}{\sqrt{2}}\right)\left(\frac{|01\rangle - |10\rangle}{\sqrt{2}}\right) = \frac{1}{2}\left(\langle 01|01\rangle - \langle 01|10\rangle + \langle 10|01\rangle - \langle 10|10\rangle\right) = 0 \\
\langle B_{11}|B_{11}\rangle &= \left(\frac{\langle 01| - \langle 10|}{\sqrt{2}}\right)\left(\frac{|01\rangle - |10\rangle}{\sqrt{2}}\right) = \frac{1}{2}\left(\langle 01|01\rangle - \langle 01|10\rangle - \langle 10|01\rangle + \langle 10|10\rangle\right) = 1
\end{aligned}
$$

so orthonormality is verified. We know show that it is a basis. Recall that we can write any vector $|\psi\rangle$ in the 2 qubit state space as:

$$
|\psi\rangle = \alpha|00\rangle + \beta|01\rangle + \gamma|10\rangle + \delta|11\rangle
$$

where $\alpha, \beta, \gamma, \delta \in \mathbb{C}$ and $|\alpha|^2 + |\beta|^2 + |\gamma|^2 + |\delta|^2 = 1$. We then observe that this is equivalent to:

$$
\left(\frac{\alpha+\delta}{\sqrt{2}}\right)|B_{00}\rangle + \left(\frac{\alpha-\delta}{\sqrt{2}}\right)|B_{01}\rangle + \left(\frac{\beta+\gamma}{\sqrt{2}}\right)|B_{10}\rangle + \left(\frac{\beta-\gamma}{\sqrt{2}}\right)|B_{11}\rangle \quad (*)
$$

as:

$$
\begin{aligned}
&\left(\frac{\alpha+\delta}{\sqrt{2}}\right)\frac{|00\rangle + |11\rangle}{\sqrt{2}} + \left(\frac{\alpha-\delta}{\sqrt{2}}\right)\frac{|00\rangle - |11\rangle}{\sqrt{2}} + \left(\frac{\beta+\gamma}{\sqrt{2}}\right)\frac{|01\rangle + |10\rangle}{\sqrt{2}} + \left(\frac{\beta-\gamma}{\sqrt{2}}\right)\frac{|01\rangle - |10\rangle}{\sqrt{2}} \\
&\qquad = \left(\frac{\alpha}{2} + \frac{\alpha}{2} + \frac{\delta}{2} - \frac{\delta}{2}\right)|00\rangle + \left(\frac{\alpha}{2} - \frac{\alpha}{2} + \frac{\delta}{2} + \frac{\delta}{2}\right)|11\rangle \\
&\qquad + \left(\frac{\beta}{2} + \frac{\beta}{2} + \frac{\gamma}{2} - \frac{\gamma}{2}\right)|01\rangle + \left(\frac{\beta}{2} - \frac{\beta}{2} + \frac{\gamma}{2} + \frac{\gamma}{2}\right)|10\rangle \\
&\qquad = \alpha|00\rangle + \beta|01\rangle + \gamma|10\rangle + \delta|11\rangle = |\psi\rangle
\end{aligned}
$$

Hence $(*)$ shows that the Bell states form a basis. □

Exercise 2.70

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

Solution

Concepts Involved: Superdense Coding, Quantum Measurement, Composite Systems

Let E be a positive operator. We have E for a single qubit can be written as a linear combination of the Pauli matrices:

$$E = a_1 I + a_2 X + a_3 Y + a_4 Z$$

To see that this is the case, consider that the vector space of linear operators acting on a single qubit has dimension 4 (one easy way to see this is that the matrix representations of these operators have 4 entries). Hence, any set of 4 linearly independent linear operators form a basis for the space. As I, X, Y, Z are linearly independent, it follows that they form a basis of the space of linear operators on one qubit. Hence any E can be written as above (Remark: the above decomposition into Paulis is possible regardless of whether E is positive or not).
We then have:

$$\begin{aligned}
\langle B_{00}|E \otimes I|B_{00}\rangle &= \left(\frac{\langle 00| + \langle 11|}{\sqrt{2}}\right) E \otimes I \left(\frac{|00\rangle + |11\rangle}{\sqrt{2}}\right) \\
&= \left(\frac{\langle 00| + \langle 11|}{\sqrt{2}}\right) (a_1 I + a_2 X + a_3 Y + a_4 Z) \otimes I \left(\frac{|00\rangle + |11\rangle}{\sqrt{2}}\right) \\
&= \left(\frac{\langle 00| + \langle 11|}{\sqrt{2}}\right) \left(a_1 \frac{|00\rangle + |11\rangle}{\sqrt{2}} + a_2 \frac{|10\rangle + |01\rangle}{\sqrt{2}} + a_3 \frac{i|10\rangle - i|01\rangle}{\sqrt{2}} + a_4 \frac{|00\rangle - |11\rangle}{\sqrt{2}}\right) \\
&= \frac{1}{2}(a_1 + a_1 + a_4 - a_4) \\
&= a_1
\end{aligned}$$

where in the second last equality we use the orthonormality of $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$. Repeating the same process for the other Bell states, we have:

$$\begin{aligned}
\langle B_{01}|E \otimes I|B_{01}\rangle &= \left(\frac{\langle 00| - \langle 11|}{\sqrt{2}}\right) (a_1 I + a_2 X + a_3 Y + a_4 Z) \otimes I \left(\frac{|00\rangle - |11\rangle}{\sqrt{2}}\right) \\
&= \left(\frac{\langle 00| - \langle 11|}{\sqrt{2}}\right) \left(a_1 \frac{|00\rangle - |11\rangle}{\sqrt{2}} + a_2 \frac{|10\rangle - |01\rangle}{\sqrt{2}} + a_3 \frac{i|10\rangle + i|01\rangle}{\sqrt{2}} + a_4 \frac{|00\rangle + |11\rangle}{\sqrt{2}}\right) \\
&= \frac{1}{2}(a_1 + a_1 + a_4 - a_4) \\
&= a_1
\end{aligned}$$

$$\begin{aligned}
\langle B_{10}|E\otimes I|B_{10}\rangle &= \left(\frac{\langle 01| + \langle 10|}{\sqrt{2}}\right)(a_1 I + a_2 X + a_3 Y + a_4 Z)\otimes I\left(\frac{|01\rangle + |10\rangle}{\sqrt{2}}\right) \\
&= \left(\frac{\langle 01| + \langle 10|}{\sqrt{2}}\right)\left(a_1\frac{|01\rangle + |10\rangle}{\sqrt{2}} + a_2\frac{|11\rangle + |00\rangle}{\sqrt{2}} + a_3\frac{i|11\rangle - i|00\rangle}{\sqrt{2}} + a_4\frac{|01\rangle - |10\rangle}{\sqrt{2}}\right) \\
&= \frac{1}{2}(a_1 + a_1 + a_4 - a_4) \\
&= a_1
\end{aligned}$$

$$\begin{aligned}
\langle B_{01}|E\otimes I|B_{01}\rangle &= \left(\frac{\langle 01| - \langle 10|}{\sqrt{2}}\right)(a_1 I + a_2 X + a_3 Y + a_4 Z)\otimes I\left(\frac{|01\rangle - |10\rangle}{\sqrt{2}}\right) \\
&= \left(\frac{\langle 01| - \langle 10|}{\sqrt{2}}\right)\left(a_1\frac{|01\rangle - |10\rangle}{\sqrt{2}} + a_2\frac{|11\rangle - |00\rangle}{\sqrt{2}} + a_3\frac{i|11\rangle + i|00\rangle}{\sqrt{2}} + a_4\frac{|01\rangle + |10\rangle}{\sqrt{2}}\right) \\
&= \frac{1}{2}(a_1 + a_1 + a_4 - a_4) \\
&= a_1
\end{aligned}$$

Now, suppose that Eve intercepts Alice's qubit. Eve cannot infer anything about which of the four possible bit strings that Alice is trying to send, as any single-qubit measurement that Eve can perform on the intercepted qubit will return the value:

$$\langle\psi|M^\dagger M\otimes I|\psi\rangle$$

Where M is the (single-qubit) measurement operator. But, $M^\dagger M$ is positive, so by the above argument, the measurement outcome will be the same regardless of which Bell state $|\psi\rangle$ is. Hence, Eve cannot obtain the information about the bit string. □

Exercise 2.71: Criterion to decide if a state is mixed or pure

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

Solution

Concepts Involved: Trace, Density Operators, Pure States, Mixed States
Recall that a density operator ρ is pure if:

$$\rho = |\psi\rangle\langle\psi|$$

for some normalized quantum state vector $|\psi\rangle$.

Since ρ is a positive operator, by the spectral decomposition we have:

$$\rho = \sum_i p_i |i\rangle\langle i|$$

where $p_i \geq 0$ (due to positivity) and $|i\rangle$ are orthonormal. Furthermore, by the property of density operators, we have $\mathrm{tr}(\rho) = 1$, hence:

$$\mathrm{tr}(\rho) = \mathrm{tr}\left(\sum_i p_i |i\rangle\langle i|\right) = \sum_i p_i \,\mathrm{tr}\big(|i\rangle\langle i|\big) = \sum_i p_i = 1$$

where in the second equality we use the linearity of the trace. We obtain that $0 \leq p_i \leq 1$ for each i. Calculating ρ^2, we have:

$$\rho^2 = \left(\sum_i p_i |i\rangle\langle i|\right)\left(\sum_{i'} p_{i'} |i'\rangle\langle i'|\right) = \sum_i \sum_{i'} p_i p_{i'} |i\rangle \langle i|i'\rangle \langle i| = \sum_i \sum_{i'} p_i p_{i'} |i\rangle\langle i'| \,\delta_{ii'} = \sum_i p_i^2 |i\rangle\langle i|$$

Hence:

$$\mathrm{tr}\Big(\rho^2\Big) = \sum_i p_i^2 \,\mathrm{tr}\big(|i\rangle\langle i|\big) = \sum_i p_i^2 \leq \sum_i p_i = 1$$

where in the inequality we use the fact that $p_i^2 \leq p_i$ as $0 \leq p_i \leq 1$. The inequality becomes an equality when $p_i^2 = p_i$, that is, when $p_i = 0$ or $p_i = 1$. In order for $\mathrm{tr}(\rho) = 1$ to hold, we have $p_i = 1$ for one i and zero for all others. Hence, ρ in this case is a pure state. Conversely, suppose ρ is a pure state. Then:

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

□

Exercise 2.72: Bloch sphere for mixed states

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

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

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

Where $\mathbf{r}$ is a real three-dimensional vector such that $\|\mathbf{r}\| \leq 1$. 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 $\|\mathbf{r}\| = 1$.

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

Solution

Concepts Involved: Trace, Density Operators, Pure States, Mixed States, Pauli Operators, Bloch Sphere

(1) Since $\{I, X, Y, Z\}$ form an basis the vector space of single-qubit linear operators, we can write (for

any ρ, regardless of whether it is a density operator or not):

$$\rho = a_1 I + a_2 X + a_3 Y + a_4 Z$$

for constants $a_1, a_2, a_3, a_4 \in \mathbb{C}$. Since ρ is a Hermitian operator, we find that each of these constants are actually real, as:

$$a_1 I + a_2 X + a_3 Y + a_4 Z = \rho = \rho^\dagger = a_1^* I^\dagger + a_2^* X^\dagger + a_3^* Y^\dagger + a_4^* Z^\dagger = a_1^* I + a_2^* X + a_3^* Y + a_4^* Z$$

Now, we require that $\operatorname{tr}(\rho) = 1$ for any density operator, hence:

$$\operatorname{tr}(\rho) = \operatorname{tr}(a_1 I + a_2 X + a_3 Y + a_4 Z) = a_1 \operatorname{tr}(I) + a_2 \operatorname{tr}(X) + a_3 \operatorname{tr}(Y) + a_4 \operatorname{tr}(Z) = 2a_1 = 1$$

from which we obtain that $a_1 = \frac{1}{2}$. Note that in the second equality we use the linearity of the trace, and in the third equality we use that $\operatorname{tr}(I) = 2$ and $\operatorname{tr}(\sigma_i) = 0$ for $i \in \{1, 2, 3\}$ (Exercise 2.36). Calculating ρ^2, we have:

$$\begin{aligned} \rho^2 = &\frac{1}{4} I + \frac{a_2}{2} X + \frac{a_3}{2} Y + \frac{a_4}{2} Z + \frac{a_2}{2} X + a_2^2 X^2 + a_2 a_3 XY + a_2 a_4 XZ \\ &+ \frac{a_3}{2} Y + a_3 a_2 YX + a_3^2 Y^2 + a_3 a_4 YZ + \frac{a_4}{2} Z + a_4 a_2 ZX + a_4 a_3 ZY + a_4^2 Z^2 \end{aligned}$$

Now, using that $\{\sigma_i, \sigma_j\} = 0$ for $i, j \in \{1, 2, 3\}$, $i \neq j$ and that $\sigma_i^2 = I$ for any $i \in \{1, 2, 3\}$ (Exercise 2.41), the above simplifies to:

$$\rho^2 = \left(\frac{1}{4} + a_2^2 + a_3^2 + a_4^2\right) I + a_2 X + a_3 Y + a_4 Z$$

Taking the trace of ρ^2 we have:

$$\operatorname{tr}\left(\rho^2\right) = \left(\frac{1}{4} + a_2^2 + a_3^2 + a_4^2\right) \operatorname{tr}(I) + a_2 \operatorname{tr}(X) + a_3 \operatorname{tr}(Y) + a_4 \operatorname{tr}(Z) = 2\left(\frac{1}{4} + a_2^2 + a_3^2 + a_4^2\right)$$

From the previous exercise (Exercise 2.71) we know that $\operatorname{tr}(\rho^2) \leq 1$, so:

$$2\left(\frac{1}{4} + a_2^2 + a_3^2 + a_4^2\right) \leq 1 \implies a_2^2 + a_3^2 + a_4^2 \leq \frac{1}{4} \implies \sqrt{a_2^2 + a_3^2 + a_4^2} \leq \frac{1}{2}$$

Hence we can write:

$$\rho = \frac{I + r_x X + r_y Y + r_z Z}{2} = \frac{I + \mathbf{r} \cdot \boldsymbol{\sigma}}{2}$$

with $\|\mathbf{r}\| \leq 1$.

(2) The Bloch representation for the state $\rho = \frac{I}{2}$ is the above form with $\mathbf{r} = \mathbf{0}$. This vector corresponds to the center of the Bloch sphere, which is a maximally mixed state ($\operatorname{tr}(\rho^2)$ is minimized, with $\operatorname{tr}(\rho^2) = \frac{1}{2}$).

(3) From the calculation in part (1), we know that for any ρ:

$$\operatorname{tr}\left(\rho^2\right) = 2\left(\frac{1 + r_x^2 + r_y^2 + r_z^2}{4}\right) = \frac{1 + r_x^2 + r_y^2 + r_z^2}{2}$$

if $\|\mathbf{r}\| = 1$, then $r_x^2 + r_y^2 + r_z^2 = 1$. Hence, $\operatorname{tr}\left(\rho^2\right) = 1$ and ρ is pure by Exercise 2.71. Conversely, suppose ρ is pure. Then, $\operatorname{tr}\left(\rho^2\right) = 1$, so:

$$\frac{1 + r_x^2 + r_y^2 + r_z^2}{2} = 1 \implies r_x^2 + r_y^2 + r_z^2 = 1 \implies \|\mathbf{r}\| = 1.$$

(4) In section 1.2, we looked at states that lie on the surface of the Bloch sphere, which we parameterized as:

$$|\psi\rangle = \cos\left(\frac{\theta}{2}\right)|0\rangle + e^{i\varphi}\sin\left(\frac{\theta}{2}\right)|1\rangle.$$

Calculating the density operator corresponding to $|\psi\rangle$, we have:

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

Conversely, we have (in the computational basis) our proposed form of $\rho = \frac{I + \mathbf{r}\cdot\boldsymbol{\sigma}}{2}$ can be represented as:

$$\rho = \frac{1 + r_z}{2} |0\rangle\langle 0| + \frac{r_x - ir_y}{2} |0\rangle\langle 1| + \frac{r_x + ir_y}{2} |1\rangle\langle 0| + \frac{1 - r_z}{2} |1\rangle\langle 1|$$

Solving for r_x, r_y, r_z by equating the two expressions for ρ (using Euler's formula and $\sin(2\theta) = 2\cos(\theta)\sin(\theta)$), we have:

$$r_x = \cos(\varphi)\sin(\theta), \quad r_y = \sin(\varphi)\sin(\theta), \quad r_z = 2\cos^2\left(\frac{\theta}{2}\right) - 1 = \cos(\theta)$$

Calculating $\|\mathbf{r}\|$ we have:

$$\begin{aligned}
\|\mathbf{r}\| = \sqrt{r_x^2 + r_y^2 + r_z^2} &= \sqrt{\cos^2(\varphi)\sin^2(\theta) + \sin^2(\varphi)\cos^2(\theta) + \cos^2(\theta)} \\
&= \sqrt{\sin^2(\theta) + \cos^2(\theta)} \\
&= 1
\end{aligned}$$

so we see that indeed, the two definitions coincide for pure states (as $\|\mathbf{r}\| = 1$).

□

Exercise 2.73

($*$) Let ρ be a density operator. A *minimal ensemble* for ρ is an ensemble $\{p_i, |\psi_i\rangle\}$ containing a number of elements equal to the rank of ρ. Let $|\psi\rangle$ 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 $|\psi\rangle$, 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.)

Solution

Concepts Involved: Density Operators, Spectral Decomposition

Below, we will use the unitary freedom in the ensemble for density matrices which is also known as Uhlmann's theorem. Specifically recall that $\rho = \sum_i p_i|\psi_i\rangle\langle\psi_i| = \sum_j q_j|\varphi_j\rangle\langle\varphi_j|$ for ensembles $\{p_i, |\psi_i\rangle\}$ and $\{q_j, |\varphi_j\rangle\}$ if and only if

$$\sqrt{p_i}\,|\psi_i\rangle = \sum_j u_{ij}\sqrt{q_j}\,|\varphi_j\rangle$$

for some unitary matrix u_{ij}.

Using the spectral decomposition of the density matrix we have

$$\rho = \sum_{k=1}^{r}\lambda_k|k\rangle\langle k| \quad \text{with } \lambda_k > 0$$

where all the eigenvectors with eigenvalue 0 have been removed. Thus, the set of vectors $S = \{|k\rangle\}_{k=1}^{r}$ forms a spanning set set for the support of ρ. An element in the support of ρ can thus be decomposed as

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

Assuming that $|\psi_i\rangle$ occurs with probability p_i, we can use the Uhlmann's theorem quoted above to arrive at the relation

$$\sqrt{p_i}|\psi_i\rangle \stackrel{?}{=} \sum_k u_{ik}\sqrt{\lambda_k}|k\rangle = \sqrt{p_i}\sum_k\langle k|\psi_i\rangle|k\rangle,$$

which allows us relate the elements of one of the columns (i th) of the unitary matrix to

$$u_{ik}\sqrt{\lambda_k} \stackrel{?}{=} \sqrt{p_i}\langle k|\psi_i\rangle.$$

Such a relation can always be satisfied for a unitary matrix with dimension r. As u is unitary, we have

$$\begin{aligned}\sum_k |u_{ik}|^2 = 1 \implies 1 &= \sum_k \frac{p_i\langle\psi_i|k\rangle\langle k|\psi_i\rangle}{\lambda_k} \\ \implies p_i &= \sum_k \frac{\lambda_k}{\langle\psi_i|k\rangle\langle k|\psi_i\rangle} \\ &= \frac{1}{\langle\psi_i|\sum_k \frac{1}{\lambda_k}|k\rangle\langle k|\psi_i\rangle} \\ &= \frac{1}{\langle\psi_i|\rho^{-1}|\psi_i\rangle}\,.\end{aligned}$$

$$\begin{aligned}|\psi_j\rangle &= \rho\rho^{-1}|\psi_i\rangle \\ : &= \sum_{i=1}^{r} p_i|\psi_i\rangle\langle\psi_i|\rho^{-1}|\psi_j\rangle \\ &= \sum_{i=1}^{r} p_i\langle\psi_i|\rho^{-1}|\psi_j\rangle|\psi_i\rangle.\end{aligned}$$

But now note that $\{|\psi_j\rangle\}_{j=1}^{r}$ are linearly independent and $|\psi_j\rangle := \sum_{i=1}^{r}\delta_{ij}|\psi_i\rangle$.

$$\implies p_i\langle\psi_i|\rho^{-1}|\psi_i\rangle = 1.$$

Thus, the probability associated with the state $|\psi_i\rangle$ in the ensemble is given by

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

□

Exercise 2.74

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

Solution

Concepts Involved: Density Operators, Reduced Density Operators, Partial Trace, Pure States, Composite Systems

Suppose we have $|a\rangle|b\rangle \in A\otimes B$. Then, the density operator of the combined system is given as $\rho^{AB} = (|a\rangle|b\rangle)(\langle a|\langle b|) = |a\rangle\langle a| \otimes |b\rangle\langle b|$. Calculating the reduced density operator of system A by tracing out system B, we have

$$\rho^A = \mathrm{tr}_B(\rho_{AB}) = \mathrm{tr}_B(\,|a\rangle\langle a| \otimes |b\rangle\langle b|) = |a\rangle\langle a|\,\mathrm{tr}\big(\,|b\rangle\langle b|\big) = |a\rangle\langle a|\,\langle b|b\rangle = |a\rangle\langle a|\,.$$

Hence we find that $\rho^A = |a\rangle\langle a|$ is indeed a pure state. $\square$

Exercise 2.75

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

Solution

Concepts Involved: Density Operators, Reduced Density Operators, Partial Trace, Composite Systems

For the bell state $|B_{00}\rangle$, we have the density operator:

$$\rho = \left(\frac{|00\rangle + |11\rangle}{\sqrt{2}}\right)\left(\frac{\langle 00| + \langle 11|}{\sqrt{2}}\right) = \frac{|00\rangle\langle 00| + |11\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 11|}{2}$$

Obtaining the reduced density operator for qubit A, we have:

$$\begin{aligned}
\rho^A = \operatorname{tr}_B(\rho) &= \frac{\operatorname{tr}_B(|00\rangle\langle 00|) + \operatorname{tr}_B(|00\rangle\langle 11|) + \operatorname{tr}_B(|11\rangle\langle 00|) + \operatorname{tr}_B(|11\rangle\langle 11|)}{2} \\
&= \frac{|0\rangle\langle 0| \langle 0|0\rangle + |0\rangle\langle 1| \langle 1|0\rangle + |1\rangle\langle 0| \langle 0|1\rangle + |1\rangle\langle 1| \langle 1|1\rangle}{2} \\
&= \frac{|0\rangle\langle 0| + |1\rangle\langle 1|}{2} \\
&= \frac{I}{2}
\end{aligned}$$

Obtaining the reduced density operator for qubit B, we have:

$$\begin{aligned}
\rho^B = \operatorname{tr}_A(\rho) &= \frac{\operatorname{tr}_A(|00\rangle\langle 00|) + \operatorname{tr}_A(|00\rangle\langle 11|) + \operatorname{tr}_A(|11\rangle\langle 00|) + \operatorname{tr}_A(|11\rangle\langle 11|)}{2} \\
&= \frac{\langle 0|0\rangle |0\rangle\langle 0| + \langle 1|0\rangle |0\rangle\langle 1| + \langle 0|1\rangle |1\rangle\langle 0| + \langle 1|1\rangle |1\rangle\langle 1|}{2} \\
&= \frac{|0\rangle\langle 0| + |1\rangle\langle 1|}{2} \\
&= \frac{I}{2}
\end{aligned}$$

We repeat a similar process for the other four bell states. For $|B_{01}\rangle$, we have:

$$\begin{aligned}
\rho &= \left(\frac{|00\rangle - |11\rangle}{\sqrt{2}}\right)\left(\frac{\langle 00| - \langle 11|}{\sqrt{2}}\right) = \frac{|00\rangle\langle 00| - |00\rangle\langle 11| - |11\rangle\langle 00| + |11\rangle\langle 11|}{2} \\
\rho^A &= \frac{\mathrm{tr}_B(|00\rangle\langle 00|) - \mathrm{tr}_B(|00\rangle\langle 11|) - \mathrm{tr}_B(|11\rangle\langle 00|) + \mathrm{tr}_B(|11\rangle\langle 11|)}{2} \\
&= \frac{|0\rangle\langle 0| \langle 0|0\rangle - |0\rangle\langle 1| \langle 1|0\rangle - |1\rangle\langle 0| \langle 0|1\rangle + |1\rangle\langle 1| \langle 1|1\rangle}{2} \\
&= \frac{I}{2} \\
\rho^B &= \frac{\mathrm{tr}_A(|00\rangle\langle 00|) - \mathrm{tr}_A(|00\rangle\langle 11|) - \mathrm{tr}_A(|11\rangle\langle 00|) + \mathrm{tr}_A(|11\rangle\langle 11|)}{2} \\
&= \frac{\langle 0|0\rangle |0\rangle\langle 0| + \langle 1|0\rangle |0\rangle\langle 1| + \langle 0|1\rangle |1\rangle\langle 0| + \langle 1|1\rangle |1\rangle\langle 1|}{2} \\
&= \frac{I}{2}
\end{aligned}$$

For $|B_{10}\rangle$, we have:

$$\begin{aligned}
\rho &= \left(\frac{|01\rangle + |10\rangle}{\sqrt{2}}\right)\left(\frac{\langle 01| + \langle 10|}{\sqrt{2}}\right) = \frac{|01\rangle\langle 01| + |01\rangle\langle 10| + |10\rangle\langle 01| + |10\rangle\langle 10|}{2} \\
\rho^A &= \frac{\mathrm{tr}_B(|01\rangle\langle 01|) + \mathrm{tr}_B(|01\rangle\langle 10|) + \mathrm{tr}_B(|10\rangle\langle 01|) + \mathrm{tr}_B(|10\rangle\langle 10|)}{2} \\
&= \frac{|0\rangle\langle 0| \langle 1|+\rangle |0\rangle\langle 1| \langle 0|1\rangle + |1\rangle\langle 0| \langle 1|0\rangle + |1\rangle\langle 1| \langle 0|0\rangle}{2} \\
&= \frac{I}{2} \\
\rho^B &= \frac{\mathrm{tr}_A(|01\rangle\langle 01|) + \mathrm{tr}_A(|01\rangle\langle 10|) + \mathrm{tr}_A(|10\rangle\langle 01|) + \mathrm{tr}_A(|10\rangle\langle 10|)}{2} \\
&= \frac{\langle 0|0\rangle |1\rangle\langle 1| + \langle 1|0\rangle |1\rangle\langle 0| + \langle 0|1\rangle |0\rangle\langle 1| + \langle 1|1\rangle |1\rangle\langle 1|}{2} \\
&= \frac{I}{2}
\end{aligned}$$

Finally, for $|B_{11}\rangle$ we have:

$$\rho = \left(\frac{|01\rangle - |10\rangle}{\sqrt{2}}\right)\left(\frac{\langle 01| - \langle 10|}{\sqrt{2}}\right) = \frac{|01\rangle\langle 01| - |01\rangle\langle 10| - |10\rangle\langle 01| + |10\rangle\langle 10|}{2}$$

$$\rho^A = \frac{\mathrm{tr}_B(|01\rangle\langle 01|) - \mathrm{tr}_B(|01\rangle\langle 10|) - \mathrm{tr}_B(|10\rangle\langle 01|) + \mathrm{tr}_B(|10\rangle\langle 10|)}{2}$$

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

$$= \frac{I}{2}$$

$$\rho^B = \frac{\mathrm{tr}_A(|01\rangle\langle 01|) - \mathrm{tr}_A(|01\rangle\langle 10|) - \mathrm{tr}_A(|10\rangle\langle 01|) + \mathrm{tr}_A(|10\rangle\langle 10|)}{2}$$

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

$$= \frac{I}{2}$$

□

Exercise 2.76

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

Solution

Concepts Involved: Schmidt Decomposition, Singular Value Decomposition, Composite Systems
Note that for this problem we will use a more general form of the Singular Value Decomposition than proven in Nielsen and Chuang (that may have been encountered in a linear algebra course). Given an arbitrary $m \times n$ rectangular matrix A, there exists an $m \times m$ unitary matrix U and $n \times n$ unitary matrix V such that $A = U\Sigma V$ where Σ is a $m \times n$ rectangular diagonal matrix with non-negative reals on the diagonal (see `https://en.wikipedia.org/wiki/Singular_value_decomposition`).

Let $|m\rangle, |n\rangle$ be orthonormal bases for A and B. We can then write:

$$A = \sum_{mn} a_{mn}|m\rangle|n\rangle$$

for some $m \times n$ matrix of complex numbers a. Using the generalized SVD, we can write:

$$A = \sum_{min} u_{mi}d_{ii}v_{in}|m\rangle|n\rangle$$

where d_{ii} is a rectangular diagonal matrix. We can then define $|i_A\rangle = \sum_m u_{mi}|m\rangle$, $|i_B\rangle = \sum_n u_{in}|n\rangle$, and $\lambda_i = d_{ii}$ to yield the Schmidt decomposition. Note that we take $i = \min(m, n)$ and our sum only has as many terms as the dimensionality of the smaller space. □

Exercise 2.77

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

$$|\psi\rangle = \sum_i \lambda_i |i_A\rangle|i_B\rangle|i_C\rangle$$

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

Solution

Concepts Involved: Linear Algebra, Schmidt Decomposition, Composite Systems

Consider the state:

$$|\psi\rangle = |0\rangle \otimes |B_{00}\rangle = \frac{|000\rangle + |011\rangle}{\sqrt{2}}$$

we claim that this state cannot be written in the form:

$$|\psi\rangle = \sum_i \lambda_i |i_A\rangle|i_B\rangle|i_C\rangle$$

for orthonormal bases $|i_A\rangle, |i_B\rangle, |i_C\rangle$. Suppose for the sake of contradiction that we could write it in this form. We then make the observation that:

$$\rho^A = \mathrm{tr}_{BC}(\,|\psi\rangle\langle\psi|) = \sum_i \lambda_i^2 \,|i_A\rangle\langle i_A|$$
$$\rho^B = \mathrm{tr}_{AC}(\,|\psi\rangle\langle\psi|) = \sum_i \lambda_i^2 \,|i_B\rangle\langle i_B|$$
$$\rho^C = \mathrm{tr}_{AB}(\,|\psi\rangle\langle\psi|) = \sum_i \lambda_i^2 \,|i_C\rangle\langle i_C|\,.$$

From this, we conclude that if it is possible to write $|\psi\rangle$ in such a form, then the eigenvalues of the reduced density matrices must all agree and be equal to λ_i^2. Computing the density matrix of the proposted $|\psi\rangle = |0\rangle \otimes |B_{00}\rangle$, we have:

$$\rho = \frac{|000\rangle\langle 000| + |000\rangle\langle 011| + |011\rangle\langle 000| + |011\rangle\langle 011|}{2}$$

Computing the reduced density matrices ρ^A and ρ^B, we find that:

$$\rho^A = \mathrm{tr}_{BC}(\rho) = |0\rangle\langle 0|$$

$$\rho^B = \mathrm{tr}_{AC}(\rho) = \frac{|0\rangle\langle 0| + |1\rangle\langle 1|}{2}.$$

However, the former reduced density matrix has eigenvalues $\lambda_1^2 = 1, \lambda_2^2 = 0$, and the latter has $\lambda_1^2 = \frac{1}{2}$, $\lambda_2^2 = \frac{1}{2}$. This contradicts the fact that the λ_i^2s must match. $\square$

Remark: Necessary and Sufficient conditions for the tripartite (and higher order) Schmidt decompositions can be found here `https://arxiv.org/pdf/quant-ph/9504006.pdf`.

Exercise 2.78

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

Solution

Concepts Involved: Schmidt Decomposition, Schmidt Number, Reduced Density Operators, Composite Systems

Suppose $|\psi\rangle$ is a product state. Then, $|\psi\rangle = |0_A\rangle|0_B\rangle$ for some $|0_A\rangle, |0_B\rangle$, and we therefore have $|\psi\rangle$ has Schmidt number 1 (it is already written in Schmidt decomposition form, and has one nonzero λ). Conversely, suppose $|\psi\rangle$ has Schmidt number 1. Then, $|\psi\rangle = 1|0_A\rangle|0_B\rangle + 0|1_A\rangle|1_B\rangle$ when writing $|\psi\rangle$ in its Schmidt decomposition. Therefore, $|\psi\rangle = |i_A\rangle|i_B\rangle$ and $|\psi\rangle$ is a product state.
Next, take any $|\psi\rangle$ and write out its Schmidt decomposition. We then get:

$$|\psi\rangle = \sum_i \lambda_i |i_A\rangle|i_B\rangle.$$

Hence:

$$\rho = \sum_i \lambda_i^2 \, |i_A\rangle\langle i_A| \otimes |i_B\rangle\langle i_B| \, .$$

Taking the partial trace of ρ to obtain ρ^A, we have:

$$\rho^A = \mathrm{tr}_B(\rho) = \sum_i \lambda_i^2 \, \mathrm{tr}_B\big(|i_A\rangle\langle i_A| \otimes |i_B\rangle\langle i_B|\big) = \sum_i \lambda_i^2 \, |i_A\rangle\langle i_A| \, \mathrm{tr}\big(|i_B\rangle\langle i_B|\big) = \sum_i \lambda_i^2 \, |i_A\rangle\langle i_A| \, .$$

Identically:

$$\rho^B = \mathrm{tr}_A(\rho) = \sum_i \lambda_i^2 \, |i_B\rangle\langle i_B| \, .$$

Now, suppose that $|\psi\rangle$ is a product state. Then, $|\psi\rangle$ has Schmidt number 1. Hence, only one of λ_1, λ_2 is nonzero. Hence, $\rho^A = |i_A\rangle\langle i_A|$ and $\rho^B = |i_B\rangle\langle i_B|$, so ρ^A, ρ^B are pure. Conversely, suppose ρ^A, ρ^B are pure. Then, we have $\rho^A = |i_A\rangle\langle i_A|$ and $\rho^B = |i_B\rangle\langle i_B|$, so it follows that one of λ_1, λ_2 in the above equations for ρ^A, ρ^B must be zero. Therefore, $|\psi\rangle$ has Schmidt number 1, and is hence a product state. □

Exercise 2.79

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

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

Solution

Concepts Involved: Singular Value Decomposition, Schmidt Decomposition, Composite Systems

For the first two expressions, by inspection we find that:

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

For the third expression, we require a little more work. We start by identifying:

$$\frac{|00\rangle + |01\rangle + |10\rangle}{\sqrt{3}} = \sum_{i,j=0}^{1} a_{ij} |i\rangle |j\rangle$$

thus we have the matrix:

$$A = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

this has the singular value decomposition:

$$A = UDV = \begin{bmatrix} \frac{2}{\sqrt{10-2\sqrt{5}}} & \frac{2}{\sqrt{10+2\sqrt{5}}} \\ \frac{2}{\sqrt{10+2\sqrt{5}}} & -\frac{2}{\sqrt{10-2\sqrt{5}}} \end{bmatrix} \begin{bmatrix} \frac{3+\sqrt{5}}{6} & 0 \\ 0 & \frac{3-\sqrt{5}}{6} \end{bmatrix} \begin{bmatrix} \frac{2}{\sqrt{10-2\sqrt{5}}} & \frac{2}{\sqrt{10+2\sqrt{5}}} \\ -\frac{2}{\sqrt{10+2\sqrt{5}}} & \frac{2}{\sqrt{10-2\sqrt{5}}} \end{bmatrix}$$

The diagonal entries D (the singular values of A) yield the Schmidt coefficients, and the columns/rows of U/V yield the Schmidt basis, so defining:

$$\lambda_1^A = \frac{3+\sqrt{5}}{6}, \quad \lambda_2^A = \frac{3-\sqrt{5}}{6}$$

$$|1_A\rangle = \frac{2}{\sqrt{10-2\sqrt{5}}}|0\rangle + \frac{2}{\sqrt{10+2\sqrt{5}}}|1\rangle, \quad |2_A\rangle = \frac{2}{\sqrt{10+2\sqrt{5}}}|0\rangle - \frac{2}{\sqrt{10-2\sqrt{5}}}|1\rangle$$

we find the Schmidt decomposition:

$$\frac{|00\rangle + |01\rangle + |10\rangle}{\sqrt{3}} = \lambda_1^A |1_A\rangle |1_A\rangle + \lambda_2^A |2_A\rangle (-|2_A\rangle)$$

□

Exercise 2.80

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

Solution

Concepts Involved: Schmidt Decomposition, Unitary Operators, Composite Systems

---

We first prove a Lemma. Suppose we have two (orthonormal) bases $\{|i\rangle\}, \{|i'\rangle\}$ of a (n-dimensional) vector space A. We claim that the change of basis transformation U where $|i'\rangle = U|i\rangle$ is unitary.

To see this is the case, let $U = \sum_i |i'\rangle\langle i|$. By orthonormality, we see that $U|i\rangle = |i'\rangle$ as desired. Computing $U^\dagger$, we have $U^\dagger = \sum_i (|i'\rangle\langle i|)^\dagger = \sum_i |i\rangle\langle i'|$. By orthonormality, we then see that $U^\dagger U = \sum_i |i\rangle\langle i| = I$ and hence U is unitary.

We now move onto the actual problem. By assumption, we can write $|\psi\rangle = \sum_i \lambda_i |i_A\rangle|i_B\rangle$ and $|\varphi\rangle = \sum_j \lambda_j |j_A\rangle|j_B\rangle$ where $\lambda_i = \lambda_j$ if $i = j$. By the lemma, there exists unitary change-of-basis matrices U, V such that $|i_A\rangle = U|j_A\rangle$ and $|i_B\rangle = V|j_B\rangle$. Hence, we have:

$$|\psi\rangle = \sum_i \lambda_i |i_A\rangle|i_B\rangle = \sum_j \lambda_j (U|j_A\rangle)(V|j_B\rangle) = (U \otimes V)\sum_j \lambda_j |j_A\rangle|j_B\rangle = (U \otimes V)|\varphi\rangle$$

which is what we wanted to prove. □

Exercise 2.81: Freedom in purifications

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

Solution

Concepts Involved: Schmidt Decomposition, Purification, Unitary Operators, Composite Systems

---

Let $|AR_1\rangle, |AR_2\rangle$ be two purifications of ρ^A to a composite system AR. We can write the orthonormal decomposition of ρ^A as $\rho^A = \sum_i p_i |i^A\rangle\langle i^A|$, from which it follows that we can write:

$$|AR_1\rangle = \sum_i \sqrt{p_i}|i^A\rangle|i^R\rangle$$

$$|AR_2\rangle = \sum_i \sqrt{p_i}|i^A\rangle|i'^R\rangle$$

for some bases $\{|i\rangle\}, \{|i'\rangle\}$ of R. By the Lemma proven in the previous exercise, the transformation U_R

such that $|i\rangle = U_R|i'\rangle$ is unitary, so hence:

$$\begin{aligned}|AR_1\rangle = \sum_i \sqrt{p_i}|i^A\rangle|i^R\rangle = \sum_i \sqrt{p_i}|i^A\rangle(U_R|i'^R\rangle) &= \sum_i \sqrt{p_i}(I_A|i^A\rangle)(U_R|i'^R\rangle) \\ &= (I_A \otimes U_R)\sum_i \sqrt{p_i}|i^A\rangle|i'^R\rangle \\ &= (I_A \otimes U_R)|AR_2\rangle\end{aligned}$$

which proves the claim. □

Exercise 2.82

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

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

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

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

Solution

Concepts Involved: Purification, Schmidt Decomposition

---

(1) To verify that $\sum_i \sqrt{p_i}|\psi_i\rangle|i\rangle$ is a purification, we see that:

$$\begin{aligned}\operatorname{tr}_R\left(\left(\sum_i \sqrt{p_i}|\psi_i\rangle|i\rangle\right)\left(\sum_j \sqrt{p_j}\langle\psi_j|\langle j|\right)\right) &= \sum_i\sum_j \sqrt{p_i p_j}\,|\psi_i\rangle\langle\psi_j|\operatorname{tr}_R(|i\rangle\langle j|) \\ &= \sum_i\sum_j \sqrt{p_i p_j}|\psi_i\rangle\langle\psi_j|\delta_{ij} \\ &= \sum_i \sqrt{p_i^2}\,|\psi_i\rangle\langle\psi_i| \\ &= \sum_i p_i\,|\psi_i\rangle\langle\psi_i| \\ &= \rho\end{aligned}$$

(2) We measure the observable $M_i = I_A \otimes \sum_i P_i = I_A \otimes \sum_i |i\rangle\langle i|$. The probability of obtaining outcome i is given by $p(i) = \langle AR|(I_A \otimes P_i)|AR\rangle$ (where $|AR\rangle = \sum_i \sqrt{p_i}|\psi_i\rangle|i\rangle$), which we can calculate to

be:

$$\begin{aligned}
p(i) &= \langle AR|(I_A \otimes |i\rangle\langle i|)|AR\rangle \\
&= \left(\sum_j \sqrt{p_j}\langle\psi_j|\langle j|\right)(I_A \otimes |i\rangle\langle i|)\left(\sum_k \sqrt{p_k}|\psi_k\rangle|k\rangle\right) \\
&= \sum_j \sum_j \sqrt{p_j}\sqrt{p_k}\langle\psi_j|\psi_k\rangle\delta_{ji}\delta_{ik} \\
&= p_i
\end{aligned}$$

The post measurement state is given by:

$$\begin{aligned}
\frac{(I_A \otimes P_i)|AR\rangle}{\sqrt{p(i)}} &= \frac{(I_A \otimes |i\rangle\langle i|)\sum_j \sqrt{p_j}|\psi_j\rangle|j\rangle}{\sqrt{p_i}} \\
&= \frac{\sum_j \sqrt{p_j}|\psi_j\rangle|j\rangle\delta_{ij}}{\sqrt{p_i}} \\
&= \frac{\sqrt{p_i}|\psi_i\rangle|i\rangle}{\sqrt{p_i}} \\
&= |\psi_i\rangle|i\rangle
\end{aligned}$$

so the corresponding state of system A is $|\psi_i\rangle$.

(3) Let $|AR\rangle$ be any purification of ρ to the combined system AR. We then have $|AR\rangle$ has Schmidt Decomposition:

$$|AR\rangle = \sum_i \lambda_i |i_A\rangle|i_R\rangle$$

for orthonormal bases $|i_A\rangle, |i_R\rangle$ of A and R respectively. Define a linear transformation U such that $\lambda_i|i_A\rangle = \sum_j U_{ij}p_j|\psi_j\rangle$. We then have:

$$|AR\rangle = \sum_i \left(\sum_j U_{ij}p_j|\psi_j\rangle\right)|i_R\rangle = \sum_j p_j|\psi_j\rangle \sum_i U_{ij}|i_R\rangle.$$

We note that we can move the U_{ij} to system R as R has the same state space as A by construction. Letting $|j\rangle = \sum_i U_{ij}|i_R\rangle$ be our orthonormal basis of R, the claim follows (by part (2) of the question.

□

Problem 2.1: Functions of the Pauli matrices

Let $f(\cdot)$ be any function from complex numbers to complex numbers. Let $\mathbf{n}$ be a normalized vector in three dimensions, and let θ be real. Show

$$f(\theta\mathbf{n}\cdot\boldsymbol{\sigma}) = \frac{f(\theta)+f(-\theta)}{2}I + \frac{f(\theta)-f(-\theta)}{2}\mathbf{n}\cdot\boldsymbol{\sigma}$$

Solution

Concepts Involved: Linear Algebra, Spectral Decomposition, Operator Functions.

From Exercise 2.35, we recall that $\mathbf{n}\cdot\boldsymbol{\sigma}$ has spectral decomposition $\mathbf{n}\cdot\boldsymbol{\sigma} = |n_+\rangle\langle n_+| - |n_-\rangle\langle n_-|$. We then have (by the definition of operator functions):

$$f(\theta\mathbf{n}\cdot\boldsymbol{\sigma}) = f\left(\theta(|n_+\rangle\langle n_+| - |n_-\rangle\langle n_-|)\right) = f(\theta)\,|n_+\rangle\langle n_+| + f(-\theta)\,|n_-\rangle\langle n_-|\,.$$

We then use the fact proven in the solution to Exercise 2.60 that we can write the projectors $P_\pm = |n_\pm\rangle\langle n_\pm|$ in terms of the operator $\mathbf{n}\cdot\boldsymbol{\sigma}$ as:

$$|n_\pm\rangle\langle n_\pm| = \frac{I\pm\mathbf{n}\cdot\boldsymbol{\sigma}}{2}.$$

Hence making this substitution we have:

$$f(\theta\mathbf{n}\cdot\boldsymbol{\sigma}) = f(\theta)\left(\frac{I+\mathbf{n}\cdot\boldsymbol{\sigma}}{2}\right) + f(-\theta)\left(\frac{I-\mathbf{n}\cdot\boldsymbol{\sigma}}{2}\right).$$

Grouping terms, we obtain the desired relation:

$$f(\theta\mathbf{n}\cdot\boldsymbol{\sigma}) = \frac{f(\theta)+f(-\theta)}{2}I + \frac{f(\theta)-f(-\theta)}{2}\mathbf{n}\cdot\boldsymbol{\sigma}.$$

□

Remark:
Arguably, the most used application of the above identity in quantum information is when $f(\theta\mathbf{n}\cdot\boldsymbol{\sigma}) = \exp\{i(\theta/2)\mathbf{n}\cdot\boldsymbol{\sigma}\}$. In this case (as in Exercise 2.35), we have

$$\begin{aligned}\exp\{i(\theta/2)\mathbf{n}\cdot\boldsymbol{\sigma}\} &= \frac{\exp\{\theta/2\}+\exp\{-\theta/2\}}{2}I + \frac{\exp\{\theta/2\}-\exp\{-\theta/2\}}{2}\mathbf{n}\cdot\boldsymbol{\sigma} \\ &= \cos\left(\frac{\theta}{2}\right)I + i\sin\left(\frac{\theta}{2}\right)\mathbf{n}\cdot\boldsymbol{\sigma} \quad .\end{aligned}$$

Problem 2.2: Properties of Schmidt numbers

Suppose $|\psi\rangle$ is a pure state of a composite system with components A and B.

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

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

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

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

Solution

Concepts Involved: Schmidt Decomposition, Schmidt Number, Reduced Density Operators, Composite Systems

(1) We write the Schmidt decomposed $|\psi\rangle$, and therefore the density matrix ρ_ψ as:

$$|\psi\rangle = \sum_i \lambda_i |i_A\rangle |i_B\rangle \implies |\psi\rangle\langle\psi| = \sum_{ii'} \lambda_i^2 |i_A\rangle\langle i_A| \otimes |i_B\rangle\langle i_B|$$

Taking the partial trace of subsystem B in the $|i_B\rangle$ basis, we obtain the reduced density matrix ρ_A to be:

$$\rho_A = \mathrm{tr}_B(|\psi\rangle\langle\psi|) = \sum_i \lambda_i^2 |i_A\rangle\langle i_A|$$

$\mathrm{Sch}(\psi)$ of the λ_is are nonzero, and therefore ρ_A has $\mathrm{Sch}(\psi)$ nonzero eigenvalues - therefore the rank of its support is $\mathrm{Sch}(\psi)$.

(2) Suppose for the sake of contradiction that some decomposition $|\psi\rangle = \sum_{j=1}^N |\alpha_j\rangle |\beta_j\rangle$ had less terms than the Schmidt decomposition of $|\psi\rangle$, i.e. $N <$ Sch(ψ).

The density matrix of $|\psi\rangle$ is:

$$\rho_\psi = |\psi\rangle\langle\psi| = \sum_{j=1,k=1}^N |\alpha_j\rangle\langle\alpha_k| \otimes |\beta_j\rangle\langle\beta_k|$$

Tracing out subsystem B, we obtain the reduced density matrix of subsystem A:

$$\rho_A = \mathrm{Tr}_B(\rho_\psi) = \sum_{j=1,k=1}^N |\alpha_j\rangle\langle\alpha_k| \langle\beta_j|\beta_k\rangle$$

where we have used that $\mathrm{Tr}\big(|\beta_1\rangle\langle\beta_2|\big) = \langle\beta_1|\beta_2\rangle$. From the above, it is clear that ρ_A has rank at most N, as the support of ρ_A is spanned by $\big\{|\alpha_1\rangle, \ldots, |\alpha_N\rangle\big\}$. But then the rank of ρ_A is less than Sch(ψ), which contradicts our finding in part (a).

(3) If $\mathrm{Sch}(\varphi) = \mathrm{Sch}(\gamma)$ then there is nothing to prove as $\mathrm{Sch}(\psi)$ is non-negative by definition. Suppose then that $\mathrm{Sch}(\varphi) \neq \mathrm{Sch}(\gamma)$. WLOG suppose $\mathrm{Sch}(\varphi) > \mathrm{Sch}(\gamma)$. We can then write:

$$|\varphi\rangle = \frac{\beta}{\alpha}|\gamma\rangle - \frac{1}{\alpha}|\psi\rangle$$

If we Schmidt decompose $|\varphi\rangle$ and $|\psi\rangle$, we have written $|\varphi\rangle$ as the sum of $\mathrm{Sch}(\gamma) + \mathrm{Sch}(\psi)$ (unnormalized) bipartite states. Applying the result from part (2) of this problem, we then have:

$$\mathrm{Sch}(\varphi) \leq \mathrm{Sch}(\gamma) + \mathrm{Sch}(\psi)$$

which we rearrange to obtain:

$$\mathrm{Sch}(\psi) \geq \mathrm{Sch}(\varphi) - \mathrm{Sch}(\gamma) = \big|\mathrm{Sch}(\varphi) - \mathrm{Sch}(\gamma)\big|$$

which proves the claim.

□

Problem 2.3: Tsirelson's inequality

Suppose $Q = \mathbf{q} \cdot \boldsymbol{\sigma}$, $R = \mathbf{r} \cdot \boldsymbol{\sigma}$, $S = \mathbf{s} \cdot \boldsymbol{\sigma}$, $T = \mathbf{t} \cdot \boldsymbol{\sigma}$, where $\mathbf{q}, \mathbf{r}, \mathbf{s}$, and $\mathbf{t}$ are real unit vectors in three dimensions. Show that

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

Use this result to prove that

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

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

Solution

Concepts Involved: Tensor Products, Commutators, Composite Systems

We first show that $N^2 = I$ for any $N = \mathbf{n} \cdot \boldsymbol{\sigma}$ where $\mathbf{n}$ is a unit vector in three dimensions. We have:

$$\begin{aligned} N^2 &= (\sum_{i=1}^{3} n_i \sigma_i)^2 \\ &= n_1^2\sigma_1^2 + n_2^2\sigma_2^2 + n_3^2\sigma_3^2 + n_1n_2(\sigma_1\sigma_2 + \sigma_2\sigma_1) + n_1n3(\sigma_1\sigma_3 + \sigma_3\sigma_1) + n_2n_3(\sigma_2\sigma_3 + \sigma_3\sigma_2) \end{aligned}$$

By Exercise 2.41, $\sigma_i^2 = I$ and $\{\sigma_i, \sigma_j\} = 0$ for $i \neq j$, so the above reduces to:

$$N^2 = n_1^2 I + n_2^2 I + n_3^2 I = (n_1^2 + n_2^2 + n_3^2)I = I$$

where we use the fact that $\mathbf{n}$ is of unit length. Using this fact, we have:

$$\begin{aligned} (Q \otimes S + R \otimes S + R \otimes T - Q \otimes T)^2 &= Q^2 \otimes S^2 + QR \otimes S^2 + QR \otimes ST - Q^2 \otimes ST \\ &+ RQ \otimes S^2 + R^2 \otimes S^2 + R^2 \otimes ST - RQ \otimes ST \\ &+ RQ \otimes TS + R^2 \otimes TS + R^2 \otimes T^2 - RQ \otimes T^2 \\ &- Q^2 \otimes TS - QR \otimes TS - QR \otimes T^2 + Q^2 \otimes T^2 \\ &= I \otimes I + QR \otimes I + QR \otimes ST - I \otimes ST \\ &+ RQ \otimes I + I \otimes I + I \otimes ST - RQ \otimes ST \\ &+ RQ \otimes TS + I \otimes TS + I \otimes I - RQ \otimes I \\ &- I \otimes TS - QR \otimes TS - QR \otimes I + I \otimes I \\ &= 4I \otimes I + RQ \otimes TS - RQ \otimes ST + QR \otimes ST - QR \otimes TS \\ &= 4I + QR \otimes (ST - TS) - RQ \otimes (ST - TS) \\ &= 4I + [Q, R] \otimes [S, T] \end{aligned}$$

which proves the first equation. We have $\langle 4I \rangle = 4 \langle I \rangle = 4$. Since each of Q, R, S, T have eigenvalues ± 1 (Exercise 2.35), we also ave that $\left\langle [Q, R] \otimes [S, T] \right\rangle \leq 4$ as the tensor product of commutators consists of 4 terms, each of which has expectation less than or equal to 1. We therefore have by the linearity of expectation (Exercise **??**) that:

$$\left\langle (Q \otimes S + R \otimes S + R \otimes T - Q \otimes T)^2 \right\rangle = \left\langle 4I + [QR] \otimes [S, T] \right\rangle \leq 8.$$

Furthermore, we have:

$$\left\langle (Q \otimes S + R \otimes S + R \otimes T - Q \otimes T) \right\rangle^2 \leq \left\langle (Q \otimes S + R \otimes S + R \otimes T - Q \otimes T)^2 \right\rangle$$

so combining the two inequalities we obtain:

$$\left\langle (Q \otimes S + R \otimes S + R \otimes T - Q \otimes T) \right\rangle^2 \leq 8.$$

Taking square roots on both sides, we have:

$$\left\langle (Q \otimes S + R \otimes S + R \otimes T - Q \otimes T) \right\rangle \leq 2\sqrt{2}$$

and again by the linearity of expectation:

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

which is the desired inequality. □

4 Quantum circuits

Exercise 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 Bloch sphere which correspond to the normalized eigenvectors of the different Pauli matrices.

Solution

Concepts Involved: Eigenvalues, Eigenvectors, Pauli Operators, Bloch Sphere

Recall that a single qubit in the state $|\psi\rangle = a|0\rangle + b|1\rangle$ can be visualized as a point (θ, φ) on the Bloch sphere, where $a = \cos(\theta/2)$ and $b = e^{i\varphi}\sin(\theta/2)$.

We recall from 2.11 that Z (and I) has eigenvectors $|0\rangle, |1\rangle$, X has eigenvectors $|+\rangle = \frac{|0\rangle+|1\rangle}{\sqrt{2}}, |-\rangle = \frac{|0\rangle-|1\rangle}{\sqrt{2}}$, and Y has eigenvectors $|y_+\rangle = \frac{|0\rangle+i|1\rangle}{\sqrt{2}}, |y_-\rangle = \frac{|0\rangle-i|1\rangle}{\sqrt{2}}$. Expressing these vectors as points on the Bloch sphere (using spherical coordinates), we have:

$$|0\rangle \cong (0,0)\,; |1\rangle \cong (\pi, 0)\,; |+\rangle \cong \left(\frac{\pi}{2}, 0\right);$$
$$|-\rangle \cong \left(\frac{\pi}{2}, \pi\right); |y_+\rangle \cong \left(\frac{\pi}{2}, \frac{\pi}{2}\right); |y_-\rangle \cong \left(\frac{\pi}{2}, \frac{3\pi}{2}\right).$$

□

Exercise 4.2

Let x be a real number and A a matrix such that $A^2 = I$. Show that

$$\exp(iAx) = \cos(x)I + i\sin(x)A$$

Use this result to verify Equations (4.4) through (4.6).

Solution

Concepts Involved: Operator Functions

Let $|v\rangle$ be an eigenvector of A with eigenvalue λ. It then follows that $A^2|v\rangle = \lambda^2|v\rangle$, and furthermore we have $A^2|v\rangle = I|v\rangle = |v\rangle$ by assumption. We obtain that $\lambda^2 = 1$ and therefore the only possible eigenvalues of A are $\lambda = \pm 1$. Let $|v_1\rangle, \dots, |v_k\rangle$ be the eigenvectors with eigenvalue 1 and $|v_{k+1}\rangle, \dots, |v_n\rangle$ be the eigenvectors with eigenvalue -1. By the spectral decomposition, we can write:

$$A = \sum_{i=1}^{k} |v_i\rangle\langle v_i| - \sum_{i=k+1}^{n} |v_i\rangle\langle v_i|$$

so by the definition of operator functions we have:

$$\exp(iAx) = \sum_{i=1}^{k} \exp(ix) |v_i\rangle\langle v_i| + \sum_{i=k+1}^{n} \exp(-ix) |v_i\rangle\langle v_i| .$$

By Euler's identity we have:

$$\exp(iAx) = \sum_{i=1}^{k} \big(\cos(x) + i\sin(x)\big) |v_i\rangle\langle v_i| + \sum_{i=k+1}^{n} \big(\cos(x) - i\sin(x)\big) |v_i\rangle\langle v_i| .$$

Grouping terms, we obtain:

$$\exp(iAx) = \cos(x) \sum_{i=1}^{n} |v_i\rangle\langle v_i| + i\sin(x) \left(\sum_{i=1}^{k} |v_i\rangle\langle v_i| - \sum_{i=k+1}^{n} |v_i\rangle\langle v_i| \right) .$$

Using the spectral decomposition and definition of I, we therefore obtain the desired relation:

$$\exp(iAx) = \cos(x)I + i\sin(x)A.$$

Since all of the Pauli matrices satisfy $A^2 = I$ (Exercise 2.41), for $\theta \in \mathbb{R}$ we can apply this obtained relation to obtain:

$$\exp\big(-i\theta X/2\big) = \cos\left(\frac{\theta}{2}\right) I - i\sin\left(\frac{\theta}{2}\right) X = \begin{bmatrix} \cos\frac{\theta}{2} & -i\sin\frac{\theta}{2} \\ -i\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{bmatrix}$$

$$\exp\big(-i\theta Y/2\big) = \cos\left(\frac{\theta}{2}\right) I - i\sin\left(\frac{\theta}{2}\right) Y = \begin{bmatrix} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{bmatrix}$$

$$\exp\big(-i\theta Z/2\big) = \cos\left(\frac{\theta}{2}\right) I - i\sin\left(\frac{\theta}{2}\right) Z = \begin{bmatrix} \cos\frac{\theta}{2} - i\sin\frac{\theta}{2} & 0 \\ 0 & \cos\frac{\theta}{2} + i\sin\frac{\theta}{2} \end{bmatrix} = \begin{bmatrix} e^{-i\theta/2} & 0 \\ 0 & e^{i\theta/2} \end{bmatrix}$$

which verifies equations (4.4)-(4.6). □

Exercise 4.3

Show that, up to a global phase, the $\pi/8$ gate satisfies $T = R_z(\pi/4)$

Solution

Concepts Involved: Rotations

Recall that the T gate is defined as:

$$T = \begin{bmatrix} 1 & 0 \\ 0 & \exp\big(i\pi/4\big) \end{bmatrix}$$

We observe that:

$$R_z(\pi/4) = \begin{bmatrix} e^{-i\pi/8} & 0 \\ 0 & e^{i\pi/8} \end{bmatrix} = e^{-i\pi/8}\begin{bmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{bmatrix} = e^{-i\pi/8}T.$$

$\square$

Exercise 4.4

Express the Hadamard gate H as a product of R_x and R_z rotations and $e^{i\varphi}$ for some φ.

Solution

Concepts Involved: Linear algebra, Quantum Gates

We claim that $H = R_z(\pi/2)R_x(\pi/2)R_z(\pi/2)$ up to a global phase of $e^{-i\pi/2}$. Doing a computation to verify this claim, we see that:

$$\begin{aligned}
R_z(\pi/2)R_x(\pi/2)R_z(\pi/2) &= \begin{bmatrix} e^{-i\pi/4} & 0 \\ 0 & e^{i\pi/4} \end{bmatrix}\begin{bmatrix} \cos\frac{\pi}{4} & -i\sin\frac{\pi}{4} \\ -i\sin\frac{\pi}{4} & \cos\frac{\pi}{4} \end{bmatrix}\begin{bmatrix} e^{-i\pi/4} & 0 \\ 0 & e^{i\pi/4} \end{bmatrix} \\
&= \begin{bmatrix} e^{-i\pi/4} & 0 \\ 0 & e^{i\pi/4} \end{bmatrix}\begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{i}{\sqrt{2}} \\ -\frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}\begin{bmatrix} e^{-i\pi/4} & 0 \\ 0 & e^{i\pi/4} \end{bmatrix} \\
&= \frac{1}{\sqrt{2}}\begin{bmatrix} e^{-i\pi/4} & 0 \\ 0 & e^{i\pi/4} \end{bmatrix}\begin{bmatrix} e^{-i\pi/4} & -ie^{i\pi/4} \\ -ie^{-i\pi/4} & e^{i\pi/4} \end{bmatrix} \\
&= \frac{1}{\sqrt{2}}\begin{bmatrix} e^{-i\pi/2} & -i \\ -i & e^{i\pi/2} \end{bmatrix} \\
&= \frac{1}{\sqrt{2}}\begin{bmatrix} e^{-i\pi/2} & -e^{-i\pi/2} \\ -e^{-i\pi/2} & e^{i\pi/2} \end{bmatrix} \\
&= \frac{e^{-i\pi/2}}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \\
&= e^{-i\pi/2}H
\end{aligned}$$

$\square$

Remark: If you are more algebraically minded, the following may appeal to you.

$$\begin{aligned} R_z(\pi/2)R_x(\pi/2)R_z(\pi/2) &= \frac{1}{2\sqrt{2}}(1-iZ)(1-iX)(1-iZ) \\ &= \frac{1}{2\sqrt{2}}(1-iZ-iX-ZX)(1-iZ) \\ &= \frac{1}{2\sqrt{2}}(1-iZ-iX-ZX-iZ-XZ-1+iZXZ) \\ &= \frac{1}{2\sqrt{2}}(-2iX-2iZ) \quad \text{(using } ZXZ=-X\text{)} \\ &=: -iH \end{aligned}$$

Exercise 4.5

Prove that $(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})^2 = I$, and use this to verify Equation (4.8)

Solution

Concepts Involved: Rotations, Pauli Operators

Expanding out the expression, we see that:

$$\begin{aligned} (\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})^2 &= (n_xX+n_yY+n_zZ)^2 \\ &= n_x^2X^2+n_y^2Y^2+n_z^2Z^2+n_xn_y(XY+YX)+n_xn_z(XZ+ZX)+n_yn_z(YZ+ZY) \end{aligned}$$

Using the result from Exercise 2.41 that $\{\sigma_i,\sigma_j\}=0$ if $i\neq j$ and $\sigma_i^2=I$, we have:

$$(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})^2 = (n_x^2+n_y^2+n_z^2)I = I$$

where we use the fact that $\hat{\mathbf{n}}$ is a vector of unit length. With this shown, we can use the result of Exercise 4.2 to conclude that:

$$\exp(-i\theta\hat{\mathbf{n}}\cdot\boldsymbol{\sigma}/2) = \cos\left(\frac{\theta}{2}\right) - i\sin\left(\frac{\theta}{2}\right)(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})$$

which verifies equation (4.8). □

Exercise 4.6: Bloch sphere interpretation of rotations

($*$) One reason why the $R_{\hat{\mathbf{n}}}(\theta)$ operators are referred to as rotation operators is the following fact, which you are to prove. Suppose a single qubit has a state represented by the Bloch vector $\boldsymbol{\lambda}$. Then, the effects of the rotation $R_{\hat{\mathbf{n}}}(\theta)$ on the state is to rotate it by an angle θ about the $\hat{\mathbf{n}}$ axis of the Bloch sphere. This fact explains the rather mysterious looking factors of two in the definition of the rotation matrices.

Solution

Concepts Involved: Rotations, Density Operators, Pauli Operators, Bloch Sphere

Let $\boldsymbol{\lambda}$ be an arbitrary Bloch vector. WLOG, we can express $\boldsymbol{\lambda}$ in a coordinate system such that $\hat{\mathbf{n}}$ is aligned with the $\hat{\mathbf{z}}$ axis, so it suffices to consider how the state behaves under application $R_z(\theta)$. Let $\boldsymbol{\lambda} = (\lambda_x, \lambda_y, \lambda_z)$ be the vector expressed in this coordinate system. By Exercise 2.72, the density operator corresponding to this Bloch vector is given by:

$$\rho = \frac{I + \boldsymbol{\lambda} \cdot \boldsymbol{\sigma}}{2}$$

We now observe how ρ transforms under conjugation by $R_z(\theta)$:

$$\begin{aligned}
R_z(\theta)\rho R_z(\theta)^\dagger &= R_z(\theta)\rho R_z(-\theta) \\
&= R_z(\theta)\left(\frac{I + \lambda_x X + \lambda_y Y + \lambda_z Z}{2}\right) R_z(-\theta)
\end{aligned}$$

Using that $XZ = -ZX$ from Exercise 2.41, we make the observation that:

$$\begin{aligned}
R_z(\theta)X &= \left(\cos\left(\frac{\theta}{2}\right)I - i\sin\left(\frac{\theta}{2}\right)Z\right)X \\
&= X\left(\cos\left(\frac{\theta}{2}\right)I + i\sin\left(\frac{\theta}{2}\right)Z\right) \\
&= X\left(\cos\left(\frac{-\theta}{2}\right)I - i\sin\left(\frac{-\theta}{2}\right)Z\right) \\
&= XR_z(-\theta)
\end{aligned}$$

Similarly, we find that $R_z(\theta)Y = R_z(-\theta)Y$ (same anticommutation) and that $R_z(\theta)Z = ZR_z(\theta)$ (all terms commute). With this, the expression for $R_z(\theta)\rho R_z(\theta)^\dagger$ simplifies to:

$$\begin{aligned}
R_z(\theta)\rho R_z(\theta)^\dagger &= R_z(\theta)\left(\frac{I + \lambda_x X + \lambda_y Y + \lambda_z Z}{2}\right) R_z(-\theta) \\
&= \left(\frac{IR_z(\theta) + \lambda_x XR_z(-\theta) + \lambda_y YR_z(-\theta) + \lambda_z ZR_z(\theta)}{2}\right) R_z(-\theta) \\
&= \frac{I + \lambda_x XR_z(-2\theta) + \lambda_y YR_z(-2\theta) + \lambda_z Z}{2}
\end{aligned}$$

Calculating each of the terms in the above expression, we have:

$$\begin{aligned}
XR_z(-2\theta) &= X\left(\cos\left(\frac{-2\theta}{2}\right) - i\sin\left(\frac{-2\theta}{2}\right)Z\right) \\
&= X\left(\cos(\theta) + i\sin(\theta)Z\right) \\
&= \cos(\theta)X + i\sin(\theta)XZ \\
&= \cos(\theta)X + i\sin(\theta)(-iY) \\
&= \cos(\theta)X + \sin(\theta)Y
\end{aligned}$$

$$
\begin{aligned}
YR_z(-2\theta) &= Y\left(\cos(\theta) + i\sin(\theta)Z\right) \\
&= \cos(\theta)Y + i\sin(\theta)YZ \\
&= \cos(\theta)Y + i\sin(\theta)(iX) \\
&= \cos(\theta)Y - \sin(\theta)X.
\end{aligned}
$$

Plugging these back into the expression for $R_z(\theta)\rho R_z(\theta)^\dagger$ and collecting like terms, we have:

$$
R_z(\theta)\rho R_z(\theta)^\dagger = \frac{I + (\lambda_x\cos(\theta) - \lambda_y\sin(\theta))X + (\lambda_x\sin(\theta) + \lambda_y\cos(\theta))Y + \lambda_z Z}{2}.
$$

From this expression, we can read off the new Bloch vector $\boldsymbol{\lambda}'$ after conjugation by $R_z(\theta)$ to be:

$$
\boldsymbol{\lambda}' = (\lambda_x\cos(\theta) - \lambda_y\sin(\theta), \lambda_x\sin(\theta) + \lambda_y\cos(\theta), \lambda_z).
$$

Alternatively, suppose we apply the 3-dimensional rotation matrix $A_z(\theta)$ to the original bloch vector $\boldsymbol{\lambda}$. We have:

$$
A_z(\theta)\boldsymbol{\lambda} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \lambda_x \\ \lambda_y \\ \lambda_z \end{bmatrix} = \begin{bmatrix} \lambda_x\cos\theta - \lambda_y\sin\theta \\ \lambda_x\sin\theta + \lambda_y\cos\theta \\ \lambda_z \end{bmatrix}.
$$

We see that we end up with the same resulting vector $\boldsymbol{\lambda}'$. We conclude that the conjugation of ρ under $R_z(\theta)$ has the equivalent effect to rotating the Bloch vector by θ about the $\hat{\mathbf{z}}$-axis, and hence the effect of $R_{\hat{\mathbf{n}}}(\theta)$ on a one qubit state is to rotate it by an angle θ about $\hat{\mathbf{n}}$. □

Exercise 4.7

Show that $XYX = -Y$ and use this to prove that $XR_y(\theta)X = R_y(-\theta)$.

Solution

Concepts Involved: Rotations, Pauli Operators

For the first claim, we use that $XY = -YX$ and $X^2 = I$ (Exercise 2.41) to obtain that:

$$
XYX = -YXX = -YI = -Y.
$$

Using this, we have:

$$
\begin{aligned}
XR_y(\theta)X &= X\left(\cos\left(\frac{\theta}{2}\right)I - i\sin\left(\frac{\theta}{2}\right)Y\right)X \\
&= \cos\left(\frac{\theta}{2}\right)XIX - i\sin\left(\frac{\theta}{2}\right)XYX \\
&= \cos\left(\frac{\theta}{2}\right)I + i\sin\left(\frac{\theta}{2}\right)Y \\
&= \cos\left(-\frac{\theta}{2}\right)I - i\sin\left(-\frac{\theta}{2}\right)Y \\
&= R_y(-\theta).
\end{aligned}
$$

□

Exercise 4.8

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

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

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

1. Prove this fact.

2. Find values for α, θ, and $\hat{\mathbf{n}}$ giving the Hadamard gate H.

3. Find values for α, θ, and $\hat{\mathbf{n}}$ giving the phase gate

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

Solution

Concepts Involved: Unitary Operators

---

1. By definition, for any unitary operator U we have $U^\dagger U = I$, so for any state vector $\langle\psi|\psi\rangle = \langle\psi|U^\dagger U|\psi\rangle$. Therefore, all unitary Us are norm-preserving, and hence for a single qubit correspond to some reflection/rotation in 3-dimensional space (up to a global phase factor). Hence, we can write $U = \exp(i\alpha)R_{\hat{\mathbf{n}}}(\theta)$ for some $\hat{\mathbf{n}}$ (rotation axis), θ (rotation angle) and α (global phase).

2. Using the fact that $H = \frac{X+Z}{\sqrt{2}}$, and that modulo a factor of i that X/Z correspond to rotations

$R_x(\pi)$ and $R_z(\pi)$, we find that:

$$H = \frac{iR_x(\pi) + iR_z(\pi)}{\sqrt{2}} = i\left(\frac{2\cos(\frac{\pi}{2})I - i\sin(\frac{\pi}{2})X - i\sin(\frac{\pi}{2})Z}{\sqrt{2}}\right)$$
$$= i\left(\cos\left(\frac{\pi}{2}\right)I - i\sin\left(\frac{\pi}{2}\right)\left(\frac{1}{\sqrt{2}}X + 0Y + \frac{1}{\sqrt{2}}Z\right)\right)$$
$$= e^{i\pi/2}\left(\cos\left(\frac{\pi}{2}\right)I - i\sin\left(\frac{\pi}{2}\right)\left(\frac{1}{\sqrt{2}}X + 0Y + \frac{1}{\sqrt{2}}Z\right)\right)$$

Note that in the second last equality we use that $\cos(\frac{\pi}{2}) = 0$ and hence $\frac{2}{\sqrt{2}}\cos(\frac{\pi}{2}) = \cos(\frac{\pi}{2})$. From the last expression, we can read off using the definition of $R_{\hat{\mathbf{n}}}(\theta)$ that $\hat{\mathbf{n}} = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right)$, $\theta = \pi$, and $\alpha = \frac{\pi}{2}$.

3. We observe that:

$$R_z\left(\frac{\pi}{2}\right) = \begin{bmatrix} e^{-i\pi/4} & 0 \\ 0 & e^{i\pi/4} \end{bmatrix} = e^{-i\pi/4}\begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix}$$

Hence:

$$S = e^{i\pi/4}R_z\left(\frac{\pi}{2}\right)$$

from which we obtain that $\hat{\mathbf{n}} = \hat{\mathbf{z}} = (0, 0, 1)$, $\theta = \frac{\pi}{2}$, and $\alpha = \frac{\pi}{4}$.

□

Remark: For part (2), one just can use the definition

$$R_{\hat{n}}(\theta) \equiv \exp\bigl(-i\theta\hat{n}\cdot\vec{\sigma}/2\bigr) = \cos\left(\frac{\theta}{2}\right)I - i\sin\left(\frac{\theta}{2}\right)\left(n_xX + n_yY + n_zZ\right),$$

and the fact $H = (X + Z)/\sqrt{2}$, to arrive at $\cos\left(\frac{\theta}{2}\right) = 0$, $n_x = n_z = \frac{1}{\sqrt{2}}, n_y = 0$.

Exercise 4.9

Explain why any single qubit unitary operator may be written in the form (4.12).

Solution

Concepts Involved: Unitary Operators, Rotations, Gate Decomposition
Recall that (4.12) states that we can write any single qubit unitary U as:

$$U = \begin{bmatrix} e^{i(\alpha-\beta/2-\delta/2)}\cos\frac{\gamma}{2} & -e^{i(\alpha-\beta/2+\delta/2)}\sin\frac{\gamma}{2} \\ e^{i(\alpha+\beta/2-\delta/2)}\sin\frac{\gamma}{2} & e^{i(\alpha+\beta/2+\delta/2)}\cos\frac{\gamma}{2} \end{bmatrix}$$

where $\alpha, \beta, \gamma, \delta \in \mathbb{R}$.

Let U be a single qubit unitary operator. We then have $U^\dagger U = I$, so identifying:

$$U = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix}$$

we obtain that:

$$\begin{bmatrix} |a|^2 + |c|^2 & a^*b + c^*d \\ ab^* + cd^* & |b|^2 + |d|^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

From the diagonal entries we obtain that $|\mathbf{v}_1| = |\mathbf{v}_2| = 1$ and from the off diagonal entries we obtain that $\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = 0$ and hence the columns of U are orthonormal. From the fact that $|v|_1$ is normalized, we can parameterize the magnitude of the entries with $\gamma \in \mathbb{R}$ such that:

$$|a| = \cos\frac{\gamma}{2}, \quad |c| = \sin\frac{\gamma}{2}.$$

From the orthogonality, we further obtain that $b = -c^*$ and $d = a^*$, from which we have $|b| = |c|$ and $|d| = |a|$. Furthermore, (also from the orthogonality) we can parameterize $\arg(a) = -\frac{\beta}{2} - \frac{\delta}{2}$ and $\arg(b) = \frac{\beta}{2} - \frac{\delta}{2}$ For $\beta, \delta \in \mathbb{R}$. Finally, multiplying U by a complex phase $e^{i\alpha}$ for $\alpha \in \mathbb{R}$ preserves the unitarity of U and the orthonormality of the colums. Combining these facts gives the form of (4.12) as desired. □

Exercise 4.10: $X - Y$ decomposition of rotations

Give a decomposition analogous to Theorem 4.1 but using R_x instead of R_z.

Solution

Concepts Involved: Unitary Operators, Rotations, Gate Decomposition

By Theorem 4.1, we have the decomposition of the single-qubit unitary HUH for $\alpha, \beta, \gamma.\delta \in \mathbb{R}$:

$$HUH = e^{i\alpha} R_z(\beta) R_y(\gamma) R_z(\delta)$$

Conjugating both sides by H and inserting $H^2 = I$ in between each of the rotations on the RHS, we obtain:

$$H^2 U H^2 = U = e^{i\alpha}(HR_z(\beta)H)(HR_y(\gamma)H)(HR_z(\delta)H)$$

Now using the result of Ex. 4.13, $HR_x(\theta)H = R_z(\theta)$ and $HR_y(\theta)H = R_y(-\theta)$ and so:

$$U = e^{i\alpha} R_x(\beta) R_y(-\gamma) R_x(\delta)$$

which is the desired decomposition of U. □

Exercise 4.11

($*$) Suppose $\hat{\mathbf{m}}$ and $\hat{\mathbf{n}}$ are non-parallel real unit vectors in three dimensions. Use Theorem 4.1 to show that an arbitrary single qubit unitary U may be written

$$U = e^{i\alpha} R_{\hat{\mathbf{n}}}(\beta) R_{\hat{\mathbf{m}}}(\gamma) R_{\hat{\mathbf{n}}}(\delta)$$

The stated verison of this exercise is incorrect. The above form only holds if $\hat{\mathbf{n}}, \hat{\mathbf{m}}$ are orthogonal. Otherwise, the correct decomposition is

$$U = e^{i\alpha} R_{\hat{\mathbf{n}}}(\beta_1) R_{\hat{\mathbf{m}}}(\gamma_1) R_{\hat{\mathbf{n}}}(\beta_2) R_{\hat{\mathbf{m}}}(\gamma_2) \ldots$$

Solution

Concepts Involved: Unitary Operators, Rotations, Gate Decomposition

Any $U \in \mathrm{U}(2)$ can be written as $U = e^{i\alpha}V$, where $V \in \mathrm{SU}(2)$, so it suffices to express arbitrary $V \in \mathrm{SU}(2)$ using $R_{\hat{n}}$ and $R_{\hat{m}}$.
Define

$$W(\gamma) := R_{\hat{n}}(\pi) R_{\hat{m}}(\gamma) R_{\hat{n}}(\pi) R_{\hat{m}}(-\gamma)$$

We show that $W(\gamma) = R_{\hat{q}}(\theta)$ for some axis $\hat{q} \perp \hat{n}$. Compute

$$\begin{aligned}
\mathrm{Tr}(W(\gamma)\, \hat{n}\cdot\vec{\sigma}) &= \mathrm{Tr}(R_{\hat{m}}(\gamma) R_{\hat{n}}(\pi) R_{\hat{m}}(-\gamma)\, \hat{n}\cdot\vec{\sigma}\, R_{\hat{n}}(\pi)) \\
&= -\mathrm{Tr}(R_{\hat{m}}(\gamma)(\hat{n}\cdot\vec{\sigma}) R_{\hat{m}}(-\gamma)) \\
&= -\mathrm{Tr}(\hat{n}\cdot\vec{\sigma}) = 0
\end{aligned}$$

This implies $\hat{q}\cdot\hat{n} = 0$, so $W(\gamma)$ is a rotation about a direction orthogonal to $\hat{n}$. Denote this axis by $\hat{n}_\perp$. If we choose γ such that the resulting rotation angle θ satisfies $\theta/\pi \notin \mathbb{Q}$, then the sequence $\{W(\gamma)^k\}$ is dense in $\{R_{\hat{n}_\perp}(\varphi)\}$, so arbitrary $R_{\hat{n}_\perp}(\varphi)$ can be implemented using only $R_{\hat{n}}$ and $R_{\hat{m}}$.
Now recall the Euler decomposition

$$V = R_z(\alpha) R_x(\beta) R_z(\gamma)$$

Let V be any unitary such that

$$V(\hat{n}\cdot\vec{\sigma})V^\dagger = Z, \quad V(\hat{n}_\perp\cdot\vec{\sigma})V^\dagger = X$$

Then

$$V^\dagger R_z(\alpha) R_x(\beta) R_z(\gamma) V = R_{\hat{n}}(\alpha) R_{\hat{n}_\perp}(\beta) R_{\hat{n}}(\gamma)$$

and each $R_{\hat{n}_\perp}$ factor can be expressed as a product of $R_{\hat{n}}$ and $R_{\hat{m}}$, as shown above.
Therefore, every $U \in \mathrm{U}(2)$ can be decomposed into a finite sequence of rotations about any two non-parallel real axes $\hat{n}$ and $\hat{m}$. $\square$

Exercise 4.12

Give $A, B, C,$ and α for the Hadamard gate.

Solution

Concepts Involved: Gate Decomposition

Recall that any single qubit unitary U can be written as $U = e^{i\alpha}AXBXC$ where $ABC = I$ and $\alpha \in \mathbb{R}$.

First, observe that we can write:

$$H = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = i\begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix}\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = e^{i\pi/2}R_z(\pi)R_y(-\pi/2)R_z(0)$$

so defining A, B, C according to the proof of Corollary 4.2, we have:

$$A = R_z(\pi)R_y(-\pi/4)$$
$$B = R_y(\pi/4)R_z(-\pi/2)$$
$$C = R_z(-\pi/2)$$

and $\alpha = \frac{\pi}{2}$. □

Exercise 4.13: Circuit identities

It is useful to be able to simplify circuits by inspection, using well-known identities. Prove the following three identities:

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

Solution

Concepts Involved: Pauli Operators

By computation, we find:

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

$$HYH = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 0 & -2i \\ 2i & 0 \end{bmatrix} = -\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} = -Y$$

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

□

Remark: Notice once we have proved $HXH = Z$, we can directly say $HZH = H(HXH)H = X$ as $H^2 = I$. If one wants to prove everything algebraically, the following calculation suffices.

$$HXH := \frac{1}{2}(X+Z)X(X+Z) = \frac{1}{2}(I+ZX)(X+Z) = \frac{1}{2}(X+Z+Z+XZX) = Z$$

$$HYH := \frac{1}{2}(X+Z)Y(X+Z) = \frac{1}{2}(XY+ZY)(X+Z) = \frac{1}{2}(XYX+ZXY+ZYX+ZYZ) = -Y$$

Exercise 4.14

Use the previous exercise to show that $HTH = R_x(\pi/4)$, up to a global phase.

Solution

Concepts Involved: Rotations

From Exercise 4.3, we know that $T = R_z(\pi/4)$ up to a global phase $e^{-i\pi/8}$. We hence have:

$$\begin{aligned}
HTH &= e^{-i\pi/8} H R_z(\pi/4) H \\
&= e^{-i\pi/8} H \left(\cos\left(\frac{\pi}{8}\right) I - i\sin\left(\frac{\pi}{8}\right) Z\right) H \\
&= e^{-i\pi/8} \left(\cos\left(\frac{\pi}{8}\right) I - i\sin\left(\frac{\pi}{8}\right) X\right) \\
&= e^{-i\pi/8} R_x(\pi/4)
\end{aligned}$$

where in the second last equality we use the previous exercise, as well as the fact that $HIH = H^2 = I$ from Exercise 2.52. □

Exercise 4.15: Composition of single qubit operations

The Bloch representation gives a nice way to visualize the effect of composing two rotations.

(1) Prove that if a rotation through an angle β_1 about the axis $\hat{\mathbf{n}}_1$ is followed by a rotation through an angle β_2 about an axis $\hat{\mathbf{n}}_2$, then the overall rotation is through an angle β_{12} about an axis $\hat{\mathbf{n}}_{12}$ given by

$$\begin{aligned}
c_{12} &= c_1 c_2 - s_1 s_2 \hat{\mathbf{n}}_1 \cdot \hat{\mathbf{n}}_2 \\
s_{12}\hat{\mathbf{n}}_{12} &= s_1 c_2 \hat{\mathbf{n}}_1 + c_1 s_2 \hat{\mathbf{n}}_2 - s_1 s_2 \hat{\mathbf{n}}_2 \times \hat{\mathbf{n}}_1,
\end{aligned}$$

where $c_i = \cos(\beta_i/2)$, $s_i = \sin(\beta_i/2)$, $c_{12} = \cos(\beta_{12}/2)$, and $s_{12} = \sin(\beta_{12}/2)$.

(2) Show that if $\beta_1 = \beta_2$ and $\hat{\mathbf{n}}_1 = \hat{\mathbf{z}}$ these equations simplify to

$$\begin{aligned}
c_{12} &= c^2 - s^2 \hat{\mathbf{z}} \cdot \hat{\mathbf{n}}_2 \\
s_{12}\hat{\mathbf{n}}_{12} &= sc(\hat{\mathbf{z}} + \hat{\mathbf{n}}_2) - s^2 \hat{\mathbf{n}}_2 \times \hat{\mathbf{z}}
\end{aligned}$$

Solution

Concepts Involved: Rotations, Bloch Sphere

(1) It suffices to show that $R_{\hat{\mathbf{n}}_2}(\beta_2)R_{\hat{\mathbf{n}}_1}(\beta_1)$ is equivalent to $R_{\hat{\mathbf{n}}_{12}}(\beta_{12})$.

$$\begin{aligned} R_{\hat{\mathbf{n}}_2}(\beta_2)R_{\hat{\mathbf{n}}_1}(\beta_1) &= (c_2 I - i s_2 \hat{\mathbf{n}}_2 \cdot \boldsymbol{\sigma}) \cdot (c_1 I - i s_1 \hat{\mathbf{n}}_1 \cdot \hat{\sigma}) \\ &= c_2 c_1 I - i\left(c_1 s_2 \hat{\mathbf{n}}_2 \cdot \boldsymbol{\sigma} + c_2 s_1 \hat{\mathbf{n}}_1 \cdot \boldsymbol{\sigma}\right) - s_2 s_1 \underbrace{(\hat{\mathbf{n}}_2 \cdot \boldsymbol{\sigma}) \cdot (\hat{\mathbf{n}}_1 \cdot \boldsymbol{\sigma})}_{(\hat{\mathbf{n}}_2 \cdot \hat{\mathbf{n}}_1) I + i (\hat{\mathbf{n}}_2 \times \hat{\mathbf{n}}_1) \cdot \boldsymbol{\sigma}} \\ &= \left[c_2 c_1 - s_2 s_1 \left(\hat{\mathbf{n}}_2 \cdot \hat{\mathbf{n}}_1\right)\right] I - i\left[c_1 s_2 \hat{\mathbf{n_2}} + c_2 s_1 \hat{\mathbf{n_1}} + s_2 s_1 \left(\hat{\mathbf{n}}_2 \times \hat{\mathbf{n}}_1\right)\right] \cdot \boldsymbol{\sigma} \end{aligned}$$

Identifying this operation to a single rotation $R_{\hat{\mathbf{n}}_{12}}(\beta_{12}) \equiv c_{12} I - i s_{12} \hat{\mathbf{n}}_{12} \cdot \boldsymbol{\sigma}$, we arrive at the required relations (up to a presumable typesetting error)

$$\begin{aligned} c_{12} &= c_2 c_1 - s_2 s_1 \left(\hat{\mathbf{n}}_2 \cdot \hat{\mathbf{n}}_1\right) \\ s_{12} \hat{\mathbf{n}}_{12} &= c_1 s_2 \hat{\mathbf{n_2}} + c_2 s_1 \hat{\mathbf{n_1}} + s_2 s_1 \left(\hat{\mathbf{n}}_2 \times \hat{\mathbf{n}}_1\right) \end{aligned}$$

(2) Setting $\beta_1 = \beta_2$ and $\hat{\mathbf{n}}_1 = \hat{\mathbf{z}}$ in the formulas proven above combined with the fact that $c = c_1 = \cos\big(\beta_1/2\big) = \cos\big(\beta_2/2\big) = c_2$ (and similiarly $s = s_1 = s_2$), we have:

$$\begin{aligned} c_{12} &= c^2 - s^2 \hat{\mathbf{z}} \cdot \hat{\mathbf{n}}_2 \\ s_{12} \hat{\mathbf{n}}_{12} &= sc\hat{\mathbf{z}} + cs\hat{\mathbf{n}}_2 - s^2 \hat{\mathbf{n}}_2 \times \hat{\mathbf{z}} = sc(\hat{\mathbf{z}} + \hat{\mathbf{n}}_2) - s^2 \hat{\mathbf{n}}_2 \times \hat{\mathbf{z}}. \end{aligned}$$

□

Remark: For the sake of completeness, we provide a proof of the identity used in part 1 of the solution. First note the familiar Pauli matrix relation $\sigma_i \sigma_j = \delta_{ij} I + i\epsilon_{ijk}\sigma_k$ (Exercise 2.43). Now massaging this equation gives

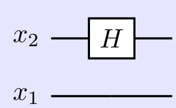
$$\begin{aligned} a_i \sigma_i b_j \sigma_j &= a_i b_j \delta_{ij} + i\left(a_i b_j \epsilon_{ijk}\right)\sigma_k \\ &= (\mathbf{a} \cdot \mathbf{b})\, I + i\,(\mathbf{a} \times \mathbf{b})_k\, \sigma_k, \end{aligned}$$

where we have used standard Einstein index notation. Thus in matrix form, we have

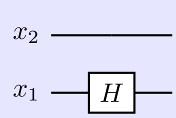
$$(\mathbf{a} \cdot \boldsymbol{\sigma}) \cdot (\mathbf{b} \cdot \boldsymbol{\sigma}) = (\mathbf{a} \cdot \mathbf{b})\, I \,+\, i\,(\mathbf{a} \times \mathbf{b}) \cdot \boldsymbol{\sigma}\,.$$

Exercise 4.16

What is the 4×4 unitary matrix for the circuit



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



Solution

Concepts Involved: Unitary Operators, Tensor Products.

The unitary matrix for the first circuit is given by:

$$I_1 \otimes H_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \otimes \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix}.$$

The unitary matrix for the second circuit is given by:

$$H_1 \otimes I_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix}$$

□

Exercise 4.17: Building a CNOT from controlled-Z gates

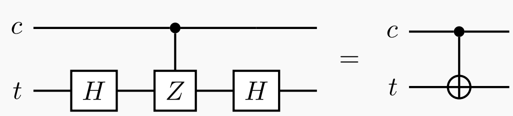
Construct a CNOT gate from one controlled-Z gate, that is, the gate whose action in the computational basis is specified by the unitary matrix

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

Solution

Concepts Involved: Controlled Operations

We showed in Exercise 4.13 that $HZH = X$. Hence, to obtain a CNOT gate from a single controlled Z gate, we can conjugate the target qubit with Hadamard gates:



We can verify this via matrix multiplication, using the result from the previous exercise:

$$\begin{aligned}
(I_1 \otimes H_2)(CZ_{1,2})(I_1 \otimes H_2) &= \frac{1}{2}\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix} \\
&= \frac{1}{2}\begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 2 & 0 \end{bmatrix} \\
&= CX_{1,2}
\end{aligned}$$

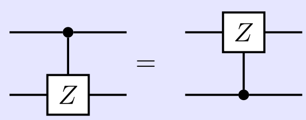
□

Remark:

$$\begin{aligned}
CX_{1,2} :&= |0\rangle\langle 0| \otimes I + |0\rangle\langle 0| \otimes X \\
&= |0\rangle\langle 0| \otimes HH + |0\rangle\langle 0| \otimes HZH \\
&=: (I \otimes H)(CZ_{1,2})(I \otimes H).
\end{aligned}$$

Exercise 4.18

Show that



Solution

Concepts Involved: Controlled Operations

It suffices to verify that the two gates have the same effect on the 2-qubit computational basis states (as it will then follow by linearity that they will have the same effect on any such superposition of the basis states). Checking the 8 necessary cases, we then have:

$$\begin{aligned}
&CZ_{1,2}(|0\rangle_1 \otimes |0\rangle_2) = |0\rangle_1 \otimes |0\rangle_2 \\
&CZ_{2,1}(|0\rangle_1 \otimes |0\rangle_2) = |0\rangle_1 \otimes |0\rangle_2 \\
&CZ_{1,2}(|1\rangle_1 \otimes |0\rangle_2) = |1\rangle_1 \otimes Z|0\rangle_2 = |1\rangle_1 \otimes |0\rangle_2 \\
&CZ_{2,1}(|1\rangle_1 \otimes |0\rangle_2) = |1\rangle_1 \otimes |0\rangle_2 \\
&CZ_{1,2}(|0\rangle_1 \otimes |1\rangle_2) = |0\rangle_1 \otimes |1\rangle_2 \\
&CZ_{2,1}(|0\rangle_1 \otimes |1\rangle_2) = Z|0\rangle_1 \otimes |1\rangle_2 = |0\rangle_1 \otimes |1\rangle_2 \\
&CZ_{1,2}(|1\rangle_1 \otimes |1\rangle_2) = |1\rangle_1 \otimes Z|1\rangle_2 = |1\rangle_1 \otimes -|1\rangle_2 = -(|1\rangle_1 \otimes |1\rangle_1) \\
&CZ_{2,1}(|1\rangle_1 \otimes |1\rangle_2) = Z|1\rangle_1 \otimes |1\rangle_2 = -|1\rangle_1 \otimes |1\rangle_2 = -(|1\rangle_1 \otimes |1\rangle_1)
\end{aligned}$$

from which we observe equality for each. The claim follows. □

Remark: More compactly, we have $CZ_{1,2}|b_1b_2\rangle = |b_1\rangle \otimes Z^{b_1}|b_2\rangle = (-1)^{b_1 \cdot b_2}|b_1b_2\rangle$ for computational basis states $b_1, b_2 \in \{0,1\}$.
Using this form we can write

$$\begin{aligned} CZ_{1,2}|b_1b_2\rangle &= (-1)^{b_1 \cdot b_2}|b_1b_2\rangle \\ &= (-1)^{b_2 \cdot b_1}|b_1b_2\rangle \\ &= Z^{b_2}|b_1\rangle \otimes |b_2\rangle \\ &=: CZ_{2,1}|b_1b_2\rangle. \end{aligned}$$

Exercise 4.19: `CNOT` action on unitary 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.

Solution

Concepts Involved: Controlled Operations, Density Operators

Let ρ be an arbitrary density matrix corresponding to a 2-qubit state. In the computational basis, we can write ρ as:

$$\rho \cong \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}.$$

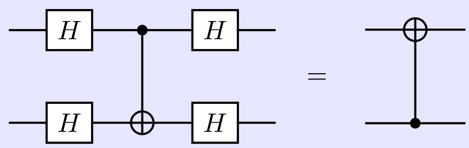
Studying the action of the `CNOT` gate on this density matrix, we calculate:

$$\begin{aligned} CX_{1,2}\,\rho\, CX_{1,2} &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{41} & a_{42} & a_{43} & a_{34} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} a_{11} & a_{12} & a_{14} & a_{13} \\ a_{21} & a_{22} & a_{24} & a_{23} \\ a_{41} & a_{42} & a_{44} & a_{33} \\ a_{31} & a_{32} & a_{34} & a_{33} \end{bmatrix} \end{aligned}$$

□

Exercise 4.20: `CNOT` basis transformations

Unlike ideal classical gates, ideal quantum gates do not have (as electrical engineers say) 'high-impedance' 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:

$$
\begin{aligned}
|+\rangle|+\rangle &\mapsto |+\rangle|+\rangle \\
|-\rangle|+\rangle &\mapsto |-\rangle|+\rangle \\
|+\rangle|-\rangle &\mapsto |-\rangle|-\rangle \\
|-\rangle|-\rangle &\mapsto |+\rangle|-\rangle.
\end{aligned}
$$

Thus, with respect to 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!

Solution

Concepts Involved: Controlled Operations

First, we have:

$$
H_1 \otimes H_2 = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \otimes \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}
$$

Now conjugating $\mathsf{CNOT}_{1,2}$ under $H_1 \otimes H_2$, we have:

$$\begin{aligned}
(H_1 \otimes H_2)CX_{1,2}(H_1 \otimes H_2) &= \frac{1}{4}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \\
&= \frac{1}{4}\begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 \\ 0 & 0 & 4 & 0 \\ 0 & 4 & 0 & 0 \end{bmatrix} \\
&= CX_{2,1}
\end{aligned}$$

which proves the circuit identity. We know already that:

$$\begin{aligned}
CX_{2,1}|0\rangle|0\rangle &= |0\rangle|0\rangle \\
CX_{2,1}|1\rangle|0\rangle &= |1\rangle|0\rangle \\
CX_{2,1}|0\rangle|1\rangle &= |1\rangle|1\rangle \\
CX_{2,1}|1\rangle|1\rangle &= |0\rangle|1\rangle
\end{aligned}$$

so using the proven circuit identity and the fact that $H|0\rangle = |+\rangle, H|1\rangle = |-\rangle$, we obtain the map:

$$\begin{aligned}
|+\rangle|+\rangle &\mapsto |+\rangle|+\rangle \\
|-\rangle|+\rangle &\mapsto |-\rangle|+\rangle \\
|+\rangle|-\rangle &\mapsto |-\rangle|-\rangle \\
|-\rangle|-\rangle &\mapsto |+\rangle|-\rangle
\end{aligned}$$

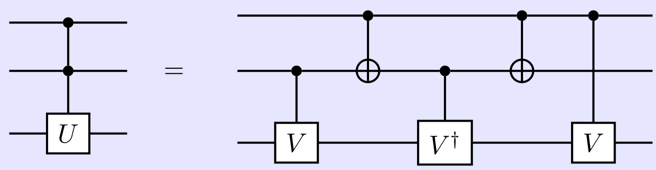
which is exactly what we wanted to prove. □

Remark: Algebraically,

$$\begin{aligned}
(H \otimes H)CX_{1,2}(H \otimes H) &= (H \otimes H)(I \otimes H)(CZ_{1,2})(I \otimes H)(H \otimes H) \\
&= (H \otimes I)(CZ_{1,2})(H \otimes I) \\
&= (H \otimes I)(CZ_{2,1})(H \otimes I) \\
&= CX_{2,1}.
\end{aligned}$$

Exercise 4.21

Verify that Figure 4.8 implements the $C^2(U)$ operation.



Solution

Concepts Involved: Controlled Operations, Gate Decomposition

Let us analyse the second circuit. First, note that circuit in symmetric in the top two control registers. If $c_1 = 0$, the transformation applied to the final register is either I or $VV^\dagger = I$. The same is the case if $c_2 = 0$. When $c_1 = c_2 = 1$, the unitary applied to the final qubit is $VV = U$. $\square$

Remark: Note that the $V^\dagger$ in not applied for $c_1 = c_2 = 1$ because of the first `CNOT` gate. This is the clever part of the construction.

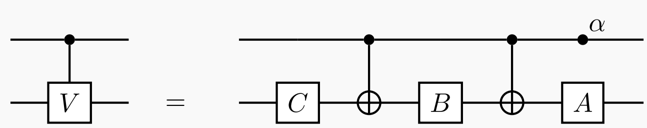
Exercise 4.22

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-`NOT`s.

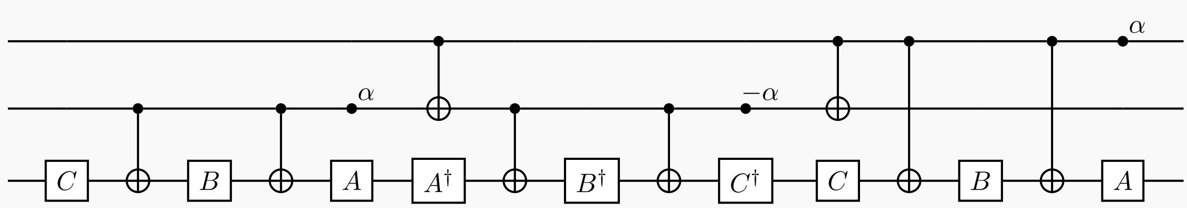
Solution

Concepts Involved: Unitary Operators, Controlled Operations, Gate Decomposition

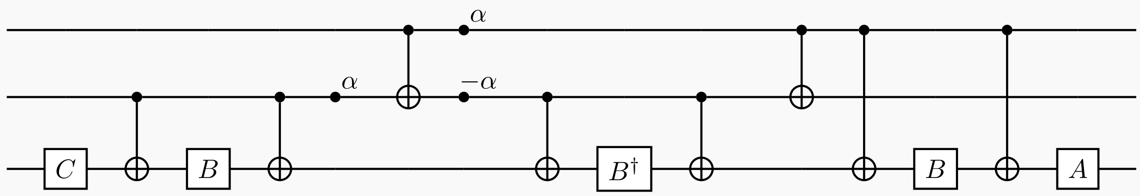
We use Figure 4.8 as our starting point (as depicted in Ex. 4.21). We then replace the controlled-V operations with the construction of Fig. 4.6:



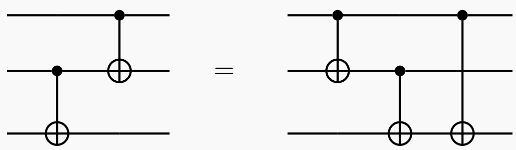
with the decomposition $V = \exp(i\alpha)AXBXC$ and $V^\dagger = \exp(-i\alpha)C^\dagger XB^\dagger XA^\dagger$. This gives the $C^2(U)$ circuit:



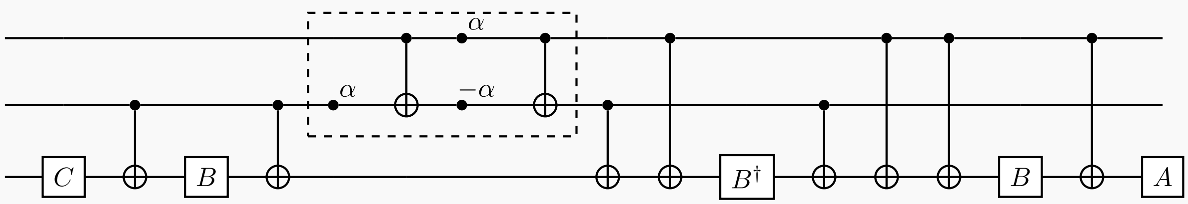
On the last qubit we have $AA^\dagger = C^\dagger C = I$, which removes these gates. Further, the α-phase gates commute through the controls, so let us move these towards the front of the circuit:



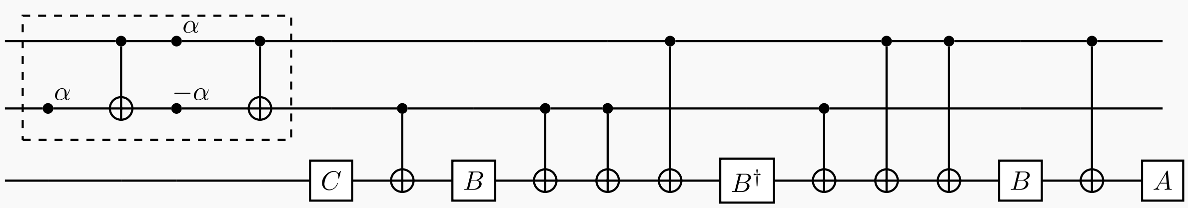
Then we make use of the identity:



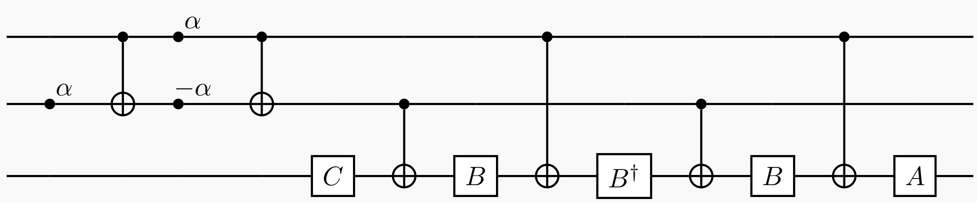
which follows from the fact that both circuits leave the first qubit unchanged, both circuits apply X to the second qubit only if the first qubit is $|1\rangle$, and both circuits apply X to the third qubit only if the second qubit is $|1\rangle$.
With this identity, we move the sixth CNOT past the fifth/fourth:



The part of the circuit in the dashed box commutes with controls on the second qubit, so we can move it to the start of the circuit:



The fifth/sixth and ninth/tenth CNOTs mutually cancel, leaving us with:



which is an implementation of $C^2(U)$ only using eight one-qubit gates and six CNOTs, as claimed. $\square$

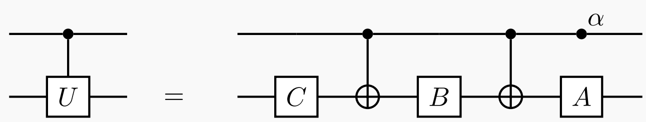
Exercise 4.23

Construct a $C^1(U)$ gate for $U = R_x(\theta)$ and $U = R_y(\theta)$, using only `CNOT` and single qubit gates. Can you reduce the number of single qubit gates needed in the construction from three to two?

Solution

Concepts Involved: Unitary Operators, Controlled Operations

We again apply the construction of Fig. 4.6 (see Ex. 4.22:



with $U = e^{i\alpha}AXBXC$ with $ABC = I$.
We start with $R_y(\theta)$, which we can write as:

$$R_y(\theta) = e^{i\cdot 0}R_z(0)R_y(\theta)R_z(0) = e^{i\cdot 0}R_y(\theta/2)XR_y(-\theta/2)XI$$

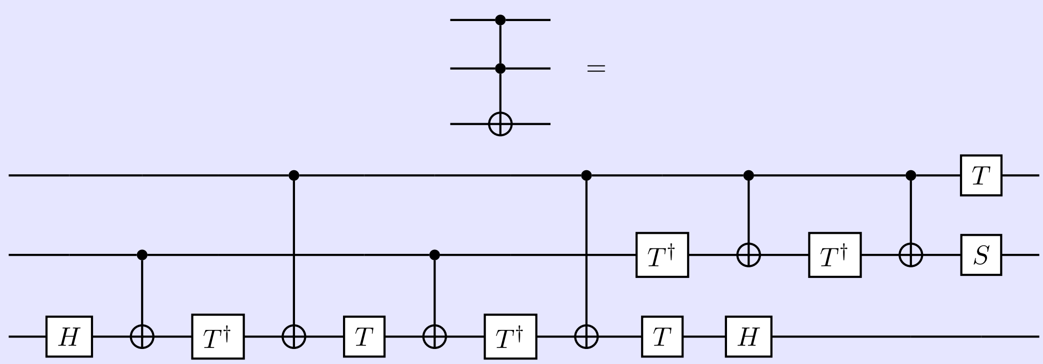
So we have the $C^1(R_y(\theta))$ gate as depicted above with $\alpha = 0$, $A = R_y(\theta/2)$, $B = R_y(-\theta/2)$, $C = I$. This is a construction with only 2 single qubit gates required.
For $R_x(\theta)$, we can write:

$$R_x(\theta) = R_z(\frac{\pi}{2})R_y(\theta)R_z(-\frac{\pi}{2}) = e^{i0}R_z(\frac{\pi}{2})R_y(\frac{\theta}{2})XR_y(-\frac{\theta}{2})R_z(-\frac{\pi}{2})XI$$

So we we have the $C^1(R_x(\theta))$ gate as depicted above with $\alpha = 0, A = R_z(\frac{\pi}{2})R_y(\frac{\theta}{2}), B = R_y(-\frac{\theta}{2})R_z(-\frac{\pi}{2}), C = I$. This again is a construction with only 2 single qubit gates required. $\square$

Exercise 4.24

Verify that Figure 4.9 implements the Toffoli gate.



Solution

Concepts Involved: Controlled Operations

As usual, by linearity it suffices to analyze the action of the circuit on computational basis states.
First, let us analyze the action of the circuit on the first two qubits. If the first qubit is in state $|0\rangle$, then the CNOTs drop out, and the T gate leaves the first qubit invariant. The action on the second qubit is $T^\dagger T^\dagger S = S^\dagger S = I$ and so it is left invariant.
If the second qubit is in state $|1\rangle$, then the CNOTs apply. The first qubit transforms as $T|1\rangle = e^{i\pi/4}|1\rangle$. The action on the second qubit is:

$$SXT^\dagger XT^\dagger = Se^{-i\pi/4}TT^\dagger = e^{-i\pi/4}S$$

where we use the identity:

$$XT^\dagger X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & e^{-i\pi/4} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} e^{-i\pi/4} & 0 \\ 0 & 1 \end{bmatrix} = e^{-i\pi/4} \begin{bmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{bmatrix} = e^{-i\pi/4}T$$

The $e^{-i\pi/4}$ cancels the phase on the first qubit, leaving just the action $S|0\rangle = |0\rangle$ if the second qubit is $|0\rangle$ or $S|1\rangle = i|1\rangle$ if the second qubit is $|1\rangle$. In summary, the action on the first two qubits is $|11\rangle \to i|11\rangle$ for $|11\rangle$ and the other basis states are left invariant.
Next, we analyze the action of the circuit on the last qubit. If the first qubit is in $|0\rangle$, then the second/fourth CNOTs drop out, leading to cancellation of the $T^\dagger T$s on the last qubit, followed by the cancellations of the first/third CNOTs and the Hadamards. Similarly, if the second qubit is in $|0\rangle$, the first/third CNOTs drop out, leading to a similar cascade of cancellations of all gates on the last qubit. The only interesting case is thus when both the first and second qubit are in the $|1\rangle$ state, in which case the resulting gate on the third qubit is:

$$HTXT^\dagger XTXT^\dagger XH = HTe^{-i\pi/4}TTe^{-i\pi/4}TH = e^{-i\pi/2}HZH = -iX$$

where we have used the T conjugation identity and $T^4 = Z$. Thus, if the first two qubits are in the $|11\rangle$ state, we apply $-iX$ to the last qubit, the phase of which is cancelled by the phase i on the second qubit. In summary, the given circuit applies X to the third qubit if the first two qubits are in $|11\rangle$, and does nothing otherwise. This is exactly the Toffoli. □

Exercise 4.25: Fredkin gate construction

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

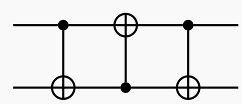
$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

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?

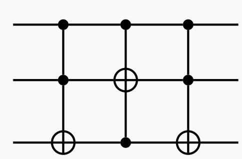
Solution

Concepts Involved: Controlled Operations

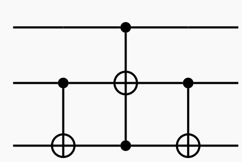
1. We recall the SWAP gate composed of three `CNOT` gates (Figure 1.7 in N&C):



Since the Fredkin gate is simply the SWAP gate controlled by an additional qubit, we can simply control each of the three gates in the above construction, as suggested by the hint. In particular we swap qubits 2/3 if the first qubit is in the $|1\rangle$ state, so:

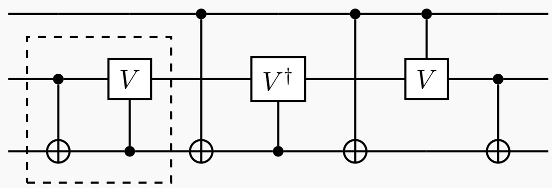


2. We claim that:



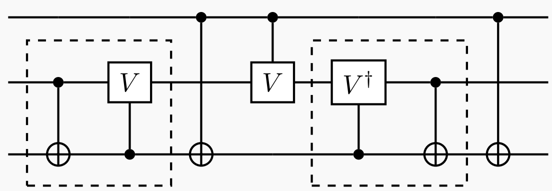
is also the Fredkin gate. In both circuits, the first qubit is left invariant (only acts as a control). The action on the second qubit is unchanged by removing the two Toffolis. The action on the third qubit requires a little more thought. If the first qubit is $|0\rangle$, then in the original circuit all three gates do not fire, in the updated circuit the middle Toffoli drops out, leading the two CNOTs to cancel. If the first qubit is $|1\rangle$ state, in both the original and updated circuit we are left with the three-CNOT SWAP gate on the remaining two qubits. Hence we have argued that the action of the two circuits on the computational basis states, and hence all states, are the same.

3. We use the construction of Fig. 4.8 (See Ex. 4.21) to rewrite the remaining central Toffoli in the Fredkin as:



where $V = \frac{(1-i)(I+iX)}{2}$. Viewing the dashed box as a single two-qubit gate, we see that this construction only requires six two qubit-gate.

4. Noting that the last controlled-V and CNOT gate can be commuted freely to the center of the circuit, we obtain:

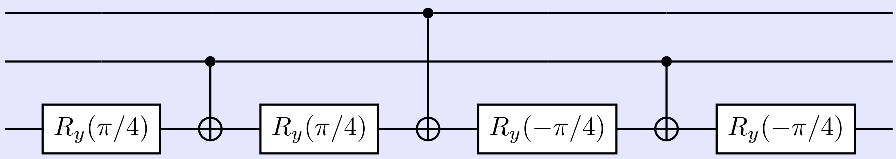


The second dashed box can also be viewed as a single two-qubit gate, and thus we can reduce the two-qubit gate count to five.

□

Exercise 4.26

Show that the circuit:



differs by a Toffoli gate only by relative phases. That is, the circuit that takes $|c_1, c_2, t\rangle$ to $e^{i\theta(c_1,c_2,t)}|c_1, c_2, t \oplus c_1 \cdot c_2\rangle$, where $e^{i\theta(c_1,c_2,t)}$ is some relative phase factor. Such gates can be sometimes be useful in experimental implementations, where it may be much easier to implement a gate that is the same as the Toffoli gate up to relative phases than it is to do the Toffoli directly.
Note: Many printings of the book are missing that the rotation angles should be $\pi/4$.

Solution

Concepts Involved: Controlled Operations, Rotations

First, much like the Toffoli, the above circuit leaves the first/second qubit invariant. What is left to check is the action on the third qubit.
If the first qubit is $|0\rangle$, then the central `CNOT` drops out, and then the inner Y-rotations cancel, followed by the remaining `CNOT`s and the outer Y-rotations. If the first qubit is $|1\rangle$ and the second qubit is $|0\rangle$, then the first/last `CNOT`s drop out, leaving us with a X conjugated by $R_y(\pi/2)$:

$$R_y(\pi/2)XR_y(-\pi/2) = R_y(\pi/2)R_y(\pi/2)X = R_y(\pi)X = -iYX = -Z$$

So the action on the third qubit is $-Z$, which applies a relative phase $\mp$ to a computational basis state. If both qubit 1 and 2 are in $|1\rangle$, then we have:

$$R_y(\pi/4)XR_y(\pi/4)XR_y(-\pi/4)XR_y(-\pi/4) = R_y(\pi/4)R_y(-\pi/4)X^2R_y(-\pi/4)R_y(\pi/4)X = X$$

we see the Y-rotations cancel, leaving just a single X-gate on the third register, as is required for a Toffoli.

□

Exercise 4.27

($*$) Using just CNOTs and Toffoli gates, construct a quantum circuit to perform the transformation

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

This kind of partial cyclic permutation operation will be useful later, in Chapter 7.

Solution

Concepts Involved: Controlled Operations, Permutations

The key observation is that CNOT and Toffoli gates are permutation matrices in the computational basis. Let us write permutations in cyclic notation, e.g. for $n = 8$ the expression $(123)(45)$ denotes the permutation which sends $1 \to 2 \to 3 \to 1$, $4 \to 5 \to 4$ and 6/7/8 to themselves. The multiplication of permutations via composition. Labeling the computational basis states by $1 = |000\rangle$, $2 = |001\rangle$, $3 = |010\rangle \ldots$ in their binary representation, the $CX_{1,2}$ controlled gate (with qubit 1 as control and qubit 2 as the target) represents the permutation:

$$CX_{1,2} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \cong (57)(68)$$

We can construct the analogous expressions for all possible CNOT and Toffoli gates on three qubits:

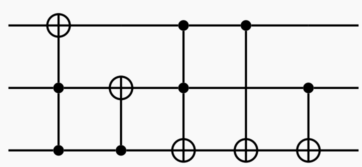
$$\begin{aligned} CX_{1,3} &= (56)(78) \\ CX_{2,1} &= (37)(48) \\ CX_{2,3} &= (34)(78) \\ CX_{3,1} &= (26)(48) \\ CX_{3,2} &= (24)(68) \end{aligned}$$

$$\begin{aligned}\text{Toffoli}_1 &= (48)\\ \text{Toffoli}_2 &= (68)\\ \text{Toffoli}_3 &= (78)\end{aligned}$$

The transformation given in the question is the permutation (2345678) and now we can reason from the permutations directly. In particular, observe that:

$$\begin{aligned}(2345678) &= (34)(56)(78)(2468)\\ &= (34)(56)(78)(24)(68)(48)\\ &= (34)(56)(78)^2(78)(24)(68)(48)\\ &= \big[(34)(78)\big]\,\big[(56)(78)\big]\,\big[(78)\big]\,\big[(24)(68)\big]\,\big[(48)\big]\\ &= CX_{2,3}CX_{1,3}\text{Toffoli}_3 CX_{3,2}\text{Toffoli}_1\end{aligned}$$

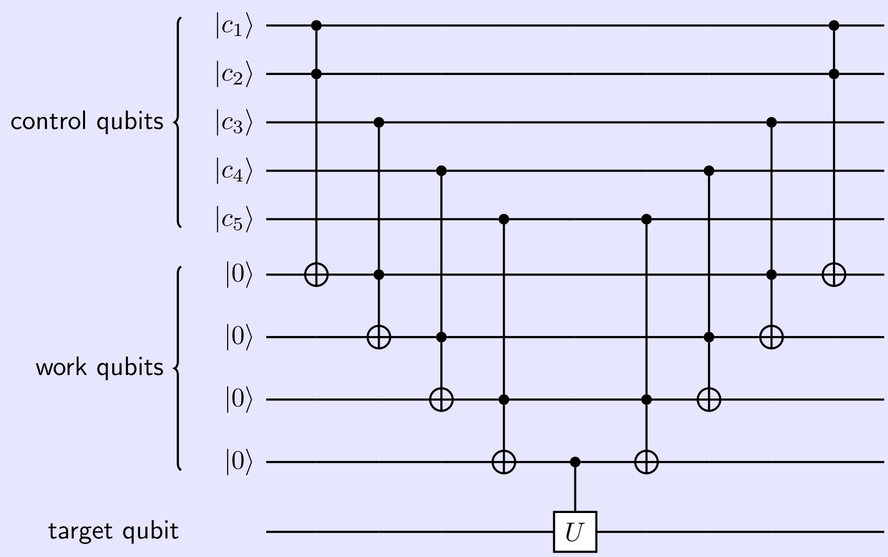
where the first two lines follow by permutation composition, the third follows by the fact that cycles square to the identity, and the fourth by the fact that disjoint cycles can be freely commuted. As a circuit:



□

Exercise 4.28

($*$) For $U = V^2$ with V unitary, construct a $C^5(U)$ gate analogous to that in Figure 4.10 (reproduced below), but using no work qubits. You may use controlled-V and controlled-$V^\dagger$ gates.

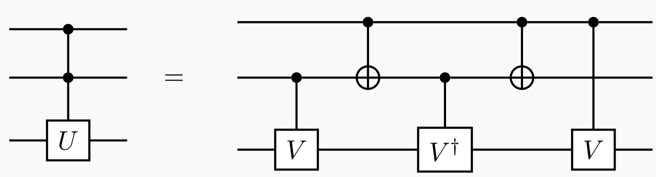


This question is impossible as stated - a proof based on determinants of the available gates (see this stackexchange post for details) shows that it is not possible.

Solution

Concepts Involved: Controlled Operations, Gate Decomposition

Given that the exercise is incorrect as stated, we solve two alternative versions. Both generalize the $C^2(U)$ operation of Fig 4.8 (pictured again below) to n controls, using two different methods.



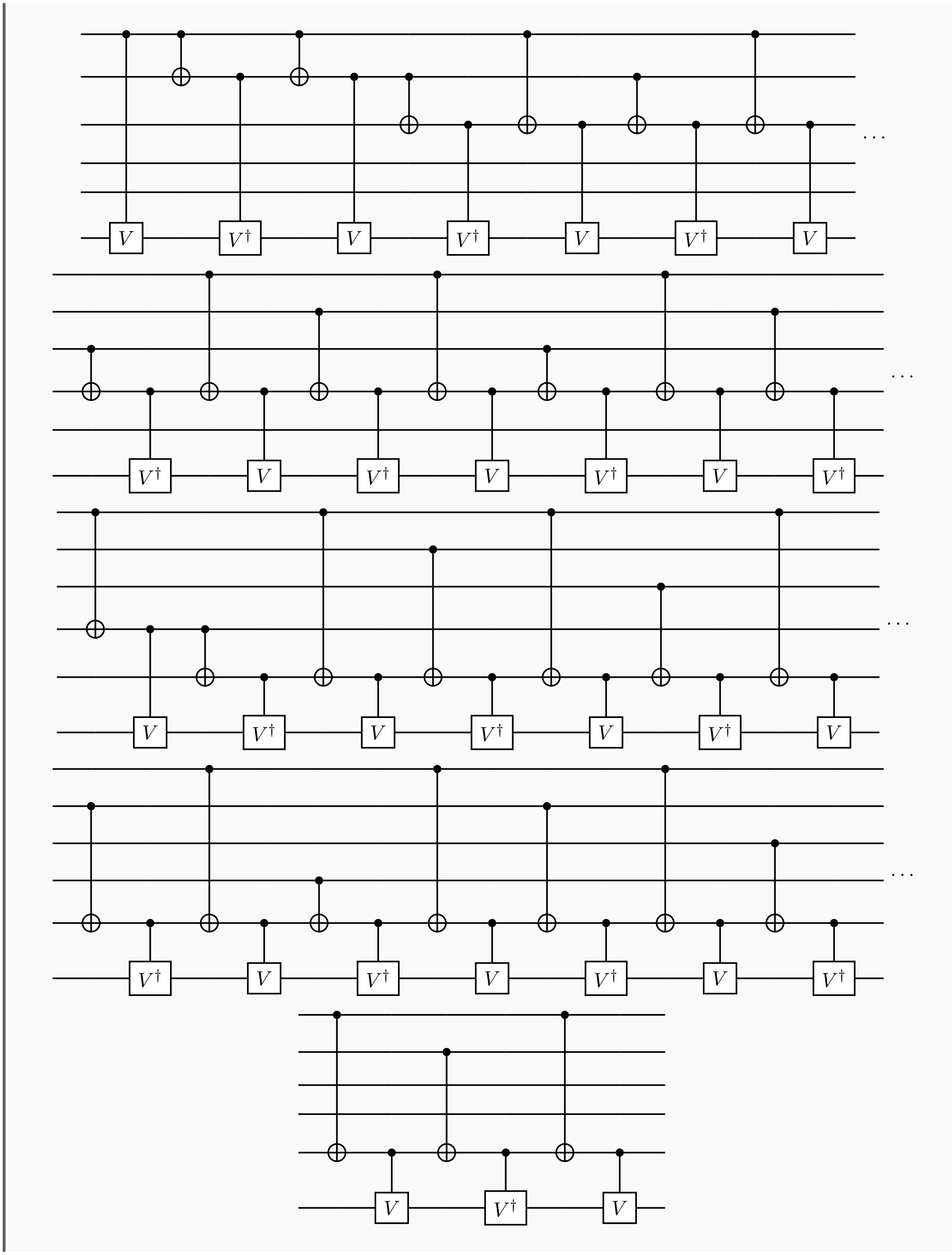
In the first version, we assume we have access to $C^1(X), C^1(V)$, and $C^1(V^\dagger)$ gates for $V^{2^{n-1}} = U$ (rather than $V^2 = U$. Then, by using these gates to create a quantum circuit for the logical identity

$$\underbrace{\sum_{k_1} x_{k_1}}_{V} - \underbrace{\sum_{k_1<k_2} (x_{k_1} \oplus x_{k_2})}_{V^\dagger} + \underbrace{\sum_{k_1<k_2<k_3} (x_{k_1} \oplus x_{k_2} \oplus x_{k_3})}_{V} - \ldots + (-1)^{n-1}(x_1 \oplus x_2 \oplus \ldots \oplus x_n) = 2^{n-1}(x_1 \wedge x_2 \wedge \ldots \wedge x_n)$$

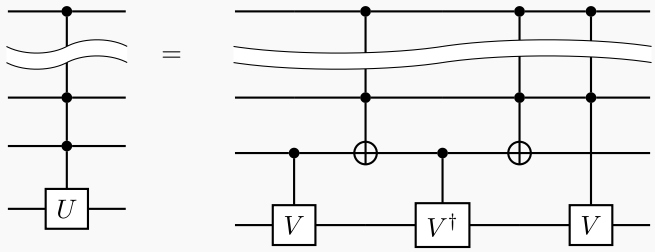
(where each positive term in the sum corresponds to V and each negative term to $V^\dagger$) we obtain a $C^n(U)$ gate. In particular we use a Gray code sequence where each step of the circuit differs from the previous in one bit only. This is not a requirement, but ensures the implementation is as efficient as possible in the number of CNOT gates - in particular allowing us to implement the operation in 2^{n-1} $C^1(V)$, $2^{n-1}-1$ $C^1(V^\dagger)$, and 2^n-2 CNOT gates. Fig 4.8 is the $n=2$ version of this identity, where the first step realizes $+x_2$ (implementing V), the second step realizes $-x_1 \oplus x_2$ (implementing $V^\dagger$) and the last step realizes $+x_1$ (implementing V_1). The higher n construction can be viewed as the nested version of the $n=2$ case. Although it is tedious, let us show in explicit detail the constructed circuit for $n=5$. Consider V such that $V^{2^{5-1}} = V^{16} = U$, and we implement the Gray code sequence:

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

Using the following circuit with 16 $C^1(V)$, 15 $C^1(V^\dagger)$, and 30 CNOT gates:


V
V†
...

The second version is where we assume we have access to $C^{n-1}(X), C^{n-1}(V)$, and $C^{n-1}(V^\dagger)$ gates with $V^2 = U$. Then, we claim that the below circuit implements the $C^n(U)$ operation, generalizing the circuit of Fig. 4.8 via



The proof is analagous to Ex. 4.21. First we observe that the first $n-1$ control qubits are left invariant (controls) and the nth control qubit is also left invariant due to $X^2 = I$. For the target qubit, if any of the first $n-1$ qubits are in $|0\rangle$, then the $C^{n-1}(X)$ and $C^{n-1}(V)$ gates drop out, leaving $C^1(V)C^1(V^\dagger) = I$. If the first $n-1$ qubits are in $|1\rangle$ while nth qubit is in $|0\rangle$, then the first $C^1(V)$ gate drops out, while the $C^1(V^\dagger)$ and $C^{n-1}(V)$ gates activate, again leading to $V^\dagger V = I$. When all n qubits are in $|1\rangle$, the middle $C^1(V^\dagger)$ gate drops out while the $C^1(V)$ and $C^{n-1}(V)$ gates fire, having the action of $V^2 = U$ on the last qubit. This is precisely the desired action of a $C^n(U)$ gate. Taking this construction for $n = 5$ yields the $C^5(U)$ in terms of $C^4(X), C^4(V)$, and $C^1(V)/C^1(V^\dagger)$ gates. □

Exercise 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.

Solution

Concepts Involved: Controlled Operations

This is just a special case of Ex. 4.30 with $U = X$. □

Remark: This special case of an ancilla-free $C^n(X)$ has been further explored in the literature, with this series of blog posts by Craig Gidney presenting an $O(n)$ construction, and $O(\text{poly}(n))$ constructions being presented in arXiv:2402.05053 and Nature Communications 15, 5886.

Exercise 4.30

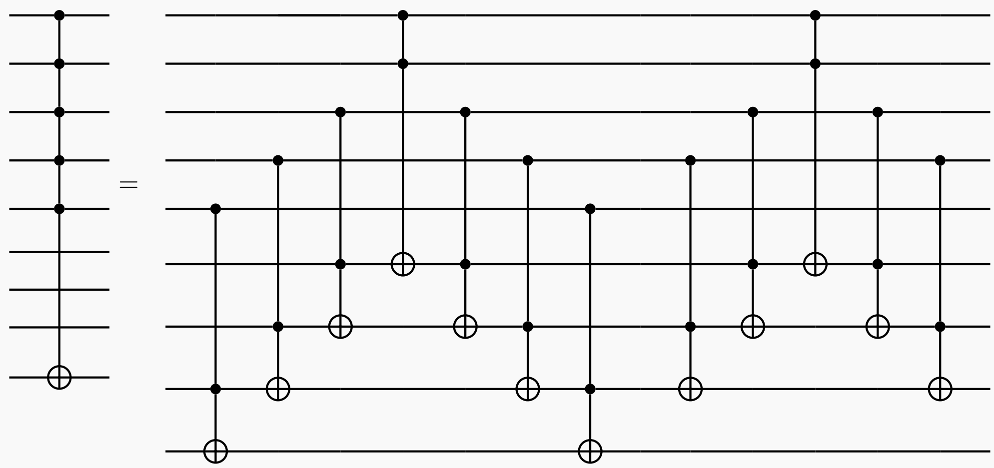
$(*)$ Suppose U is a single qubit unitary operation. Find a circuit containing $O(n^2)$ Toffoli, `CNOT` and single qubit gates which implements a $C^n(U)$ gate (for $n > 3$), using no work qubits.

Solution

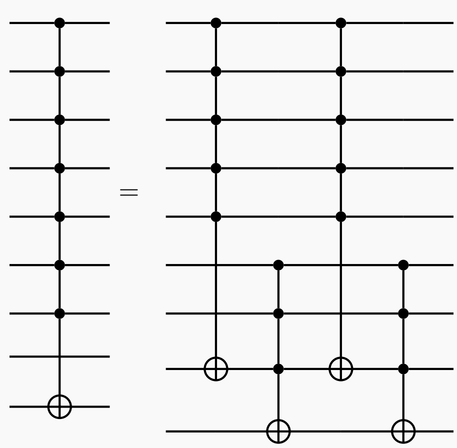
Concepts Involved: Controlled Operations, Gate Decomposition

The construction goes in three steps, following arXiv:9503016. The first step is a $C^n(X)$ construction that requires $n-2$ work qubits (and hence n control + $n-2$ work qubits + 1 control qubit = $2n-1$

qubits total); for this we modify the construction of Fig. 4.10 (as seen in Ex. 4.28). The circuit is shown below for $n = 5$:

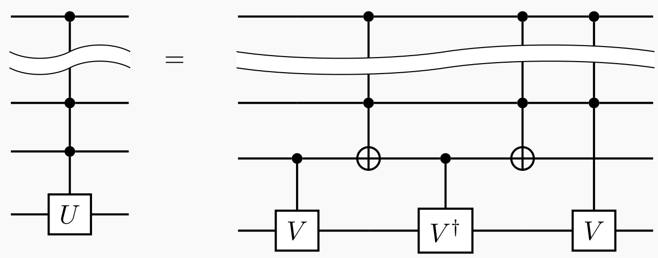


Looking at the first part of the circuit on the LHS, we can see that for $m \in \{1, \dots n-2\}$, the $n+m$th bit is flipped if the first m bits are 1. Hence if the first n bits are 1 then this correctly flips the target $2n-1$th qubit (the desired $C^n(X)$ action), at the expense of flipping work qubits $n+1 \dots 2n-2$. The second part of the circuit then undoes the flips on these work qubits.
The Toffoli count of this circuit is $O(n)$, and we note that the work qubits can be initialized arbitarily and are reset at the end of the circuit.
The second step is to improve the work qubit count for the $C^n(X)$ gate. Letting $n \geq 3$ and $m = \lceil \frac{n}{2} \rceil + 1$, we observe that the $C^n(X)$ gate can be written as a circuit that requires 1 work qubit only (and hence n control + 1 work qubit + 1 target qubit = $n+2$ qubits total), via the composition of two $C^m(X)$ gates and two $C^{n-m+1}(X)$ gates (shown below for $n = 7$):



This circuit has the desired operation as if any of the first n qubits are in the 0 state, the $C^n(X)$s on the RHS drop out or annihilate, and only if all first n qubits are in the 1 state is the $n+2$th qubit flipped.
For the two $C^m(X)$ and $C^{n-m-1}(X)$ gates in the construction, we can use the construction from step 1 - these appear to require $m - 2 = \lceil \frac{n}{2} \rceil - 1$ and $(n-m+1) - 2 = n - \lceil \frac{n}{2} \rceil$ work qubits. However, we do not require any additional ancilla, as we can just use the other qubits in the circuit that are idle during

the $C^m(X)/C^{n-m+1}(X)$ gates (as the construction from step 1 works with work qubits of arbitrary intial state, and leaves the state of the work qubits unchanged).
The construction of step 2 requires $2O(m) + 2(n-m-1) = 4O(n) = O(n)$ Toffoli gates.
The third and final step is the construction for the $C^n(U)$ operation in Ex. 4.28:



The $C^1(V)/C^1(V^\dagger)$ gates can be implemented by the construction of Fig. 4.8 using $O(1)$ elementary gates (4 single qubit gates and 2 CNOTs). For the two $C^{n-1}(X)$ gates, we use the construction of step 2, requiring $O(n)$ elementary gates/Toffolis and a single work qubit; instead of using an additional ancilla we can simply use the last qubit in the circuit as a temporary work qubit.
The cost of the $C^n(U)$ operation can then be determined inductively. Also note that we have a work-qubit free construction for $C^1(V), C^2(V)$ and via the inductive hypothesis, we may assume that $C^{n-1}(V)$ is constructed without work qubits. If we denote the gate complexity of $C^n(U)$ as $\mathcal{C}(n)$, we have the recursive relation:

$$\mathcal{C}(n) = \underbrace{O(1)}_{C^1(V)/C^1(V^\dagger)} + \underbrace{O(n)}_{C^{n-1}(X)} + \underbrace{\mathcal{C}(n-1)}_{C^{n-1}(V)} = O(n) + \mathcal{C}(n-1)$$

By induction we can see that:

$$\mathcal{C}(n) = O(n) + O(n-1) + O(n-2) + \ldots + O(1) = n \cdot O(n) = O(n^2)$$

and hence the number of elementary operations required to implement $C^n(U)$ is $O(n^2)$. This construction required no work qubits, completing the induction. □

Exercise 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 qubit and qubit 2 the target qubit. Prove the following identities:

$$\begin{aligned}
CX_1C &= X_1X_2 \\
CY_1C &= Y_1X_2 \\
CZ_1C &= Z_1 \\
CX_2C &= X_2 \\
CY_2C &= Z_1Y_2 \\
CZ_2C &= Z_1Z_2 \\
R_{z,1}(\theta)C &= CR_{z,1}(\theta) \\
R_{x,2}(\theta)C &= CR_{x,2}(\theta).
\end{aligned}$$

Solution

Concepts Involved: Controlled Operations, Rotations, Pauli Operators

- In the computation basis, we have

$$\begin{aligned} CX_1C|q_1, q_2\rangle &= CX_1|q_1, q_1 \oplus q_2\rangle \\ &= C|\widetilde{q_1}, q_1 \oplus q_2\rangle \\ &= |\widetilde{q_1}, \widetilde{q_1} \oplus q_1 \oplus q_2\rangle \\ &= |\widetilde{q_1}, \widetilde{q_2}\rangle \\ &= X_1X_2|q_1, q_2\rangle. \end{aligned}$$

- Similarly,

$$\begin{aligned} CZ_1C|q_1, q_2\rangle &= CZ_1|q_1, q_1 \oplus q_2\rangle \\ &= C(-1)^{q_1}|q_1, q_1 \oplus q_2\rangle \\ &= (-1)^{q_1}|q_1, q_1 \oplus q_1 \oplus q_2\rangle \\ &= (-1)^{q_1}|q_1, q_2\rangle \\ &= Z_1|q_1, q_2\rangle. \end{aligned}$$

- $$CY_1C = iCX_1Z_1C = iCX_1(CC)Z_1C = i(CX_1C)(CZ_1C) = iX_1X_2Z_1 = Y_1X_2.$$

- $$CR_{z,1}(\theta)C = C(aI_1 + bZ_1)C = aI_1 + bCZ_1C = aI_1 + bZ_1 = R_{z,1}(\theta)$$

□

Exercise 4.32

Let ρ be the density matrix describing a two qubit system. Suppose we perform a projective measurement in the computational basis of the second qubit. Let $P_0 = |0\rangle\langle 0|$ and $P_1 = |1\rangle\langle 1|$ be the projectors onto the $|0\rangle$ and the $|1\rangle$ states of the second qubit, respectively. Let ρ' be the density matrix which would be assigned to the system after the measurement by an observer who did not learn the measurement result. Show that

$$\rho' = P_0\rho P_0 + P_1\rho P_1$$

Also show that the reduced density matrix for the first qubit is not affected by the measurement, that is $\mathrm{tr}_2(\rho) = \mathrm{tr}_2(\rho')$.

Solution

Concepts Involved: Density Operators, Projective Measurement

The density matrix for the system after the measurement is given by

$$\begin{aligned}\rho' &= p_0\rho_0 + p_1\rho_1 && (1)\\ &= \operatorname{Tr}(\rho P_0)\frac{P_0\rho P_0}{\operatorname{Tr}(\rho P_0)} + \operatorname{Tr}(\rho P_1)\frac{P_1\rho P_1}{\operatorname{Tr}(\rho P_1)} && (2)\\ &= P_0\rho P_0 + P_1\rho P_1 && (3)\end{aligned}$$

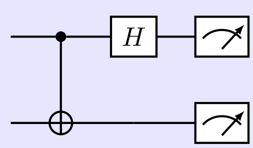
Finally, we have

$$\begin{aligned}\operatorname{Tr}_2(\rho') &= \operatorname{Tr}_2(P_0\rho P_0 + P_1\rho P_1)\\ &= \operatorname{Tr}_2(\rho P_0^2) + \operatorname{Tr}_2(\rho P_1^2)\\ &= \operatorname{Tr}_2(\rho P_0) + \operatorname{Tr}_2(\rho P_1)\\ &= \operatorname{Tr}_2(\rho(P_0 + P_1))\\ &= \operatorname{Tr}_2(\rho)\end{aligned}$$

□

Exercise 4.33: Measurement in the Bell basis

The measurement model we have specified for the quantum circuit model is that measurements are performed only in the computational basis. However, often we want to perform a measurement in some other basis, defined by a complete set of orthonormal states. To perform this measurement, simply unitarily transform from the basis we wish to perform the measurement in to the computational basis, then measure. For example, show that the circuit



performs a measurement in the basis of the Bell states. More precisely, show that this circuit results in a measurement being performed with corresponding POVM elements the four projectors onto the Bell states. What are the corresponding measurement operators?

Solution

Concepts Involved: Quantum Measurement, POVM Measurement

The full action of the circuit is equivalent to the following four operators

$$M'_{00} = |00\rangle \left(\langle 00| \left(H_1 \otimes I_2\right) \texttt{CNOT}\right) = \frac{1}{\sqrt{2}}|00\rangle(\langle 00| + \langle 11|) = |00\rangle\langle\beta_{00}|$$
$$M'_{01} = |01\rangle \left(\langle 01| \left(H_1 \otimes I_2\right) \texttt{CNOT}\right) = \frac{1}{\sqrt{2}}|01\rangle(\langle 01| + \langle 10|) = |01\rangle\langle\beta_{01}|$$
$$M'_{10} = |10\rangle \left(\langle 10| \left(H_1 \otimes I_2\right) \texttt{CNOT}\right) = \frac{1}{\sqrt{2}}|10\rangle(\langle 00| - \langle 11|) = |10\rangle\langle\beta_{10}|$$
$$M'_{11} = |11\rangle \left(\langle 11| \left(H_1 \otimes I_2\right) \texttt{CNOT}\right) = \frac{1}{\sqrt{2}}|11\rangle(\langle 01| - \langle 10|) = |11\rangle\langle\beta_{11}|$$

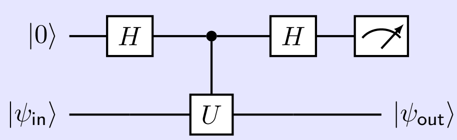
Therefore, it is easy to see that $\left(M'_k\right)^\dagger M'_k$ are projections onto the Bell states

$$\left(M'_{00}\right)^\dagger M_{00} = |\beta_{00}\rangle\langle\beta_{00}|$$
$$\left(M'_{01}\right)^\dagger M_{01} = |\beta_{01}\rangle\langle\beta_{01}|$$
$$\left(M'_{10}\right)^\dagger M_{10} = |\beta_{10}\rangle\langle\beta_{10}|$$
$$\left(M'_{11}\right)^\dagger M_{11} = |\beta_{11}\rangle\langle\beta_{11}|\,.$$

□

Exercise 4.34: Measuring an operator

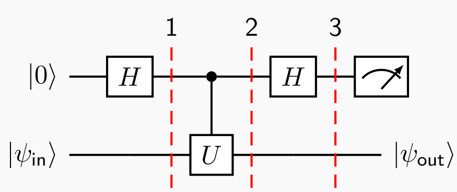
Suppose we have a single qubit operator U with eigenvalues ± 1, so that U is both Hermitian and unitary, so it can be regarded as both an observable and a quantum gate. Suppose we wish to measure the observable U. That is, we desire to obtain a measurement result indicating one of the two eigenvalues, and leaving a post-measurement state which is the corresponding eigenvector. How can this be implemented by a quantum circuit? Show that the following circuit implements a measurement of U:



Solution

Concepts Involved: Quantum Measurement

First, let us determine the action of this circuit on the eigenstates of U, $|\psi_{\text{in}}\rangle = |u_\pm\rangle$.



$$|\Psi(t_1)\rangle = \frac{1}{\sqrt{2}}\big(|0\rangle + |1\rangle\big)\,|u_\pm\rangle \quad .$$

$$\begin{aligned}
|\Psi(t_2)\rangle &= \frac{1}{\sqrt{2}}\big(|0\rangle\,|u_\pm\rangle + |1\rangle U|u_\pm\rangle\big) \\
&= \frac{1}{\sqrt{2}}\big(|0\rangle\,|u_\pm\rangle \pm |1\rangle|u_\pm\rangle\big) \\
&= |\pm\rangle|u_\pm\rangle
\end{aligned}$$

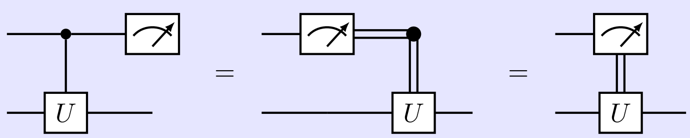
$$|\Psi(t_3)\rangle = |0/1\rangle|u_\pm\rangle$$

This shows that the above circuit implements a measurement of U for input states, $|\psi_{\text{in}}\rangle = |u_\pm\rangle$. Now, as U is also Hermitian, $|u_\pm\rangle$ forms a full basis and thereby, we can extend the result to arbitrary input states simply via linearity.

□

Exercise 4.35: Measurement commutes with controls

A consequence of the principle of deferred measurement is that measurements commute with quantum gates when the qubit being measured is a control qubit, that is:



(Recall that the double lines represent classical bits in this diagram.) Prove the first equality. The rightmost circuit is simply a convenient notation to depict the use of a measurement result to classically control a quantum gate.

Solution

Concepts Involved: Quantum Measurement, Controlled Operations

Let A and B be two qubits, where A is the control for a controlled-U gate acting on B. Suppose the joint state is

$$|\Psi\rangle = \alpha|0\rangle_A \otimes |\varphi_0\rangle_B + \beta|1\rangle_A \otimes |\varphi_1\rangle_B.$$

Applying the controlled-U gate yields:

$$|\Psi'\rangle = \alpha|0\rangle_A \otimes |\varphi_0\rangle_B + \beta|1\rangle_A \otimes U|\varphi_1\rangle_B.$$

Now, measuring qubit A in the computational basis, we obtain:

- With probability $|\alpha|^2$, outcome 0, post-measurement state:

$$|0\rangle_A \otimes |\varphi_0\rangle_B.$$

- With probability $|\beta|^2$, outcome 1, post-measurement state:

$$|1\rangle_A \otimes U|\varphi_1\rangle_B.$$

Now consider measuring A first, obtaining classical outcome $x \in \{0, 1\}$, and then applying the gate U^x to B. The resulting state is:

$$|x\rangle_A \otimes U^x|\varphi_x\rangle_B.$$

In both scenarios, the final joint state and measurement statistics are identical. Therefore, the controlled-U gate commutes with measurement on the control qubit, and the coherent control may be replaced with classical control:

$$\text{C-}U \quad \Longleftrightarrow \quad \text{Measure } A, \text{ then apply } U^x \text{ to } B.$$

This is a direct consequence of the *principle of deferred measurement*. □

Exercise 4.36

Construct a quantum circuit to add two two-bit numbers x and y modulo 4. That is, the circuit should perform the transformation $|x, y\rangle \mapsto |x, x + y \bmod 4\rangle$.

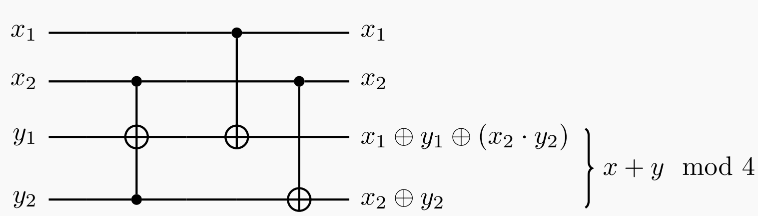
Solution

Concepts Involved: Controlled Operations

Writing $x = x_1x_2, y = y_1y_2$ in binary, we have:

$$x + y \bmod 4 = x_1x_2 + y_1y_2 \bmod 4 = [x_1 \oplus y_1 \oplus (x_2 \cdot y_2)][x_2 \oplus y_2]$$

so realizing the mod 2 addition via controlled gates, we have:



where the first two registers are preserved as they only act as controls. □

Exercise 4.37

Provide a decomposition of the transform

$$\frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{bmatrix}$$

into a product of two-level unitaries. This is a special case of the quantum Fourier transform, which we study in more detail in the next chapter.

Solution

Concepts Involved: Quantum Fourier Transform, QR Decomposition, Two-level Unitaries

We provide the standard QFT decomposition

$$F_4 = \text{SWAP} \cdot (I \otimes H) \cdot \text{CPhase}\left(\tfrac{\pi}{2}\right) \cdot (H \otimes I)$$

where each gate in the decomposition is a two-level unitary. □

Remark: For a general construction for arbitrary unitaries, one has to make full use of the QR decomposition. For a $n \times n$ dense unitary matrix, one needs $n(n-1)/2$ two level unitaries in the decomposition.

Exercise 4.38

Prove that there exist a $d \times d$ unitary matrix U which cannot be decomposed as a product of fewer than $d-1$ two-level unitary matrices.

Solution

Concepts Involved: Unitary Operators

A two-level unitary is a $d \times d$ unitary matrix that acts nontrivially only on a 2-dimensional subspace; that is, it modifies only two components of any vector it acts upon.
Let $e_1 = (1, 0, \ldots, 0)^T \in \mathbb{C}^d$ be the first standard basis vector. Suppose we apply k two-level unitaries $V_1, V_2, \ldots, V_k$ to e_1 to obtain $v = V_1 V_2 \cdots V_k e_1$. Since each V_j acts non-trivially on at most two coordinates, and e_1 has only one nonzero entry, the number of nonzero entries in v is at most $k+1$. Thus, if $k < d-1$, v cannot have all d entries nonzero.
Now consider a unitary matrix $U \in U(d)$ whose first column has all entries nonzero. Such matrices exist, as they form an open dense subset of $U(d)$. If U were a product of fewer than $d-1$ two-level unitaries, then its first column would violate the bound above, leading to a contradiction. Therefore, any such U requires at least $d-1$ two-level unitaries in its decomposition. □

Exercise 4.39

Find a quantum circuit using single qubit operations and CNOTs to implement the transformation

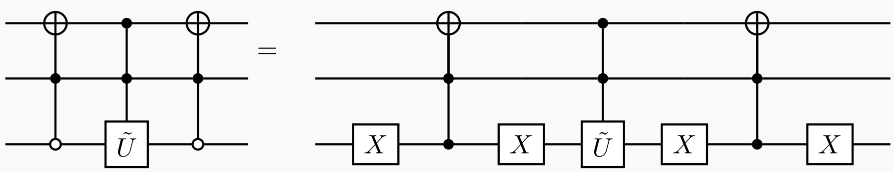
$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & a & 0 & 0 & 0 & 0 & c \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & b & 0 & 0 & 0 & 0 & d \end{bmatrix}$$

where $\tilde{U} = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$ is an arbitrary 2×2 unitary matrix.

Solution

Concepts Involved: Controlled Operations, Gate Decomposition

The following quantum circuit that acts on the span$(\left|010\right\rangle, \left|111\right\rangle)$ subspace via $\tilde{U}$ does the trick:



The Toffoli construction of Figure 4.9 (see Ex. 4.24) and the $C^2(U)$ construction of Figure 4.8 (see Ex. 4.21) then reduces this circuit to single qubit gates, CNOT, and $C^1(V)/C^1(V^\dagger)$ operations where $V^2 = \tilde{U}$. Writing $V = e^{i\alpha}AXBXC$, the construction of Fig 4.6 then reduces the circuit to be solely composed of single qubit gates and CNOTs. □

Exercise 4.40

For arbitrary α and β show that

$$E(R_n(\alpha), R_n(\alpha+\beta)) = \left|1 - \exp\left(i\beta/2\right)\right|,$$

and use this to justify (4.76).

Solution

Concepts Involved: Rotations

We start with the rotation operators

$$R_n(\alpha) = e^{-i\frac{\alpha}{2}\mathbf{n}\cdot\sigma},\ R_n(\alpha+\beta) = e^{-i\frac{\alpha+\beta}{2}\mathbf{n}\cdot\sigma}.$$

The error between them is

$$\begin{aligned} E(R_n(\alpha), R_n(\alpha+\beta)) &= \|R_n(\alpha) - R_n(\alpha+\beta)\|_{\text{op}}. \\ &= \|R_n(\alpha) - R_n(\alpha)R_n(\beta)\|_{\text{op}} \\ &= \|I - R_n(\beta)\|_{\text{op}} \end{aligned}$$

where we have used $R_n(\alpha+\beta) = R_n(\alpha)R_n(\beta)$ and the fact that operator norms are invariant under unitary operators.
We now notice that $I - R_n(\beta)$ is normal, i.e. $(I - R_n(\beta))(I - R_n(\beta))^\dagger = (I - R_n(\beta))^\dagger(I - R_n(\beta))$. For normal operators, the operator norm coincides with the largest eigenvalue.[Ref]
The eigenvalues of $I - R_n(\beta)$ are given by

$$1 - e^{i\frac{\beta}{2}} \quad \text{and} \quad 1 - e^{-i\frac{\beta}{2}}.$$

Thus, the operator norm is

$$\begin{aligned} \|I - R_n(\beta)\|_{\text{op}} &= \max\left(|1 - e^{i\frac{\beta}{2}}|, |1 - e^{-i\frac{\beta}{2}}|\right) \\ &= |1 - e^{i\frac{\beta}{2}}| \end{aligned}$$

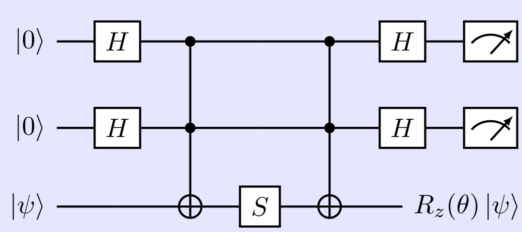
since both are equal. Thus, the error

$$E(R_n(\alpha), R_n(\alpha+\beta)) = |1 - e^{i\frac{\beta}{2}}|.$$

□

Exercise 4.41

This and the next two exercises develop a construction showing that the Hadamard, phase, controlled-`NOT` and Toffoli gates are universal. Show that the circuit in Figure 4.17 (reproduced below) applies the operation $R_z(\theta)$ to the third (target) qubit if the measurement outcomes are both 0, where $\cos\theta = 3/5$, and otherwise applies Z to the target qubit. Show that the probability of both measurement outcomes being 0 is $5/8$, and explain how repeated use of this circuit and $Z = S^2$ gates may be used to apply a $R_z(\theta)$ gate with probability approaching 1.



Solution

Concepts Involved: Rotations, Controlled Operations, Quantum Measurement

The circuit in Fig. 4.17 uses an ancilla state $|\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, applies a Toffoli gate with the ancilla qubits as control and the target qubit as target, then measures the ancillas after Hadamard gates. If the measurement outcomes are both 0, the resulting operation on the third qubit is

$$R_z(\theta) = \begin{bmatrix} e^{-i\theta/2} & 0 \\ 0 & e^{i\theta/2} \end{bmatrix}, \quad \text{with } \cos\theta = \frac{3}{5}.$$

In other cases, the outcome is a correctable variant: $ZR_z(\theta)$, $R_z(-\theta)$, or $ZR_z(-\theta)$. The success probability for obtaining outcome 00 is given by

$$P(00) = \left| \langle 00| H^{\otimes 2} \cdot \text{Toffoli} \, |\psi\rangle \right|^2 = \frac{5}{8}.$$

To implement $R_z(\theta)$ deterministically, repeat the process upon failure. Since the measurement outcomes are known, Pauli corrections (e.g., $Z = S^2$) can be applied conditionally. Repeating the process n times yields

$$P_{\text{success}} = 1 - \left(\frac{3}{8}\right)^n \to 1 \quad \text{as } n \to \infty.$$

□

Exercise 4.42: Irrationality of θ

Suppose $\cos\theta = 3/5$. We give a proof by contradiction that θ is an irrational multiple of 2π.

(1) Using the fact that $e^{i\theta} = (3+4i)/5$, show that if θ is rational, then there must exist a positive integer m such that $(3+4i)^m = 5^m$.

(2) Show that $(3+4i)^m = 3+4i \pmod 5$ for all $m > 0$, and conclude that no m such that $(3+4i)^m = 5^m$ can exist.

Solution

Concepts Involved: Induction

If θ is a rational multiple of 2π, then there exist a positive integer m for which

$$\begin{aligned} & e^{im\theta} = 1 \\ \implies & \frac{(3+4i)^m}{5^m} = 1 \\ \implies & (3+4i)^m = 5^m. \end{aligned}$$

We want to prove that $(3+4i)^m \equiv 3+4i \pmod 5$ for all $m > 0$. **Base Case** ($m = 1$):

$$(3+4i)^1 = 3+4i \pmod 5.$$

Inductive Step: Assume $(3+4i)^k \equiv 3+4i \pmod 5$ for some k. We show this holds for $k+1$:

$$(3+4i)^{k+1} = (3+4i)^k \cdot (3+4i).$$

Using the inductive hypothesis:

$$(3+4i)^{k+1} \equiv (3+4i)\cdot(3+4i) \pmod 5.$$

Now, compute:

$$(3+4i)^2 = 9 + 24i + 16i^2 = -7 + 24i \equiv 3 + 4i \pmod 5.$$

Thus, $(3+4i)^{k+1} \equiv 3+4i \pmod 5$.
By induction, $(3+4i)^m \equiv 3+4i \pmod 5$ for all $m > 0$.

□

Exercise 4.43

Use the results of the previous two exercises to show that the Hadamard, phase, controlled-NOT and Toffoli gates are universal for quantum computation.

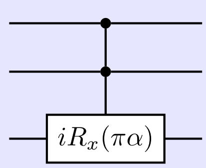
Solution

Concepts Involved: Universality

Ex. 4.41 gave a construction for $R_z(\theta)$ for $\cos\theta = 3/5$ using Hadamard, phase, and Toffoli gates. Ex. 4.42 then showed that θ is an irrational multiple of 2π. Since this rotation is dense in the subgroup of z-rotations of $SU(2)$, we can obtain any z-rotation in this manner. Since we also have access to Hadamard gates, we can obtain any x-rotation by conjugation, and arbitrary single-qubit rotations via the Euler decomposition. Combining this with CNOTs, we have a universal gateset (as shown in the text). □

Exercise 4.44

$(*)$ Show that the three qubit gate G defined by the circuit:



is universal for quantum computation whenever α is irrational.

Solution

Concepts Involved: Universality, Rotations, Controlled Operations

Let $G = \text{CC}(iR_x(\pi\alpha))$ be the 3-qubit gate that applies the unitary $iR_x(\pi\alpha) = ie^{-i\frac{\pi\alpha}{2}X}$ to the third qubit if and only if the first two qubits are in the state $|11\rangle$. We show that G is universal for quantum computation whenever $\alpha \notin \mathbb{Q}$. We assume that we have free access to computational basis states $\{|0\rangle, |1\rangle\}$.

(1) The unitary $iR_x(\pi\alpha) \in \text{SU}(2)$ is a non-Clifford single-qubit gate. Since α is irrational, the subgroup generated by $R_x(\pi\alpha)$ is dense in the subgroup of $\text{SU}(2)$ corresponding to X-rotations. To isolate this single-qubit rotation from G, fix the two control qubits in state $|1\rangle$, so that G acts as $iR_x(\pi\alpha)$

on the target. Hence we have access to arbitrary single-qubit X-rotations.

(2) To generate entangling operations, use G with one control qubit in state $|1\rangle$, and vary the second control. This allows us to simulate a controlled-$R_x(\pi\alpha)$ gate. By composing such gates, we can construct a `CNOT` (or any controlled-X rotation) entangling gate to arbitrary accuracy.

(3) We show that this gate allows us to produce $|-\rangle$ states. Using the ingridients of steps 1/2, we can produce/approximate the gate sequence $R_{x,1}(-\frac{\pi}{2})CX_{1,2}R_{x,1}(\frac{\pi}{2})$, which applying to the $|00\rangle$ state:

$$\begin{aligned}R_{x,1}(-\frac{\pi}{2})CX_{1,2}R_{x,1}(\frac{\pi}{2})|00\rangle &= R_{x,1}(-\frac{\pi}{2})CX_{1,2}\left(\frac{|0\rangle - i|1\rangle}{\sqrt{2}}\right)\otimes|0\rangle \\ &= R_{x,1}(-\frac{\pi}{2})\frac{|00\rangle - i|11\rangle}{\sqrt{2}} \\ &= \frac{(|0\rangle + i|1\rangle)|0\rangle - i(|1\rangle + i|0\rangle)|1\rangle}{2} \\ &= \frac{|00\rangle + i|10\rangle + |01\rangle - i|11\rangle}{2} \\ &= \frac{|0+\rangle + i|1-\rangle}{\sqrt{2}}\end{aligned}$$

by measuring the first qubit in the computational basis, we have a 50% probability that the second qubit is in $|-\rangle$, which (when repeated) gives an arbitrarily high probability to construct $|-\rangle$.

(4) By Ex. 4.11, if we can use this gate to generate (arbitrary) single-qubit Z-rotations, any single-qubit unitary can be generated by interleaving z and x. Suppose we act the gate on the state $|1\psi-\rangle$ for $|\psi\rangle = a|0\rangle + b|1\rangle$, then:

$$\begin{aligned}C^2(iR_x(\pi\alpha))|1\psi-\rangle &= |1\rangle\otimes C^1(iR_x(\pi\alpha))|\psi-\rangle \\ &= |1\rangle\otimes\left[a|0\rangle\otimes|-\rangle + b|1\rangle\otimes iR_x(\pi\alpha)|-\rangle\right] \\ &= |1\rangle\otimes\left[a|0\rangle\otimes|-\rangle + b|1\rangle\otimes ie^{i\frac{\pi\alpha}{2}}|-\rangle\right] \\ &= |1\rangle\otimes\left[a|0\rangle + ie^{i\frac{\pi\alpha}{2}}b|1\rangle\right]\otimes|-\rangle \\ &= |1\rangle\otimes e^{i(\frac{\pi(\alpha+1)}{4})}\left[ae^{-i(\frac{\pi(\alpha+1)}{4})}|0\rangle + e^{i(\frac{\pi(\alpha+1)}{4})}b|1\rangle\right]\otimes|-\rangle \\ &= |1\rangle\otimes e^{i(\frac{\pi(\alpha+1)}{4})}R_z(\frac{\pi(\alpha+1)}{2})|\psi\rangle\otimes|-\rangle\end{aligned}$$

Since α is irrational, the subgroup generated by $R_z(\frac{\pi(\alpha+1)}{2})$ is dense in the subgroup of $SU(2)$ corresponding to Z-rotations. Hence having access to both arbitrary X and Z rotations, we obtain arbitrary approximation of single-qubit unitaries using just G and ancillas.

(5) Since we can approximate arbitrary single-qubit unitaries and implement an entangling two-qubit gate using only G, the gate G alone is universal for quantum computation, by standard universality results.

$\Rightarrow$ G is universal for quantum computation whenever $\alpha \notin \mathbb{Q}$.

□

Exercise 4.45

Suppose U is a unitary transform implemented by an n qubit quantum circuit constructed from H, S, CNOT and Toffoli gates. Show that U is of the form $2^{-k/2}M$ for some integer k, where M is a $2^n \times 2^n$ matrix with only complex integer entries. Repeat this exercise with the Toffoli gate replaced by the $\pi/8$ gate.

Solution

Concepts Involved: Unitary Operators, Universality

- **Case 1: Gate set** $\{H, S, \texttt{CNOT}, \texttt{Toffoli}\}$
  - The gates S, CNOT, and Toffoli have matrix entries in $\mathbb{Z}[i]$.
  - The Hadamard gate H introduces a factor of $1/\sqrt{2}$ per application.
  - Therefore, any circuit built from this gate set yields a unitary of the form:

$$U = 2^{-k/2}M, \quad \text{where } M \in \mathbb{Z}[i]^{2^n \times 2^n},\ k \in \mathbb{N}.$$

- **Case 2: Gate set** $\{H, S, \texttt{CNOT}, T\}$
  - The T-gate introduces the 8th root of unity $\omega = e^{i\pi/4}$, so entries lie in $\mathbb{Z}[\omega]$.
  - The Hadamard gate again introduces dyadic denominators.
  - Thus, the unitary matrix has the form:

$$U = 2^{-k/2}M, \quad \text{where } M \in \mathbb{Z}[\omega]^{2^n \times 2^n},\ k \in \mathbb{N}.$$

□

Exercise 4.46: Exponential complexity growth of quantum systems

Let ρ be a density matrix describing the state of n qubits. Show that describing ρ requires $4^n - 1$ independent real numbers.

Solution

Concepts Involved: Density Operators

To show that a density matrix ρ for n qubits requires $4^n - 1$ independent real parameters, we proceed as follows:

1. Dimension of the Hilbert Space

The state space for n qubits is $\mathbb{C}^{2^n}$, which has dimension 2^n.

2. Structure of the Density Matrix

A density matrix ρ is:

$$\begin{aligned}&\text{Hermitian: } \rho = \rho^\dagger,\\&\text{Trace One: } \text{Tr}(\rho) = 1,\\&\text{Positive Semidefinite: } \langle v|\rho|v\rangle \geq 0 \text{ for all vectors } |v\rangle.\end{aligned}$$

3. Counting Parameters

A. **Hermitian Condition:** A $2^n \times 2^n$ Hermitian matrix has:

2^n real parameters for the diagonal,

$\dfrac{(2^n)(2^n-1)}{2}$ complex parameters for the off-diagonal elements, contributing

$(2^n)(2^n-1)$ real parameters.

Total:

$$2^n + (2^n)(2^n-1) = 2^{2n}.$$

B. **Trace Condition:** The trace condition reduces the number of independent parameters by 1:

$$2^{2n} - 1.$$

Thus, a density matrix ρ for n qubits indeed requires $4^n - 1$ independent real numbers. □

Exercise 4.47

For $H = \sum_{k=1}^{L} H_k$, prove that $e^{-iHt} = e^{-iH_1 t}e^{-iH_2 t}\dots e^{-iH_L t}$ for all i if $[H_j, H_k] = 0$, for all j, k.

Solution

Concepts Involved: Commutators, Operator Functions

We proceed by induction. For $L = 2$ we have

$$e^{-i(H_1+H_2)t} = \sum_{n=0}^{\infty} \frac{([-it]^n(H_1+H_2)^n}{n!}$$

Using the fact that $[H_1, H_2] = 0$, we can then write:

$$e^{-i(H_1+H_2)t} = \sum_{n=0}^{\infty}\frac{1}{n!}\sum_{k=0}^{n}\binom{n}{k}H_1^k H_2^{n-k} = \sum_{n=0}^{\infty}(-it)^n\sum_{k=0}^{n}\frac{H_1^k H_2^{n-k}}{k!(n-k)!} = \sum_{n=0}^{\infty}\frac{(-itH_1)^n}{n!}\sum_{m=0}^{\infty}\frac{(-itH_2)^m}{m!}$$

and so:

$$e^{-i(H_1+H_2)t} = e^{-iH_1 t}e^{-iH_2 t}$$

Suppose the claim holds for $L-1$. Then letting $\sum_{k=1}^{L-1} H_k$ take the place of H_1 and H_L take the place of H_2 in the above argument (since $[\sum_{k=1}^{L-1} H_k, H_L] = 0$), we find:

$$e^{-i\sum_{k=1}^{L} H_k t} = e^{-i(\sum_{k=1}^{L-1} H_k + H_L)t} = e^{-i\sum_{k=1}^{L-1} H_k t} e^{-iH_L t} = \prod_{k=1}^{L-1} e^{-iH_k t} e^{-iH_L t} = \prod_{k=1}^{L} e^{-iH_k t} \tag{4}$$

where in the second-to-last equality we use the inductive hypothesis. Thus the L case holds and the claim is proven via induction. □

Exercise 4.48

Show that the restriction of H_k to at most c particles *implies* that in the sum (4.97), L is upper bounded by a polynomial in n.

Solution

Concepts Involved: Counting

Each term H_k acts nontrivially on at most c particles out of n. The number of such subsets is

$$L \leq \sum_{j=1}^{c} \binom{n}{j} = O(n^c),$$

since c is a constant. Therefore, the total number of c-local terms is upper bounded by a polynomial in n. □

Remark: This reflects the physical constraint that a c-local Hamiltonian cannot contain more than $O(n^c)$ interaction terms, which ensures efficient simulability of such systems.

Exercise 4.49: Baker–Campbell–Hausdorf formula

Prove that

$$e^{(A+B)\Delta t} = e^{A\Delta t} e^{B\Delta t} e^{-\frac{1}{2}[A,B]\Delta t^2} + O(\Delta t^3)$$

and also prove Equations (4.103) and (4.104).

Solution

Concepts Involved: Operator Functions

To prove that

$$e^{(A+B)\Delta t} = e^{A\Delta t} e^{B\Delta t} e^{-\frac{1}{2}[A,B]\Delta t^2} + O(\Delta t^3),$$

we use the Taylor expansion of the exponential

$$e^{(A+B)\Delta t} = I + (A+B)\Delta t + \frac{(A+B)^2(\Delta t)^2}{2} + O(\Delta t^3).$$

Calculating

$$(A+B)^2 = A^2 + AB + BA + B^2 = A^2 + B^2 + 2AB - [A, B].$$

Thus, we have

$$\begin{aligned} e^{(A+B)\Delta t} &= I + (A+B)\Delta t + \left(\frac{A^2+B^2}{2} + AB - \frac{[A,B]}{2}\right)\Delta t^2 + O(\Delta t^3) \\ &= \left(I + A\Delta t + \frac{A^2(\Delta t)^2}{2}\right)\left(I + B\Delta t + \frac{B^2(\Delta t)^2}{2}\right)\left(I - \frac{1}{2}[A,B]\Delta t^2\right) + O(\Delta t^3) \\ &= e^{A\Delta t}e^{B\Delta t}e^{-\frac{1}{2}[A,B]\Delta t^2} + O(\Delta t^3). \end{aligned}$$

This completes the proof.

□

Exercise 4.50

Let $H = \sum_k^L H_k$, and define

$$U_{\Delta t} = \left[e^{-iH_1\Delta t}e^{-iH_2\Delta t}\dots e^{-iH_L\Delta t}\right]\left[e^{-iH_L\Delta t}e^{-iH_{L-1}\Delta t}\dots e^{-iH_1\Delta t}\right]$$

(a) Prove that $U_{\Delta t} = e^{-2iH\Delta t} + O(\Delta t^3)$

(b) Use the results in Box 4.1 to prove that for a positive integer m,

$$E(U^m_{\Delta t}, e^{-2miH\Delta t}) \leq m\alpha\Delta t^3,$$

for some constant α.

Solution

Concepts Involved: Unitary Operators, Operator Functions

(a) We expand the symmetric product

$$U_{\Delta t} = \left(\prod_{k=1}^{L} e^{-iH_k\Delta t}\right)\left(\prod_{k=L}^{1} e^{-iH_k\Delta t}\right)$$

using the second-order Suzuki–Trotter formula. For small Δt, define

$$S_{\Delta t} = \prod_{k=1}^{L} e^{-iH_k \Delta t}.$$

Then

$$U_{\Delta t} = S_{\Delta t} S_{\Delta t}^\dagger + [S_{\Delta t}, S_{\Delta t}^\dagger] + \text{higher-order terms}.$$

But since $S_{\Delta t}^\dagger = \prod_{k=L}^{1} e^{iH_k \Delta t}$, this is precisely the adjoint of $S_{\Delta t}$, so their product is:

$$U_{\Delta t} = e^{-iH\Delta t} e^{-iH\Delta t} + O(\Delta t^3) = e^{-2iH\Delta t} + O(\Delta t^3),$$

where the error arises from nested commutators like $[H_j, [H_k, H_\ell]]$, and the constant in the $O(\Delta t^3)$ term depends on the operator norms of these commutators.

(b) Let

$$U_{\Delta t} = e^{-2iH\Delta t} + R(\Delta t), \quad \text{with } \|R(\Delta t)\| \leq \alpha \Delta t^3.$$

Consider

$$U_{\Delta t}^m = \left(e^{-2iH\Delta t} + R(\Delta t)\right)^m.$$

We expand:

$$U_{\Delta t}^m = e^{-2imH\Delta t} + \sum_{k=1}^{m} \binom{m}{k} e^{-2iH\Delta t(m-k)} R(\Delta t)^k.$$

Taking norms:

$$\|U_{\Delta t}^m - e^{-2imH\Delta t}\| \leq \sum_{k=1}^{m} \binom{m}{k} \|R(\Delta t)\|^k.$$

For small Δt, the dominant term is $k = 1$, so we obtain:

$$\|U_{\Delta t}^m - e^{-2imH\Delta t}\| \leq m\|R(\Delta t)\| + O(m\Delta t^6) \leq m\alpha\Delta t^3 + O(m\Delta t^6).$$

Hence,

$$E(U_{\Delta t}^m, e^{-2imH\Delta t}) \leq m\alpha\Delta t^3.$$

□

Exercise 4.51

Construct a quantum circuit to simulate the Hamiltonian

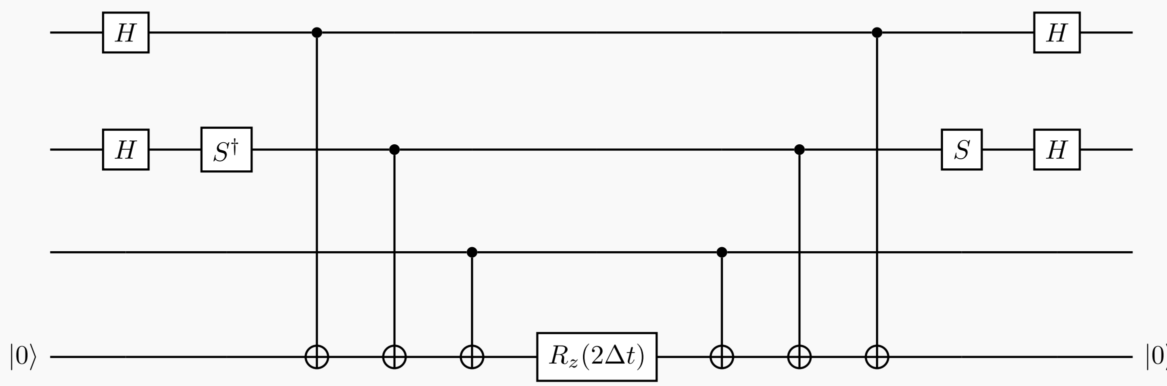
$$H = X_1 \otimes Y_2 \otimes Z_3$$

performing the unitary transform $e^{-i\Delta t H}$ for any Δt.

Solution

Concepts Involved: Operator Functions, Gate Decomposition

Below is the quantum circuit for simulating $U = e^{-i\Delta t(X_1 \otimes Y_2 \otimes Z_3)}$:



It is easily obtained by starting with the simulation circuit of Figure 4.19 for $H = Z_1 \otimes Z_2 \otimes Z_3$ and conjugating qubits 1/2 appropriately to transform $Z_1 \to X_1$ and $Z_2 \to Y_2$. □

Problem 4.1: Computable phase shifts

Let m and n be positive integers. Suppose $f : \{0, \ldots, 2^m - 1\} \mapsto \{0, \ldots, 2^n - 1\}$ is a classical function from m to n bits which may be computed reversibly using T Toffoli gates, as described in Section 3.2.5. That is, the function $(x, y) \mapsto (x, y \oplus f(x))$ may be implemented using T Toffoli gates. Give a quantum circuit using $2T + n$ (or fewer) one, two and three qubit gates to implement the unitary operation defined by

$$|x\rangle \mapsto \exp\left(\frac{-2i\pi f(x)}{2^n}\right)|x\rangle$$

Solution

Concepts Involved:

We construct a quantum circuit using

- a reversible circuit for $f(x)$,
- an ancilla register prepared in a Fourier basis state,
- phase kickback via modular addition.

Let $|\tilde{1}\rangle \in \mathbb{C}^{2^n}$ denote the quantum Fourier transform of $|1\rangle$, given by

$$|\tilde{1}\rangle := \frac{1}{\sqrt{2^n}} \sum_{y=0}^{2^n-1} e^{2\pi i y/2^n} |y\rangle.$$

We initialize the system in the joint state

$$|\psi_0\rangle := |x\rangle \otimes |\tilde{1}\rangle.$$

Let U_f be the reversible map

$$U_f : |x\rangle|y\rangle \mapsto |x\rangle|y \oplus f(x)\rangle.$$

Then

$$\begin{aligned}
U_f|\psi_0\rangle &= |x\rangle \otimes \left(\frac{1}{\sqrt{2^n}} \sum_{y=0}^{2^n-1} e^{2\pi i y/2^n} |y \oplus f(x)\rangle \right) \\
&= |x\rangle \otimes \left(\frac{1}{\sqrt{2^n}} \sum_{z=0}^{2^n-1} e^{2\pi i (z \ominus f(x))/2^n} |z\rangle \right) \\
&= e^{-2\pi i f(x)/2^n} |x\rangle \otimes \left(\frac{1}{\sqrt{2^n}} \sum_{z=0}^{2^n-1} e^{2\pi i z/2^n} |z\rangle \right) \\
&= e^{-2\pi i f(x)/2^n} |x\rangle \otimes |\tilde{1}\rangle.
\end{aligned}$$

Thus, the modular addition $y \mapsto y \oplus f(x)$ induces a phase $e^{-2\pi i f(x)/2^n}$ on $|x\rangle$ when the ancilla is in $|\tilde{1}\rangle$. Now apply U_f^{-1}. Since the second register is unchanged, we obtain:

$$U_f^{-1} \cdot U_f(|x\rangle \otimes |\tilde{1}\rangle) = e^{-2\pi i f(x)/2^n} |x\rangle \otimes |\tilde{1}\rangle.$$

Lemma. Let $|\tilde{1}\rangle = \frac{1}{\sqrt{2^n}} \sum_{y=0}^{2^n-1} e^{2\pi i y/2^n} |y\rangle$. Then for all $a \in \mathbb{Z}_{2^n}$,

$$\sum_y e^{2\pi i y/2^n} |y \oplus a\rangle = e^{2\pi i a/2^n} |\tilde{1}\rangle.$$

Conclusion. The overall effect is

$$|x\rangle \mapsto e^{-2\pi i f(x)/2^n} |x\rangle,$$

while restoring the ancilla to $|\tilde{1}\rangle$. This completes the implementation of the desired unitary.

Gate Count

- Applying U_f costs T Toffoli gates.
- Applying U_f^{-1} costs another T Toffoli gates.
- Preparing $|\tilde{1}\rangle = \mathsf{QFT}_n|1\rangle$ can be done using at most n one- and two-qubit gates.

Thus the total number of one-, two-, and three-qubit gates is at most

$$2T + n.$$

□

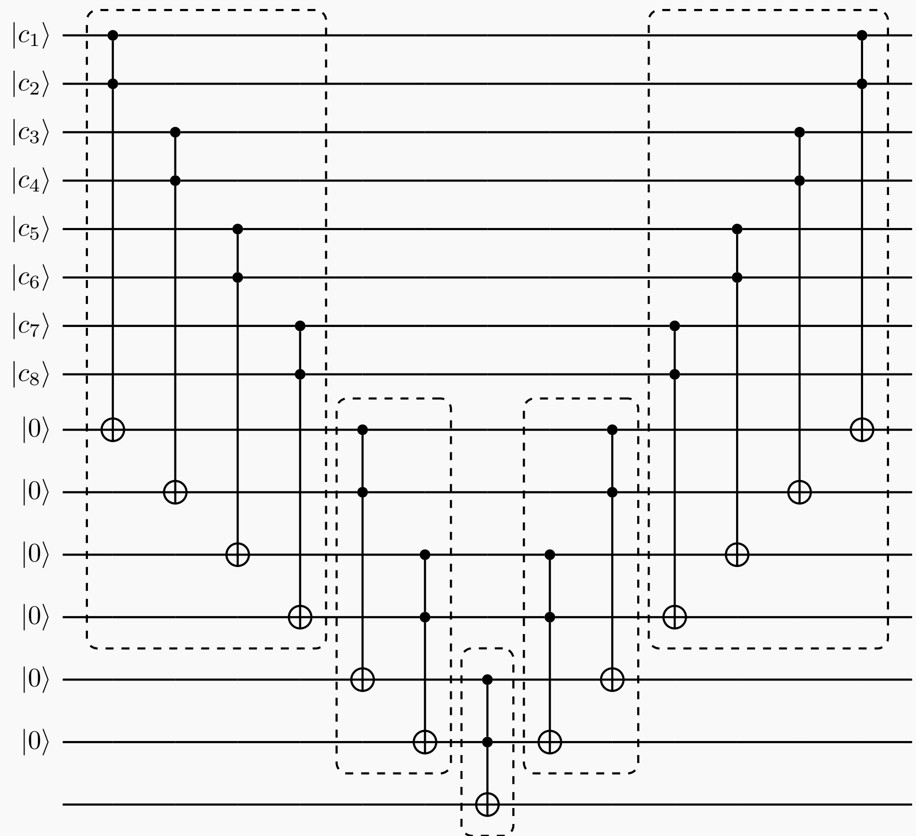
Problem 4.2

Find a depth $O(\log n)$ construction for the $C^n(X)$ gate. (*Comment:* The depth of a circuit is the number of distinct time steps at which gates are applied; the point of this problem is that it is possible to parallelize the $C^n(X)$ construction by applying many gates in parallel during the same timestep.)

Solution

Concepts Involved: Controlled Operators, Gate Decomposition

A simple construction exists for $n = 2^m$, assuming access to $n - 2$ work qubits (initialized to $|0\rangle$), as in Fig. 4.10 (see Ex. 4.28). As the comment suggests, the strategy is to parallelize. We provide the example for $n = 8$:



By inspection we can see that this circuit has the desired action - the target/last qubit is only flipped if all the control qubits are set to 1 (if any of the controls are 0, some subset of the work qubits are not toggled, and hence the target qubit is not flipped). The second half of the circuit is solely for resetting the work qubits to the $|0\rangle$ state.

We observe that each dashed box in the circuit above consists of gates acting on different subsets of qubits, and hence can be performed in a single timestep. With this observation, we note that each timestep of the above construction successively reduces the number of involved qubits in half, and generally for $n = 2^m$ control qubits we require $2m - 1$ timesteps to run the circuit (m halvings/timesteps to perform the gate, and $m - 1$ timesteps to reset the work qubits). Hence the circuit depth is $O(m) = O(\log n)$. The simplest (if slightly wasteful) way to generalize the construction for the case that n is not a power of two is to simply append $|1\rangle$ ancilla qubits to the control register until the number of controls becomes a power of two. □

Remark: Note that there is no aspect of this construction that is specific to the X-gate in particular, and indeed it works for an arbitrary single-qubit U.

Problem 4.3: Alternate universality construction

Suppose U is a unitary matrix on n qubits. Define $H \equiv i\ln(U)$. Show that

(1) H is Hermitian, with eigenvalues in the range 0 to 2π.

(2) H can be written

$$H = \sum_g h_g g,$$

where h_g are real numbers and the sum is over all n-fold tensor products g of the Pauli matrices $\{I, X, Y, Z\}$.

(3) Let $\Delta = 1/k$, for some positive integer k. Explain how the unitary operation $\exp\big(-ih_g g\Delta\big)$ may be implemented using $O(n)$ one and two qubit operations.

(4) Show that

$$\exp(-iH\Delta) = \prod_g \exp\big(-ih_g g\Delta\big) + O(4^n\Delta^2)$$

where the product is taken with respect to any fixed ordering of the n-fold tensor products of Pauli matrices, g.

(5) Show that

$$U = \left[\prod_g \exp\big(-ih_g h\Delta\big)\right]^k + O(4^n\Delta)$$

(6) Explain how to approximate U to within a distance $\epsilon > 0$ using $O(n16^n/\epsilon)$ one and two qubit unitary operations.

Solution

Concepts Involved: Unitary Operators, Universality

(1) Since U is unitary, all eigenvalues λ satisfy $|\lambda| = 1$, so $\lambda = e^{i\theta}$ with $\theta \in [0, 2\pi)$. Let $H = i\ln U$, where the logarithm is taken in the principal branch. Then

$$U = V\mathrm{diag}(e^{i\theta_1}, \dots, e^{i\theta_{2^n}})V^\dagger \Rightarrow H = V\mathrm{diag}(\theta_1, \dots, \theta_{2^n})V^\dagger,$$

which is Hermitian with eigenvalues $\theta_j \in [0, 2\pi)$.

(2) The 4^n Pauli strings $g \in \{I, X, Y, Z\}^{\otimes n}$ form a basis for Hermitian operators on n qubits. Hence,

$$H = \sum_g h_g g, \quad \text{where} \quad h_g = \frac{1}{2^n}\mathrm{Tr}(gH) \in \mathbb{R}.$$

(3) Each $g = \sigma_1 \otimes \cdots \otimes \sigma_n$ acts nontrivially on $k \leq n$ qubits. One can conjugate each $\sigma_i \in \{X, Y\}$ to Z via Clifford unitaries (e.g., $HXH = Z$, $S^\dagger HYHS = Z$). After mapping to $Z^{\otimes k}$, apply the diagonal gate $e^{-ih_g Z^{\otimes k}\Delta}$, then undo the basis change. Total cost is $O(n)$ 1- and 2-qubit gates.

(4) Let $H = \sum_g h_g g$, with each g a Pauli string. Then first-order Trotter expansion gives

$$e^{-iH\Delta} = \prod_g e^{-ih_g g\Delta} + O\left(\sum_{g<g'} \|[h_g g, h_{g'} g']\| \Delta^2\right).$$

Since each Pauli string satisfies $\|g\| = 1$ and $\|[g, g']\| \leq 2$, the total error over $\binom{4^n}{2} = O(4^{2n})$ terms gives:

$$e^{-iH\Delta} = \prod_g e^{-ih_g g\Delta} + O(4^n\Delta^2).$$

(5) Using the previous result and setting $\Delta = 1/k$,

$$U = e^{-iH} = \left(e^{-iH\Delta}\right)^k = \left(\prod_g e^{-ih_g g\Delta}\right)^k + O(k \cdot 4^n\Delta^2) = \left(\prod_g e^{-ih_g g\Delta}\right)^k + O(4^n\Delta).$$

(6) To approximate U within $\varepsilon > 0$, choose $\Delta = \varepsilon/(C \cdot 4^n)$ for some constant C, so that the total error is $O(\varepsilon)$. Then $k = 1/\Delta = O(4^n/\varepsilon)$. Each product $\prod_g e^{-ih_g g\Delta}$ has 4^n terms, each implementable with $O(n)$ gates. Total gate count is

$$O\left(\frac{4^n}{\varepsilon} \cdot 4^n \cdot n\right) = O\left(\frac{n \cdot 16^n}{\varepsilon}\right).$$

□

Problem 4.4: Minimal Toffoli construction (Research)

The following problems concern constructions of the Toffoli with some minimal number of other gates. ‘

(1) What is the smallest number of two qubit gates that can be used to implement the Toffoli gate?

(2) What is the smallest number of one qubit gates and `CNOT` gates that can be used to implement the Toffoli gate?

(3) What is the smallest number of one qubit gates and controlled-Z gates that can be used to implement the Toffoli gate?

Solution

Concepts Involved: Controlled Operators, Gate Decomposition

(1) In Phys. Rev. A 88, 010304(R) it is shown that 5 two-qubit gates are necessary.

(2) arXiv:0803.2316 Gives a construction where the minimal number of CNOTs is 6, but the number of 1-qubit gates/minimizing the CNOT + one-qubit gates appears open.

(3) A simple modification of (2) wherein each CNOT is conjugated by H on the target qubit similarly yields a minimum number of 6 CZ gates required, but this increases the single-qubit gate count by 12-likely more efficient constructions for minimizing one-qubit gate counts exist.

□

Problem 4.5: (Research)

Construct a family of Hamiltonians, $\{H_n\}$, on n qubits, such that simulating H_n requires a number of operations super-polynomial in n. (*Comment:* This problem seems to be quite difficult.)

Solution

Concepts Involved:

To the authors' knowledge, this problem remains open. If such a family exists, the Hamiltonians would have to be non-local and dense, as efficient simulation techniques for local Hamiltonians (via Trotter simulation, as discussed in the text) and sparse Hamiltonians (see arXiv:0301023) have been found (plus many subsequent improvements). □

Problem 4.6: Universality with prior entanglement

Controlled-NOT gates and single qubit gates form a universal set of quantum logic gates. Show that an alternative universal set of resources is comprised of single qubit unitaries, the ability to perform measurements of pairs of qubits in the Bell basis, and the ability to prepare arbitrary four qubit entangled states.

Solution

Concepts Involved:

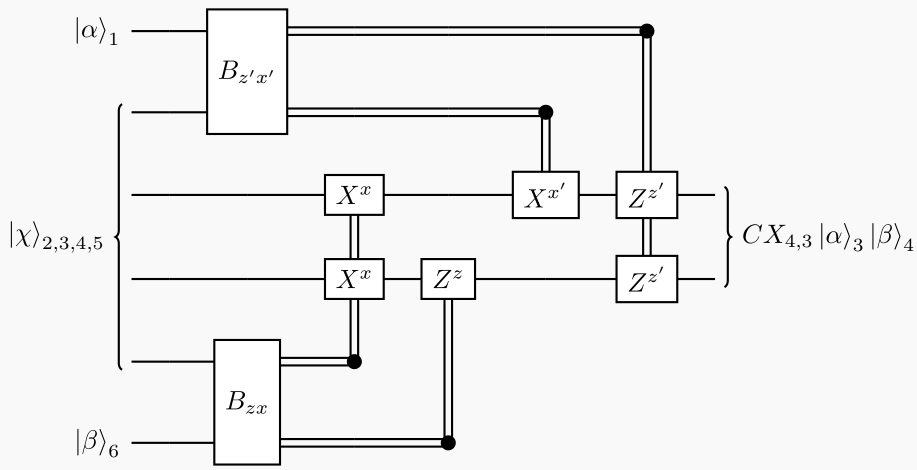
We follow the construction of arXiv:9908010 Throughout, it will be helpful to recall the definition of the bell states:

$$|B_{zx}\rangle = \frac{|0x\rangle + (-1)^z |1\bar{x}\rangle}{\sqrt{2}} = Z_1^z X_1^x |B_{00}\rangle = Z_1^z X_1^x \frac{|00\rangle + |11\rangle}{\sqrt{2}}$$

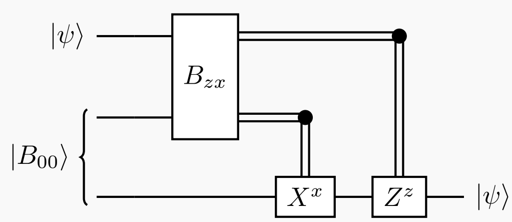
Since we have access to arbitary single qubit unitaries, for universality it suffices to show that we are able to perform a CNOT gate between two arbitrary qubits using the given resources. In particular, we consider the 4-qubit state:

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

wherein the following circuit (inspired by the quantum teleportation protocol) of Bell-measurements + single-qubit feedback allows for a CNOT gate to be performed between arbitrary single qubit states $|\alpha\rangle = a|0\rangle + b|1\rangle$ and $|\beta\rangle = c|0\rangle + d|1\rangle$:



with $|\alpha\rangle$ as the target and $|\beta\rangle$ as the control. Although the above procedure looks complicated, it is understood conceptually as teleportating a CNOT gate.
First, consider the following modified version of the quantum teleportation protocol, wherein we measure in the Bell basis to teleport an arbitrary qubit $|\psi\rangle = a|0\rangle + b|1\rangle$



This easily follows from rewriting the first two qubits of $|\psi\rangle |B_{00}\rangle$ in the Bell basis:

$$\begin{aligned}
|\psi\rangle |B_{00}\rangle &= (a|0\rangle + b|1\rangle)\frac{|00\rangle + |11\rangle}{\sqrt{2}} \\
&= \frac{a|000\rangle + a|011\rangle + b|100\rangle + b|111\rangle}{\sqrt{2}} \\
&= \frac{1}{2}|B_{00}\rangle (a|0\rangle + b|1\rangle) + \frac{1}{2}|B_{01}\rangle (a|1\rangle + b|0\rangle) \\
&+ \frac{1}{2}|B_{10}\rangle (a|0\rangle - b|11\rangle) + \frac{1}{2}|B_{11}\rangle (a|1\rangle - b|0\rangle)
\end{aligned}$$

From which we can see that the state of the teleported qubit is $Z^z X^x (a|0\rangle + b|1\rangle) = Z^z X^x |\psi\rangle$. The correction operators thus have the desired effect and the teleportation protocol works as claimed.

Next, observe the following property of the $|\chi\rangle$ state:

$$|\chi\rangle_{2,3,4,5} = CX_{4,3} |B_{00}\rangle_{2,3} |B_{00}\rangle_{4,5}$$

We can now combine the ingredients of $|\chi\rangle$ with the teleportation protocol:

$$|\alpha\rangle_1 |\chi\rangle_{2,3,4,5} |\beta\rangle_6 = CX_{4,3} |\alpha\rangle_1 |B_{00}\rangle_{2,3} |B_{00}\rangle_{4,5} |\beta\rangle_6$$

Now measuring qubits 1/2 and 5/6 in the Bell basis, the $|\alpha\rangle$, $|\beta\rangle$ states get teleported to qubits 3/4, with appropriate correction operators

$$|\alpha\rangle_1 |\chi\rangle_{2,3,4,5} |\beta\rangle_6 \overset{\text{Bell meas.}}{\mapsto} CX_{4,3} Z_3^{z'} X_3^{x'} |\alpha\rangle_3 Z_4^z X_4^x |\beta\rangle_4$$

Now using the result of Ex. 4.31, we can commute the correction operators past the `CNOT` gate

$$\begin{aligned}
(CX_{4,3})X_3 &= X_3(CX_{4,3}) \\
(CX_{4,3})X_4 &= X_3X_4(CX_{4,3}) \\
(CX_{4,3})Z_3 &= Z_3(CX_{4,3}) \\
(CX_{4,3})Z_4 &= Z_3Z_4(CX_{4,3})
\end{aligned}$$

Thus we have

$$(Z_3Z_4)^{z'} X_3^{x'} Z_4^z (X_3X_4)^x CX_{4,3} |\alpha\rangle_3 |\beta\rangle_4$$

These are precisely the correction operators as depicted in the circuit, and hence the given circuit indeed performs a `CNOT` + single-qubit corrections. Thus the construction gives us access to `CNOT` gates, and combined with single-qubit gates, universality. ☐

Remark: In the reference, it is noted that $|\chi\rangle$ can be generated from Bell measurements + single qubit operations on two GHZ states $|\text{GHZ}\rangle = \frac{|000\rangle + |111\rangle}{\sqrt{2}}$, so in fact the requirement of arbitrary four qubit states is stronger than necessary. Note that this notion of performing logical operations through teleporation can be further extended to the striking computational model known as the measurement-based, or the one-way quantum computer (see Phys. Rev. Lett. 86, 5188). There, a universal computation can be performed by generating a single entangled resource state at the outset, and thereafter only performing adaptive single-qubit measurements on the resource state.

7 Quantum computers: physical realization

Exercise 7.1

Using the fact that x and p do note commute, and that in fact $[x,p] = i\hbar$, explicitly show that $a^\dagger a = H/\hbar\omega - 1/2$.

Solution

Concepts Involved: Commutators, Creation/Annihilation Operators
Recall the definition of the $a, a^\dagger$ in terms of the position and momentum operators:

$$a = \frac{1}{\sqrt{2m\hbar\omega}}(m\omega x + ip)$$
$$a^\dagger = \frac{1}{\sqrt{2m\hbar\omega}}(m\omega x - ip)$$

as well as the definition of the Hamiltonian for a particle in a quadratic potential:

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2$$

Calculating $a^\dagger a$ we find:

$$\begin{aligned} a^\dagger a &= \frac{1}{2m\hbar\omega}(m^2\omega^2 x^2 + im\omega[x,p] + p^2) \\ &= \frac{1}{2m\hbar\omega}(m^2\omega^2 x^2 + im\omega(i\hbar) + p^2) \\ &= \frac{1}{\hbar\omega}(\frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2 - \frac{\hbar\omega}{2}) \\ &= \frac{H}{\hbar\omega} - \frac{1}{2} \end{aligned}$$

where we note the use of the commutation relation between position and momentum in the second equality. □

Exercise 7.2

Given that $[x,p] = i\hbar$, compute $[a, a^\dagger]$.

Solution

Concepts Involved: Commutators, Creation/Annihilation Operators

---

Using the linearity of the commutator:

$$\begin{aligned}
[a, a^\dagger] &= [\frac{1}{\sqrt{2m\hbar\omega}}(m\omega x + ip), \frac{1}{\sqrt{2m\hbar\omega}}(m\omega x - ip)] \\
&= \frac{1}{2m\hbar\omega}\left(m^2\omega^2[x,x] - i\frac{[x,p]}{m\omega} + i\frac{[p,x]}{m\omega} + [p,p]\right) \\
&= \frac{1}{2m\hbar\omega}\left(0 - i\frac{i\hbar}{m\omega} + i\frac{-i\hbar}{m\omega} + 0\right) \\
&= 1
\end{aligned}$$

□

Exercise 7.3

Compute $[H, a]$ and use the result to show that if $|\psi\rangle$ is an eigenstate of H with energy $E \geq n\hbar\omega$, then $a^n |\psi\rangle$ is an eigenstate with energy $E - n\hbar\omega$.

Solution

Concepts Involved: Commutators, Creation/Annihilation Operators, Eigenvalues, Eigenvectors

---

We have:

$$\begin{aligned}
[H, a] &= \hbar\omega[a^\dagger a + \frac{1}{2}, a] \\
&= \hbar\omega\left([a^\dagger a, a] + \frac{1}{2}[1, a]\right) \\
&= \hbar\omega\left(a^\dagger[a, a] + [a^\dagger, a]a\right) \\
&= \hbar\omega\,(-a) \\
&= -\hbar\omega a
\end{aligned}$$

where in the first line we use the result of Exercise 7.1 and in the second to last line we use the result of Exercise 7.2. From this, it follows that if $|\psi\rangle$ is an eigenstate of H with energy $E \geq \hbar\omega$, then:

$$\begin{aligned}
Ha\,|\psi\rangle &= ([H, a] + aH)\,|\psi\rangle \\
&= (-\hbar\omega a + aH)\,|\psi\rangle \\
&= -\hbar\omega a\,|\psi\rangle + aH\,|\psi\rangle \\
&= -\hbar\omega a\,|\psi\rangle + aE\,|\psi\rangle \\
&= (E - \hbar\omega)a\,|\psi\rangle
\end{aligned}$$

i.e. $a\,|\psi\rangle$ is an eigenstate of H with energy $E - \hbar\omega$. The argument can be repeated n-fold if $E \geq n\hbar\omega$ to conclude that $a^n\,|\psi\rangle$ is an eigenstate of H with eigenvalue $E - n\hbar\omega$. □

Exercise 7.4

Show that $|n\rangle = \frac{(a^\dagger)^n}{\sqrt{n!}} |0\rangle$.

Solution

Concepts Involved: Creation/Annihilation Operators, Eigenvalues, Eigenvectors
Eq. (7.11) in the text yields the action of the creation operator on an eigenstate of H:

$$a^\dagger |n\rangle = \sqrt{n+1} |n+1\rangle$$

Applying Eq. (7.11) n times to $|0\rangle$, we find:

$$(a^\dagger)^n |0\rangle = \sqrt{n}\sqrt{n-1}\ldots\sqrt{2}\sqrt{1} |n\rangle$$

and dividing both sides by $\sqrt{n!}$ we conclude:

$$|n\rangle = \frac{(a^\dagger)^n}{\sqrt{n!}} |0\rangle .$$

□

Exercise 7.5

Verify that equations (7.11) and (7.12) are consistent with (7.10) and the normalization condition $\langle n|n\rangle = 1$.

Solution

Concepts Involved: Creation/Annihilation Operators, Eigenvalues, Eigenvectors
Equations (7.10) - (7.12) are given as:

$$\begin{aligned} a^\dagger a |n\rangle &= n |n\rangle \\ a^\dagger |n\rangle &= \sqrt{n+1} |n+1\rangle \\ a |n\rangle &= \sqrt{n} |n-1\rangle \end{aligned}$$

Applying the lowering and raising operator successively and invoking equations (7.12), (7.11) we find:

$$\begin{aligned} \langle n| a^\dagger a |n\rangle &= \langle n| a^\dagger \sqrt{n} |n-1\rangle \\ &= \sqrt{n} \langle n| a^\dagger |n-1\rangle \\ &= \sqrt{n} \langle n| \sqrt{n} |n\rangle \\ &= n \langle n|n\rangle \\ &= n \end{aligned}$$

meanwhile if we invoke Equation (7.10):

$$\begin{aligned}\langle n| a^\dagger a |n\rangle &= \langle n| n |n\rangle \\ &= n \langle n|n\rangle \\ &= n\end{aligned}$$

so the equations are consistent. □

Exercise 7.6

Prove that a coherent state is an eigenstate of the photon annihilation operator, that is, show $a|\alpha\rangle = \lambda|\alpha\rangle$ for some constant λ.

Solution

Concepts Involved: Creation/Annihilation Operators, Eigenvalues, Eigenvectors, Coherent States

Recall the definition of the coherent state $|\alpha\rangle$:

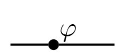
$$|\alpha\rangle = e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle$$

Applying a to $|\alpha\rangle$ and using that $a|n\rangle = \sqrt{n}|n\rangle$, we have:

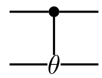
$$\begin{aligned}a|\alpha\rangle &= ae^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle \\ &= e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} a|n\rangle \\ &= e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} \sqrt{n}|n-1\rangle \\ &= e^{-|\alpha|^2/2}\alpha \sum_{n=0}^{\infty} \frac{\alpha^{n-1}}{\sqrt{(n-1)!}} |n-1\rangle \\ &= \alpha e^{-|\alpha|^2/2} \sum_{n'=0}^{\infty} \frac{\alpha^{n'}}{\sqrt{(n')!}} |n'\rangle \\ &= \alpha|\alpha\rangle\end{aligned}$$

where in the second-to-last inequality we re-index the sum. We conclude that $|\alpha\rangle$ is an eigenstate of a with eigenvalue α. □

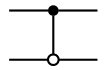
Before continuing, we introduce our drawing convention for optical circuits, which departs from that of the text. In particular, we depict phase shifters of angle φ as:



A directed beamsplitter of angle θ as:



and a 50/50 ($\theta = \pi/4$) beamsplitter as:

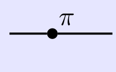


Exercise 7.7

Show that the circuit below transforms a dual-rail state by

$$|\psi_{out}\rangle = \begin{bmatrix} e^{i\pi} & 0 \\ 0 & 1 \end{bmatrix} |\psi_{in}\rangle ,$$

if we take the top wire to represent the $|01\rangle$ mode, and $|10\rangle$ the bottom mode, and the boxed π to represent a phase shift by π:



Note that in such 'optical circuits', propagation in space is explicitly represented by putting in lumped circuit elements such as in the above, to represent phase evolution. In the dual-rail representation, evolution according to (7.20) changes the logical state only by an unobservable global phase, and thus we are free to disregard it and keep only relative phase shifts.

Solution

Concepts Involved: Dual-Rail representation, Phase Shifters

From the diagram, we have that:

$$|01\rangle \mapsto e^{i\pi} |01\rangle$$

$$|10\rangle \mapsto |10\rangle$$

so in the logical dual-rail representation with $|0_L\rangle = |01\rangle$ and $|1_L\rangle = |10\rangle$ we can write the transformation matrix as:

$$\begin{bmatrix} e^{i\pi} & 0 \\ 0 & 1 \end{bmatrix}.$$

□

Exercise 7.8

Show that $P|\alpha\rangle = |\alpha e^{i\Delta}\rangle$ where $|\alpha\rangle$ is a coherent state (note that, in general, α is a complex number!)

Solution

Concepts Involved: Dual-Rail Representation, Phase Shifters, Coherent States

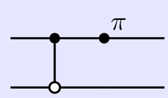
First, we note that $P = e^{i\Delta a^\dagger a}$ and so $P\left|n\right\rangle = e^{i\Delta n}$. Therefore:

$$\begin{aligned}
P\left|\alpha\right\rangle &= Pe^{-\frac{|\alpha|^2}{2}}\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}\left|n\right\rangle \\
&= e^{-\frac{|\alpha|^2}{2}}\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}P\left|n\right\rangle \\
&= e^{-\frac{|\alpha|^2}{2}}\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}e^{i\Delta n}\left|n\right\rangle \\
&= e^{-\frac{|\alpha|^2}{2}}\sum_{n=0}^{\infty}\frac{(\alpha e^{i\Delta})^n}{\sqrt{n!}}\left|n\right\rangle \\
&= e^{-\frac{\left|\alpha e^{i\alpha}\right|^2}{2}}\sum_{n=0}^{\infty}\frac{(\alpha e^{i\Delta})^n}{\sqrt{n!}}\left|n\right\rangle \\
&= \left|\alpha e^{i\Delta}\right\rangle
\end{aligned}$$

where in the second-to-last equality we note that $\left|e^{i\Delta}\right| = 1$. □

Exercise 7.9: Optical Hadamard gate

Show that the following circuit acts as a Hadamard gate on dual-rail single photon states, that is, $\left|01\right\rangle \mapsto (\left|01\right\rangle + \left|10\right\rangle)/\sqrt{2}$ and $\left|10\right\rangle \mapsto (\left|01\right\rangle - \left|10\right\rangle)/\sqrt{2}$ up to an overall phase:



The assertion of the exercise is incorrect, and requires the order of the beamsplitter and phase shifter to be flipped (or alternatively to flip the angle of the beamsplitter).

Solution

Concepts Involved: Dual-Rail Representation, Phase Shifters, Beamsplitters

The action of the 50/50 beamsplitter on the $\left|01\right\rangle, \left|10\right\rangle$ manifold of states is given by:

$$B_{\theta=\pi/4} = \begin{bmatrix}\cos(\pi/4) & -\sin(\pi/4)\\ \sin(\pi/4) & \cos(\pi/4)\end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix}1 & -1\\ 1 & 1\end{bmatrix}$$

So combining this with the phase shifter (noting $e^{i\pi} = -1$), we have:

$$U = P(\pi)B_{\theta=\pi/4} = B\begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix}1 & -1\\ 1 & 1\end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix}-1 & 1\\ 1 & 1\end{bmatrix}$$

which is not equal to a Hadamard up to a global phase. If we instead flip the order of the given optical components:

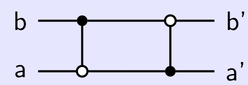
$$U = B_{\theta=\pi/4}P(\pi) = \frac{1}{\sqrt{2}}\begin{bmatrix}1 & -1\\ 1 & 1\end{bmatrix}\begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix}-1 & -1\\ -1 & 1\end{bmatrix} = -H$$

which is indeed a Hadamard (up to a global negative sign). The same can be accomplished with the original circuit, but the angle of the beamsplitter flipped from $\pi/4 \to -\pi/4$. □

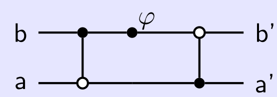
Exercise 7.10: Mach–Zehnder interferometer

Interferometers are optical tools used to measure small phase shifts, which are constructed from two beamsplitters. Their basic principle of operation can be understood by this simple exercise.

1. Note that this circuit performs the identity operation:



2. Compute the rotation operation (on dual-rail states) which this circuit performs, as a function of the phase shift φ:



Solution

Concepts Involved: Dual-Rail Representation, Beamsplitters

1. The first part is simple to check (we can do it for arbitrary θ):

$$B^\dagger B = \begin{bmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{bmatrix}\begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix} = \begin{bmatrix}\cos^2\theta + \sin^2\theta & 0\\ 0 & \cos^2\theta + \sin^2\theta\end{bmatrix} = I.$$

2. For the second part we include the φ-phase shift in between the two 50-50 beamsplitters:

$$\begin{aligned}
B^\dagger_{\theta=\pi/4} P(\varphi) B_{\theta=\pi/4} &= \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} e^{i\varphi} & 0 \\ 0 & 1 \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \\
&= \frac{1}{2}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} e^{i\varphi} & -e^{i\varphi} \\ 1 & 1 \end{bmatrix} \\
&= \frac{1}{2}\begin{bmatrix} 1+e^{i\varphi} & 1-e^{i\varphi} \\ 1-e^{i\varphi} & 1+e^{i\varphi} \end{bmatrix} \\
&= \frac{e^{i\varphi/2}}{2}\begin{bmatrix} e^{-i\varphi/2}+e^{i\varphi/2} & e^{-i\varphi/2}-e^{i\varphi/2} \\ e^{-i\varphi/2}-e^{i\varphi/2} & e^{-i\varphi/2}+e^{i\varphi/2} \end{bmatrix} \\
&= e^{i\varphi/2}\begin{bmatrix} \cos\left(\frac{\varphi}{2}\right) & -i\sin\left(\frac{\varphi}{2}\right) \\ -i\sin\left(\frac{\varphi}{2}\right) & \cos\left(\frac{\varphi}{2}\right) \end{bmatrix} \\
&= e^{i\varphi/2} R_x(\varphi)
\end{aligned}$$

thus this operation (up to a global phase) corresponds to an x-rotation of angle φ. This can also be seen more easily from the fact that the beamsplitters act as a Y-rotation, and thus here in the 50-50 case we rotate the z-axis by $-\pi/2$ about the y-axis yielding the rotation about x.

□

Exercise 7.11

What is $B\,|2,0\rangle$ for $\theta = \pi/4$?

Solution

Concepts Involved: Creation/Annihilation Operators, Dual-Rail Representation, Beamsplitters

---

The beamsplitter operator B maps the creation operator for mode a as (Eq. 7.34 of N&C):

$$B a^\dagger B^\dagger = a^\dagger \cos\theta + b^\dagger \sin\theta$$

thus we calculate:

$$
\begin{aligned}
B\,|2,0\rangle &= B\frac{(a^\dagger)^2}{\sqrt{2}}\,|0,0\rangle \\
&= \frac{1}{\sqrt{2}}Ba^\dagger B^\dagger Ba^\dagger\,|0,0\rangle \\
&= \frac{1}{\sqrt{2}}Ba^\dagger B^\dagger Ba^\dagger B\,|0,0\rangle \\
&= \frac{1}{\sqrt{2}}\left(a^\dagger\cos\theta + b^\dagger\sin\theta\right)^2|0,0\rangle \\
&= \frac{1}{\sqrt{2}}(\cos^2\theta(a^\dagger)^2 + 2\sin\theta\cos\theta a^\dagger b^\dagger + \sin^2\theta(b^\dagger)^2)\,|0,0\rangle \\
&= \cos^2\theta\,|2,0\rangle + \frac{1}{\sqrt{2}}\sin(2\theta)\,|1,1\rangle + \sin^2\theta\,|0,2\rangle
\end{aligned}
$$

where we use in the second line that $B^\dagger B = I$ and in the third line that $B\,|0,0\rangle = |0,0\rangle$. With $\theta = \pi/4$, this becomes:

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

□

Exercise 7.12: Quantum beamsplitter with classical inputs

What is $B|\alpha\rangle|\beta\rangle$ where $|\alpha\rangle$ and $|\beta\rangle$ are two coherent states as in Equation (7.16)? (*Hint:* recall $|n\rangle = \frac{(a^\dagger)^n}{\sqrt{n!}}\,|0\rangle$.

Solution

Concepts Involved: Annihilation/creation operators, Mode Mixing, Beam Splitters

---

Let $|\alpha\rangle$ and $|\beta\rangle$ be coherent states of two bosonic modes a and b, i.e.,

$$
|\alpha\rangle = D_a(\alpha)\,|0\rangle\,, \qquad |\beta\rangle = D_b(\beta)\,|0\rangle\,,
$$

where $D(\gamma) = \exp\bigl(\gamma a^\dagger - \gamma^* a\bigr)$ is the displacement operator.
The beam splitter operator B is a unitary transformation that mixes the two modes:

$$
B^\dagger aB = a\cos\theta + b\sin\theta, \qquad B^\dagger bB = -a\sin\theta + b\cos\theta.
$$

Assuming a 50:50 beam splitter with $\theta = \frac{\pi}{4}$, define the rotated modes:

$$
a' = \frac{1}{\sqrt{2}}(a+b), \qquad b' = \frac{1}{\sqrt{2}}(-a+b).
$$

Coherent states remain coherent under linear transformations of the mode operators. Therefore, the product state $|\alpha\rangle_a |\beta\rangle_b$ transforms to:

$$B|\alpha\rangle_a |\beta\rangle_b = \left|\frac{1}{\sqrt{2}}(\alpha+\beta)\right\rangle_{a'} \left|\frac{1}{\sqrt{2}}(-\alpha+\beta)\right\rangle_{b'}.$$

In terms of the original mode labels (since the names of the physical modes don't change), this becomes

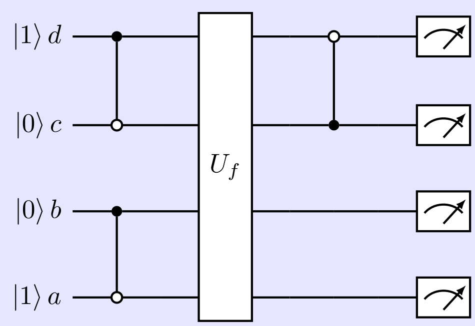
$$B|\alpha\rangle_a |\beta\rangle_b = \left|\frac{1}{\sqrt{2}}(\alpha+\beta)\right\rangle_{a} \left|\frac{1}{\sqrt{2}}(-\alpha+\beta)\right\rangle_{b}.$$

□

Remark: This transformation of coherent states under a beam splitter is classical in character — no entanglement is generated between the output modes.

Exercise 7.13: Optical Deutsch–Jozsa quantum circuit

($*$) In Section 1.4.4 (page 34), we described a quantum circuit for solving the one-bit Deutsch–Jozsa problem. Here is a version of that circuit for single photon states (in the dual-rail representation), using beamsplitters, phase shifters, and nonlinear Kerr media:



1. Construct circuits for the four possible classical functions U_f using Fredkin gates and beamsplitters.
2. Why are no phase shifters necessary in this construction?
3. For each U_f show explicitly how interference can be used to explain how the quantum algorithm works.
4. Does this implementation work if the single photon states are replaced by coherent states?

Solution

Concepts Involved: Deutsch–Jozsa Algorithm, Phase Shifters, Beamsplitters, Coherent States

1. There are only four possible functions $f : \{0,1\} \mapsto \{0,1\}$. For each, we give the unitary U_f which accomplishes $U_f|x,y\rangle = U_f|x, y + f(x)\rangle$.

   In these constructions, a useful subroutine will be the optical Fredkin gate with ancilla c in the $|1\rangle$ state. In the dual rail representation, the Fredkin gate (for $\xi = \pi$) is a CNOT gate and hence this realizes a NOT gate on the dual rail qubit.

- $f(x) = 0$. In this case, U_f is just the identity.
- $f(x) = 1$. In this case we apply X to the second qubit, so we can use the Fredkin-NOT construction described above.
- $f(x) = x$. In this case, we apply X to the second dual-rail qubit depending on the state of the first dual-rail qubit. Since $|0_L\rangle = |01\rangle$ (the top wire/d mode) and $|1_L\rangle = |10\rangle$ (the bottom wire/ $x$mode), we choose "control mode" of the optical Fredkin to be the top/d mode of the top dual-rail qubit, which then acts on the bottom dual qubit as a CNOT.
- $f(x) = x \oplus 1$. In this case, we simply compose the CNOT/optical Fredkin of the $f(x) = x$ case with an additional Fredkin-NOT.

2. No phase-shifters are necessary for this construction as we only require the use of CNOTs/Xs, which can fully be accomplished by the optical Fredkin ($+$ ancilla). This is because the transformation U_f is based on purely classical/Boolean logic and thus can be composed purely out of the classical gates of CNOT/NOTs, and does not require the quantum-mechanical gates $R_z(\varphi)/R_y(\theta)$ that a phase-shifter/individual beamsplitter would provide.

3. In the dual-rail representation, the initial state is $|+\rangle_L |-\rangle_L$. We then have (with the U_f described in the first part and a subsequent 50/50 beamsplitter/Hadamard on the first qubit):

$$|+\rangle_L |-\rangle_L \overset{U_f}{\mapsto} \begin{cases} |+\rangle_L |-\rangle_L & f(x) = 0 \\ -|+\rangle_L |-\rangle_L & f(x) = 1 \\ |-\rangle_L |-\rangle_L & f(x) = x \\ -|-\rangle_L |-\rangle_L & f(x) = x \oplus 1 \end{cases} \overset{H_1}{\mapsto} \begin{cases} \pm |0\rangle_L |-\rangle_L & f(0) = f(1) \\ \pm |1\rangle_L |-\rangle_L & f(0) \neq f(1) \end{cases}$$

Thus by measuring the first dual rail qubit, i.e. checking if the photon is in the top/d mode $|0\rangle_L = |01\rangle$ or the second/c mode $|1\rangle_L = |10\rangle$ we can verify with 100% probability whether the function f is constant or balanced. The algorithm hinges on the interference of the action of U_f on the superposition of computational basis states $|+\rangle_L$.

4. No, the implementation fails. We only need to analyze the $f(x) = 0$ case to see this - In this case $U_f = I$ and the quantum circuit reduces to a single beamsplitter acting on the a/b modes, which as deduced in Ex. 7.12 maps the input coherent states to other coherent states. We therefore have 4 coherent states in the output, wherein there is some probability to measure any possible combination of photon number. Even if hypothetically the balanced cases for f had only one possible measurement outcome (which can be shown to be not the case), since there is some probability of measuring any possible outcome for the $f(x) = 0$ case the implementation can no longer distinguish f with a single shot.

□

Exercise 7.14: Classical cross phase modulation

To see that the expected classical behavior of a Kerr medium is obtained from the definition of K, Equation (7.41), apply it to two modes, one with a coherent state and the other in state $|n\rangle$; that is, show that

$$K|\alpha\rangle|n\rangle = |\alpha e^{i\chi Ln}\rangle|n\rangle$$

Use this to compute

$$\begin{aligned}\rho_a &= \mathrm{Tr}_b\left[K|\alpha\rangle|\beta\rangle\langle\beta|\langle\alpha|K^\dagger\right] \\ &= e^{-|\beta|^2}\sum_m \frac{|\beta|^2 m}{m!}|\alpha e^{i\chi Ln}\rangle\langle\alpha e^{i\chi Ln}|\end{aligned}$$

and show that the main contribution to the sum is for $m = |\beta|^2$.

Solution

Concepts Involved: Kerr Interaction, Coherent States, Operator Functions

---

The Kerr unitary is given by

$$K = \exp\Big(i\chi L\, a^\dagger a\, b^\dagger b\Big)$$

and acts on $|\alpha\rangle_a |n\rangle_b$ as

$$K|\alpha\rangle|n\rangle = \exp\Big(i\chi L\, n\, a^\dagger a\Big)|\alpha\rangle|n\rangle = \left|\alpha e^{i\chi Ln}\right\rangle|n\rangle.$$

Now let $|\beta\rangle = e^{-|\beta|^2/2}\sum_{n=0}^\infty \frac{\beta^n}{\sqrt{n!}}|n\rangle$. Then

$$K|\alpha\rangle|\beta\rangle = e^{-|\beta|^2/2}\sum_{n=0}^\infty \frac{\beta^n}{\sqrt{n!}}\left|\alpha e^{i\chi Ln}\right\rangle|n\rangle.$$

To compute the reduced density matrix on mode a, trace out mode b:

$$\begin{aligned}\rho_a &= \mathrm{Tr}_b\left[K|\alpha\rangle|\beta\rangle\langle\beta|\langle\alpha|K^\dagger\right] \\ &= e^{-|\beta|^2}\sum_{n=0}^\infty \frac{|\beta|^{2n}}{n!}\left|\alpha e^{i\chi Ln}\right\rangle\!\left\langle\alpha e^{i\chi Ln}\right|.\end{aligned}$$

This is a mixture over rotated coherent states weighted by a Poisson distribution peaked at $n = |\beta|^2$, so the main contribution is

$$\rho_a \approx \left|\alpha e^{i\chi L|\beta|^2}\right\rangle\!\left\langle\alpha e^{i\chi L|\beta|^2}\right|.$$

□

Exercise 7.15

Plot (7.55) as a function of field detuning φ, for $R_1 = R_2 = 0.9$.

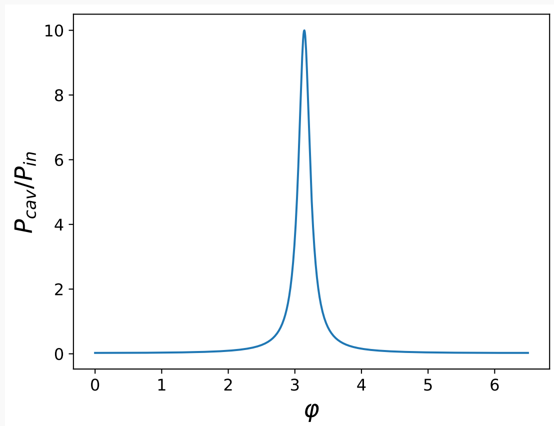
Solution

Concepts Involved: Plotting

A plot of the power of the cavity internal field:

$$\frac{P_{\text{cav}}}{P_{\text{in}}} = \frac{1-R_1}{\left|1+e^{i\varphi}\sqrt{R_1R_2}\right|^2}$$

with the given parameters yields:



We observe a peak in the power profile at detuning $\varphi = \pi$. □

Exercise 7.16: Electric dipole selection rules

($*$) Show that (7.60) is non-zero only when $m_2 - m_1 = \pm 1$ and $\Delta l = \pm 1$.

Solution

Concepts Involved: Commutators, Angular Momentum

Although the intended solution of the exercise seems to be arguing the selection rules directly from the integral over spherical harmonics:

$$\int Y^*_{l_1,m_2} Y_{1m} Y_{l_2m_2} d\Omega$$

The selection rule for l particularly seems to be quite cumbersome to argue, requiring a deep dive into Legendre functions.
Instead, we take an algebraic approach that only relies on the algebra of angular momentum operators (this will imply the integral, which is the matrix element evaluated in the angular basis, must vanish). Throughout we set $\hbar = 1$, and sum over repeated indices. First, recall the canonical commutation relations of position and momentum:

$$[r_i, r_j] = [p_i, p_j] = 0, \quad [r_i, p_j] = i\delta_{ij}$$

if we then define the angular momentum operators:

$$L_i = (\mathbf{r} \times \mathbf{p})_i = \epsilon_{ijk} r_j p_k$$

from this we can compute:

$$[r_i, L_j] = [r_i, \epsilon_{jkm} r_k p_m] = \epsilon_{jkm} r_k [r_i, p_m] = \epsilon_{jkm} r_k i \delta_{im} = i\epsilon_{jki} r_k = i\epsilon_{ijk} r_k$$

and in particular:

$$[r_x, L_z] = -ir_y$$
$$[r_y, L_z] = ir_x$$

First we show the m selection rule. Since the orbital states are eigenstates of L_z with $L_z |l, m\rangle = m |l, m\rangle$, we can consider the relation:

$$\langle l_1, m_1| [r_i, L_z] |l_2, m_2\rangle = \langle l_1, m_1| (r_i L_z - L_z r_i) |l_2, m_2\rangle = (m_1 - m_2) \langle l_1, m_1| r_i |l_2, m_2\rangle$$

which for r_x, r_y give us the two equations:

$$-i \langle l_1, m_1| r_y |l_2, m_2\rangle = (m_1 - m_2) \langle l_1, m_1| r_x |l_2, m_2\rangle$$
$$i \langle l_1, m_1| r_x |l_2, m_2\rangle = (m_1 - m_2) \langle l_1, m_1| r_y |l_2, m_2\rangle$$

Plugging the first equation into the second we find:

$$\langle l_1, m_1| r_x |l_2, m_2\rangle = (m_1 - m_2)^2 \langle l_1, m_1| r_x |l_2, m_2\rangle$$

If $\langle l_1, m_1| r_x |l_2, m_2\rangle$ is to be nonvanishing, then for the above to hold we require $(m_1 - m_2)^2 = 1$, i.e.:

$$m_1 - m_2 = \pm 1$$

which was the claimed selection rule.
We now show the l selection rule, where we will use that the orbital states are also eigenstates of $L^2 = L_x^2 + L_y^2 + L_z^2$ with $L^2 |l, m\rangle = l(l+1) |l, m\rangle$. The commutator algebra is about to get heavy, so let us note the relations:

$$[AB, CD] = AC[B, D] + A[B, C]D + C[A, D]B + [A, C]BD$$

$$\epsilon_{ijk}\epsilon_{imn} = \delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km}$$

and as a special case:

$$\epsilon_{ijk}\epsilon_{ijn} = \delta_{kn}$$

First, we evaluate some commutation rules of orbital angular momentum:

$$
\begin{aligned}
[L_i, L_j] &= [\epsilon_{inm} r_n p_m, \epsilon_{jkl} r_k p_l] \\
&= \epsilon_{inm}\epsilon_{jkl}[r_n p_m, r_k p_l] \\
&= \epsilon_{inm}\epsilon_{jkl}(r_n r_k[p_m, p_l] + r_n[p_m, r_k]p_l, r_k[r_n, p_l]p_m, [r_n, r_k]p_m p_l) \\
&= \epsilon_{inm}\epsilon_{jkl}(0 - i\delta_{mk} r_n p_l + i\delta_{nl} r_k p_m + 0) \\
&= -i\epsilon_{ink}\epsilon_{jkl} r_n p_l + i\epsilon_{inm}\epsilon_{jkn} r_k p_m \\
&= i\epsilon_{ink}\epsilon_{kjl} r_n p_l - i\epsilon_{imn}\epsilon_{njk} r_k p_m \\
&= i(\delta_{ij}\delta_{nl} - \delta_{il}\delta_{nj}) r_n p_l - i(\delta_{ij}\delta_{mk} - \delta_{ik}\delta_{mj}) r_k p_m \\
&= i\delta_{ij} r_n p_n - i r_j p_i - i\delta_{ij} r_m p_m + i r_i p_j \\
&= i(r_i p_j - r_j p_i) \\
&= i\epsilon_{ijk} L_k
\end{aligned}
$$

$$
\begin{aligned}
[L_i, L^2] &= [L_i, L_n L_n] \\
&= L_n[L_i, L_n] + [L_i, L_n]L_n \\
&= L_n(i\epsilon_{ink} L_k) + (i\epsilon_{ink} L_k)L_n \\
&= i\epsilon_{ink}(L_n L_k + L_k L_n) \\
&= 0
\end{aligned}
$$

where in the last equality we observe that ϵ_{ink} is antisymmetric under interchange $n \leftrightarrow k$ while $L_n L_k + L_k L_n$ is symmetric so the product must vanish.
Next, we prove some commutation rules between position and total angular momentum:

$$
\begin{aligned}
[r_i, L^2] &= [r_i, L_n L_n] \\
&= L_n[r_i, L_n] + [r_i, L_n]L_n \\
&= L_n(i\epsilon_{ink} r_k) + (i\epsilon_{ink} r_k)L_n \\
&= i\epsilon_{ink}(L_n r_k + r_k L_n) \\
&= i\epsilon_{ink}([r_k, L_n] + 2r_k L_n) \\
&= i\epsilon_{ink}(-i\epsilon_{knl} r_l + 2r_k L_n) \\
&= \epsilon_{ink}\epsilon_{knl} r_l + 2i\epsilon_{ink} r_k L_n \\
&= -\epsilon_{ink}\epsilon_{lnk} r_l - 2i\epsilon_{ikn} r_k L_n \\
&= -2\delta_{il} r_l - 2i\epsilon_{ikn} r_k L_n \\
&= -2(r_i + i\epsilon_{ikn} r_k L_n)
\end{aligned}
$$

which for z for example evaluates to:

$$[r_z, L^2] = -2(r_z + i r_x L y - i r_y L_x) = 2i(r_y L_x - r_x L_y + i r_z)$$

Now let us recall our previous result for $[r_i, L_j]$; we can multiply both sides by $i\epsilon_{ijn}$ to get:

$$i\epsilon_{ijn}[r_i, L_j] = i\epsilon_{ijn} i\epsilon_{ijk} r_k = -2\delta_{nk} r_k = -2r_n$$

At this point, we temporarily abandon the use of index notation because the relative positions of the L_i, r_i operators becomes important. In particular for r_x, r_y, r_z combining the above two relations we obtain:

$$[r_x, L^2] = 2i(L_y r_z - r_y L_z) = 2i(r_z L_y - L_z r_y)$$
$$[r_y, L^2] = 2i(L_z r_x - r_z L_x) = 2i(r_x L_z - L_x r_z)$$
$$[r_z, L^2] = 2i(L_x r_y - r_x L_y) = 2i(r_y L_x - L_y r_x)$$

Finally, we wish to study the nested commutator:

$$\begin{aligned}
[[r_z, L^2], L^2] &= [2i(r_y L_x - r_x L_y + i r_z), L^2] \\
&= 2i([r_y, L^2]L_x - [r_x, L^2]L_y + i[r_z, L^2]) \\
&= 2i(2i(L_z r_x - r_z L_x)L_x - 2i(r_z L_y - L_z r_y)L_y + i(r_z L^2 - L^2 r_z)) \\
&= -2(2(L_z r_x L_x - r_z L_x^2 - r_z L_y^2 + L_z r_y L_y) + r_z L^2 - L^2 r_z) \\
&= -2(2(L_z r_x L_x - r_z L_x^2 - r_z L_y^2 + L_z r_y L_y + L_z r_z L_z - r_z L_z^2) + r_z L^2 - L^2 r_z) \\
&= -2(2L_z(\mathbf{r} \cdot \mathbf{L}) - 2r_z(L_x^2 + L_y^2 + L_z^2) + r_z L^2 - L^2 r_z) \\
&= -2(-2r_z L^2 + r_z L^2 - L^2 r_z) \\
&= 2(r_z L^2 + L^2 r_z)
\end{aligned}$$

where in the first equality we use the known result for $[r_z, L^2]$, in the second equality we use that $[L_i, L^2] = 0$ so the problem reduces to the commutators with r_i, in the third equality we use the known results for $[r_x, L^2]$ and $[r_y, L^2]$, in the fourth equality we add zero in the form of $L_z r_z L_z - r_z L_z^2$ (r_z, L_z commute) and in the seventh equality we use that $\mathbf{r}, \mathbf{L}$ are orthogonal.
Identical reasoning holds for r_x, r_y and so:

$$[[r_i, L^2], L^2] = 2(r_i L^2 + L^2 r_i).$$

Note that we also obtain directly from the commutator that:

$$[[r_i, L^2], L^2] = (r_i L^2 - L^2 r_i)L^2 - L^2(r_i L^2 - L^2 r_i) = r_i L^2 L^2 - 2L^2 r_i L^2 - L^2 L^2 r_i$$

Now similar to the m selection rule we evaluate $\langle l_1, m_1| [[r_i, L^2], L^2] |l_2, m_2\rangle$, using the two expressions for it we have derived:

$$\langle l_1, m_1| 2(r_i L^2 + L^2 r_i) |l_2, m_2\rangle = \langle l_1, m_1| L^2 L^2 r_i - 2L^2 r_i L^2 + r_i L^2 L^2 |l_2, m_2\rangle$$

Using the eigenvalue relation on both sides:

$$\begin{aligned}
&2(l_1(l_1+1) + l_2(l_2+1)) \langle l_1, m_1| r_i |l_2, m_2\rangle \\
&= \left[(l_1(l_1+1))^2 - 2(l_1(l_1+1)l_2(l_2+1)) + (l_2(l_2+1))^2\right] \langle l_1, m_1| r_i |l_2, m_2\rangle \\
&= (l_1(l_1+1) - l_2(l_2+1))^2 \langle l_1, m_1| r_i |l_2, m_2\rangle
\end{aligned}$$

If the dipole moment is to be nonvanishing, then the prefactors on both sides must be equal, and therefore:

$$2(l_1(l_1+1) + l_2(l_2+1)) = (l_1(l_1+1) - l_2(l_2+1))^2$$

which rewriting both sides:

$$(l_1 - l_2)^2 + (l_1 + l_2 + 1)^2 - 1 = (l_1 - l_2)^2 (l_1 + l_2 + 1)^2$$

which we can rearrange to obtain:

$$[(l_1 - l_2)^2 - 1][(l_1 + l_2 + 1)^2 - 1] = 0$$

The second term vanishes only if $l_1 = l_2 = 0$ but in this case $m_1 = m_2 = 0$ for which the dipole moment must vanish, in contradiction to our prior assumption. Hence is must be the case that the first term vanishes, in which we find that:

$$l_1 - l_2 = \pm 1$$

as claimed. □

Exercise 7.17: Eigenstates of the Jaynes–Cummings Hamiltonian

Show that

$$|\chi_n\rangle = \frac{1}{\sqrt{2}}\left[\,|n,1\rangle + \,|n+1,0\rangle\right]$$
$$|\overline{\chi}_n\rangle = \frac{1}{\sqrt{2}}\left[\,|n,1\rangle - \,|n+1,0\rangle\right]$$

are eigenstates of the Jaynes–Cummings Hamiltonian (7.71) for $\omega = \delta = 0$, with the eigenvalues

$$H\,|\chi_n\rangle = g\sqrt{n+1}\,|\chi_n\rangle$$
$$H\,|\overline{\chi}_n\rangle = -g\sqrt{n+1}\,|\overline{\chi}_n\rangle$$

where the labels in the ket are $|\text{field,atom}\rangle$.

Solution

Concepts Involved: Jaynes–Cummings Model, Creation/Annihilation Operators, Tensor Products

The Jaynes–Cummings Hamiltonian at resonance ($\omega = \delta = 0$) is

$$H = g(a\sigma_+ + a^\dagger\sigma_-),$$

where a and $a^\dagger$ are the field annihilation and creation operators, and $\sigma_+ = |1\rangle\langle 0|\,, \sigma_- = |0\rangle\langle 1|$ act on the atomic qubit.
We compute the action of H on the basis states:

$$H\,|n,1\rangle = ga^\dagger\,|n\rangle \otimes \sigma_-\,|1\rangle = g\sqrt{n+1}\,|n+1,0\rangle\,,$$

$$H\,|n+1,0\rangle = ga\,|n+1\rangle \otimes \sigma_+\,|0\rangle = g\sqrt{n+1}\,|n,1\rangle\,.$$

Therefore,

$$H|\chi_n\rangle = \frac{1}{\sqrt{2}}\big(H|n,1\rangle + H|n+1,0\rangle\big) = g\sqrt{n+1}\,|\chi_n\rangle\,,$$

$$H|\overline{\chi}_n\rangle = \frac{1}{\sqrt{2}}\big(H|n,1\rangle - H|n+1,0\rangle\big) = -g\sqrt{n+1}\,|\overline{\chi}_n\rangle\,.$$

Thus, $|\chi_n\rangle$ and $|\overline{\chi}_n\rangle$ are eigenstates of H with eigenvalues $\pm g\sqrt{n+1}$, respectively. □

Remark: These are the dressed eigenstates of the coupled atom-field system, forming a doublet within the $\{|n,1\rangle, |n+1,0\rangle\}$ subspace with energy splitting $2g\sqrt{n+1}$.

Exercise 7.18: Rabi oscillations

Show that (7.77) is correct by using

$$e^{i\mathbf{n}\cdot\boldsymbol{\sigma}} = \sin|n| + i\hat{\mathbf{n}}\cdot\boldsymbol{\sigma}\cos|n|$$

to exponentiate H. This is an unusually simple derivation of the Rabi oscillations and the Rabi frequency; ordinarily, one solves coupled differential equations to obtain Ω, but here we obtain the essential dynamics just by focusing on the single-atom, single-photon subspace!

Solution

Concepts Involved: Jaynes–Cummings Model, Operator Functions

In the one-excitation subspace $\{|1,0\rangle, |0,1\rangle\}$, the Jaynes–Cummings Hamiltonian with $\delta = 0$ is

$$H = g\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = g\sigma_x.$$

We want to compute the time evolution operator

$$U(t) = e^{-iHt} = e^{-igt\sigma_x}.$$

Using identity (7.78),

$$e^{i\vec{n}\cdot\vec{\sigma}} = \cos|\vec{n}|\,I + i\hat{n}\cdot\vec{\sigma}\,\sin|\vec{n}|,$$

we substitute $\vec{n} = -gt\hat{x}$, so $|\vec{n}| = gt$, and $\hat{n}\cdot\vec{\sigma} = -\sigma_x$. This gives

$$U(t) = \cos(gt)I - i\sigma_x\sin(gt).$$

Acting on the initial state $\big|\psi(0)\big\rangle = |0,1\rangle$ (atom excited, no photons), we find

$$\big|\psi(t)\big\rangle = U(t)\,|0,1\rangle = \cos(gt)\,|0,1\rangle - i\sin(gt)\,|1,0\rangle\,.$$

Thus the system undergoes Rabi oscillations with frequency $\Omega = 2g$, and the excited state population is

$$P_e(t) = \cos^2(gt).$$

□

Remark: This approach derives the full dynamics algebraically using spin rotations, bypassing coupled differential equations typical in semiclassical treatments.

Exercise 7.19: Lorentzian absorption profile

Plot (7.79) for $t = 1$ and $g = 1.2$, as a function of the detuning δ, and (if you know it) the corresponding classical result. What are the oscillations due to?

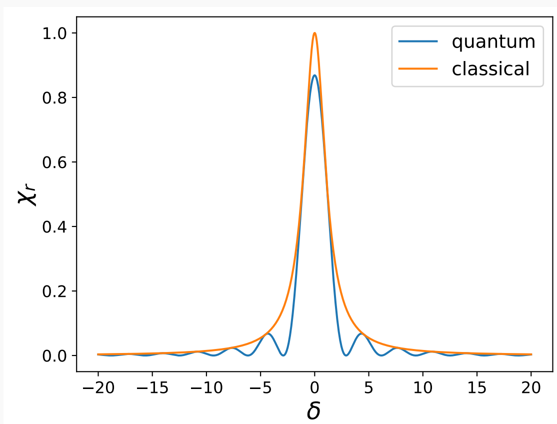
Solution

Concepts Involved: Plotting

A plot of the photon absorption probability:

$$\chi_r = \frac{g^2}{g^2 + \delta^2} \sin^2(\Omega t), \quad \Omega = \sqrt{g^2 + \delta^2}$$

with the given parameters yields:



where the classical result is the same expression without the oscillation factor. The oscillation arises from the exchange of energy between the field and the atom; the absorption probability can oscillate in time (or as a function of the detuning, for a fixed time) as the atom can give back energy to the field. □

Exercise 7.20: Single photon phase shift

Derive (7.80) from U, and plot it for $t = 1$ and $g = 1.2$, as a function of the detuning δ. Compare with δ/Ω^2.

Solution

Concepts Involved: Plotting

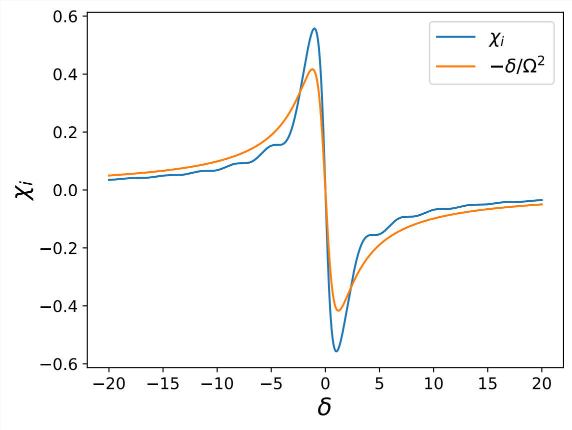
Computing the phase shift of the photon, we take the difference in the rotation angles of the $|1\rangle, |0\rangle$ states of the field, i.e.:

$$\chi_i = \arg(\langle 01| U |01\rangle) - \arg(\langle 00| U |00\rangle)$$

Reading off the matrix elements of U, we get the desired result:

$$\begin{aligned}\chi_i &= \arg(\cos\Omega t + i\frac{\delta}{\Omega}\sin\Omega t) - \arg(e^{-i\delta t}) \\ &= \arg(\cos\Omega t + i\frac{\delta}{\Omega}\sin\Omega t) + \arg(e^{i\delta t}) \\ &= \arg\left(e^{i\delta t}(\cos\Omega t + i\frac{\delta}{\Omega}\sin\Omega t)\right)\end{aligned}$$

Plotting χ_i and δ/Ω^2 for $t = 1, g = 1.2$ we find:



for which we find that $-\delta/\Omega^2$ is a good approximation to the phase shift (though this does not seem to hold for arbitrary values of t, g). □

Exercise 7.21

Explicitly exponentiate (7.82) and show that

$$\varphi_{ab} = \arg\left[e^{i\delta t}\left(\cos\Omega' t - i\frac{\delta}{\Omega'}\sin\Omega' t\right)\right],$$

where $\Omega' = \sqrt{\delta^2 + g_a^2 + g_b^2}$. Use this to compute χ_3, the nonlinear Kerr phase shift. This is a very simple way to model and understand the Kerr interaction, which sidesteps much of the complication typically involved in classical nonlinear optics.

H is given by:

$$H = \begin{bmatrix} H_0 & 0 & 0 \\ 0 & H_1 & 0 \\ 0 & 0 & H_2 \end{bmatrix}$$

where:

$$H_0 = -\delta, H_1 = \begin{bmatrix} -\delta & g_a & 0 & 0 \\ g_a & \delta & 0 & 0 \\ 0 & 0 & -\delta & g_b \\ 0 & 0 & g_b & \delta \end{bmatrix}, H_2 = \begin{bmatrix} -\delta & g_a & g_b \\ g_a & \delta & 0 \\ g_b & 0 & \delta \end{bmatrix}$$

Solution

Concepts Involved: Operator Functions, Kerr Interaction

We exponentitate (7.82) to obtain $U = \exp(iHt)$. Since H is block-diagonal, we can exponentiate block-by-block.

H_0 is simple as it is just a 1×1 matrix:

$$U_{H_0} = \exp(iH_0 t) = \exp(-i\delta t)$$

Exponentiating H_1 is also simple, as it is composed of two diagonal blocks that are identical in form to the Hamiltonian of Eq. (7.76) (with $\delta \leftrightarrow -\delta$), and hence the result can be read off from (7.77):

$$\begin{aligned} U_{H_1} &= (\cos\Omega_a t - i\frac{\delta}{\Omega_a}\sin\Omega_a t)\left|100\right\rangle\!\left\langle 100\right| + (\cos\Omega_a t + i\frac{\delta}{\Omega_a}\sin\Omega_a t)\left|001\right\rangle\!\left\langle 001\right| \\ &\quad - i\frac{g_a}{\Omega_a}\sin\Omega_a t(\left|100\right\rangle\!\left\langle 001\right| + \left|001\right\rangle\!\left\langle 100\right|) \\ &\quad + (\cos\Omega_b t - i\frac{\delta}{\Omega_b}\sin\Omega_b t)\left|010\right\rangle\!\left\langle 010\right| + (\cos\Omega_b t + i\frac{\delta}{\Omega_b}\sin\Omega_b t)\left|002\right\rangle\!\left\langle 002\right| \\ &\quad - i\frac{g_b}{\Omega_b}\sin\Omega_b t(\left|010\right\rangle\!\left\langle 002\right| + \left|002\right\rangle\!\left\langle 010\right|) \end{aligned}$$

where $\Omega_i = \sqrt{g_i^2 + \delta^2}$.

For H_2 we have to do a little more work. Asking `sympy` to diagonalize the matrix, we obtain:

$$H_2 = PDP^{-1} \cong \begin{bmatrix} 0 & \frac{-\delta-\Omega'}{g_b} & \frac{-\delta+\Omega'}{g_b} \\ -\frac{g_b}{g_a} & \frac{g_a}{g_b} & \frac{g_a}{g_b} \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \delta & 0 & 0 \\ 0 & -\Omega' & 0 \\ 0 & 0 & \Omega' \end{bmatrix} \begin{bmatrix} 0 & -\frac{1}{\frac{g_a}{g_b}+\frac{g_b}{g_a}} & \frac{g_a}{g_a+\frac{g_b^2}{g_a}} \\ -\frac{g_b}{2\Omega'} & \frac{g_a g_b(-\delta+\Omega')}{2(g_a^2+g_b^2)\Omega'} & \frac{g_b^2(-\delta+\Omega')}{2(g_a^2+g_b^2)\Omega'} \\ \frac{g_b}{2\Omega'} & \frac{g_a g_b(\delta+\Omega')}{2(g_a^2+g_b^2)\Omega'} & \frac{g_b^2(\delta+\Omega')}{2(g_a^2+g_b^2)\Omega'} \end{bmatrix}$$

P above can be made unitary by normalization of the eigenvectors appearing in the columns, but the result will be the same.
Exponentiating H_2 is done just by exponentiating the eigenvalues:

$$U_{H_2} \cong \begin{bmatrix} 0 & \frac{-\delta-\Omega'}{g_b} & \frac{-\delta+\Omega'}{g_b} \\ -\frac{g_b}{g_a} & \frac{g_a}{g_b} & \frac{g_a}{g_b} \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} e^{i\delta t} & 0 & 0 \\ 0 & e^{-i\Omega' t} & 0 \\ 0 & 0 & e^{i\Omega' t} \end{bmatrix} \begin{bmatrix} 0 & -\frac{1}{\frac{g_a}{g_b}+\frac{g_b}{g_a}} & \frac{g_a}{g_a+\frac{g_b^2}{g_a}} \\ -\frac{g_b}{2\Omega'} & \frac{g_a g_b(-\delta+\Omega')}{2(g_a^2+g_b^2)\Omega'} & \frac{g_b^2(-\delta+\Omega')}{2(g_a^2+g_b^2)\Omega'} \\ \frac{g_b}{2\Omega'} & \frac{g_a g_b(\delta+\Omega')}{2(g_a^2+g_b^2)\Omega'} & \frac{g_b^2(\delta+\Omega')}{2(g_a^2+g_b^2)\Omega'} \end{bmatrix}$$

all of the matrix elements of U can then be obtained by matrix multiplication.
For φ_{ab} we only require $\langle 110| U |110\rangle$, so let us compute this entry:

$$U_{H_2} \cong \begin{bmatrix} 0 & \frac{-\delta-\Omega'}{g_b} e^{-i\Omega' t} & \frac{-\delta+\Omega'}{g_b} e^{i\Omega' t} \\ * & * & * \\ * & * & * \end{bmatrix} \begin{bmatrix} 0 & * & * \\ -\frac{g_b}{2\Omega'} & * & * \\ \frac{g_b}{2\Omega'} & * & * \end{bmatrix} = \begin{bmatrix} -\frac{g_b}{2\Omega'}\frac{-\delta-\Omega'}{g_b} e^{-i\Omega' t} + \frac{g_b}{2\Omega'}\frac{-\delta+\Omega'}{g_b} e^{i\Omega' t} & * & * \\ * & * & * \\ * & * & * \end{bmatrix}$$

Thus:

$$\begin{aligned} \langle 110| U |110\rangle &= \frac{1}{2\Omega'}(\Omega'(e^{i\Omega' t} + e^{-i\Omega' t}) - \delta(e^{i\Omega' t} - e^{-i\Omega' t})) \\ &= \cos \Omega' t - i\frac{\delta}{\Omega'} \sin \Omega' t \end{aligned}$$

where in the last equality we use Euler's formula. Thus computing the two-photon phase shift:

$$\begin{aligned} \varphi_{ab} &= \arg[\langle 110| U |110\rangle] - \arg[\langle 000| U |000\rangle] \\ &= \arg[\cos \Omega' t - i\frac{\delta}{\Omega'} \sin \Omega' t] - \arg[e^{-i\delta t}] \\ &= \arg[\cos \Omega' t - i\frac{\delta}{\Omega'} \sin \Omega' t] + \arg[e^{i\delta t}] \\ &= \arg\left[e^{i\delta t}\left(\cos \Omega' t - i\frac{\delta}{\Omega'} \sin \Omega' t\right)\right] \end{aligned}$$

we get the claimed result. □

Exercise 7.22

Associated with the cross phase modulation is also a certain amount of loss, which is given by the probability that a photon is absorbed by the atom. Compute this probability, $1 - \langle 110|U|110\rangle$, where $U = \exp(-iHt)$ for H as in (7.82); compare with $1 - \langle 100|U|100\rangle$ as a function of δ, g_a, g_b, and t.
The formulas for the probabilities should be $1 - \left|\langle 110| U |110\rangle\right|^2$ and $1 - \left|\langle 100| U |100\rangle\right|^2$.

Solution

Concepts Involved: Operator Functions, Kerr Interaction

Computing the loss probability, we take the mod square of $\langle 110| U |110\rangle$ we found in the previous question:

$$\begin{aligned}
p_{\text{2-loss}} &= 1 - \left|\cos\Omega' t - i\frac{\delta}{\Omega'}\sin\Omega' t\right|^2 \\
&= 1 - \left(\cos^2(\Omega' t) + \frac{\delta^2}{\Omega'^2}\sin^2(\Omega' t)\right) \\
&= 1 - \left(\cos^2(\Omega' t) + \left(\frac{\delta^2 + g_a^2 + g_b^2 - g_a^2 + g_b^2}{\delta^2 + g_a^2 + g_b^2}\right)\sin^2(\Omega' t)\right) \\
&= 1 - \left(\cos^2(\Omega'^2) + \sin^2(\Omega' t) - \frac{g_a^2 + g_b^2}{\Omega'^2}\sin^2(\Omega' t)\right) \\
&= \frac{g_a^2 + g_b^2}{\Omega'^2}\sin^2(\Omega' t) \\
&= \frac{1}{1 + \frac{\delta}{g_a^2 + g_b^2}}\sin^2(\Omega' t)
\end{aligned}$$

Comparing this to the single photon loss case (also using our result from the previous exercise):

$$\begin{aligned}
p_{\text{1-loss}} &= 1 - \left|\cos\Omega_a t - i\frac{\delta}{\Omega_a}\sin\Omega_a t\right|^2 \\
&= \frac{1}{1 + \frac{\delta}{g_a^2}}\sin^2(\Omega_a t)
\end{aligned}$$

This is just the same expression with $g_b = 0$. Averaging over time (for $g_b > 0$) $p_{2-loss} > p_{1-loss}$ which makes sense, as the two-photon absorption probability should be physically higher than the one-photon absorption probability. □

Exercise 7.23

Show that the two qubit gate of (7.87) can be used to realize a controlled-NOT gate, when augmented with arbitrary single qubit operations, for any φ_a and φ_b, and $\Delta = \pi$. It turns out that for nearly any value of Δ this gate is universal when augmented with single qubit unitaries.

Solution

Concepts Involved: Controlled Operations

We take the two qubit gate of (7.87) and compose it with the phase gate diag$(1, e^{-i\varphi_b})$ on the first qubit and the phase gate diag$(1, e^{-i\varphi_a})$ on the first second, which yields:

$$\begin{aligned}
&\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & e^{i\varphi_a} & 0 & 0 \\ 0 & 0 & e^{i\varphi_b} & 0 \\ 0 & 0 & 0 & e^{i(\varphi_a+\varphi_b+\Delta)} \end{bmatrix} \cdot \left(\begin{bmatrix} 1 & 0 \\ 0 & e^{-i\varphi_b} \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 \\ 0 & e^{-i\varphi_a} \end{bmatrix} \right) \\
&= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & e^{i\varphi_a} & 0 & 0 \\ 0 & 0 & e^{i\varphi_b} & 0 \\ 0 & 0 & 0 & e^{i(\varphi_a+\varphi_b+\Delta)} \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & e^{-i\varphi_a} & 0 & 0 \\ 0 & 0 & e^{-i\varphi_b} & 0 \\ 0 & 0 & 0 & e^{-i(\varphi_a+\varphi_b)} \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & e^{i\Delta} \end{bmatrix}
\end{aligned}$$

for $\Delta = \pi$ this is just the controlled-Z gate. If we conjugate this by two Hadamard gates (which specifies the target qubit) as in Ex. 4.17, we get the desired `CNOT` gate. □

Exercise 7.24

The energy of a nuclear spin in a magnetic field is approximately $\mu_N B$, where $\mu_N = eh/4\pi m_p \approx 5 \times 10^{-27}$ joules per tesla is the nuclear Bohr magneton. Compute the energy of a nuclear spin in a $B = 10$ tesla field, and compare with the thermal energy $k_B T$ at $T = 300K$.

Solution

Concepts Involved: Numerical Estimation

The energy of a nuclear spin in a 10 Tesla field is:

$$E_{\text{mag}} \approx \mu_N B = \frac{eh}{4\pi m_p} B \approx (5 \times 10^{-27} \text{JT}^{-1}) 10\text{T} = 5 \times 10^{-26}\text{J}$$

Which compared to the thermal energy at 300K is:

$$E_{\text{therm}} = k_B T = (1.4 \times 10^{-23} \text{JK}^{-1})(300\text{K}) = 4 \times 10^{-21}\text{J}$$

Thus we find that the thermal energy is 5 orders of magnitude larger. □

Exercise 7.25

Show that the total angular momenta operators obey the commutation relations for $SU(2)$, that is, $[j_i, j_k] = i\epsilon_{ikl} j_l$.

Solution

Concepts Involved: Angular Momentum, Pauli Operators

We find:

$$\begin{aligned}
[j_i, j_k] &= [\frac{\sigma_i^1 + \sigma_i^2}{2}, \frac{\sigma_k^1 + \sigma_k^2}{2}] \\
&= \frac{1}{4}\left([\sigma_i^1, \sigma_k^1] + [\sigma_i^1, \sigma_k^2] + [\sigma_i^2, \sigma_k^1] + [\sigma_i^2, \sigma_k^2]\right) \\
&= \frac{1}{4}\left(2i\epsilon_{ikl}\sigma_l^1 + 0 + 0 + 2i\epsilon_{ikl}\sigma_l^2\right) \\
&= i\epsilon_{ikl}\frac{\sigma_l^1 + \sigma_l^2}{2} \\
&= i\epsilon_{ikl}j_l
\end{aligned}$$

where in the third equality we use the Pauli commutation relations as worked out in Ex. 2.40. □

Exercise 7.26

Verify the properties of $|j, m_j\rangle_J$ by explicitly writing the 4×4 matrices J^2 and j_z in the basis defined by $|j, m_j\rangle_J$.

The definition of the states given in the text in Eqs. (7.94), (7.96) is inconsistent with the convention that $Z|0\rangle = +|0\rangle$ (spin-up) and $Z|1\rangle = -|1\rangle$ (spin-down). Accounting for this, the basis we use is:

$$\begin{aligned}
|0, 0\rangle_J &= \frac{|01\rangle - |10\rangle}{\sqrt{2}} \\
|1, -1\rangle_J &= |11\rangle \\
|1, 0\rangle_J &= \frac{|01\rangle + |10\rangle}{\sqrt{2}} \\
|1, 1\rangle_J &= |00\rangle
\end{aligned}$$

Somewhere along in the book there seems to be a switch in convention where $Z \to -Z$ and $Y \to -Y$, but in this solution manual we will not follow this switch.

Solution

Concepts Involved: Angular Momentum, Composite Systems

We check that the provided states are indeed eigenstates of J^2 with eigenvalue $j(j+1)$, and j_z with eigenvalue m by showing that the two operators are simultaneously diagonalized in the given basis. The j_i operators are given by:

$$j_z = \frac{Z_1 \otimes I_2 + I_1 \otimes Z_2}{2} \cong \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

$$j_x = \frac{X_1 \otimes I_2 + I_1 \otimes X_2}{2} \cong \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$$

$$j_y = \frac{Y_1 \otimes I_2 + I_1 \otimes Y_2}{2} \cong \begin{bmatrix} 0 & -\frac{i}{2} & -\frac{i}{2} & 0 \\ \frac{i}{2} & 0 & 0 & -\frac{i}{2} \\ \frac{i}{2} & 0 & 0 & -\frac{i}{2} \\ 0 & \frac{i}{2} & \frac{i}{2} & 0 \end{bmatrix}$$

therein J^2 is given by:

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

Forming the columns of the basis change unitary U from the given vectors we have:

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

We can then explicitly write J^2, j_z in the given basis by conjugating by U:

$$U^\dagger J^2 U \cong \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 0(0+1) & 0 & 0 & 0 \\ 0 & 1(1+1) & 0 & 0 \\ 0 & 0 & 1(1+1) & 0 \\ 0 & 0 & 0 & 1(1+1) \end{bmatrix}$$

$$U^\dagger j_z U \cong \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

from which we confirm that $J^2 \left|j, m_j\right\rangle_J = j(j+1) \left|j, m_j\right\rangle_J$ and $j_z \left|j, m_j\right\rangle_J = m \left|j, m_j\right\rangle_J$ hold. □

Exercise 7.27: Three spin angular momenta states

Three spin-1/2 states can combine together to give states of total angular momenta with $j = 1/2$ and $j = 3/2$. Show that the states

$$\begin{aligned}
|3/2, 3/2\rangle &= |111\rangle \\
|3/2, 1/2\rangle &= \frac{1}{\sqrt{3}}\left[|011\rangle + |101\rangle + |110\rangle\right] \\
|3/2, -1/2\rangle &= \frac{1}{\sqrt{3}}\left[|100\rangle + |010\rangle + |001\rangle\right] \\
|3/2, -3/2\rangle &= |000\rangle \\
|1/2, 1/2\rangle_1 &= \frac{1}{\sqrt{2}}\left[-|001\rangle + |100\rangle\right] \\
|1/2, -1/2\rangle_1 &= \frac{1}{\sqrt{2}}\left[|110\rangle - |011\rangle\right] \\
|1/2, 1/2\rangle_2 &= \frac{1}{\sqrt{6}}\left[|001\rangle - 2|010\rangle + |100\rangle\right] \\
|1/2, -1/2\rangle_2 &= \frac{1}{\sqrt{6}}\left[-|110\rangle + 2|101\rangle - |011\rangle\right]
\end{aligned}$$

form a basis for the space, satisfying $J^2|j, m_j\rangle = j(j+1)|j, m_j\rangle$ and $j_z|j, m_j\rangle = m_j|j, m_j\rangle$, for $j_z = (Z_1 + Z_2 + Z_3)/2$ (similarly for j_x and j_y) and $J^2 = j_x^2 + j_y^2 + j_z^2$. There are sophisticated ways to obtain these states, but a straightforward brute-force method is simply to simultaneously diagonalize the 8×8 matrices J^2 and j_z.

The definition of the states given above is inconsistent with the convention that $Z|0\rangle = +|0\rangle$ (spin-up) and $Z|1\rangle = -|1\rangle$ (spin-down), for the $J = 3/2$ subspace. The corrected states we use for that subspace are:

$$\begin{aligned}
|3/2, 3/2\rangle &= |000\rangle \\
|3/2, 1/2\rangle &= \frac{1}{\sqrt{3}}\left[|100\rangle + |010\rangle + |001\rangle\right] \\
|3/2, -1/2\rangle &= \frac{1}{\sqrt{3}}\left[|011\rangle + |101\rangle + |110\rangle\right] \\
|3/2, -3/2\rangle &= |111\rangle
\end{aligned}$$

Solution

Concepts Involved: Angular Momentum, Composite Systems

We follow the hint (and our approach from the previous problem) and simply verify that these states simultaneously diagonalize J^2, j_z. The 8×8 matrices for the j_is are given by:

$$j_z = \frac{Z_1 \otimes I_2 \otimes I_3 + I_1 \otimes Z_2 \otimes I_3 + I_1 \otimes I_2 \otimes Z_3}{2} \cong \begin{bmatrix} \frac{3}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{3}{2} \end{bmatrix}$$

$$j_x = \frac{X_1 \otimes I_2 \otimes I_3 + I_1 \otimes X_2 \otimes I_3 + I_1 \otimes I_2 \otimes X_3}{2} \cong \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$$

$$j_y = \frac{Y_1 \otimes I_2 \otimes I_3 + I_1 \otimes Y_2 \otimes I_3 + I_1 \otimes I_2 \otimes Y_3}{2} \cong \begin{bmatrix} 0 & -\frac{i}{2} & -\frac{i}{2} & 0 & -\frac{i}{2} & 0 & 0 & 0 \\ \frac{i}{2} & 0 & 0 & -\frac{i}{2} & 0 & -\frac{i}{2} & 0 & 0 \\ \frac{i}{2} & 0 & 0 & -\frac{i}{2} & 0 & 0 & -\frac{i}{2} & 0 \\ 0 & \frac{i}{2} & \frac{i}{2} & 0 & 0 & 0 & 0 & -\frac{i}{2} \\ \frac{i}{2} & 0 & 0 & 0 & 0 & -\frac{i}{2} & -\frac{i}{2} & 0 \\ 0 & \frac{i}{2} & 0 & 0 & \frac{i}{2} & 0 & 0 & -\frac{i}{2} \\ 0 & 0 & \frac{i}{2} & 0 & \frac{i}{2} & 0 & 0 & -\frac{i}{2} \\ 0 & 0 & 0 & \frac{i}{2} & 0 & \frac{i}{2} & \frac{i}{2} & 0 \end{bmatrix}$$

So the matrix for J^2 is given by:

$$J^2 = j_x^2 + j_y^2 + j_z^2 \cong \begin{bmatrix} \frac{15}{4} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{7}{4} & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & \frac{7}{4} & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{7}{4} & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & \frac{7}{4} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & \frac{7}{4} & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & \frac{7}{4} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{15}{4} \end{bmatrix}$$

Forming the columns of the basis change U from the provided vectors:

$$U \cong \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{3}} & 0 & 0 & -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{6}} & 0 \\ 0 & \frac{1}{\sqrt{3}} & 0 & 0 & 0 & 0 & -\frac{2}{\sqrt{6}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{3}} & 0 & 0 & -\frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{3}} & 0 & 0 & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{6}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{3}} & 0 & 0 & 0 & 0 & \frac{2}{\sqrt{6}} \\ 0 & 0 & \frac{1}{\sqrt{3}} & 0 & 0 & \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{6}} \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

We can check that this diagonalizes J^2, j_z via matrix multiplication:

$$U^\dagger J^2 U \cong \begin{bmatrix} \frac{15}{4} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{15}{4} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{15}{4} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{15}{4} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{3}{4} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{3}{4} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{3}{4} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{3}{4} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{3}{2}(\frac{3}{2}+1) & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{3}{2}(\frac{3}{2}+1) & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{3}{2}(\frac{3}{2}+1) & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{3}{2}(\frac{3}{2}+1) & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2}(\frac{1}{2}+1) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{2}(\frac{1}{2}+1) & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2}(\frac{1}{2}+1) & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2}(\frac{1}{2}+1) \end{bmatrix}$$

$$U^\dagger j_z U \cong \begin{bmatrix} \frac{3}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{3}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{3}{2} \end{bmatrix}$$

Thus the claim that the provided states are eigenstates of J^2, j_z with the claimed eigenvalues are proven. Finally, we argue that these states form a basis. Since the Hilbert space is $\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2 = \mathbb{C}^8$, given 8 vectors (the same number as the dimension as the space) it suffices to check they are linearly independent. In fact we can do better and argue that the states are orthogonal and so the basis is orthonormal. We could check this one-by-one, but we can also observe that the orthogonality of most of the states follows from the fact that they are eigenstates of Hermitian operators J^2, j_z with distinct eigenvalues (Ex. 2.22). The only degenerate eigenstates are the pairs $\left|1/2, 1/2\right\rangle_1, \left|1/2, 1/2\right\rangle_2$ and $\left|1/2, -1/2\right\rangle_1, \left|1/2, -1/2\right\rangle_2$ but these states have no overlap by inspection and so are orthogonal.

□

Exercise 7.28: Hyperfine states

We shall be taking a look at beryllium in Section 7.6.4 – the total angular momenta states relevant there involve a nuclear spin $I = 3/2$ combining with an electron spin $S = 1/2$ to give $F = 2$ or $F = 1$. For a spin-3/2 particle, the angular momenta operators are

$$i_x = \frac{1}{2}\begin{bmatrix} 0 & \sqrt{3} & 0 & 0 \\ \sqrt{3} & 0 & 2 & 0 \\ 0 & 2 & 0 & \sqrt{3} \\ 0 & 0 & \sqrt{3} & 0 \end{bmatrix}$$

$$i_y = \frac{1}{2}\begin{bmatrix} 0 & i\sqrt{3} & 0 & 0 \\ -i\sqrt{3} & 0 & 2i & 0 \\ 0 & -2i & 0 & i\sqrt{3} \\ 0 & 0 & -i\sqrt{3} & 0 \end{bmatrix}$$

$$i_z = \frac{1}{2}\begin{bmatrix} -3 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

1. Show that i_x, i_y and i_z satisfy $SU(2)$ commutation rules.

2. Give 8×8 matrix representations of $f_z = i_z \otimes I + I \otimes Z/2$ (where I here represents the identity operator on the appropriate subspace) and similarly f_x and f_y, and, $F^2 = f_x^2 + f_y^2 + f_z^2$. Simultaneously diagonalize f_z and F^2 to obtain basis states $\left|F, m_F\right\rangle$ for which $F^2\left|F, m_F\right\rangle = F(F+1)\left|F, m_F\right\rangle$ and $f_z\left|F, m_F\right\rangle = m_F\left|F, m_F\right\rangle$.

As noted in a previous exercises, the text seems to have performed a $S_z \to -S_z, S_y \to -S_y$ switch in a sign convention. To stay internally consistent, we will not perform this switch, and so will use:

$$i_y = \frac{1}{2}\begin{bmatrix} 0 & -i\sqrt{3} & 0 & 0 \\ i\sqrt{3} & 0 & -2i & 0 \\ 0 & 2i & 0 & -i\sqrt{3} \\ 0 & 0 & i\sqrt{3} & 0 \end{bmatrix}$$

$$i_z = \frac{1}{2}\begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -3 \end{bmatrix}$$

Solution

Concepts Involved: Angular Momentum, Composite Systems

1. We compute the commutators:

$$[i_x, i_y] = \frac{1}{4}\left(\begin{bmatrix} 3i & 0 & -2\sqrt{3}i & 0 \\ 0 & i & 0 & -2\sqrt{3}i \\ 2\sqrt{3}i & 0 & -i & 0 \\ 0 & 2\sqrt{3}i & 0 & -3i \end{bmatrix} - \begin{bmatrix} -3i & 0 & -2\sqrt{3}i & 0 \\ 0 & -i & 0 & -2\sqrt{3}i \\ 2\sqrt{3}i & 0 & i & 0 \\ 0 & 2\sqrt{3}i & 0 & 3i \end{bmatrix}\right)$$
$$= i\frac{1}{2}\begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -3 \end{bmatrix}$$
$$= ii_z$$

$$[i_y, i_z] = \frac{1}{4}\left(\begin{bmatrix} 0 & -i\sqrt{3} & 0 & 0 \\ 3i\sqrt{3} & 0 & 2i & 0 \\ 0 & 2i & 0 & 3i\sqrt{3} \\ 0 & 0 & -i\sqrt{3} & 0 \end{bmatrix} - \begin{bmatrix} 0 & -3i\sqrt{3} & 0 & 0 \\ i\sqrt{3} & 0 & -2i & 0 \\ 0 & -2i & 0 & i\sqrt{3} \\ 0 & 0 & -3i\sqrt{3} & 0 \end{bmatrix}\right)$$
$$= i\frac{1}{2}\begin{bmatrix} 0 & \sqrt{3} & 0 & 0 \\ \sqrt{3} & 0 & 2 & 0 \\ 0 & 2 & 0 & \sqrt{3} \\ 0 & 0 & \sqrt{3} & 0 \end{bmatrix}$$
$$= ii_x$$

$$[i_z, i_x] = \frac{1}{4}\left(\begin{bmatrix} 0 & 3\sqrt{3} & 0 & 0 \\ \sqrt{3} & 0 & 2 & 0 \\ 0 & -2 & 0 & -\sqrt{3} \\ 0 & 0 & -3\sqrt{3} & 0 \end{bmatrix} - \begin{bmatrix} 0 & \sqrt{3} & 0 & 0 \\ 3\sqrt{3} & 0 & -2 & 0 \\ 0 & 2 & 0 & -3\sqrt{3} \\ 0 & 0 & -\sqrt{3} & 0 \end{bmatrix}\right)$$
$$= i\frac{1}{2}\begin{bmatrix} 0 & -i\sqrt{3} & 0 & 0 \\ i\sqrt{3} & 0 & -2i & 0 \\ 0 & 2i & 0 & -i\sqrt{3} \\ 0 & 0 & i\sqrt{3} & 0 \end{bmatrix}$$
$$= ii_y$$

Combining the above with $[A, B] = -[B, A]$ (Ex. 2.46) and $[A, A] = 0$ we conclude that $[i_j, i_k] = i\epsilon_{jkl}i_l$ as claimed.

2. We make the identification:

$$|i_z = \frac{3}{2}, 0\rangle \cong \begin{bmatrix}1\\0\\0\\0\\0\\0\\0\\0\end{bmatrix}, |i_z = \frac{3}{2}, 1\rangle \cong \begin{bmatrix}0\\1\\0\\0\\0\\0\\0\\0\end{bmatrix}, |i_z = \frac{1}{2}, 0\rangle \cong \begin{bmatrix}0\\0\\1\\0\\0\\0\\0\\0\end{bmatrix}, |i_z = \frac{1}{2}, 1\rangle \cong \begin{bmatrix}0\\0\\0\\1\\0\\0\\0\\0\end{bmatrix},$$

$$|i_z = -\frac{1}{2}, 0\rangle \cong \begin{bmatrix}0\\0\\0\\0\\1\\0\\0\\0\end{bmatrix}, |i_z = -\frac{1}{2}, 1\rangle \cong \begin{bmatrix}0\\0\\0\\0\\0\\1\\0\\0\end{bmatrix}, |i_z = -\frac{3}{2}, 0\rangle \cong \begin{bmatrix}0\\0\\0\\0\\0\\0\\1\\0\end{bmatrix}, |i_z = -\frac{3}{2}, 1\rangle \cong \begin{bmatrix}0\\1\\0\\0\\0\\0\\0\\1\end{bmatrix}$$

The matrix representations of the joint spin operators are then given by:

$$f_z \cong \begin{bmatrix}2 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\0 & 1 & 0 & 0 & 0 & 0 & 0 & 0\\0 & 0 & 1 & 0 & 0 & 0 & 0 & 0\\0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\0 & 0 & 0 & 0 & 0 & -1 & 0 & 0\\0 & 0 & 0 & 0 & 0 & 0 & -1 & 0\\0 & 0 & 0 & 0 & 0 & 0 & 0 & -2\end{bmatrix}$$

$$f_x \cong \begin{bmatrix}0 & \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 & 0 & 0 & 0 & 0\\\frac{1}{2} & 0 & 0 & \frac{\sqrt{3}}{2} & 0 & 0 & 0 & 0\\\frac{\sqrt{3}}{2} & 0 & 0 & \frac{1}{2} & 1 & 0 & 0 & 0\\0 & \frac{\sqrt{3}}{2} & \frac{1}{2} & 0 & 0 & 1 & 0 & 0\\0 & 0 & 1 & 0 & 0 & \frac{1}{2} & \frac{\sqrt{3}}{2} & 0\\0 & 0 & 0 & 1 & \frac{1}{2} & 0 & 0 & \frac{\sqrt{3}}{2}\\0 & 0 & 0 & 0 & \frac{\sqrt{3}}{2} & 0 & 0 & \frac{1}{2}\\0 & 0 & 0 & 0 & 0 & \frac{\sqrt{3}}{2} & \frac{1}{2} & 0\end{bmatrix}$$

$$f_y \cong \begin{bmatrix} 0 & -\frac{i}{2} & -\frac{i\sqrt{3}}{2} & 0 & 0 & 0 & 0 & 0 \\ \frac{i}{2} & 0 & 0 & -\frac{i\sqrt{3}}{2} & 0 & 0 & 0 & 0 \\ \frac{i\sqrt{3}}{2} & 0 & 0 & -\frac{i}{2} & -i & 0 & 0 & 0 \\ 0 & \frac{i\sqrt{3}}{2} & \frac{i}{2} & 0 & 0 & -i & 0 & 0 \\ 0 & 0 & i & 0 & 0 & -\frac{i}{2} & -\frac{i\sqrt{3}}{2} & 0 \\ 0 & 0 & 0 & i & \frac{i}{2} & 0 & 0 & -\frac{i\sqrt{3}}{2} \\ 0 & 0 & 0 & 0 & \frac{i\sqrt{3}}{2} & 0 & 0 & -\frac{i}{2} \\ 0 & 0 & 0 & 0 & 0 & \frac{i\sqrt{3}}{2} & \frac{i}{2} & 0 \end{bmatrix}$$

From this we can construct F^2:

$$F^2 = \begin{bmatrix} 6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & \sqrt{3} & 0 & 0 & 0 & 0 & 0 \\ 0 & \sqrt{3} & 5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 & \sqrt{3} & 0 \\ 0 & 0 & 0 & 0 & 0 & \sqrt{3} & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 6 \end{bmatrix}$$

We notice that F^2 is block diagonal, and moreover that within a given block all states have the same eigenvalue of f_z. Hence to simultaneously diagonalize the two operators, it suffices to diagonalize the three nontrivial blocks of F^2:

$$\begin{bmatrix} 3 & \sqrt{3} \\ \sqrt{3} & 5 \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 6 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix} \quad (m = 1 \text{ block})$$

$$\begin{bmatrix} 4 & 2 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 6 \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \quad (m = 0 \text{ block})$$

$$\begin{bmatrix} 5 & \sqrt{3} \\ \sqrt{3} & 3 \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 6 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix} \quad (m = -1 \text{ block})$$

We can read off the eigenvectors from the columns of the diagonalizing unitaries. Organizing the eigenvalues of F^2 in order, in the basis where F^2, f_z are simultaneously diagonal we have:

$$F^2 \cong \begin{bmatrix} 6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 6 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 6 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 6 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} 2(2+1) & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2(2+1) & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2(2+1) & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2(2+1) & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2(2+1) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1(1+1) & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1(1+1) & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1(1+1) \end{bmatrix}$$

$$f_z \cong \begin{bmatrix} 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}.$$

Where the basis vectors are:

$$\begin{aligned}
|2,2\rangle_F &= |\frac{3}{2},0\rangle \\
|2,1\rangle_F &= \frac{1}{2}\left(|\frac{3}{2},1\rangle + \sqrt{3}\,|\frac{1}{2},0\rangle\right) \\
|2,0\rangle_F &= \frac{1}{\sqrt{2}}\left(|\frac{1}{2},1\rangle + |-\frac{1}{2},0\rangle\right) \\
|2,-1\rangle_F &= \frac{1}{2}\left(-|-\frac{1}{2},0\rangle + \sqrt{3}\,|-\frac{3}{2},0\rangle\right) \\
|2,-2\rangle_F &= |-\frac{3}{2},1\rangle \\
|1,1\rangle_F &= \frac{1}{2}\left(-|\frac{1}{2},0\rangle + \sqrt{3}\,|\frac{3}{2},1\rangle\right) \\
|1,0\rangle_F &= \frac{1}{\sqrt{2}}\left(-|\frac{1}{2},1\rangle + |-\frac{1}{2},0\rangle\right) \\
|1,-1\rangle_F &= \frac{1}{2}\left(|-\frac{1}{2},1\rangle + \sqrt{3}\,|-\frac{3}{2},0\rangle\right)
\end{aligned}$$

□

Exercise 7.29: Spontaneous emission

The spontaneous emission rate (7.112) can be derived from (7.110)–(7.111) by the following steps.

1. Integrate

$$\frac{1}{(2\pi c)^3}\frac{8\pi}{3}\int_0^\infty \omega^2 p_{\text{decay}} d\omega$$

   where the $8\pi/3$ comes from summing over polarizations and integrating over the solid angle $d\Omega$, and $\omega^2/(2\pi c)^3$ comes from the mode density in three-dimensional space. (*Hint:* you may want to extend the lower limit of the integral to $-\infty$.)

2. Differentiate the result with respect to t to obtain γ_{rad}.

The form of g^2 is a result of quantum electrodynamics; taking this for granted, the remainder of the calculation as presented here really stems from just the Jaynes–Cummings interaction. Again, we see how considering its properties in the single atom, single photon regime gives us a fundamental property of atoms, without resorting to perturbation theory!

The expression for p_{decay} given in the text has a typo (no square in the arguement of the sin), and it should be:

$$p_{\text{decay}} = g^2\frac{4\sin^2(\frac{1}{2}(\omega-\omega_0)t)}{(\omega-\omega_0)^2}$$

The atom-field coupling constant is given by:

$$g^2 = \frac{\omega_0^2}{2\hbar\omega\epsilon_0 c^2}\left|\langle 0|\boldsymbol{\mu}|1\rangle\right|^2$$

Formally, the integral diverges. We discuss this subtlety in the solution.

Solution

Concepts Involved: Integration

1. Substituting the expression for the decay probability and the coupling constant into the integral for the total decay probability P_{decay}, we get:

$$\begin{aligned} P_{\text{decay}} &= \frac{1}{(2\pi c)^3}\frac{8\pi}{3}\int_0^\infty \omega^2\frac{\omega_0^2}{2\hbar\omega\epsilon_0 c^2}\left|\langle 0|\boldsymbol{\mu}|1\rangle\right|^2\frac{4\sin^2(\frac{1}{2}(\omega-\omega_0)t)}{(\omega-\omega_0)^2}d\omega \\ &= \frac{2\omega_0^2\left|\langle 0|\boldsymbol{\mu}|1\rangle\right|^2}{3\pi^2\hbar\epsilon_0 c^5}\int_0^\infty \frac{\omega\sin^2(\frac{1}{2}(\omega-\omega_0)t)}{(\omega-\omega_0)^2}d\omega \end{aligned}$$

   Formally, the above integral is of the form $\int_0^\infty \frac{\sin^2(x)}{x}$ which is divergent. This occurs due to the fact that the expression for the mode density breaks down at high frequencies (else the QED vacuum energy would be infinite). A more careful derivation of the spontaneous emission probability can be done from the stimulated emission rate + Einstein A/B coefficient method, e.g. as is found in Chapter 11 of Griffiths https://en.wikipedia.org/wiki/Introduction_to_Quantum_Mechanics_(book), where the divergence can be tamed by the Planckian correction to the mode density $\frac{1}{e^{\frac{\hbar\omega}{k_B T}}-1}$. Here we hack our way to the result with a bit of mathematical sleight of hand. First, since the integrand is sharply peaked at $\omega = \omega_0$ (and hence is practically zero for $\omega < 0$), we may replace the lower limit with $\omega = -\infty$:

$$P_{\text{decay}} \approx \frac{2\omega_0^2 \left|\langle 0| \boldsymbol{\mu} |1\rangle\right|^2}{3\pi^2 \hbar \epsilon_0 c^5} \int_{-\infty}^{\infty} \frac{\omega \sin^2(\frac{1}{2}(\omega - \omega_0)t)}{(\omega - \omega_0)^2} d\omega$$

Now let us make the substitution $x = \frac{1}{2}(\omega - \omega_0)t$:

$$\begin{aligned} P_{\text{decay}} &= \frac{2\omega_0^2 \left|\langle 0| \boldsymbol{\mu} |1\rangle\right|^2}{3\pi^2 \hbar \epsilon_0 c^5} \int_{-\infty}^{\infty} \frac{\frac{2}{t}(x + \omega_0) \sin^2(x)}{(\frac{2}{t}x)^2} d\omega \\ &= \frac{\omega_0^2 \left|\langle 0| \boldsymbol{\mu} |1\rangle\right|^2 t}{3\pi^2 \hbar \epsilon_0 c^5} \int_{-\infty}^{\infty} \frac{(x + \omega_0) \sin^2(x)}{x^2} dx \end{aligned}$$

Now comes the sleight of hand; though it is formally divergent, the $\frac{\sin^2(x)}{x}$ is odd and hence (at least via the Cauchy principle value definition) this integral vanishes. We are thus left with:

$$P_{\text{decay}} = \frac{\omega_0^3 \left|\langle 0| \boldsymbol{\mu} |1\rangle\right|^2 t}{3\pi^2 \hbar \epsilon_0 c^5} \int_{-\infty}^{\infty} \frac{\sin^2(x)}{x^2} dx$$

The integral evaluates to π, and so:

$$P_{\text{decay}} = \frac{\omega_0^3 \left|\langle 0| \boldsymbol{\mu} |1\rangle\right|^2 t}{3\pi \hbar \epsilon_0 c^5}$$

2. The above probability is linear in time, and so differentiating we obtain a constant rate of decay:

$$\gamma_{\text{rad}} = \frac{\omega_0^3 \left|\langle 0| \boldsymbol{\mu} |1\rangle\right|^2}{3\pi \hbar \epsilon_0 c^5}$$

□

Exercise 7.30: Electronic state lifetimes

A calculation similar to that for γ_{rad} can be done to estimate the lifetimes expected for electronic transitions, that is, those which involve energy level changes $\Delta n \neq 0$. For such transitions, the relevant interaction couples the atom's electric dipole moment to the electromagnetic field, giving

$$g_{\text{ed}}^2 = \frac{\omega_0^2}{2\hbar\omega\epsilon_0}\left|\langle 0|\boldsymbol{\mu}_{\text{ed}}|1\rangle\right|^2.$$

This gives a spontaneous emission rate

$$\gamma_{\text{red}}^{\text{ed}} = \frac{\omega_0^3\left|\langle 0|\boldsymbol{\mu}_{\text{ed}}|1\rangle\right|^2}{3\pi\hbar\epsilon_0 c^3}.$$

Give a value for $\gamma_{\text{red}}^{\text{ed}}$, taking $\left|\langle 0|\boldsymbol{\mu}_{\text{ed}}|1\rangle\right| \approx qa_0$, where q is the eletric charge, and a_0 the Bohr radius, and assuming $\omega_0/2\pi \approx 10^{15}$Hz. The result show how much faster electronic states can decay compared with hyperfine states.

Solution

Concepts Involved: Numerical Estimation

We numerically evaluate:

$$\gamma_{\text{red}}^{\text{ed}} = \frac{\omega_0^3\left|\langle 0|\boldsymbol{\mu}_{\text{ed}}|1\rangle\right|^2}{3\pi\hbar\epsilon_0 c^3} \approx \frac{\omega_0^3 q^2 a_0^2}{3\pi\hbar\epsilon_0 c^3}$$

evaluating this for $\omega_0 = (2\pi)\cdot 10^{15}$Hz, $q = 1.60\times10^{-19}$C, $a_0 = 5.29\times10^{-11}$m, $\hbar = 1.05\times10^{-34}\text{kgm}^2\text{s}^{-1}$, $\epsilon_0 = 8.85\times10^{-12}\text{C}^2\text{kg}^{-1}\text{m}^{-3}\text{s}^2$ and $c = 3.00\times10^8\text{ms}^{-1}$ we find:

$$\gamma_{\text{red}}^{\text{ed}} = 7.5\times10^7\text{s}^{-1}$$

In comparison the given hyperfine decay rate in the text is $\gamma_{\text{rad}} \approx 10^{-15}\text{s}^{-1}$, corresponding to electronic transitions having a decay rate that is quicker by 22 orders of magnitude. □

Exercise 7.31

Construct a Hadamard gate from R_y and R_x rotations.

Solution

Concepts Involved: Rotations, Gate Decomposition

In Ex. 4.12; we found the decomposition:

$$H = e^{i\pi/2}R_z(\pi)R_y(-\pi/2)R_z(0)$$

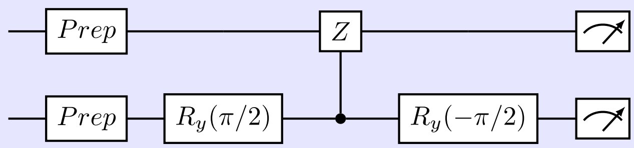
and combining this with the result of Ex. 4.10 (conjugating both sides by Hadamards) we conclude:

$$H = e^{i\pi/2}R_x(\pi)R_y(\pi/2)$$

□

Exercise 7.32

Show that the circuit in Figure 7.14 (reproduced below) is equivalent (up to relative phases) to a controlled-NOT gate, with the phonon state as the control qubit.



Solution

Concepts Involved: Controlled Operations

By matrix multiplication we have:

$$
\begin{aligned}
(I_1 \otimes R_y(\pi/2))CZ(I_1 \otimes R_y(-\pi/2)) &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \otimes \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \otimes \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \\
&= \frac{1}{2}\begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -1 & 1 \end{bmatrix} \\
&= \frac{1}{2}\begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 2 & 0 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \\
&= CX_{1,2}
\end{aligned}
$$

so indeed the given circuit is euqivalent to a `CNOT`. □

Exercise 7.33: Magnetic resonance

Show that (7.128) simplifies to become (7.129). What laboratory frame Hamiltonian gives rise to the rotating frame Hamiltonian (7.135)?

Solution

Concepts Involved: Commutators

(7.128) is the time-dependent SE:

$$i\partial_t \left|\chi(t)\right\rangle = H \left|\chi(t)\right\rangle$$

defining $\left|\varphi(t)\right\rangle = e^{i\omega tZ/2}\left|\chi(t)\right\rangle$, we have that $\left|\chi(t)\right\rangle = e^{-i\omega tZ/2}\left|\varphi(t)\right\rangle$ and so Eq. (7.128) becomes:

$$i\partial_t(e^{-i\omega tZ/2}\left|\varphi(t)\right\rangle) = H(e^{-i\omega tZ/2}\left|\varphi(t)\right\rangle)$$

Applying the product rule on the LHS:

$$\left(i(-i\omega Z/2)e^{-i\omega tZ/2} + e^{-i\omega tZ/2}i\partial_t\right)\left|\varphi(t)\right\rangle = He^{-i\omega tZ/2}\left|\varphi(t)\right\rangle$$

Commuting Z with $e^{-i\omega tZ/2}$ and shifting over the Z term to the RHS, we obtain:

$$e^{-i\omega tZ/2}i\partial_t\left|\varphi(t)\right\rangle = \left(He^{-i\omega Z/2} - e^{-i\omega tZ/2}\frac{\omega}{2}Z\right)\left|\varphi(t)\right\rangle$$

Finally multiplying both sides by $e^{i\omega tZ/2}$ we obtain:

$$i\partial_t\left|\varphi(t)\right\rangle = \left(e^{i\omega Z/2}He^{-i\omega Z/2} - \frac{\omega}{2}Z\right)\left|\varphi(t)\right\rangle$$

which is Eq. (7.129).
The laboratory frame Hamiltonian that gives rise to the rotating frame Hamiltonian of (7.135) is any Hamiltonian of the form:

$$H_{\text{lab}} = \frac{\omega}{2}Z + c_1(t)X + c_2(t)Y$$

As we can write such a Hamiltonian using the above derived transformation as:

$$\begin{aligned}
H_{\text{rot}} &= e^{i\omega Zt/2}H_{\text{lab}}e^{-i\omega Zt/2} - \frac{\omega}{2}Z \\
&= e^{i\omega Zt/2}\left(\frac{\omega}{2}Z + c_1(t)X + c_2(t)Y\right)e^{-i\omega Zt/2} - \frac{\omega}{2}Z \\
&= c_1(t)e^{i\omega Zt/2}Xe^{-i\omega Zt/2} + c_2(t)e^{i\omega Zt/2}Ye^{-i\omega Zt/2} \\
&= c_1(t)(X\cos(\omega t) - Y\sin(\omega t)) + c_2(t)(Y\cos(\omega t) + X\sin(\omega t)) \\
&= \big(c_1(t)\cos(\omega t) + c_2(t)\sin(\omega t)\big)X + \big(-c_1(t)\sin(\omega t) + c_2(t)\cos(\omega t)\big)Y
\end{aligned}$$

so defining $g_1(t) = c_1(t)\cos(\omega t) + c_2(t)\sin(\omega t)$ and $g_2(t) = -c_1(t)\sin(\omega t) + c_2(t)\cos(\omega t)$ we indeed obtain the rotating frame Hamiltonian of (7.135). □

Exercise 7.34: NMR frequencies

Starting with the nuclear Bohr magneton, compute the precession frequency of a proton in a magnetic field of 11.8 tesla. How many gauss should B_1 be to accomplish a $90°$ rotation in 10 microseconds?

Solution

Concepts Involved: Numerical Estimation

We can find the precession frequency via:

$$\hbar\omega = \mu_B B \implies \omega = \frac{\mu_B B}{\hbar} = \frac{(5\times 10^{-27}\text{JT}^{-1})(11.8\text{T})}{1.05\times 10^{-34}\text{Js}} = 562\text{rad}\cdot\text{s}^{-1}$$

and dividing by 2π gives us the precession frequency in Hz:

$$f = \frac{\omega}{2\pi} = 91\text{MHz}$$

The period of a $90°$ rotation is equal to the precession frequency due to B_1 divided by 4, so:

$$T_{\pi/2} = \frac{\frac{2\pi}{\omega_1}}{4} = \frac{\pi}{2\omega_1} = \frac{\pi}{2\left(\frac{\mu_B B_1}{\hbar}\right)} = \frac{\hbar\pi}{2\mu_B B_1} \implies B_1 = \frac{\hbar\pi}{2\mu_B T_{\pi/2}}$$

so with $T_{\pi/2} = 10\times 10^{-6}\text{s}$ we find:

$$B_1 = \frac{(1.05\times 10^{-34}\text{Js})\cdot\pi}{2\cdot(5\times 10^{-27}\text{JT}^{-1})\cdot(1\times 10^{-5}\text{s})} = 3.3\times 10^{-3}\text{T} = 33\text{Gauss}$$

□

Exercise 7.35: Motional narrowing

Show that the spherical average of $H^D_{1,2}$ over $\hat{\mathbf{n}}$ is zero.

Solution

Concepts Involved: Integration

We work in the high-field approximation, wherein $\boldsymbol{\sigma}_1, \boldsymbol{\sigma_2}$ are approximately polarized in the same direction (that of the external magnetic field). Thus we have that $\boldsymbol{\sigma}_1\cdot\boldsymbol{\sigma}_2 \approx 1$ and $\boldsymbol{\sigma}_1\cdot\hat{\mathbf{n}} \approx \boldsymbol{\sigma}_2\cdot\hat{\mathbf{n}} = \cos\theta$ where θ is the angle between the spin(s) and the vector joining the two nuclei. Thus the Hamiltonian becomes:

$$H^D_{1,2} = \frac{\gamma_1\gamma_2\hbar}{4r^3}\left[1 - 3\cos^2\theta\right]$$

In a low viscosity liquid, the orientation of $\hat{\mathbf{n}}$ (and hence θ) relative to the static field varies rapidly, and hence we may assess the contribution as a spherical average over all possible orientations:

$$\begin{aligned}\left\langle H_{1,2}^D \right\rangle &= \frac{\gamma_1\gamma_2\hbar}{4r^3}\frac{1}{4\pi}\int d\Omega[1 - 3\cos^2\theta] \\ &= \frac{\gamma_1\gamma_2\hbar}{4r^3}\frac{1}{4\pi}\int_0^{2\pi} d\varphi \int_0^{\pi} \sin\theta d\theta[1 - 3\cos^2\theta] \\ &= \frac{\gamma_1\gamma_2\hbar}{8r^3}\left((-\cos\theta|_0^\pi) + \left(\cos^3\theta\Big|_0^\pi\right)\right) \\ &= \frac{\gamma_1\gamma_2\hbar}{8r^3}(2-2) \\ &= 0\end{aligned}$$

hence the claim is proven. □

Exercise 7.36: Thermal equilibrium NMR state

For $n = 1$ show that the thermal equilibrium state is

$$\rho \approx 1 - \frac{\hbar\omega}{2k_B T}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

and for $n = 2$ (and $\omega_A \approx 4\omega_B$),

$$\rho \approx 1 - \frac{\hbar\omega_B}{4k_B T}\begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & -5 \end{bmatrix}$$

The above formulas are incorrect - by inspection they violate the normalization condition $\mathrm{Tr}(\rho) = 1$. The corrected formula for $n = 1$ is:

$$\rho \approx \frac{1}{2}\left(1 - \frac{\hbar\omega}{2k_B T}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\right)$$

and the corrected formula for $n = 2$ is:

$$\rho \approx \frac{1}{4}\left(1 - \frac{\hbar\omega_B}{2k_B T}\begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & -5 \end{bmatrix}\right)$$

Solution

Concepts Involved: Density Operators, Composite Systems

We work in the high-temperature approximation where the thermal equilibrium state takes the form:

$$\rho \approx 2^{-n}\left(1 - \beta H\right).$$

for $n = 1$ we have the single-qubit Hamiltonian $H = \frac{\hbar\omega}{2}Z$, so the thermal equilibrium state is:

$$\rho \approx \frac{1}{2}\left(1 - \frac{\hbar\omega}{2k_B T}\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}\right) \tag{5}$$

as claimed. For $n = 2$ the Hamiltonian (without coupling) takes the form (for $\omega_A \approx 4\omega_B$)

$$H = \frac{\hbar\omega_A}{2}Z_1 \otimes I_2 + \frac{\hbar\omega_B}{2}I_1 \otimes Z_2 = \frac{\hbar 4\omega_B}{2}\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & -1 & 0\\ 0 & 0 & 0 & -1\end{bmatrix} + \frac{\hbar\omega_B}{2}\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & -1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & -1\end{bmatrix}$$

Hence the thermal equilibrium state is:

$$\rho \approx \frac{1}{4}\left(1 - \frac{\hbar\omega_B}{2k_B T}\begin{bmatrix}5 & 0 & 0 & 0\\ 0 & 3 & 0 & 0\\ 0 & 0 & -3 & 0\\ 0 & 0 & 0 & -5\end{bmatrix}\right)$$

□

Exercise 7.37: NMR spectrum of coupled spins

Calculate $V(t)$ for $H = JZ_1Z_2$ and $\rho = e^{i\pi Y_1/4}\frac{1}{4}\left[1 - \beta\hbar\omega_0(Z_1 + Z_2)\right]e^{-i\pi Y_1/4}$. How many lines would there be in the spectrum of the first spin if the Hamiltonian were $H = JZ_1(Z_2 + Z_3 + Z_4)$ (with a similar initial density matrix) and what would their relative magnitudes be?

Solution

Concepts Involved: Trace, Commutators, Composite Systems, Pauli Operators

Since H, ρ are symmetric in the two spins, WLOG let us study the spectrum of the first spin. We wish to compute

$$\begin{aligned}
V(t) &= V_0 \operatorname{Tr}\left[e^{iHt}\rho e^{-iHt}(iX_1 + Y_1)\right]\\
&= V_0 \operatorname{Tr}\left[e^{iJZ_1Z_2t}e^{i\pi Y_1/4}\frac{1}{4}[1 - \beta\hbar\omega_0(Z_1 + Z_2)]e^{-i\pi Y_1/4}e^{-iJZ_1Z_2t}(iX_1 + Y_1)\right]
\end{aligned}$$

The identity term in ρ has no contribution to $V(t)$, as it trivially commutes past $e^{iJZ_1Z_2t}e^{i\pi Y_1/4}$ which then cancels with $e^{-i\pi Y_1/4}e^{iJZ_1Z_2t}$, leaving only $\sim \operatorname{Tr}\big[(iX_1 + Y_1)\big]$ which vanishes as the Pauli matrices are traceless. Similarly, the Z_2 term also commutes with $e^{iJZ_1Z_2t}e^{i\pi Y_1/4}$ (the first term since it is composed of Zs, the second term since it only acts on the first qubit) and similarly vanishes. Thus we are left with just the Z_1 term:

$$V(t) = -\frac{V_0 \beta \hbar \omega_0}{4} \operatorname{Tr}\Big[e^{iJZ_1Z_2t} e^{i\pi Y_1/4} Z_1 e^{-i\pi Y_1/4} e^{-iJZ_1Z_2t}(iX_1 + Y_1)\Big]$$

Since Y, Z anticommute we have that

$$e^{i\pi Y_1/4} Z_1 e^{-i\pi Y_1/4} = e^{i\pi Y_1/4} e^{i\pi Y_1/4} Z_1 = e^{i\pi Y_1/2} Z_1 = \left(\cos\left(\frac{\pi}{2}\right) I_1 + i \sin\left(\frac{\pi}{2}\right) Y_1\right) Z_1 = iY_1 Z_1 = -X_1$$

Now we study $e^{iJZ_1Z_2t} X_1 e^{-iJZ_1Z_2t}$, which is composed of four terms after expanding the exponentials into sines/cosines

$$e^{iJZ_1Z_2t} X_1 e^{-iJZ_1Z_2t} = \cos^2(Jt) X_1 + i\cos(Jt)\sin(Jt)(Z_1Z_2X_1 - X_1Z_1Z_2) + \sin^2(Jt) Z_1Z_2X_1Z_1Z_2$$

The two "cross-terms" vanish as there is a persisting Z_2 which is traceless. The last term we can use that Z, X anticommute to rewrite

$$Z1 Z_2 X_1 Z_1 Z_2 = Z_1 Z_2 (-Z_1) Z_2 X_1 = -X_1$$

Hence in the trace expression we are left with:

$$V(t) = \frac{V_0 \beta \hbar \omega_0}{4} \operatorname{Tr}\Big[(\cos^2(Jt) - \sin^2(Jt)) X_1 (iX_1 + Y_1)\Big]$$

The X_1 term gives $\operatorname{Tr}\big[X_1^2\big] = \operatorname{Tr}[I] = 4$ and the Y_1 term gives $\operatorname{Tr}[X_1 Y_1] = \operatorname{Tr}[iZ_1] = 0$, and so with a double angle identity we conclude:

$$V(t) = iV_0 \beta \hbar \omega_0 \cos(2Jt) = \frac{iV_0 \beta \hbar \omega_0}{2}(e^{i2Jt} + e^{-i2Jt})$$

Which corresponds to two (equal magnitude) peaks in the spectra at $\pm 2J$.
For the four-spin case, the density matrix takes the form

$$\rho = e^{i\pi Y_1/4} \frac{1}{16}[1 - \beta \hbar \omega_0 (Z_1 + Z_2 + Z_3 + Z_4)] e^{-i\pi Y_1/4}$$

and the expression to compute becomes

$$V(t) = \frac{V_0}{16} \operatorname{Tr}\Big[e^{iJZ_1(Z_2+Z_3+Z_4)t} e^{i\pi Y_1/4}[1 - \beta\hbar\omega_0(Z_1 + Z_2 + Z_3 + Z_4)] e^{-i\pi Y_1/4} e^{-iJZ_1(Z_2+Z_3+Z_4)t}(iX_1 + Y_1)\Big]$$

as in the 2-spin case, the identity, Z_2, Z_3, Z_4 terms of ρ drop out from the trace, leaving us with the Z_1 term that gets rotated to X_1:

$$V(t) = \frac{V_0 \beta \hbar \omega_0}{16} \operatorname{Tr}\Big[e^{iJZ_1(Z_2+Z_3+Z_4)t} X_1 e^{-iJZ_1(Z_2+Z_3+Z_4)t}(iX_1 + Y_1)\Big]$$

since each of the terms in the Hamiltonian commute, we can write:

$$e^{iJZ_1(Z_2+Z_3+Z_4)t} = e^{iJZ_1Z_2}e^{iJZ_1Z_3}e^{iJZ_1Z_4}$$

and expand the individual exponentials into sines/cosines, giving a total of $2^3 \cdot 2^3 = 64$ terms. Of these, only the terms where the $Z_{2/3/4}$s mutually cancel contribute to the trace, leaving 8 terms only. The resulting calculation is very similar to the 2-spin case

$$\begin{aligned}
V(t) = \frac{V_0\beta\hbar\omega_0}{16} \operatorname{Tr}\Bigg[\Big(&\cos^6(Jt)X_1 + \cos^4(Jt)\sin^2(Jt)(Z_1Z_2X_1Z_1Z_2 + Z_1Z_3X_1Z_1Z_3 + Z_1Z_4X_1Z_1Z_4) \\
&+ \cos^2(Jt)\sin^4(Jt)(Z_1Z_2Z_1Z_3X_1Z_1Z_2Z_1Z_3 + Z_1Z_2Z_1Z_4X_1Z_1Z_2Z_1Z_4 + Z_1Z_3Z_1Z_4X_1Z_1Z_3Z_1Z_4) \\
&+ \sin^6(Jt)(Z_1Z_2Z_1Z_3Z_1Z_4)X_1(Z_1Z_2Z_1Z_3Z_1Z_4)\Big)(iX_1+Y_1)\Bigg] \\
&= \frac{V_0\beta\hbar\omega_0}{16}(\cos^6(Jt) - 3\cos^4(Jt)\sin^2(Jt) + 3\cos^2(Jt)\sin^4(Jt) - \sin^6(Jt)) \operatorname{Tr}\big[(X_1(iX_1+Y_1)\big] \\
&= \frac{V_0\beta\hbar\omega_0}{16}(\cos^6(Jt) - 3\cos^4(Jt)\sin^2(Jt) + 3\cos^2(Jt)\sin^4(Jt) - \sin^6(Jt)) \operatorname{Tr}[iI + iZ_1] \\
&= \frac{V_0\beta\hbar\omega_0}{16}(\cos^6(Jt) - 3\cos^4(Jt)\sin^2(Jt) + 3\cos^2(Jt)\sin^4(Jt) - \sin^6(Jt))(i16+0) \\
&= iV_0\beta\hbar\omega_0(\cos^6(Jt) - 3\cos^4(Jt)\sin^2(Jt) + 3\cos^2(Jt)\sin^4(Jt) - \sin^6(Jt)) \\
&= iV_0\beta\hbar\omega_0(\cos^2(Jt) - \sin^2(Jt))^3 \\
&= iV_0\beta\hbar\omega_0\cos^3(2Jt) \\
&= iV_0\beta\hbar\omega_0\left(\frac{e^{i2Jt}+e^{-i2Jt}}{2}\right)^3 \\
&= \frac{iV_0\beta\hbar\omega_0}{8}(e^{i6Jt} + 3e^{i2Jt} + 3e^{-i2Jt} + e^{-i6Jt})
\end{aligned}$$

This corresponds to four peaks in the spectrum at $\pm 6J$ and $\pm 2J$, with the peaks at $\pm 2J$ being three times larger. □

Exercise 7.38: Refocusing

Explicitly show that (7.150) is true (use the anti-commutativity of the Pauli matrices).
The expression is missing a global minus sign, the correct relation to show should be:

$$R_{x1}^2 e^{-iaZ_1t} R_{x1}^2 = -e^{iaZ_1t}$$

though of course such a phase is physically irrelevant.

Solution

Concepts Involved: Commutators, Rotations

Looking at the LHS of (7.150), we have:

$$R_{x1}^2 e^{-iaZ_1t} R_{x1}^2 = e^{-i\pi X_1/2}e^{-iaZ_1t}e^{-i\pi X_1/2}$$

Now using Ex. 2.35 we can write the exponentials as sines/cosines:

$$\begin{aligned}
R_{x1}^2 e^{-iaZ_1 t} R_{x1}^2 &= \left(\cos\left(\frac{\pi}{2}\right)I - i\sin\left(\frac{\pi}{2}\right)X\right)\left(\cos(at)I - i\sin(at)Z\right)\left(\cos\left(\frac{\pi}{2}\right)I - i\sin\left(\frac{\pi}{2}X\right)\right) \\
&= (-iX)\big(\cos(at)I - i\sin(at)Z\big)(-iX) \\
&= \big(\cos(at)I + i\sin(at)Z\big)(-iX)(-iX) \\
&= -\big(\cos(at)I + i\sin(at)Z\big) \\
&= -e^{iaZ_1 t}.
\end{aligned}$$

In the third-to-last equality we have used the anti-commutativity of the Pauli matrices (Ex. 2.41), in the second-to-last equality we use $X^2 = I$ and in the last equality we package up the two terms back into an exponential. The claim is thus proven. □

Exercise 7.39: Three-dimensional refocusing

$(*)$ What set of pulses can be used to refocus evolution under *any* single spin Hamiltonian $H^{\text{sys}} = \sum_k c_k \sigma_k$?

Solution

Concepts Involved: Commutators, Rotations

First, let us write:

$$H = \sum_k c_k \sigma_k = c\hat{\mathbf{n}} \cdot \boldsymbol{\sigma} = c\begin{bmatrix} \cos\theta & e^{-i\varphi}\sin\theta \\ e^{i\varphi}\sin\theta & \cos\theta \end{bmatrix}$$

where $\hat{\mathbf{n}} = (n_x, n_y, n_z) = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$. The eigenstates of this Hamiltonian are:

$$|+\hat{\mathbf{n}}\rangle = \begin{bmatrix} \cos\frac{\theta}{2} \\ e^{i\varphi}\sin\frac{\theta}{2} \end{bmatrix}, \quad |-\hat{\mathbf{n}}\rangle = \begin{bmatrix} e^{-i\varphi}\sin\frac{\theta}{2} \\ -\cos\frac{\theta}{2} \end{bmatrix}$$

with eigenvalues $\pm c$. Therein we can write:

$$e^{-iHt} = e^{-ict}\,|+\hat{\mathbf{n}}\rangle\langle+\hat{\mathbf{n}}| + e^{ict}\,|-\hat{\mathbf{n}}\rangle\langle-\hat{\mathbf{n}}|$$

Now define the unitary $U = |+\hat{\mathbf{n}}\rangle\langle-\hat{\mathbf{n}}| + |+\hat{\mathbf{n}}\rangle\langle-\hat{\mathbf{n}}|$. Then we have:

$$Ue^{-iHt}U^\dagger = e^{ict}\,|+\hat{\mathbf{n}}\rangle\langle+\hat{\mathbf{n}}| + e^{-ict}\,|-\hat{\mathbf{n}}\rangle\langle-\hat{\mathbf{n}}| = e^{iHt}$$

which successfully refocuses the Hamiltonian. In terms of rotations, we can write $U = R_{\hat{\mathbf{m}}}(\pi)$ where $\hat{\mathbf{m}}$ is a vector orthogonal to $\hat{\mathbf{n}}$, or if we are restricted to x/y pulses we may use the $X - Y$ Euler decomposition of Ex. 4.10. □

Exercise 7.40: Refocusing dipolar interactions

($*$) Give a sequence of pulses which can be used to turn two spin dipolar coupling $H^D_{1,2}$ into the much simpler form of (7.138).

Solution

Concepts Involved: Commutators, Rotations

The dipolar coupling has the form:

$$
\begin{aligned}
H^D_{1,2} &= \frac{\gamma_1\gamma_2\hbar}{4r^3}\left[\boldsymbol{\sigma}_1\cdot\boldsymbol{\sigma}_2 - 3(\boldsymbol{\sigma}_1\cdot\hat{\mathbf{n}})(\boldsymbol{\sigma}_2\cdot\hat{\mathbf{n}})\right] \\
&= \frac{\gamma_1\gamma_2\hbar}{4r^3}\Big[(1-3n_x^2)\sigma_x^{(1)}\sigma_x^{(2)} + (1-3n_y^2)\sigma_y^{(1)}\sigma_y^{(2)} + (1-3n_z^2)\sigma_z^{(1)}\sigma_z^{(2)} \\
&- 3n_xn_y(\sigma_x^{(1)}\sigma_y^{(2)} + \sigma_y^{(1)}\sigma_x^{(2)}) - 3n_xn_z(\sigma_x^{(1)}\sigma_z^{(2)} + \sigma_z^{(1)}\sigma_x^{(2)}) - 3n_yn_z(\sigma_y^{(1)}\sigma_z^{(2)} + \sigma_z^{(1)}\sigma_y^{(2)})\Big]
\end{aligned}
$$

Which is a specific case of a general class of Hamiltonians:

$$
H = \sum_{i\in\{x,y,z\}}\sum_{j\in\{x,y,z\}} H_{ij}\sigma_i^{(1)}\sigma_j^{(2)}
$$

for which we will demonstrate the refocusing result to the simpler ZZ-coupling of Eq. (7.138). We consider the 180° z-rotation (on the ith qubit) R_{zi}^2, for which we observe that (similar to Ex. 7.38):

$$
R_{zi}^2 e^{-iaX_it}R_{zi}^2 = (-iZ)e^{-iaX_it}(-iZ) = e^{iaX_it}(-iZ)^2 = -e^{iaX_it}
$$

$$
R_{zi}^2 e^{-iaY_it}R_{zi}^2 = (-iZ)e^{-iaY_it}(-iZ) = e^{iaX_it}(-iZ)^2 = -e^{iaY_it}
$$

where we have used the anticommutativity of the Pauli matrices. Of course, $R_{zi}^2 e^{-iaZ_it}R_{zi}^2 = -e^{-iaZ_it}$ as Z commutes with itself.
Thus, we find:

$$
\begin{aligned}
&e^{-iHt}R_{z1}^2e^{-iHt}R_{z1}^2 \\
&= e^{-iHt}R_{z1}^2 e^{-i(\sum_{i\in\{x,y,z\}}\sum_{j\in\{x,y,z\}}H_{ij}\sigma_i^{(1)}\sigma_j^{(2)})t}R_{z1}^2 \\
&= -e^{-iHt}e^{+i(\sum_{i\in\{x,y\}}\sum_{j\in\{x,y,z\}}H_{ij}\sigma_i^{(1)}\sigma_j^{(2)})t}e^{-i(\sum_{j\in\{x,y,z\}}H_{zj}\sigma_z^{(1)}\sigma_j^{(2)})t} \\
&= -e^{-i(\sum_{i\in\{x,y,z\}}\sum_{j\in\{x,y,z\}}H_{ij}\sigma_i^{(1)}\sigma_j^{(2)})t}e^{+i(\sum_{i\in\{x,y\}}\sum_{j\in\{x,y,z\}}H_{ij}\sigma_i^{(1)}\sigma_j^{(2)})t}e^{-i(\sum_{j\in\{x,y,z\}}H_{zj}\sigma_z^{(1)}\sigma_j^{(2)})t} \\
&= -e^{-i(\sum_{j\in\{x,y,z\}}H_{zj}\sigma_z^{(1)}\sigma_j^{(2)})2t}
\end{aligned}
$$

and analogously for R_{z2}^2. Therefore, successive z-rotations can refocus the Hamiltonian, retaining only the ZZ term:

$$\begin{aligned}&(e^{-iHt/4}R_{z1}^2e^{-iHt/4}R_{z1}^2)R_{z2}^2(e^{-iHt/4}R_{z1}^2e^{-iHt/4}R_{z1}^2)R_{z2}^2\\&=(-e^{-i(\sum_{j\in\{x,y,z\}}H_{zj}\sigma_z^{(1)}\sigma_j^{(2)})t/2})R_{z2}^2(-e^{-i(\sum_{j\in\{x,y,z\}}H_{zj}\sigma_z^{(1)}\sigma_j^{(2)})t/2})R_{z2}^2\\&=-e^{-iH_{zz}\sigma_z^{(1)}\sigma_z^{(2)}}\end{aligned}$$

We can write R_z in terms of x/y pulses as $R_z = R_y\bar{R}_x\bar{R}_y$ and so:

$$R_z^2 = R_y\bar{R}_x\bar{R}_yR_y\bar{R}_x\bar{R}_y = R_y\bar{R}_x^2\bar{R}_y = R_yR_x^2\bar{R}_y.$$

Thus, we conclude that the pulse sequence that gives us the desired refocusing is:

$$\begin{aligned}&e^{-iH_{zz}Z_1Z_2t}\\&\cong e^{-iHt/4}R_{y1}R_{x1}^2\bar{R}_{y1}e^{-iHt/4}R_{y1}R_{x1}^2\bar{R}_{y1}R_{y2}R_{x2}^2\bar{R}_{y2}e^{-iHt/4}R_{y1}R_{x1}^2\bar{R}_{y1}e^{-iHt/4}R_{y1}R_{x1}^2\bar{R}_{y1}R_{y2}R_{x2}^2\bar{R}_{y2}\end{aligned}$$

□

Exercise 7.41: NMR controlled-NOT

Give an explicit sequence of single qubit rotations which realize a controlled-NOT between two spins evolving under the Hamiltonian of (7.147). You may start with (7.46), but the result can be simplified to reduce the number of single qubit rotations.

Solution

Concepts Involved: Controlled Operations, Rotations

We are given the relation:

$$\sqrt{i}e^{iZ_1Z_2\pi/4}e^{-iZ_1\pi/4}e^{-iZ_2\pi/4} = U_{CZ}$$

So if we consider time evolution:

$$U(t) = e^{iHt} = e^{i(aZ_1+bZ_2+cZ_1Z_2)t}$$

with $t = \frac{\pi}{4c}$ (having set $\hbar = 1$) we have:

$$U(\frac{\pi}{4c}) = e^{i(aZ_1+bZ_2+cZ_1Z_2)\frac{\pi}{4c}} = e^{iZ_1Z_2\pi a/4}e^{iZ_1\pi/4c}e^{iZ_2\pi b/4c}$$

Thus if we combine the above with the single-qubit rotations $e^{iZ_1(-\frac{\pi}{4}(\frac{a}{c}+1))}, e^{iZ_2(-\frac{\pi}{4}(\frac{b}{c}+1))}$ we get U_{CZ} up to a global phase. Now conjugating with $I\otimes H$ as in Eq. (7.46) to obtain the CNOT gate, we have:

$$CX_{1,2} \cong (I_1\otimes H_2)U(\frac{\pi}{4c})e^{iZ_1(-\frac{\pi}{4}(\frac{a}{c}+1))}e^{iZ_2(-\frac{\pi}{4}(\frac{b}{c}+1))}(I_1\otimes H_2)$$

From Ex. 4.8 we know that the Hadamard can be written as the rotation:

$$H = e^{i\pi/2}R_{\hat{\mathbf{n}}}(\pi)$$

for $\hat{\mathbf{n}} = (\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}})$, so:

$$CX_{1,2} \cong R_{\hat{\mathbf{n}},2}(\pi)U(\frac{\pi}{4c})R_{z,1}(-\frac{\pi}{2}(\frac{a}{c}+1))R_{z,2}(-\frac{\pi}{2}(\frac{b}{c}+1))R_{\hat{\mathbf{n}},2}(\pi)$$

The last simplification is we can use Ex. 4.15 to compose the last two rotations on the second qubit into a single rotation:

$$CX_{1,2} \cong R_{\hat{\mathbf{n}},2}(\pi)U(\frac{\pi}{4c})R_{z,1}(-\frac{\pi}{2}(\frac{a}{c}+1))R_{\hat{\mathbf{m}},2}(\beta)$$

where:

$$\begin{aligned}
\beta &= 2\arccos\left(\cos\left(-\frac{\pi}{4}(\frac{b}{c}+1)\right)\cos\left(\frac{\pi}{2}\right) - \sin\left(-\frac{\pi}{4}(\frac{b}{c}+1)\right)\sin\left(\frac{\pi}{2}\right)\hat{\mathbf{z}}\cdot\hat{\mathbf{n}}\right) \\
&= 2\arccos\left(\frac{1}{\sqrt{2}}\sin\left(\frac{\pi}{4}(\frac{b}{c}+1)\right)\right)
\end{aligned}$$

and:

$$\begin{aligned}
\hat{\mathbf{m}} &= \frac{\sin\left(-\frac{\pi}{2}(\frac{b}{c}+1)\right)\cos(\frac{\pi}{2})\hat{\mathbf{z}} + \cos\left(-\frac{\pi}{2}(\frac{b}{c}+1)\right)\sin(\frac{\pi}{2})\hat{\mathbf{n}} - \sin\left(-\frac{\pi}{2}(\frac{b}{c}+1)\right)\sin(\frac{\pi}{2})\hat{\mathbf{n}}\times\hat{\mathbf{z}}}{\sin(\beta/2)} \\
&= \frac{\cos\left(\frac{\pi}{2}(\frac{b}{c}+1)\right)\hat{\mathbf{n}} + \sin\left(\frac{\pi}{2}(\frac{b}{c}+1)\right)\frac{\hat{\mathbf{y}}}{\sqrt{2}}}{\sin(\beta/2)}
\end{aligned}$$

□

Exercise 7.42: Permutations for temporal labeling

Give a quantum circuit to accomplish the permutations P and $P^\dagger$ necessary to transform ρ_1 of (7.153) to ρ_2 of (7.154).

Solution

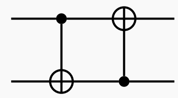
Concepts Involved: Permutations, Controlled Operations, Density Operators

---

We saw in Ex. 4.19 that the action of a CNOT (with target as the second qubit) on the diagonal elements of the density matrix are to interchange $\rho_{33} \leftrightarrow \rho_{44}$ (from interchange of $|10\rangle \leftrightarrow |11\rangle$ states). Analogously, a CNOT with a target as the first qubit has the action (on the diagonal elements) of interchanging $\rho_{22} \leftrightarrow \rho_{44}$ (from the interchange of $|01\rangle \leftrightarrow |11\rangle$ states). Thus, to go from:

$$\rho_1 = \text{diag}(a, b, c, d) \mapsto \rho_2 = P\rho_1 P^\dagger = \text{diag}(a, c, d, b)$$

we first exchange $c \leftrightarrow d$ ($\rho_{33} \leftrightarrow \rho_{44}$) which puts the d in the correct position, followed by $b \leftrightarrow c$ ($\rho_{22} \leftrightarrow \rho_{44}$), corresponding to the circuit of two CNOTS (the first with target as the second qubit, the second with target as the first qubit):



□

Exercise 7.43: Permutations for logical labeling

Give a quantum circuit to accomplish the permutations P necessary to transform ρ of (7.163) to ρ' of (7.165).

There is an error in the equation reference, and it should be the ρ' of (7.164).

Solution

Concepts Involved: Permutations, Controlled Operations, Density Operators

Since things are again getting a tad more complicated, let us work in permutation notation again, where we recall from Ex. 4.27 that:

$$\begin{aligned} CX_{1,2} &= (57)(68) \\ CX_{1,3} &= (56)(78) \\ CX_{2,1} &= (37)(48) \\ CX_{2,3} &= (34)(78) \\ CX_{3,1} &= (26)(48) \\ CX_{3,2} &= (24)(68) \end{aligned}$$

Out of these we want to construct a permutation taking:

$$\text{diag}(6, 2, 2, -2, 2, -2, -2, -6) \mapsto \text{diag}(6, -2, -2, -2, -6, 2, 2, 2)$$

since some of the entries are the same, there is not a unique permutation that the above corresponds to, but one such permutation would be:

$$\underbrace{(58)}_{-2\leftrightarrow 8}\underbrace{(26)}_{-2\leftrightarrow 2}\underbrace{(37)}_{-2\leftrightarrow 2}$$

We note then that:

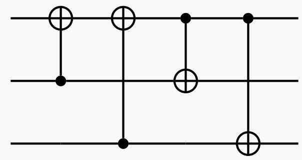
$$(26)(37) = (26)(48)^2(37) = (26)(48)(37)(48) = CX_{3,1}CX_{2,1}$$

and further that:

$$CX_{1,3}CX_{1,2} = (56)(78)(57)(68) = (58)(67)$$

this is almost what we want, save for the (67) permutation, but since positions 6, 7 have already been swapped to the same value (2) by the (26)(37) permutation, its action is trivial. Hence, our desired circuit/permutation is:

$$CX_{1,3}CX_{1,2}CX_{3,1}CX_{2,1}$$



□

Exercise 7.44: Logical labeling for n spins

$(*)$ Suppose we have a system of n nearly identical spins of Zeeman frequency $\hbar\omega$ in thermal equilibrium at temperature T with state ρ. What is the largest effective pure state that you can construct from ρ using logical labeling? (*Hint:* take advantage of states whose labels have Hamming weight of $n/2$.)

Solution

Concepts Involved: Counting, Stirling's Approximation, Density Operators, Pure States, Mixed States

In order to to construct a logical pure state on k qubits, we require a block of ρ which has $2^k - 1$ zero eigenvalues/diagonal entries, or $2^k - 1$ identical eigenvalues λ (which can then be subtracted off via $-\lambda I$, as is done in Eq. (7.165)).
The largest degeneracy of eigenvalues/diagonal entries comes from states that have Hamming weight $n/2$ - these have equal populations of up/down spins and have eigenvalue zero. The number of such states is given by:

$$\binom{n}{n/2} = \frac{n!}{(\frac{n}{2})!(\frac{n}{2})!} \approx \frac{\sqrt{2\pi n}\left(\frac{n}{e}\right)^n}{\left(\sqrt{2\pi n/2}\left(\frac{n/2}{e}\right)^{n/2}\right)^2} = \frac{2^n}{\sqrt{\frac{\pi n}{2}}}$$

where we have used Stirling's approximation. We require the number of such states to be greater than $2^k - 1$, so:

$$2^k - 1 \leq 2^k \leq \binom{n}{n/2} \approx \frac{2^n}{\sqrt{\frac{\pi n}{2}}}$$

Taking the logarithm of both sides:

$$k \leq n - O(\log n)$$

which tells us that (asymptotically) we can construct a n-qubit pure state given an ensemble of n nearly identical spins. □

Exercise 7.45: State tomography with NMR

Let the voltage measurement $V_k(t) = V_0 \operatorname{tr}\left[e^{-iHt} M_k \rho M_k^\dagger e^{iHt}(iX_k + Y_k)\right]$ be the result of experiment k. Show that for two spins, nine experiments, with $M_0 = I$, $M_1 = R_{x1}$, $M_2 = R_{y1}$, $M_3 = R_{x2}$, $M_4 = R_{x2}R_{x1}$, $M_5 = R_{x2}R_{y1}$ etc. provide sufficient data from which ρ can be reconstructed.

The notation given here doesn't quite make sense, as there should be distinct labels for experiment number and the X/Y operators; we thus denote:

$$V_k(t) = V_0 \operatorname{Tr}\left[e^{-iHt} M_k \rho M_k^\dagger e^{iHt}(iX_n + Y_n)\right]$$

where k denotes experiment number and $n \in \{1, 2\}$ denotes the spectrum of the nth spin. In particular for a single experiment we can read out the spectra of both spins, so we can write:

$$V_k(t) = V_0 \operatorname{Tr}\left[e^{-iHt} M_k \rho M_k^\dagger e^{iHt} \sum_{n=1}^{2}(iX_n + Y_n)\right]$$

Solution

Concepts Involved: Density Operators, Tensor Products, Pauli Operators, Commutators, Rotations

For two spins, a general density matrix can be written as:

$$\rho = \sum_{i,j \in \{I,x,y,z\}} \rho_{ij} \sigma_i^{(1)} \sigma_j^{(2)}$$

with $\sigma_I = I$. The trace condition $\operatorname{Tr}(\rho) = 1$ fixes $\rho_{II} = \frac{1}{4}$, so to reconstruct ρ we require all other 15 coefficents ρ_{ij}.

The system Hamiltonian takes the form:

$$H = aZ_1 + bZ_2 + cZ_1Z_2$$

where all three terms mutually commute. We can use the cyclicity of the trace to observe:

$$V_k(t) = V_0 \operatorname{Tr}\left[\rho M_k^\dagger e^{iHt} \sum_{n=1}^{2}(iX_n + Y_n) e^{-iHt} M_k\right]$$

Now, let's look at the $n = 1$ term:

$$\begin{aligned} e^{iHt}(iX_1 + Y_1)e^{-iHt} &= e^{i(aZ_1 + bZ_2 + cZ_1Z_2)t}(iX_1 + Y_1)e^{-i(aZ_1 + bZ_2 + cZ_1Z_2)t} \\ &= e^{i(aZ_1 + bZ_2 + cZ_1Z_2)t} e^{-i(-aZ_1 + bZ_2 + -cZ_1Z_2)t}(iX_n + Y_n) \\ &= e^{i(2aZ_1 + 2cZ_1Z_2)t}(iX_n + Y_n) \end{aligned}$$

where we have used the anticommutation properties of the Paulis. Then expanding out the exponential:

$$\begin{aligned}e^{iHt}(iX_1+Y_1)e^{-iHt} &= (\cos(2a)+i\sin(2a)Z_1)(\cos(2c)+i\sin(2c)Z_1Z_2)(iX_1+Y_1)\\ &= \cos(2a)\cos(2c)(iX_1+Y_1)+i\sin(2a)\cos(2c)(-Y_1-iX_1)\\ &+ i\cos(2a)\sin(2c)(-Y_1Z_2-iX_1Z_2)-\sin(2a)\sin(2c)(iX_1Z_2+Y_1Z_2)\end{aligned}$$

Let us be cavalier with the prefactors and simply denote what Pauli terms are contained in the above:

$$e^{iHt}(iX_1+Y_1)e^{-iHt} \sim X_1+Y_1+Y_1Z_2+X_1Z_2$$

For the second spin, the argument is analogous, and the terms appearing are the same as above with $1 \leftrightarrow 2$:

$$e^{iHt}(iX_2+Y_2)e^{-iHt} \sim X_2+Y_2+Z_1Y_2+Z_1X_2 \tag{6}$$

Thus for $M_0 = I$ we have:

$$V_0(t) \sim \mathrm{Tr}\big(\rho(X_1+Y_1+Y_1Z_2+X_1Z_2+X_2+Y_2+Z_1Y_2+Z_1X_2)\big)$$

And hence for the first experiment we measure the $\rho_{XI}, \rho_{Y1}, \rho_{YZ}, \rho_{XZ}, \rho_{IX}, \rho_{IY}, \rho_{ZY}, \rho_{ZX}$ terms. A single experiment already provides us with 8 of the coefficients!

Now, let's consider the second experiment with $M_1 = R_{x1}$. We note from the anticommutativity of the Paulis that:

$$\bar{R}_x X R_x = e^{i\pi X/4} X e^{-i\pi X/4} = X$$

$$\bar{R}_x Y R_x = e^{i\pi X/4} Y e^{-i\pi X/4} = e^{i\pi X/2} Y = iXY = Z$$

$$\bar{R}_x Z R_x = e^{i\pi X/4} Z e^{-i\pi X/4} = e^{i\pi X/2} Z = iXZ = -Y$$

and hence:

$$\begin{aligned}\bar{R}_{x1}e^{iHt}\sum_{n=1}^{2}(iX_n+Y_n)e^{-iHt}R_{x1} &\sim \bar{R}_{x1}(X_1+Y_1+Y_1Z_2+X_1Z_2+X_2+Y_2+Z_1Y_2+Z_1X_2)R_{x1}\\ &= (X_1+Z_1+Z_1Z_2+X_1Z_2+X_2+Y_2+Y_1Y_2+Y_1X_2)\end{aligned}$$

Thus for $M_1 = R_{x1}$ we have:

$$V_1(t) \sim \mathrm{Tr}\big(\rho((X_1+Z_1+Z_1Z_2+X_1Z_2+X_2+Y_2+Y_1Y_2+Y_1X_2))\big)$$

and hence we measure the $\rho_{XI}, \rho_{ZI}, \rho_{ZZ}, \rho_{XZ}, \rho_{IX}, \rho_{IY}, \rho_{YY}, \rho_{YX}$ terms. Of these, the $\rho_{ZI}, \rho_{ZZ}, \rho_{YY}, \rho_{ZX}$ terms are new, and we have obtained 12 out of the 15 coefficients.

For $M_2 = R_{y1}$, we note again from the anticommutativity of the Paulis that:

$$\bar{R}_y X R_y = e^{i\pi Y/4} X e^{-i\pi Y/4} = e^{i\pi Y/2} X = iYX = -Z$$

$$\bar{R}_y Y R_y = e^{i\pi Y/4} Y e^{-i\pi Y/4} = Y$$

$$\bar{R}_y Z R_y = e^{i\pi Y/4} Z e^{-i\pi Y/4} = e^{i\pi Y/2} Z = iYZ = X$$

and hence:

$$\begin{aligned}\bar{R}_{y1} e^{iHt} \sum_{n=1}^{2} (iX_n + Y_n) e^{-iHt} R_{y1} &\sim \bar{R}_{y1}(X_1 + Y_1 + Y_1Z_2 + X_1Z_2 + X_2 + Y_2 + Z_1Y_2 + Z_1X_2) R_{y1} \\ &= (Z_1 + Y_1 + Y_1Z_2 + Z_1Z_2 + X_2 + Y_2 + X_1Y_2 + X_1X_2)\end{aligned}$$

Thus for $M_2 = R_{y1}$ we have:

$$V_2(t) \sim \mathrm{Tr}\big(\rho((Z_1 + Y_1 + Y_1Z_2 + Z_1Z_2 + X_2 + Y_2 + X_1Y_2 + X_1X_2))\big)$$

and hence we measure the $\rho_{ZI}, \rho_{YI}, \rho_{YZ}, \rho_{ZZ}, \rho_{IX}, \rho_{IY}, \rho_{XY}, \rho_{XX}$ terms. Out of these, the ρ_{XY}, ρ_{XX} terms are new, and so we have obtained 14 out of the 15 coefficeints.
For $M_3 = R_{x2}$ we have:

$$\begin{aligned}\bar{R}_{x2} e^{iHt} \sum_{n=1}^{2} (iX_n + Y_n) e^{-iHt} R_{x2} &\sim \bar{R}_{x2}(X_1 + Y_1 + Y_1Z_2 + X_1Z_2 + X_2 + Y_2 + Z_1Y_2 + Z_1X_2) R_{x2} \\ &= (X_1 + Y_1 + Y_1Y_2 + X_1Y_2 + X_2 + Z_2 + Z_1Z_2 + Z_1X_2)\end{aligned}$$

and hence:

$$V_3(t) \sim \mathrm{Tr}\big(\rho((X_1 + Y_1 + Y_1Y_2 + X_1Y_2 + X_2 + Z_2 + Z_1Z_2 + Z_1X_2))\big)$$

and hence we measure the $\rho_{XI}, \rho_{Y1}, \rho_{YY}, \rho_{XY}, \rho_{IX}, \rho_{IZ}, \rho_{ZZ}, \rho_{ZX}$ terms. Out of these, the ρ_{IZ} term is new, and hence we have obtained 15 out of 15 coefficents.
Thus, in fact the four experiments $M_0 = I, M_1 = R_{x1}, M_2 = R_{y1}, M_3 = R_{x2}$ are sufficient to reconstruct ρ. This automatically implies that the full suite of nine experiments (with all possible combinations of X/Y pulses) are sufficient. □

Exercise 7.46

($*$) How many experiments are sufficient for three spins? Necessary?

Solution

Concepts Involved: Density Operators, Tensor Products, Pauli Operators, Commutators, Rotations, Composite Systems

---

For three spins, the general density matrix takes the form:

$$\rho = \sum_{i,j,k \in \{I,x,y,z\}} \rho_{ijk} \sigma_i^{(1)} \sigma_j^{(2)} \sigma_k^{(3)}.$$

The trace condition $\mathrm{Tr}(\rho) = 1$ fixes $\rho_{II} = \frac{1}{8}$, so to reconstruct ρ we require all 63 other coefficients ρ_{ij}. The system Hamiltonian takes the form:

$$H = aZ_1 + bZ_2 + cZ_3 + dZ_1Z_2 + eZ_1Z_3 + fZ_2Z_3$$

Given this, we now ask the necessary/sufficient k to measure the 63 coefficeints from the FID signal:

$$V_k(t) = V_0 \,\mathrm{Tr}\left[\rho M_k^\dagger e^{iHt} \sum_{n=1}^{3} (iX_n + Y_n) e^{-iHt} M_k\right]$$

Again let's study the $n = 1$ term, for which:

$$\begin{aligned}
&e^{iHt}(iX_1 + Y_1)e^{-iHt} \\
&= e^{i(aZ_1+bZ_2+cZ_3+dZ_1Z_2+eZ_1Z_3+fZ_2Z_3)t}(iX_1 + Y_1)e^{-i(aZ_1+bZ_2+cZ_3+dZ_1Z_2+eZ_1Z_3+fZ_2Z_3)t} \\
&= e^{i(aZ_1+bZ_2+cZ_3+dZ_1Z_2+eZ_1Z_3+fZ_2Z_3)t}(iX_1 + Y_1)e^{-i(-aZ_1+bZ_2+cZ_3+-dZ_1Z_2+-eZ_1Z_3+fZ_2Z_3)t} \\
&= e^{i(2aZ_1+2dZ_1Z_2+2eZ_1Z_3)t}(iX_1 + Y_1)
\end{aligned}$$

Which if we expand out in terms of sines/cosines, we find the terms:

$$\begin{aligned}
&e^{iHt}(iX_1 + Y_1)e^{-iHt} \\
&= (\cos(2a)I + i\sin(2a)Z_1)(\cos(2d)I + i\sin(2d)Z_1Z_2)(\cos(2e)I + i\sin(2e)Z_1Z_3)(iX_1 + Y_1) \\
&\sim X_1 + Y_1 + X_1Z_2 + Y_1Z_2 + X_1Z_3 + Y_1Z_3 + X_1Z_2Z_3 + Y_1Z_2Z_3
\end{aligned}$$

Repeating the same procedure for $n = 2, 3$ we have:

$$e^{iHt}(iX_2 + Y_2)e^{-iHt} \sim X_2 + Y_2 + Z_1X_2 + Z_1Y_2 + X_2Z_3 + Y_2Z_3 + Z_1X_2Z_3 + Z_1Y_2Z_3$$

$$e^{iHt}(iX_3 + Y_3)e^{-iHt} \sim X_3 + Y_3 + Z_1X_3 + Z_1Y_3 + Z_2X_3 + Z_2Y_3 + Z_1Z_2X_3 + Z_1Z_2Y_3$$

Hence with a first experiment $M_0 = I$ we would measure the above 24 terms/corresponding entries of ρ. We can already derive a necessary number of experiments from this one calculation. Of the 63 entries of ρ_{ijk}, 9 are single-Pauli terms, 27 are two-Pauli terms, and 27 are three-Pauli terms. As seen from the identity case above, a single experiment measures 6 single-Pauli terms, 12 two-Pauli terms, and 6 three-Pauli terms. The latter is what sees the least amount of coverage per experiment, and thus sets our bound. Assuming that we were able to construct a set of experiments where all three-Pauli terms were distinct, we would require $\lceil 27//6 \rceil = 5$ experiments in order to measure all three-Pauli terms.
Thus we have 5 experiments as necessary for full state tomography of 3 spins. Evidently, the suite of 27 experiments with all possible combinations of x/y rotations on the three qubits should be sufficient, but as in the 2-spin case we expect that there is overlap between data obtained from each experiment, and so much fewer should be sufficient. As in the previous exercise, we conjugate $e^{iHt}\sum_n (iX_n + Y_n)e^{-iHt}$ via the possible pulses to obtain the terms that are measured in each experiment. We then sweep over all possible $\binom{27}{5}$ combinations of experiments via computer program to see what the minimal sufficient set of experiments is. Doing so, we find that in fact no combination of 5 experiments is sufficient for the reconstruction. We find a similar result sweeping over all $\binom{27}{6}$ combinations of 6 experiments. The

minimal number of experiments appears to be 7, where there are many possible sufficient sets; one such set is:

$$M_0 = I, M_1 = R_{x1}, M_2 = R_{x2}, M_3 = R_{y1}, M_4 = R_{x1}R_{x2}, M_5 = R_{y1}R_{y2}R_{x3}, M_6 = R_{y1}R_{y2}R_{y3}$$

To prove that this is sufficient, we provide the measured terms in each experiment, highlighting the new terms in blue (which total to 63/all possible terms). For $M_0 = I$ (already calculated):

$$\begin{aligned} V_0 \sim \mathrm{Tr}\big(\rho\,(&X_1 + Y_1 + X_1Z_2 + Y_1Z_2 + X_1Z_3 + Y_1Z_3 + X_1Z_2Z_3 + Y_1Z_2Z_3 \\ &+ X_2 + Y_2 + Z_1X_2 + Z_1Y_2 + X_2Z_3 + Y_2Z_3 + Z_1X_2Z_3 + Z_1Y_2Z_3 \\ &+\ X_3 + Y_3 + Z_1X_3 + Z_1Y_3 + Z_2X_3 + Z_2Y_3 + Z_1Z_2X_3 + Z_1Z_2Y_3)\big) \end{aligned}$$

For $M_1 = R_{x1}$:

$$\begin{aligned} V_1 \sim \mathrm{Tr}\big(\rho\,(&X_1 + Z_1 + X_1Z_2 + Z_1Z_2 + X_1Z_3 + Z_1Z_3 + X_1Z_2Z_3 + Z_1Z_2Z_3 \\ &+ X_2 + Y_2 + Y_1X_2 + Y_1Y_2 + X_2Z_3 + Y_2Z_3 + Y_1X_2Z_3 + Y_1Y_2Z_3 \\ &+\ X_3 + Y_3 + Y_1X_3 + Y_1Y_3 + Z_2X_3 + Z_2Y_3 + Y_1Z_2X_3 + Y_1Z_2Y_3)\big) \end{aligned}$$

For $M_2 = R_{x2}$:

$$\begin{aligned} V_2 \sim \mathrm{Tr}\big(\rho\,(&X_1 + Y_1 + X_1Y_2 + Y_1Y_2 + X_1Z_3 + Y_1Z_3 + X_1Y_2Z_3 + Y_1Y_2Z_3 \\ &+ X_2 + Z_2 + Z_1X_2 + Z_1Z_2 + X_2Z_3 + Z_2Z_3 + Z_1X_2Z_3 + Z_1Z_2Z_3 \\ &+\ X_3 + Y_3 + Z_1X_3 + Z_1Y_3 + Y_2X_3 + Y_2Y_3 + Z_1Y_2X_3 + Z_1Y_2Y_3)\big) \end{aligned}$$

For $M_3 = R_{y1}$:

$$\begin{aligned} V_3 \sim \mathrm{Tr}\big(\rho\,(&Z_1 + Y_1 + Z_1Z_2 + Y_1Z_2 + Z_1Z_3 + Y_1Z_3 + Z_1Z_2Z_3 + Y_1Z_2Z_3 \\ &+ X_2 + Y_2 + X_1X_2 + X_1Y_2 + X_2Z_3 + Y_2Z_3 + X_1X_2Z_3 + X_1Y_2Z_3 \\ &+\ X_3 + Y_3 + X_1X_3 + X_1Y_3 + Z_2X_3 + Z_2Y_3 + X_1Z_2X_3 + X_1Z_2Y_3)\big) \end{aligned}$$

For $M_4 = R_{x1}R_{x2}$:

$$\begin{aligned} V_4 \sim \mathrm{Tr}\big(\rho\,(&X_1 + Z_1 + X_1Y_2 + Z_1Y_2 + X_1Z_3 + Z_1Z_3 + X_1Y_2Z_3 + Z_1Y_2Z_3 \\ &+ X_2 + Z_2 + Y_1X_2 + Y_1Z_2 + X_2Z_3 + Z_2Z_3 + Y_1X_2Z_3 + Y_1Z_2Z_3 \\ &+\ X_3 + Y_3 + Y_1X_3 + Y_1Y_3 + Y_2X_3 + Y_2Y_3 + Y_1Y_2X_3 + Y_1Y_2Y_3)\big) \end{aligned}$$

For $M_5 = R_{y1}R_{y2}R_{x3}$:

$$\begin{aligned} V_5 \sim \mathrm{Tr}\big(\rho\,(&Z_1 + X_1 + Z_1X_2 + Y_1X_2 + Z_1Y_3 + Y_1Y_3 + Z_1X_2Y_3 + Y_1X_2Y_3 \\ &+ Z_2 + Y_2 + X_1Z_2 + X_1Y_2 + Z_2Y_3 + Y_2Y_3 + X_1Z_2Y_3 + X_1Y_2Y_3 \\ &+\ X_3 + Z_3 + X_1X_3 + X_1Z_3 + X_2X_3 + X_2Z_3 + X_1X_2X_3 + X_1X_2Z_3)\big) \end{aligned}$$

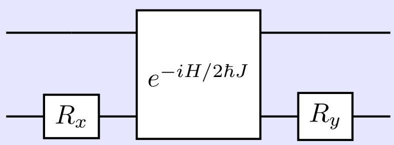
Finally for $M_6 = R_{y1}R_{y2}R_{y3}$:

$$V_6 \sim \text{Tr}\left(\rho\left(Z_1 + Y_1 + Z_1X_2 + Y_1X_2 + Z_1X_3 + Y_1Z_3 + Z_1X_2X_3 + Y_1X_2X_3 \right.\right.$$
$$+ Z_2 + Y_2 + X_1Z_2 + X_1Y_2 + Z_2Y_3 + Y_2X_3 + X_1Z_2X_3 + X_1Y_2X_3$$
$$\left.\left. + \ Z_3 + Y_3 + X_1Z_3 + X_1Y_3 + X_2Z_3 + X_2Y_3 + X_1X_2Z_3 + X_1X_2Y_3\right)\right)$$

Thus, we conclude that (certain sets of) 7 experiments are sufficient. □

Exercise 7.47: NMR controlled-NOT gate

Verify that the circuit shown in the top left of Figure 7.19 performs a controlled-NOT gate, up to single qubit phases; that is, it acts properly on classical input states, and furthermore can be turned into a proper controlled-NOT gate by applying additional single qubit R_z rotations. Give another circuit using the same building blocks to realize a proper CNOT gate.



There is an error in this exercise, as using the provided expression for the Hamiltonian:

$$H = 2\pi\hbar J Z_1 Z_2$$

yields $e^{-iJ/2\hbar J} = e^{-i\pi Z_1 Z_2} = -I$ which has trivial action and cannot possibly generate entanglement. We use the corrected expression for the Hamiltonian:

$$H = \frac{\pi\hbar J}{2} Z_1 Z_2$$

and also perform the opposite-sign time evolution to what is given in the question, with $e^{+iH/2\hbar J}$.

Solution

Concepts Involved: Controlled Operations, Rotations, Gate Decomposition

Throughout, R_{ik} denotes a $90°$ rotation about the ith axis on the kth qubit.
Computing the time evolution operator, we have:

$$U = e^{iH/2\hbar J} = e^{i\pi Z_1 Z_2/4} = \cos\left(\frac{\pi}{4}\right)I + i\sin\left(\frac{\pi}{4}\right)Z_1Z_2 = \text{diag}(e^{i\pi/4}, e^{-i\pi/4}, e^{-i\pi/4}, e^{i\pi/4})$$

For convenience, let us factor out a global phase:

$$U \cong \text{diag}(1, -i, -i, 1)$$

We can then calculate:

$$R_{y2}UR_{x2} = (\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} \otimes \frac{1}{\sqrt{2}}\begin{bmatrix}1 & -1\\1 & 1\end{bmatrix})\mathsf{diag}(1, -i, -i, 1)(\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} \otimes \frac{1}{\sqrt{2}}\begin{bmatrix}1 & -i\\-i & 1\end{bmatrix})$$
$$= \begin{bmatrix}1 & 0 & 0 & 0\\0 & -i & 0 & 0\\0 & 0 & 0 & -1\\0 & 0 & -i & 0\end{bmatrix}$$

which we can see has the action:

$$|00\rangle \mapsto |00\rangle$$
$$|01\rangle \mapsto -i\,|01\rangle$$
$$|10\rangle \mapsto -i\,|11\rangle$$
$$|11\rangle \mapsto -\,|10\rangle$$

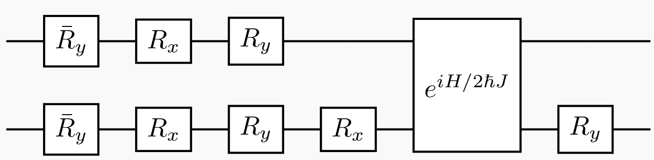
i.e up to single qubit phases we indeed have a `CNOT` gate. We require the use one-qubit Z rotations that have the net action:

$$|00\rangle \mapsto |00\rangle$$
$$|01\rangle \mapsto i\,|01\rangle$$
$$|10\rangle \mapsto i\,|10\rangle$$
$$|11\rangle \mapsto -\,|11\rangle$$

before the faulty `CNOT` gate in order to cancel the phases. We observe that $R_z = \mathsf{diag}(e^{-i\pi/4}, e^{i\pi/4}) \cong \mathsf{diag}(1, i)$ acting on both qubits has precisely this action, so the proper `CNOT` is given by:

$$CX_{1,2} = R_{y2}UR_{x2}R_{z1}R_{z2}$$

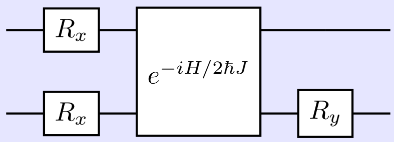
Noticing that $R_y\bar{R}_x\bar{R}_y = \mathsf{diag}(e^{-i\pi/4}, e^{i\pi/4}) = R_z$ (with $\bar{R} = R^\dagger$), we thus have that the circuit below has the action (up to a global phase) of a `CNOT`:



□

Exercise 7.48

Verify that the circuit shown in the bottom left of Figure 7.19 creates the Bell state $(|00\rangle - |11\rangle)/\sqrt{2}$ as advertised.



The same corrections apply here as with Ex. 7.47

Solution

Concepts Involved: Controlled Operations, Rotations, Gate Decomposition

Calculating the action of the circuit, again with $U = e^{+iH/2\hbar J} \cong \text{diag}(1, -i, -i, 1)$ we have:

$$(I_1 \otimes R_{y2}(\pi/2))U(R_{x1}(\pi/2) \otimes I_2)(I_1 \otimes R_{x2}(\pi/2)) = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 & -i & 0 \\ 0 & -i & 0 & -1 \\ 0 & i & 0 & -1 \\ -1 & 0 & -i & 0 \end{bmatrix}$$

From which we can see that:

$$(I_1 \otimes R_{y2}(\pi/2))U(R_{x1}(\pi/2) \otimes I_2)(I_1 \otimes R_{x2}(\pi/2))\,|00\rangle = \frac{|00\rangle - |11\rangle}{\sqrt{2}}$$

as claimed. □

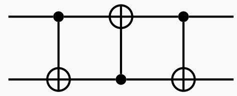
Exercise 7.49: NMR swap gate

An important chemical application of NMR is measurement of connectivity of spins, i.e. what protons, carbons, and phosphorus atoms are nearest neighbors in a molecule. One pulse sequence to do this is known as INADEQUATE (incredible natural abundance double quantum transfer experiment – the art of NMR is full of wonderfully creative acronyms). In the language of quantum computation, it can be understood as simply trying to apply a `CNOT` between any two resonances; if the `CNOT` works, the two nuclei must be neighbors. Another building block which is used in sequences such as TOCSY (total correlation spectroscopy) is a swap operation, but not quite in the perfect form we can describe simply with quantum gates! Construct a quantum circuit using only $e^{-iH/2\hbar J}$, R_x, and R_y operations to implement a swap gate (you may start from the circuit in Figure 1.7).
The same corrections apply here as with Ex. 7.47

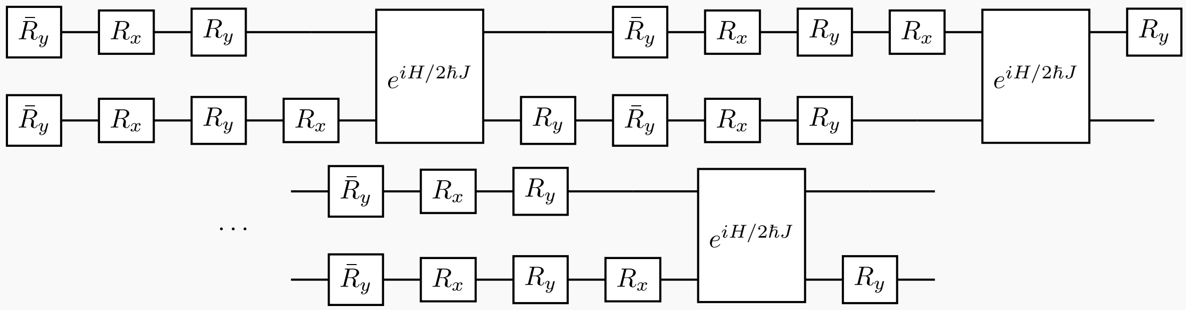
Solution

Concepts Involved: Controlled Operations, Rotations, Gate Decomposition

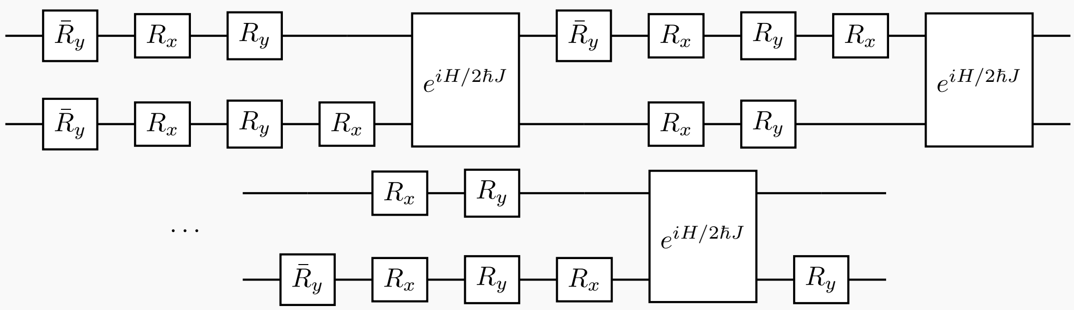
We can use the circuit of Figure 1.7:



and then replace each CNOT with the construction of Ex. 7.47:



Two pairs of intermediate R_y gates cancel, and we are left with:



□

Exercise 7.50

Find quantum circuits using just single qubit rotations and $e^{-iH/2\hbar J}$ to implement the oracle O for $x_0 = 0, 1, 2$.

The same corrections apply here as with Ex. 7.47

Solution

Concepts Involved: Gate Decomposition, Grover Algorithm

Now using the notation that $\tau = e^{iH/2\hbar J}$, we have $\tau \cong \text{diag}(1, -i, -i, 1)$ as our two-qubit gate as before. We will use z-rotations to construct the desired diagonal phase gates. Then, we will use an alternative decomposition of z-rotation from Ex. 7.47, namely that $R_y \bar{R}_x \bar{R}_y = \text{diag}(e^{-i\pi/4}, e^{i\pi/4}) = R_z \cong \text{diag}(1, i)$ to convert this into quantum circuits only involving x/y pulses.
With this, we obtain:

$$\bar{R}_{z1}\bar{R}_{z2}\tau = R_{y1}R_{x1}\bar{R}_{y1}R_{y2}R_{x2}\bar{R}_{y2}\tau = \text{diag}(1, -1, -1, -1) = P \cong \text{diag}(-1, 1, 1, 1) = O_0$$

$$R_{z1}\bar{R}_{z2}\tau = R_{y1}\bar{R}_{x1}\bar{R}_{y1}R_{y2}R_{x2}\bar{R}_{y2}\tau = \text{diag}(1,-1,1,1) = O_1$$

$$\bar{R}_{z1}R_{z2}\tau = R_{y1}R_{x1}\bar{R}_{y1}R_{y2}\bar{R}_{x2}\bar{R}_{y2}\tau = \text{diag}(1,1,-1,1) = O_2$$

$$R_{z1}R_{z2}\tau = R_{y1}\bar{R}_{x1}\bar{R}_{y1}R_{y2}\bar{R}_{x2}\bar{R}_{y2}\tau = \text{diag}(1,1,1,-1) = O_3$$

□

Remark: We choose a different decomposition as this choice will be more convenient for the cancellation of intermediate gates in the proceeding exercise.

Exercise 7.51

Show that the Grover iteration can be simplified, by canceling adjacent single qubit rotations appropriately, to obtain

$$G = \begin{cases} \bar{R}_{x1}\bar{R}_{y1}\bar{R}_{x2}\bar{R}_{y2}\tau R_{x1}\bar{R}_{y1}R_{x2}\bar{R}_{y2}\tau & (x_0 = 3) \\ \bar{R}_{x1}\bar{R}_{y1}\bar{R}_{x2}\bar{R}_{y2}\tau R_{x1}\bar{R}_{y1}\bar{R}_{x2}\bar{R}_{y2}\tau & (x_0 = 2) \\ \bar{R}_{x1}\bar{R}_{y1}\bar{R}_{x2}\bar{R}_{y2}\tau\bar{R}_{x1}\bar{R}_{y1}R_{x2}\bar{R}_{y2}\tau & (x_0 = 1) \\ \bar{R}_{x1}\bar{R}_{y1}\bar{R}_{x2}\bar{R}_{y2}\tau\bar{R}_{x1}\bar{R}_{y1}\bar{R}_{x2}\bar{R}_{y2}\tau & (x_0 = 0) \end{cases}$$

for the four possible cases of x_0.
The same corrections apply here as with Ex. 7.47. Also, note that the problem statement has switched the results for $x_0 = 2$ and $x_0 = 1$.

Solution

Concepts Involved: Rotations, Grover Algorithm

Using the expressions for O, P from the previous exercise, and $G = H^{\otimes 2}PH^{\otimes 2}O$ with $H_i = R_{xi}^2\bar{R}_{yi}$ we can go through the simplifications. Starting with $x_0 = 3$:

$$\begin{aligned} G_3 &= H^{\otimes 2}PH^{\otimes 2}O_3 \\ &= (R_{x1}^2\bar{R}_{y1}R_{x2}^2\bar{R}_{y2})R_{y1}R_{x1}\bar{R}_{y1}R_{y2}R_{x2}\bar{R}_{y2}\tau(R_{x1}^2\bar{R}_{y1}R_{x2}^2\bar{R}_{y2})R_{y1}\bar{R}_{x1}\bar{R}_{y1}R_{y2}\bar{R}_{x2}\bar{R}_{y2}\tau \\ &= R_{x1}^3\bar{R}_{y1}R_{x2}^3\bar{R}_{y2}\tau R_{x1}\bar{R}_{y1}R_{x2}\bar{R}_{y2}\tau \\ &= \bar{R}_{x1}\bar{R}_{y1}\bar{R}_{x2}\bar{R}_{y2}\tau R_{x1}\bar{R}_{y1}R_{x2}\bar{R}_{y2}\tau \end{aligned}$$

where in the second line we cancel out all coloured pairs of rotations and in the last line we use that 3 $\pi/2$ rotations is equivalent to a $-\pi/2$ rotation.
For the remaining entries, the simplification of $H^{\otimes 2}P$ is identical, so we only need to consider the simplification of $H^{\otimes 2}O_i$. For $x_0 = 2$:

$$\begin{aligned} H^{\otimes 2}O_1 &= (R_{x1}^2\bar{R}_{y1}R_{x2}^2\bar{R}_{y2})R_{y1}R_{x1}\bar{R}_{y1}R_{y2}\bar{R}_{x2}\bar{R}_{y2}\tau \\ &= R_{x1}^3\bar{R}_{y1}R_{x2}\bar{R}_{y2}\tau \\ &= \bar{R}_{x1}^3\bar{R}_{y1}R_{x2}\bar{R}_{y2}\tau \end{aligned}$$

and so:

$$G_2 = \bar{R}_{x1}\bar{R}_{y1}\bar{R}_{x2}\bar{R}_{y2}\tau\bar{R}_{x1}\bar{R}_{y1}R_{x2}\bar{R}_{y2}\tau$$

For $x_0 = 1$:

$$\begin{aligned} H^{\otimes 2}O_1 &= (R_{x1}^2\bar{R}_{y1}R_{x2}^2\bar{R}_{y2})R_{y1}\bar{R}_{x1}\bar{R}_{y1}R_{y2}R_{x2}\bar{R}_{y2}\tau \\ &= R_{x1}\bar{R}_{y1}R_{x2}^3\bar{R}_{y2}\tau \\ &= R_{x1}\bar{R}_{y1}\bar{R}_{x2}\bar{R}_{y2}\tau \end{aligned}$$

and so:

$$G_1 = \bar{R}_{x1}\bar{R}_{y1}\bar{R}_{x2}\bar{R}_{y2}\tau R_{x1}\bar{R}_{y1}\bar{R}_{x2}\bar{R}_{y2}\tau$$

For $x_0 = 0$ we know that $P \cong O_0$, so we just use the $H^{\otimes 2}P$ simplification result twice:

$$G_0 = \bar{R}_{x1}\bar{R}_{y1}\bar{R}_{x2}\bar{R}_{y2}\tau\bar{R}_{x1}\bar{R}_{y1}\bar{R}_{x2}\bar{R}_{y2}\tau$$

□

Exercise 7.52: Universality of Heisenberg Hamiltonian

Show that a swap operation U can be implemented by turning on $J(t)$ for an appropriate amount of time in the Heisenberg coupling Hamiltonian of (7.174), to obtain $U = \exp(-i\pi\mathbf{S}_1 \cdot \mathbf{S}_2)$. The '$\sqrt{\texttt{SWAP}}$' gate obtained by turning on the interaction for half this time is univeral; compute this transform and show how to obtained a controlled-NOT gate by composing it with single qubit rotations.

Solution

Concepts Involved: Universality, Controlled Operations, Pauli Operators, Operator Functions

We have the Hamiltonian:

$$H(t) = J(t)\mathbf{S}_1 \cdot \mathbf{S}_2 \overset{\text{turn on J(t)}}{=} \frac{1}{4}[X_1X_2 + Y_1Y_2 + Z_1Z_2]$$

By turning on $J(t)$ for time $t = \pi$, we obtain:

$$U = \exp(-i\pi\mathbf{S}_1 \cdot \mathbf{S}_2) = \exp\big(-i\pi X_1X_2/4\big)\exp\big(-i\pi Y_1Y_2/4\big)\exp\big(-i\pi Z_1Z_2/4\big)$$

where we have used that X_1X_1, Y_1Y_2, Z_1Z_2 mutually commute to split up the exponential. Then rewriting this in terms of sines/cosines:

$$\begin{aligned}
U &= \left(\cos\left(\frac{\pi}{4}\right)I - i\sin\left(\frac{\pi}{4}\right)X_1X_2\right)\left(\cos\left(\frac{\pi}{4}\right)I - i\sin\left(\frac{\pi}{4}\right)Y_1Y_2\right)\left(\cos\left(\frac{\pi}{4}\right)I - i\sin\left(\frac{\pi}{4}\right)Z_1Z_2\right) \\
&= \frac{1}{2\sqrt{2}}(I - iX_1X_2)(I - iY_1Y_2)(I - iZ_1Z_2) \\
&= \begin{bmatrix} e^{-i\pi/4} & 0 & 0 & 0 \\ 0 & 0 & e^{-i\pi/4} & 0 \\ 0 & e^{-i\pi/4} & 0 & 0 \\ 0 & 0 & 0 & e^{-i\pi/4} \end{bmatrix} \\
&\cong \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}
\end{aligned}$$

where in the last line we neglect the overall global phase. Thus we see that U indeed realizes the swap operation.
If we turn on $J(t)$ for time $t = \pi/2$ instead, we have:

$$\begin{aligned}
\sqrt{\texttt{SWAP}} &= \exp\bigl(-i\pi\mathbf{S}_1 \cdot \mathbf{S}_2/2\bigr) \\
&= \exp\bigl(-i\pi X_1X_2/8\bigr)\exp\bigl(-i\pi Y_1Y_2/8\bigr)\exp\bigl(-i\pi Z_1Z_2/8\bigr) \\
&= \left(\cos\left(\frac{\pi}{8}\right)I - i\sin\left(\frac{\pi}{8}\right)X_1X_2\right)\left(\cos\left(\frac{\pi}{8}\right)I - i\sin\left(\frac{\pi}{8}\right)Y_1Y_2\right)\left(\cos\left(\frac{\pi}{8}\right)I - i\sin\left(\frac{\pi}{8}\right)Z_1Z_2\right) \\
&= \begin{bmatrix} e^{i\pi/8} & 0 & 0 & 0 \\ 0 & e^{i\pi/8}\frac{1+i}{2} & e^{i\pi/8}\frac{1-i}{2} & 0 \\ 0 & e^{i\pi/8}\frac{1-i}{2} & e^{i\pi/8}\frac{1+i}{2} & 0 \\ 0 & 0 & 0 & e^{i\pi/8} \end{bmatrix} \\
&\cong \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1+i}{2} & \frac{1-i}{2} & 0 \\ 0 & \frac{1-i}{2} & \frac{1+i}{2} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}
\end{aligned}$$

We construct a `CNOT` using this and single-qubit gates. We first note that we can use two $\sqrt{\texttt{SWAP}}$s to make a diagonal gate:

$$\sqrt{\texttt{SWAP}}Z_1\sqrt{\texttt{SWAP}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

to get a CZ-gate we can multiply this by $R_{z1}(\pi/2)R_{z2}(-\pi/2)$. This is because:

$$\begin{aligned}
R_{z1}(\pi/2)R_{z2}(-\pi/2) &= \begin{bmatrix} e^{-i\pi/4} & 0 \\ 0 & e^{i\pi/4} \end{bmatrix} \otimes \begin{bmatrix} e^{i\pi/4} & 0 \\ 0 & e^{-i\pi/4} \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & e^{-i\pi/2} & 0 & 0 \\ 0 & 0 & e^{i\pi/2} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \\ 0 & 0 & i & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}
\end{aligned}$$

and so:

$$\begin{aligned}
R_{z1}(\pi/2)R_{z2}(-\pi/2)\sqrt{\texttt{SWAP}}Z_1\sqrt{\texttt{SWAP}} &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & e^{-i\pi/2} & 0 & 0 \\ 0 & 0 & e^{i\pi/2} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \\
&= CZ
\end{aligned}$$

Then as per Ex. 4.17 we can conjugate the desired target qubit by Hadamard gates to obtain a CNOT from CZ, and hence:

$$\texttt{CNOT}_{1,2} = H_2 R_{z1}(\pi/2)R_{z2}(-\pi/2)\sqrt{\texttt{SWAP}}Z_1\sqrt{\texttt{SWAP}}H_2$$

□

Problem 7.1: Efficient temporal labeling

Can you construct efficient circuits (which require only $O(\text{poly}(n))$ gates) to cyclically permute all diagonal elements in a $2^n \times 2^n$ diagonal density matrix except the $|0^n\rangle\langle 0^n|$ term?

Solution

Concepts Involved: Permutations, Controlled Operations, LU-decomposition

---

We follow the construction of `https://qcg.engin.umich.edu/wp-content/uploads/sites/393/2018/06/ApplicableAlgebra.pdf`. We make the observation that each permutation matrix (does not necessarily have to be cyclic) which permutes the diagonal entries of a $2^n \times 2^n$ diagonal density matrix ρ has the action of permuting the 2^n computational basis states, i.e. bitstrings of length n. We can

thus associate such permutations with invertible $n \times n$ matrices A acting on $(\mathbb{Z}_2)^n$, the vector space of n-dimensional vectors with elements in $\mathbb{Z}_2 \in \{0,1\}$ (equipped with mod 2 addition).
We now make the observation that any square matrix A can be decomposed as:

$$A = PLU$$

where P is a permutation matrix, L is a lower triangular matrix, and U is an upper triangular matrix. This is simply the matrix equation representing Gaussian elimination - U is the row echelon form of A, L corresponds to the elementary row operations/additions, and P corresponds to the row permutations.
Let us perform this decomposition on A. We then observe that the permutation matrix P appearing in the decomposition corresponds to a permutation of the n qubits for our desired quantum circuit. Such a permutation requires the use of n `SWAP` gates, with one swap per relabelling of a qubit. This corresponds to $3n$ `CNOT` gates, so $O(n)$ gates.
Now let us study L. The action of L on the basis states $\mathbf{e}_j = (0, 0, \ldots, 1, \ldots, 0)^T$ of $(\mathbb{Z}_2)^n$ are to map $L\mathbf{e}_j = \sum_{i \geq j} L_{ij}\mathbf{e}_i$ (with the sum restricted to $i \geq j$ as L is lower diagonal). Since A is invertible, L has 1 all across the diagonal, and hence $L\mathbf{e}_j$ will have an jth component equal to 1. Hence, since the action of the jth column of L is to preserve e_j while adding 1 to/flipping the $i > j$th components whenever $L_{ij} = 1$, in the desired quantum circuit this column corresponds to the composition of `CNOT` operations $\prod_{i>j} L_{ij} CX_{j,i}$ (with the jth qubit the control and ith the target). The total action of L is over all of its columns:

$$\prod_{j=1}^{n} \prod_{i>j} L_{ij} CX_{j,i}$$

We can bound the number of gates appearing in the above by the number of entries in the lower diagonal, which is $\frac{n \times n}{2} = O(n^2)$ `CNOT`s.
The analysis for U follows in the exact same way, with the appropriate modification that U is upper triangular as opposed to lower triangular. It thus also has $O(n^2)$ `CNOT` gate cost.
Combining the circuits for P, L, U, we find that the total cost of constructing a permutation of the diagonal elements of a density matrix (except for the $|0^n\rangle\langle 0^n|$ term - note that all circuits appearing in our construction are composed of `CNOT`s and so the $|0^n\rangle\langle 0^n|$ entry must be left invariant) is:

$$\underbrace{O(n)}_{P} + \underbrace{O(n^2)}_{L} + \underbrace{O(n^2)}_{U} = O(\text{poly}(n)).$$

□

Problem 7.2

In performing quantum computation with single photons, suppose that instead of the dual rail representation of Section 7.4.1 we use a *unary* representation of states, where $|00\ldots01\rangle$ is 0, $|00\ldots010\rangle$ is 1, $|00\ldots0100\rangle$ is 2, and so on, up to $|10\ldots0\rangle$ being $2^n - 1$.

1. Show that an arbitrary unitary transformation on these states can be constructed completely from just beamsplitters and phase shifters (and no nonlinear media).
2. Construct a circuit of beamsplitters and phase shifters to perform the one qubit Deutsch–Jozsa algorithm.

3. Construct a circuit of beamsplitters and phase shifters to perform the two qubit quantum search algorithm.

4. Prove that an arbitrary unitary transform will, in general, require an exponential number (in n) of components to realize.

Solution

Concepts Involved: Beamsplitters, Phase Shifters, Deutsch-Jozsa Algorithm, Grover Search

1. We construct effective qubit operations in the unary representations. If we can show the ability to perform arbitrary single qubit gates and `CNOT` gates, by the universality constructions of Chapter 4 we can construct an arbitrary unitary transformation.

   - Z-rotations: For the effective qubit we wish to perform the Z-rotation on, identify the 2^{n-1} unary states/wires that correspond to the qubit in the $|0\rangle$ state and apply the phase shift $P(-\theta/2) = e^{-i\theta/2}$. Identify the 2^{n-1} unary states/wires corresponding to the qubit in the $|1\rangle$ state and apply the phase shift $P(\theta/2) = e^{i\theta/2}$. This is the action of the $R_z(\theta)$ gate.
   - Y-rotations: For the effective qubit we wish to perform the Y-rotation on, identify the 2^{n-2} pairs of unary states/wires corresponding to the qubit in the $\big|0/1\big\rangle$ state with all other states in the same basis state. Then, between all such pairs apply the beamsplitter operation $B_{\theta/2} = \begin{bmatrix} \cos(\theta/2) & -\sin(\theta/2) \\ \sin(\theta/2) & \cos(\theta/2) \end{bmatrix}$. This is the action of the $R_y(\theta)$ gate. Arbitrary single-qubit unitaries follow from the Euler decomposition of rotations of Theorem 4.1.
   - `CNOT` gates: It will be simpler to perform a CZ gate (which can then be conjugated via Hadamards to yield a `CNOT`). For the effective two qubits we wish to perform the CZ between, we identify the 2^{n-2} pairs of states/wires in the unary representation that correspond to the qubit state $|11\rangle$. To these unary states we apply the phase shift $P(\pi) = e^{i\pi} = -1$, which is the CZ action.

   Let us give a concrete example for three qubits. The mapping between the three-qubit computational basis states and the unary states are given by:

$$
\begin{aligned}
|000\rangle &\leftrightarrow |00000001\rangle = |0\rangle \\
|001\rangle &\leftrightarrow |00000010\rangle = |1\rangle \\
|010\rangle &\leftrightarrow |00000100\rangle = |2\rangle \\
|011\rangle &\leftrightarrow |00001000\rangle = |3\rangle \\
|100\rangle &\leftrightarrow |00010000\rangle = |4\rangle \\
|101\rangle &\leftrightarrow |00100000\rangle = |5\rangle \\
|110\rangle &\leftrightarrow |01000000\rangle = |6\rangle \\
|111\rangle &\leftrightarrow |10000000\rangle = |7\rangle
\end{aligned}
$$

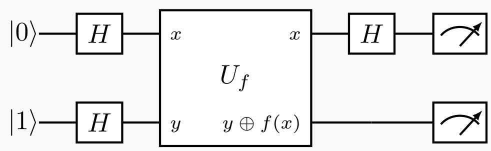
   To apply an effective Z-rotation on the second qubit, we would apply phase shifters of $P(-\theta/2)$ to $|0\rangle, |1\rangle, |4\rangle, |5\rangle$ and $P(\theta/2)$ to $|2\rangle, |3\rangle, |6\rangle, |7\rangle$. To apply an effective Y-rotation on the second qubit, we would apply beamsplitters B_θ between pairs $|0\rangle$ / $|2\rangle$, $|1\rangle$ / $|3\rangle$, $|4\rangle$ / $|6\rangle$, $|5\rangle$ / $|7\rangle$. To apply a CZ gate between qubits 1 and 2, we would apply phase shifters $P(\pi)$ on states $|3\rangle$ and $|7\rangle$.

2. For this problem, we consider 2-qubit algorithms, where we make the identification:

$$|00\rangle \leftrightarrow |0001\rangle$$
$$|01\rangle \leftrightarrow |0010\rangle$$
$$|10\rangle \leftrightarrow |0100\rangle$$
$$|11\rangle \leftrightarrow |1000\rangle$$

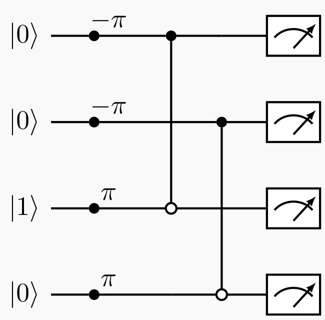
and denote $|0001\rangle$ as the top optical line down to $|1000\rangle$ as the bottom optical line.

The 1-qubit Deutsch-Jozsa circuit to reproduce looks like (in the binary representation):

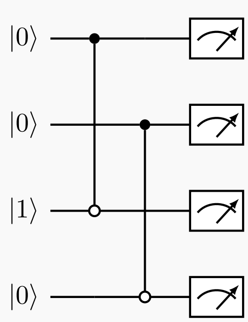


It will also be useful to recall the decomposition of the Hadamard gate as $H = R_y(\pi/2)Z = R_y(\pi/2)R_z(\pi)$. Let us look at the four possible cases of $f(x)$.

- $f(x) = 0$. in this case, U_f is the identity, so the Hadamard on the top qubit cancels, leaving just the single Hadamard on the bottom qubit. In unary, the resulting circuit takes the form:



- $f(x) = 1$. In this case $U_f = X_2$. Again the top two Hadamards cancel, and on the second qubit we have $XH = (HZH)H = HZ = R_y(\pi/2)ZZ = R_y(\pi/2)$. Thus in the unary representation:



TODO: draw remaining two circuits for the other two $f(x)$

3. TODO: more drawing
4. This already follows from our constructions in step 1. Since even a effective single-qubit unitary requires $O(2^n)$ optical components, a generic unitary transform will have this component cost.

□

Problem 7.3: Control via Jaynes–Cummings interactions

Robust and accurate control of small quantum systems – via an external *classical* degree of freedom – is important to the ability to perform quantum computation. It is quite remarkable that atomic states can be controlled by applying optical pulses, without causing superpositions of atomic states to decohere very much! In this problem, we see what approximations are necessary for this to be the case. Let us begin with the Jaynes–Cummings Hamiltonian for a single atom coupled to a single mode of an electromagnetic field,

$$H = a^\dagger \sigma_- + a\sigma_+,$$

where $\sigma_\pm$ act on the atom, and $a, a^\dagger$ act on the field.

1. For $U = e^{i\theta H}$, compute

$$A_n = \langle n|U|\alpha\rangle$$

where $|\alpha\rangle$ and $|n\rangle$ are coherent states and number eigenstates of the field, respectively; A_n is an *operator* on atomic states, and you should obtain

$$A_n = e^{-|\alpha|^2}\frac{|\alpha|^2}{n!}\begin{bmatrix} \cos(\theta\sqrt{n}) & \frac{i\sqrt{n}}{\alpha}\sin(\theta\sqrt{n}) \\ \frac{i\alpha}{\sqrt{n+1}}\sin\left(\theta\sqrt{n+1}\right) & \cos\left(\theta\sqrt{n+1}\right)\end{bmatrix}$$

The results of Exercise 7.17 may be helpful.

There is an error in the above expression, the correct expression should be:

$$A_n = e^{-|\alpha|^2/2}\frac{\alpha^n}{\sqrt{n!}}\begin{bmatrix} \cos(\theta\sqrt{n}) & \frac{i\sqrt{n}}{\alpha}\sin(\theta\sqrt{n}) \\ \frac{i\alpha}{\sqrt{n+1}}\sin\left(\theta\sqrt{n+1}\right) & \cos\left(\theta\sqrt{n+1}\right)\end{bmatrix}$$

2. It is useful to make an approximation that α is large (without loss of generality, we may choose α real). Consider the probability distribution

$$p_n = e^{-x}\frac{x^n}{n!}$$

which has mean $\langle n\rangle = x$ and standard deviation $\sqrt{\langle n^2\rangle - \langle n\rangle^2} = \sqrt{x}$. Now change variables to

$n = x - L\sqrt{x}$, and use Stirling's approximation

$$n! \approx \sqrt{2\pi n} n^n e^{-n}$$

to obtain

$$p_L \approx \frac{e^{-L^2/2}}{\sqrt{2\pi}}$$

3. The most important term is A_n for $n = |\alpha|^2$. Define $n = \alpha^2 + L\alpha$, and for

$$a = y\sqrt{\frac{1}{y^2} + \frac{L}{y}} \text{ and } b = y\sqrt{\frac{1}{y^2} + \frac{L}{y} + 1},$$

where $y = 1/\alpha$, show that

$$A_L \approx \frac{e^{-L^2/4}}{(2\pi)^{1/4}} \begin{bmatrix} \cos a\varphi & ia \sin a\varphi \\ (i/b) \sin b\varphi & \cos b\varphi \end{bmatrix},$$

using $\theta = \varphi/\alpha$. Also verify that

$$\int_{-\infty}^{\infty} A_L^\dagger A_L dL = I$$

as expected.

4. The ideal unitary transform which occurs to the atom is

$$U = \begin{bmatrix} \cos \alpha\theta & i \sin \alpha\theta \\ i \sin \alpha\theta & \cos \alpha\theta \end{bmatrix}$$

. How close is A_L to U? See if you can estimate the *fidelity*

$$\mathcal{F} = \min_{|\psi\rangle} \int_{-\infty}^{\infty} \left| \langle\psi|U^\dagger A_L|\psi\rangle \right|^2 dL$$

as a Taylor series in y.

Solution

Concepts Involved: Jaynes–Cummings Model, Operator Functions, Coherent States, Stirling's Approximation, Fidelity

1. We recall from Exercise 7.17 that:

$$H\left|\chi_n\right\rangle = \sqrt{n+1}\left|\chi_n\right\rangle$$
$$H\left|\overline{\chi}_n\right\rangle = -\sqrt{n+1}\left|\overline{\chi}_n\right\rangle$$

where $|\chi_n\rangle, |\overline{\chi}_n\rangle$ are the eigenstates

$$|\chi_n\rangle = \frac{1}{\sqrt{2}}[|n,1\rangle + |n+1,0\rangle]$$
$$|\overline{\chi}_n\rangle = \frac{1}{\sqrt{2}}[|n,1\rangle - |n+1,0\rangle]$$

We can invert these to write:

$$|n,1\rangle = \frac{1}{\sqrt{2}}[|\chi_n\rangle + |\overline{\chi}_n\rangle]$$
$$|n,0\rangle = \frac{1}{\sqrt{2}}[|\chi_{n-1}\rangle - |\overline{\chi}_{n-1}\rangle]$$

We also note from the definition of the coherent state $|\alpha\rangle = e^{-|\alpha|^2/2}\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}|n\rangle$ and the orthogonality of the number eigenstates

$$\begin{aligned}\langle\chi_n|\alpha\rangle &= \frac{1}{\sqrt{2}}[\langle 1|\langle n|\alpha\rangle + \langle 0|\langle n+1|\alpha\rangle]\\ &= \frac{1}{\sqrt{2}}[\langle 1| e^{-|\alpha|^2/2}\frac{\alpha^n}{\sqrt{n!}} + \langle 0| e^{-|\alpha|^2/2}\frac{\alpha^{n+1}}{\sqrt{(n+1)!}}]\\ &= \frac{e^{-|\alpha|^2/2}\alpha^n}{\sqrt{2n!}}[\langle 1| + \langle 0|\frac{\alpha}{\sqrt{(n+1)}}]\end{aligned}$$

$$\langle\overline{\chi}_n|\alpha\rangle = \frac{e^{-|\alpha|^2/2}\alpha^n}{\sqrt{2n!}}[\langle 1| - \langle 0|\frac{\alpha}{\sqrt{(n+1)}}]$$

So combining these two observations, we can evaluate

$$\begin{aligned}&\langle n,1| U |\alpha\rangle\\ &= \frac{1}{\sqrt{2}}[\langle\chi_n| + \langle\overline{\chi}_n|]e^{i\theta H}|\alpha\rangle\\ &= \frac{1}{\sqrt{2}}[\langle\chi_n| e^{i\theta\sqrt{n+1}} + \langle\overline{\chi}_n| e^{-i\theta\sqrt{n+1}}]|\alpha\rangle\\ &= \frac{1}{\sqrt{2}}[\langle\chi_n|\alpha\rangle e^{i\theta\sqrt{n+1}} + \langle\overline{\chi}_n|\alpha\rangle e^{-i\theta\sqrt{n+1}}]\\ &= \frac{1}{\sqrt{2}}[\left(\frac{e^{-|\alpha|^2/2}\alpha^n}{\sqrt{2n!}}[\langle 1| + \langle 0|\frac{\alpha}{\sqrt{(n+1)}}]\right)e^{i\theta\sqrt{n+1}} + \left(\frac{e^{-|\alpha|^2/2}\alpha^n}{\sqrt{2n!}}[\langle 1| - \langle 0|\frac{\alpha}{\sqrt{(n+1)}}]\right)e^{-i\theta\sqrt{n+1}}]\\ &= \frac{e^{-|\alpha|^2/2}\alpha^n}{2\sqrt{n!}}\left[2\cos\left(\theta\sqrt{n+1}\right)\langle 1| + \frac{\alpha}{\sqrt{n+1}}2i\sin\left(\theta\sqrt{n+1}\right)\langle 0|\right]\\ &= \frac{e^{-|\alpha|^2/2}\alpha^n}{\sqrt{n!}}\left[\cos\left(\theta\sqrt{n+1}\right)\langle 1| + \frac{i\alpha}{\sqrt{n+1}}\sin\left(\theta\sqrt{n+1}\right)\langle 0|\right]\end{aligned}$$

$$\begin{aligned}
&\langle n,0|\, U \,|\alpha\rangle \\
&= \frac{1}{\sqrt{2}}[\langle\chi_{n-1}| - \langle\overline{\chi}_{n-1}|]e^{i\theta H}\,|\alpha\rangle \\
&= \frac{1}{\sqrt{2}}[\langle\chi_{n-1}|\, e^{i\theta\sqrt{n}} - \langle\overline{\chi}_{n-1}|\, e^{-i\theta\sqrt{n}}]\,|\alpha\rangle \\
&= \frac{1}{\sqrt{2}}[\langle\chi_{n-1}|\alpha\rangle\, e^{i\theta\sqrt{n}} - \langle\overline{\chi}_{n-1}|\alpha\rangle\, e^{-i\theta\sqrt{n}}] \\
&= \frac{1}{\sqrt{2}}[\left(\frac{e^{-|\alpha|^2/2}\alpha^{n-1}}{\sqrt{2(n-1)!}}[\langle 1| + \langle 0|\,\frac{\alpha}{\sqrt{n}}]\right)e^{i\theta\sqrt{n}} - \left(\frac{e^{-|\alpha|^2/2}\alpha^{n-1}}{\sqrt{2(n-1)!}}[\langle 1| - \langle 0|\,\frac{\alpha}{\sqrt{n}}]\right)e^{-i\theta\sqrt{n}}] \\
&= \frac{e^{-|\alpha|^2/2}\alpha^{n}}{2\sqrt{n!}}\left[2i\frac{\sqrt{n}}{\alpha}\sin\!\left(\theta\sqrt{n}\right)\langle 1| + \frac{\alpha}{\sqrt{n}}\frac{\sqrt{n}}{\alpha}2\cos\!\left(\theta\sqrt{n}\right)\langle 0|\right] \\
&= \frac{e^{-|\alpha|^2/2}\alpha^{n}}{\sqrt{n!}}\left[\frac{i\sqrt{n}}{\alpha}\sin\!\left(\theta\sqrt{n}\right)\langle 1| + \cos\!\left(\theta\sqrt{n}\right)\langle 0|\right]
\end{aligned}$$

and from these we can read off the matrix elements

$$\langle n|\, U \,|\alpha\rangle = \frac{e^{-|\alpha|^2/2}\alpha^{n}}{\sqrt{n!}}\begin{bmatrix} \cos\!\left(\theta\sqrt{n}\right) & \frac{i\sqrt{n}}{\alpha}\sin\!\left(\theta\sqrt{n}\right) \\ \frac{i\alpha}{\sqrt{n+1}}\sin\!\left(\theta\sqrt{n+1}\right) & \cos\!\left(\theta\sqrt{n+1}\right) \end{bmatrix}$$

2. Throughout, we consider the limit of large x. Using Stirling's approximation on the Poisson distribution

$$p_n \approx e^{-x}\frac{x^n}{\sqrt{2\pi n}n^n e^{-n}}$$

we then change variables to $n = x - L\sqrt{x}$, multiplying the distribution by $\sqrt{x}$ to retain normalization:

$$\begin{aligned}
p_L &\approx e^{-x}\frac{\sqrt{x}x^{x-L\sqrt{x}}}{\sqrt{2\pi(x-L\sqrt{x})}(x-L\sqrt{x})^{x-L\sqrt{x}}e^{-(x-L\sqrt{x})}} \\
&= \frac{e^{-L\sqrt{x}}}{\sqrt{2\pi}}\sqrt{\frac{x}{x-L\sqrt{x}}}\left(\frac{x}{x-L\sqrt{x}}\right)^{x-L\sqrt{x}}
\end{aligned}$$

In the limit of large x the prefactor $\sqrt{\frac{x}{x-L\sqrt{x}}}$ approaches 1 and so can be dropped.

$$\begin{aligned}
p_L &\approx \frac{e^{-L\sqrt{x}}}{\sqrt{2\pi}}\left(\frac{x}{x-L\sqrt{x}}\right)^{x-L\sqrt{x}} \\
&= \frac{e^{-L\sqrt{x}}}{\sqrt{2\pi}}\left(\frac{x-L\sqrt{x}+L\sqrt{x}}{x-L\sqrt{x}}\right)^{x-L\sqrt{x}}
\end{aligned}$$

Now using that $\exp\big(\log(a)\big) = a$ and log laws we obtain

$$p_L \approx \frac{e^{-L\sqrt{x}}}{\sqrt{2\pi}} \exp\left(\log\left(\left(\frac{x}{x - L\sqrt{x}}\right)^{x - L\sqrt{x}}\right)\right)$$
$$= \frac{1}{\sqrt{2\pi}} \exp\Big(-L\sqrt{x} + (x - L\sqrt{x})[\log(x) - \log\big(x - L\sqrt{x})]\Big)$$

We can now write

$$\log\big(x - L\sqrt{x}\big) = \log\left(x(1 - \frac{L}{\sqrt{x}})\right) = \log(x) + \log\left(1 - \frac{L}{\sqrt{x}}\right) \approx \log(x) - \frac{L}{\sqrt{x}} - \frac{1}{2}\frac{L^2}{x}$$

where in the last step we Taylor expand to second order (valid as $L \ll \sqrt{x}$. Thus p_L becomes

$$p_L \approx \frac{1}{\sqrt{2\pi}} \exp\left(-L\sqrt{x} + (x - L\sqrt{x})[\log(x) - (\log(x) - \frac{L}{\sqrt{x}} - \frac{1}{2}\frac{L^2}{x})]\right)$$
$$= \frac{1}{\sqrt{2\pi}} \exp\left(-L\sqrt{x} + (x - L\sqrt{x})[\frac{L}{\sqrt{x}} + \frac{1}{2}\frac{L^2}{x})]\right)$$
$$= \frac{1}{\sqrt{2\pi}} \exp\left(-L\sqrt{x} + L\sqrt{x} + \frac{1}{2}L^2 - L^2 - \frac{1}{2}\frac{L^3}{\sqrt{x}}\right)$$
$$= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{L^2}{2} - \frac{1}{2}\frac{L^3}{\sqrt{x}}\right)$$

Since $L \ll \sqrt{x}$ we may drop the last term, and thus we conclude

$$p_L \approx \frac{e^{-L^2/2}}{\sqrt{2\pi}}$$

3. First studying the prefactor to A_n, we find (dropping the $|\alpha|$ as α is assumed to be real)

$$e^{-\alpha^2/2}\frac{\alpha^n}{\sqrt{n!}} = \sqrt{e^{-\alpha^2}\frac{\alpha^{2n}}{n!}} \approx \sqrt{\frac{e^{-L^2/2}}{\sqrt{2\pi}}} = \frac{e^{-L^2/4}}{(2\pi)^{1/4}}$$

where in the approximations step we observe that the expression is just a Poisson distribution with parameter α^2, so in the large α limit we can apply the result of the previous part.

Now we study the matrix elements. With the given definitions, we note that

$$\frac{\sqrt{n}}{\alpha} = a, \quad \frac{\sqrt{n+1}}{\alpha} = b$$

and so:

$$\begin{bmatrix} \cos\big(\theta\sqrt{n}\big) & \frac{i\sqrt{n}}{\alpha}\sin\big(\theta\sqrt{n}\big) \\ \frac{i\alpha}{\sqrt{n+1}}\sin\Big(\theta\sqrt{n+1}\Big) & \cos\Big(\theta\sqrt{n+1}\Big) \end{bmatrix} = \begin{bmatrix} \cos(\theta a\alpha) & ia\sin(\theta a\alpha) \\ (i/b)\sin(\theta b\alpha) & \cos(\theta b\alpha) \end{bmatrix}$$

and so with $\theta = \varphi/\alpha$ we conclude:

$$A_L \approx \frac{e^{-L^2/4}}{(2\pi)^{1/4}} \begin{bmatrix} \cos(a\varphi) & ia\sin(a\varphi) \\ (i/b)\sin(b\varphi) & \cos(b\varphi) \end{bmatrix}$$

as claimed. Evaluating $A_L^\dagger A_L$ we have:

$$\begin{aligned} &A_L^\dagger A_L \\ &= \frac{e^{-L^2/4}}{(2\pi)^{1/4}} \begin{bmatrix} \cos(a\varphi) & -(i/b)\sin(b\varphi) \\ -ia\sin(a\varphi) & \cos(b\varphi) \end{bmatrix} \frac{e^{-L^2/4}}{(2\pi)^{1/4}} \begin{bmatrix} \cos(a\varphi) & ia\sin(a\varphi) \\ (i/b)\sin(b\varphi) & \cos(b\varphi) \end{bmatrix} \\ &= \frac{e^{-L^2/2}}{\sqrt{2\pi}} \begin{bmatrix} \cos^2(a\varphi) + \frac{1}{b^2}\sin^2(b\varphi) & ia\cos(a\varphi)\sin(a\varphi) - (i/b)\cos(b\varphi)\sin(b\varphi) \\ -ia\cos(a\varphi)\sin(a\varphi) + (i/b)\cos(b\varphi)\sin(b\varphi) & \cos^2(b\varphi) + a^2\sin^2(a\varphi) \end{bmatrix} \end{aligned}$$

In the limit where $L \ll \alpha$ we have that:

$$a = y\sqrt{\frac{1}{y^2} + \frac{L}{y}} \approx y\sqrt{\frac{1}{y^2}} = 1, \quad b = y\sqrt{\frac{1}{y^2} + \frac{L}{y} + 1} \approx y\sqrt{\frac{1}{y^2}} = 1$$

and so $A_L^\dagger A_L$ becomes:

$$\begin{aligned} A_L^\dagger A_L &\approx \frac{e^{-L^2/2}}{\sqrt{2\pi}} \begin{bmatrix} \cos^2(\varphi) + \sin^2(\varphi) & i\cos(\varphi)\sin(\varphi) - i\cos(\varphi)\sin(\varphi) \\ -i\cos(\varphi)\sin(\varphi) + i\cos(\varphi)\sin(\varphi) & \cos^2(\varphi) + \sin^2(\varphi) \end{bmatrix} \\ &= \frac{e^{-L^2/2}}{\sqrt{2\pi}} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \end{aligned}$$

Now, we note the Gaussian integral identity

$$\int_{-\infty}^{\infty} e^{-L^2/2} dL = \sqrt{2\pi}$$

which allows us to conclude:

$$\int_{-\infty}^{\infty} A_L^\dagger A_L dL \approx \int_{-\infty}^{\infty} \frac{e^{-L^2/2}}{\sqrt{2\pi}} I dL = \frac{\sqrt{2\pi}}{\sqrt{2\pi}} I = I$$

4. **Setup and notation.** $y := 1/\alpha$, $n = \alpha^2 + L\alpha = y^{-2} + Ly^{-1}$ with $L = O(1)$; $a := \sqrt{n}/\alpha = \sqrt{1+Ly}$, $b := \sqrt{n+1}/\alpha = \sqrt{1+Ly+y^2}$; $U = e^{i\varphi\sigma_x} = \begin{pmatrix} \cos\varphi & i\sin\varphi \\ i\sin\varphi & \cos\varphi \end{pmatrix}$, $\varphi = \alpha\theta$; $M_L = \begin{pmatrix} \cos(a\varphi) & i\,a\,\sin(a\varphi) \\ i\,b^{-1}\sin(b\varphi) & \cos(b\varphi) \end{pmatrix}$; $K_L := U^\dagger M_L - I$; $|g(L)|^2 = (2\pi)^{-1/2} e^{-L^2/2}$.

Series to $O(y^2)$.

$$a = 1 + \tfrac{L}{2}y - \tfrac{L^2}{8}y^2 + O(y^3), \quad b = 1 + \tfrac{L}{2}y + \Big(\tfrac{1}{2} - \tfrac{L^2}{8}\Big)y^2 + O(y^3), \quad b^{-1} = 1 - \tfrac{L}{2}y + \Big(-\tfrac{1}{2} + \tfrac{3L^2}{8}\Big)y^2 + O(y^3).$$

$$\sin\big((1+\eta)\varphi\big) = \sin\varphi + \eta\varphi\cos\varphi - \tfrac{1}{2}\eta^2\varphi^2\sin\varphi + O(\eta^3\varphi^3),$$

$$\cos\big((1+\eta)\varphi\big) = \cos\varphi - \eta\varphi\sin\varphi - \tfrac{1}{2}\eta^2\varphi^2\cos\varphi + O(\eta^3\varphi^3).$$

Deviations $\Delta := M_L - U$ (keep only terms that feed $O(y^2\varphi^2)$ in the fidelity).

$$\Delta_{11} = \Delta_{22} = -\Big(\tfrac{L}{2}y\Big)\varphi\sin\varphi - \Big(\tfrac{L^2}{8}y^2\Big)\varphi^2\cos\varphi + O(y^3,\ y^2\varphi^3,\ y^2\varphi),$$

$$\Delta_{12} = i\,[a\sin(a\varphi) - \sin\varphi] = i\,(Ly\,\varphi) + i\Big(\tfrac{L^2}{4}y^2\varphi\cos\varphi\Big) + O(y\varphi^3,\ y^2\varphi^2),$$

$$\Delta_{21} = i\,[b^{-1}\sin(b\varphi) - \sin\varphi] = i\cdot O(y\varphi^3) + O(y^2\varphi^2).$$

Form $K_L = U^\dagger\Delta$, drop pieces odd in L or beyond $O(y^2\varphi^2)$.

$$K_{12} = (\cos\varphi)\Delta_{12} + (-i\sin\varphi)\Delta_{22} = i\,(Ly\,\varphi) + O(y\varphi^2)_{\text{odd in } L} + O(y^2\varphi),$$

$$K_{21} = (-i\sin\varphi)\Delta_{11} + (\cos\varphi)\Delta_{21} = i\cdot O(y\varphi^3),$$

$$K_{11} = K_{22} = (\cos\varphi)\Delta_{11} + (-i\sin\varphi)\Delta_{21} = -\Big(\tfrac{L^2}{8}y^2\varphi^2\Big)\cos^2\varphi + O(y\varphi^2)_{\text{odd in } L},$$

$$K_L = \begin{pmatrix} -\dfrac{L^2y^2}{8}\,\varphi^2\cos^2\varphi & i\,(Ly\,\varphi) \\ 0 & -\dfrac{L^2y^2}{8}\,\varphi^2\cos^2\varphi \end{pmatrix} + \ \text{(all other terms are odd in } L \text{ or higher order).}$$

Fidelity as a variance.

$$s_L := \langle\psi|(I+K_L)|\psi\rangle, \quad |s_L|^2 = 1 - \Big(\langle K_L^\dagger K_L\rangle_\psi - |\langle K_L\rangle_\psi|^2\Big) + O(y^3\varphi^3),$$

with $|\psi\rangle = \cos t\,|0\rangle + e^{i\varphi_0}\sin t\,|1\rangle$, $t_L := Ly\,\varphi$, $d_L := (L^2y^2/8)\varphi^2\cos^2\varphi$.

$$\langle K_L\rangle_\psi = -\,d_L + i\,t_L\cos t\,\sin t\,e^{-i\varphi_0}, \qquad \langle K_L^\dagger K_L\rangle_\psi = t_L^2\,\sin^2 t + d_L^2 + O(y^3\varphi^3),$$

$$\Rightarrow \quad \langle K_L^\dagger K_L\rangle_\psi - |\langle K_L\rangle_\psi|^2 = t_L^2\,\sin^4 t, \qquad |s_L|^2 = 1 - (Ly\,\varphi)^2\,\sin^4 t + O(y^3\varphi^3).$$

Worst case and Gaussian averaging . Keep only terms that can contribute at order $y^2\varphi^2$ and drop everything odd in L (zero under the even Gaussian). Then

$$K_L \ = \ -d_L I \ + \ N_L \ + \ O(y^3\varphi^3), \qquad N_L = \begin{pmatrix} 0 & i\,t_L \\ i\,u_L & 0 \end{pmatrix}, \quad u_L = O(y^2\varphi).$$

The diagonal shift $-d_L I$ cancels out of the variance exactly:

$$|s_L|^2 \ = \ 1 - \Big(\langle K_L^\dagger K_L\rangle_\psi - |\langle K_L\rangle_\psi|^2\Big) + O(y^3\varphi^3) = 1 - \Big(\langle N_L^\dagger N_L\rangle_\psi - |\langle N_L\rangle_\psi|^2\Big) + O(y^3\varphi^3).$$

Let $\rho = \frac{1}{2}(I + \mathbf{r}\cdot\boldsymbol{\sigma})$ be the pure state ($|\mathbf{r}| = 1$). A short Pauli-algebra calculation for $N_L = \left(\begin{smallmatrix} 0 & it_L \\ iu_L & 0 \end{smallmatrix}\right)$ gives,

$$\langle N_L^\dagger N_L\rangle_\psi - |\langle N_L\rangle_\psi|^2 = \frac{1}{2}(t_L^2+u_L^2) + \frac{1}{2}(u_L^2-t_L^2)\,r_z - \frac{1}{4}(t_L+u_L)^2 r_x^2 - \frac{1}{4}(t_L-u_L)^2 r_y^2.$$

This is a quadratic form $Q_L(\mathbf{r}) = \mathbf{r}^\top A_L\,\mathbf{r} + \mathrm{const}$ on the unit sphere with

$$A_L = \frac{1}{4}\begin{pmatrix} -(t_L+u_L)^2 & 0 & 0 \\ 0 & -(t_L-u_L)^2 & 0 \\ 0 & 0 & 2(u_L^2-t_L^2) \end{pmatrix}.$$

Its maximum over $|\mathbf{r}|=1$ is the top eigenvalue of $-A_L$ with a minus sign (since the constant term is independent of $\mathbf{r}$). Because $u_L = O(y^2\varphi) \ll t_L = O(y\varphi)$, we may set $u_L = 0$ in the eigenvalues without changing the value at order $y^2\varphi^2$. Thus,

$$\lambda_{\max}(-A_L) \;=\; \frac{t_L^2}{2} \qquad\Longrightarrow\qquad \max_{|\psi\rangle}\big(\langle N_L^\dagger N_L\rangle_\psi - |\langle N_L\rangle_\psi|^2\big) \;=\; \frac{t_L^2}{2} \;+\; O(y^3\varphi^3).$$

Equivalently, for the worst–case state at fixed L,

$$1-|s_L|^2 \;=\; \frac{1}{2}\,(Ly\varphi)^2 \;+\; O(y^3\varphi^3).$$

To connect this quadratic–form bound with the small *unitary* misrotation, note that the off–diagonal of $U^\dagger M_L$ at order $y\varphi$ equals it_L in the upper entry while the lower entry is $iu_L = O(y^2\varphi)$; the Hermitian generator that reproduces this (to the same order in the *variance*) is $i\varepsilon_L\sigma_x$ with

$$\varepsilon_L = \frac{t_L}{2} \quad \big(\text{since } i\varepsilon_L\sigma_x \text{ has } (0,1) \text{ and } (1,0) \text{ entries } i\varepsilon_L\big).$$

For a small unitary $e^{i\varepsilon_L\sigma_x}$, the worst–case pure-state overlap is $1-\varepsilon_L^2+O(\varepsilon_L^4)$. Substituting $\varepsilon_L = t_L/2$ into the previous line yields the additional factor $1/4$, and the per–L worst–case infidelity becomes

$$1-|s_L|^2 \;=\; \frac{t_L^2}{2}\;\times\;\frac{1}{4} \;+\; O(y^3\varphi^3) \;=\; \frac{1}{8}\,(Ly\varphi)^2 \;+\; O(y^3\varphi^3).$$

Finally, average with the even Gaussian

$$1-F = \int |g(L)|^2\,[1-|s_L|^2]\,dL = \frac{1}{8}\,\varphi^2 \underbrace{\int |g|^2 (Ly)^2\,dL}_{y^2=\alpha^{-2}} + O\Big(\frac{\varphi^4}{\alpha^2},\frac{\varphi^3}{\alpha^3}\Big) = \boxed{\frac{\varphi^2}{8\alpha^2} + O\Big(\frac{\varphi^4}{\alpha^2},\frac{\varphi^3}{\alpha^3}\Big)}.$$

□

Remark: Odd-in-L pieces vanish under the even Gaussian; the leading contribution is a variance built from the linear-in-L off-diagonal $K_{12} = i(Ly)\varphi$; the diagonal $O(L^2y^2\varphi^2)$ enforces variance (not mean) and does not contribute at leading order after cancellation; any symmetric pointer with the same $\mathbb{E}[L^2]$ yields the same α^{-2} scaling; shaping to cancel the $O(y)$ off-diagonal slope pushes $1-F$ to $O(\alpha^{-4})$.

Problem 7.4: Ion trap logic with two-level atoms

The controlled-NOT gate described in Section 7.6.3 used a three-level atom for simplicity. It is possible to do without this third level, with some extra complication, as this problem demonstrates.

Let $\mathcal{Y}_{\hat{\mathbf{n}}}^{\text{blue, j}}(\theta)$ denote the operation accomplished by pulsing light into the sideband frequency, $\omega = \Omega + \omega_z$, of the jth particle for time $\theta\sqrt{N}/\eta\Omega$, and similarly for the red sideband. $\hat{\mathbf{n}}$ denotes the axis of the rotation in the $\hat{\mathbf{x}} - \hat{\mathbf{y}}$ plane, controlled by setting the phase of the incident light. The superscript j may be omitted when it it clear which ion is being addressed. Specifically,

$$\mathcal{Y}_{\hat{\mathbf{n}}}^{\text{blue}}(\theta) = \exp\left[\left(e^{i\varphi}\left|00\right\rangle\!\left\langle 11\right| + e^{-i\varphi}\left|11\right\rangle\!\left\langle 00\right| + e^{i\varphi}\sqrt{2}\left|01\right\rangle\!\left\langle 12\right| + e^{-i\varphi}\left|12\right\rangle\!\left\langle 01\right| + \dots\right)\frac{i\theta}{2}\right],$$

where $\hat{\mathbf{n}} = \hat{\mathbf{x}}\cos\varphi + \hat{\mathbf{y}}\sin\varphi$, and the two labels in the ket represent the internal and the motional states, respectively, from left to right. The $\sqrt{2}$ factor comes from the fact that $a^\dagger\left|n\right\rangle = \sqrt{n+1}\left|n+1\right\rangle$ for bosonic states.

(1) Show that $S^j = \mathcal{Y}_{\hat{\mathbf{n}}}^{\text{red}}(\pi)$ performs a swap between the internal and motional states of ion j when the motional state is initially $\left|0\right\rangle$.

(2) Find a value of θ such that $\mathcal{Y}_{\hat{\mathbf{n}}}^{\text{blue}}(\theta)$ acting on any state in the computational subspace, spanned by $\left|00\right\rangle, \left|01\right\rangle, \left|10\right\rangle$, and $\left|11\right\rangle$, leaves it in that subspace. This should work for any axis $\hat{\mathbf{n}}$.

(3) Show that if $\mathcal{Y}_{\hat{\mathbf{n}}}^{\text{blue}}(\varphi)$ stays within the computational subspace, then

$$U = \mathcal{Y}_{\alpha}^{\text{blue}}(-\beta)\mathcal{Y}_{\hat{\mathbf{n}}}^{\text{blue}}(\theta)\mathcal{Y}_{\alpha}^{\text{blue}}(\beta)$$

also stays within the computational subspace, for any choice of rotation angle β and axis α.

(4) Find values of α and β such that U is diagonal. Specifically, it is useful to obtain an operator such as

$$\begin{bmatrix} e^{-i\pi/\sqrt{2}} & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & e^{i\pi/\sqrt{2}} \end{bmatrix}$$

.

(5) Show that (7.187) describes a non-trivial gate, in that a controlled-NOT gate between the internal states of two ions can be constructed from it and single qubit operations. Can you come up with a composite pulse sequence for performing a CNOT without requiring the motional state to initially be $\left|0\right\rangle$?

Solution

Concepts Involved: Controlled Operations, Operator Functions, Rotations

(1) The red sideband couples $\left|1,n\right\rangle \leftrightarrow \left|0,n+1\right\rangle$; with motion initially $\left|0\right\rangle$ the active subspace is $\mathcal{S} =$

$\mathrm{span}\{|1,0\rangle, |0,1\rangle\}$ and $|0,0\rangle$ is dark. In the ordered basis $\{|1,0\rangle, |0,1\rangle\}$ the generator is

$$A = e^{+i\varphi}\,|0,1\rangle\langle 1,0| + e^{-i\varphi}\,|1,0\rangle\langle 0,1| = \begin{pmatrix} 0 & e^{-i\varphi} \\ e^{+i\varphi} & 0 \end{pmatrix}, \qquad A^2 = I.$$

Thus a red-sideband pulse of area θ acts as

$$\mathcal{Y}^{\mathrm{red}}_{\hat{n}}(\theta)\big|_{\mathcal{S}} = \exp\Big(\tfrac{i\theta}{2}A\Big) = \cos\tfrac{\theta}{2}\,I + i\sin\tfrac{\theta}{2}\,A.$$

For a π pulse,

$$S^j \equiv \mathcal{Y}^{\mathrm{red},\,j}_{\hat{n}}(\pi)\big|_{\mathcal{S}} = iA = i\begin{pmatrix} 0 & e^{-i\varphi} \\ e^{+i\varphi} & 0 \end{pmatrix},$$

so that

$$|1,0\rangle \mapsto -i\,e^{+i\varphi}\,|0,1\rangle, \qquad |0,1\rangle \mapsto -i\,e^{-i\varphi}\,|1,0\rangle, \qquad |0,0\rangle \mapsto |0,0\rangle.$$

Hence, when the motion starts in $|0\rangle$, S^j swaps the internal excitation with one phonon, up to a global $(-i)$ and known Z-axis phases set by φ (removable by a virtual R_z or by choosing $\varphi = \pi/2$).

(2) On the blue sideband, the dynamics decomposes into disjoint SU(2) blocks $\{|0,n\rangle, |1,n{+}1\rangle\}$, with rotation angle $\theta\sqrt{n{+}1}$ about an in–plane axis $\hat{n}$ set by the laser phase φ. Within $\mathcal{H}_c = \mathrm{span}\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ the only possible leakage is $|01\rangle \leftrightarrow |12\rangle$ (the $\sqrt{2}$ factor comes from $a^\dagger|1\rangle = \sqrt{2}\,|2\rangle$). Explicitly,

$$\begin{aligned}
|00\rangle &\mapsto \cos\frac{\theta}{2}\,|00\rangle + i\,e^{i\varphi}\sin\frac{\theta}{2}\,|11\rangle,\\
|11\rangle &\mapsto \cos\frac{\theta}{2}\,|11\rangle + i\,e^{-i\varphi}\sin\frac{\theta}{2}\,|00\rangle,\\
|01\rangle &\mapsto \cos\frac{\sqrt{2}\,\theta}{2}\,|01\rangle + i\,e^{i\varphi}\sin\frac{\sqrt{2}\,\theta}{2}\,|12\rangle,\\
|10\rangle &\mapsto |10\rangle.
\end{aligned}$$

To ensure the image of *every* state in $\mathcal{H}_c$ remains in $\mathcal{H}_c$, we require the leakage amplitude to vanish

$$\sin\Big(\frac{\sqrt{2}\,\theta}{2}\Big) = 0 \iff \boxed{\theta = m\,\pi\sqrt{2}\,, \quad m \in \mathbb{Z}}.$$

This choice works for any axis $\hat{n}$ (the axis only affects phases within each SU(2) block, not the rotation angles). The minimal nontrivial choice is $\theta = \pi\sqrt{2}$, which returns $|01\rangle$ (up to a -1 phase), mixes $\{|00\rangle, |11\rangle\}$ within $\mathcal{H}_c$, and leaves $|10\rangle$ unchanged.

(3) Let $\mathcal{H}_c = \mathrm{span}\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ and, for the blue sideband, decompose the Hilbert space into disjoint two-level blocks

$$\mathcal{B}_n = \mathrm{span}\{|0,n\rangle, |1,n+1\rangle\}, \qquad n = 0,1,2,\ldots,$$

on which any $\mathcal{Y}^{\mathrm{blue}}_{\gamma}(\cdot)$ acts block-diagonally as an $\mathrm{SU}(2)$ rotation. Assume $\mathcal{Y}^{\mathrm{blue}}_{\hat{n}}(\theta)$ leaves $\mathcal{H}_c$ invariant (e.g., by part (b), take $\theta = m\pi\sqrt{2}$, so on $\mathcal{B}_1 = \mathrm{span}\{|01\rangle, |12\rangle\}$ it is $(-1)^m I$). For arbitrary axis α and angle β, define

$$W = \mathcal{Y}_{\alpha}^{\text{blue}}(\beta), \qquad U = W^{\dagger}\,\mathcal{Y}_{\hat{n}}^{\text{blue}}(\theta)\,W.$$

Because W is also block-diagonal over the same $\{\mathcal{B}_n\}$, we have

$$U\big|_{\mathcal{B}_1} = W^{\dagger}\big((-1)^m I\big)\,W = (-1)^m I,$$

so starting from $|01\rangle \in \mathcal{H}_c$ no $|12\rangle$ component can be produced. On $\mathcal{B}_0 = \text{span}\{|00\rangle, |11\rangle\}$, each factor maps $\mathcal{B}_0$ to itself, hence U does too; and $|10\rangle$ is fixed by every blue pulse. Therefore $U\,\mathcal{H}_c \subseteq \mathcal{H}_c$ for any α, β.

(4) Use the leakage-free choice from part (b), $\theta = \pi\sqrt{2}$, so that on $\mathcal{B}_1 = \text{span}\{|01\rangle, |12\rangle\}$ the blue pulse is a scalar $-I$ (hence $|01\rangle \mapsto -|01\rangle$ and no leakage), while on $\mathcal{B}_0 = \text{span}\{|00\rangle, |11\rangle\}$ it is an $\text{SU}(2)$ rotation by angle $\pi\sqrt{2}$ about some in-plane axis. Choose

$$U \;=\; \mathcal{Y}_{\hat{y}}^{\text{blue}}\big(-\tfrac{\pi}{2}\big)\;\mathcal{Y}_{\hat{x}}^{\text{blue}}\Big(\pi\sqrt{2}\Big)\;\mathcal{Y}_{\hat{y}}^{\text{blue}}\big(\tfrac{\pi}{2}\big)\,.$$

By axis conjugation, on $\mathcal{B}_0$ this equals a Z-rotation of angle $\pi\sqrt{2}$

$$R_z(\pi\sqrt{2}) = \exp\Big(-i\,\tfrac{\pi\sqrt{2}}{2}\,\sigma_z\Big),$$

so $|00\rangle \mapsto e^{-i\pi/\sqrt{2}}\,|00\rangle$ and $|11\rangle \mapsto e^{+i\pi/\sqrt{2}}\,|11\rangle$. On $\mathcal{B}_1$, the middle pulse is $-I$ and conjugation by the side pulses leaves it unchanged, hence $|01\rangle \mapsto -|01\rangle$. The blue sideband leaves $|10\rangle$ invariant. Therefore, in the computational basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$,

$$\boxed{U = \text{diag}\,\big(e^{-i\pi/\sqrt{2}},\ -1,\ 1,\ e^{+i\pi/\sqrt{2}}\big)}$$

is diagonal as required.

(5) Let

$$U = \text{diag}\big(e^{-i\gamma},\,-1,\,1,\,e^{+i\gamma}\big), \qquad \gamma = \tfrac{\pi}{\sqrt{2}},$$

acting on (internal$\otimes$bus) in the basis $\{|0,0\rangle, |0,1\rangle, |1,0\rangle, |1,1\rangle\}$.

With the shared mode prepared in $|0\rangle$, define the two-ion diagonal gate

$$G \;=\; S^{(1)}\,U^{(2)}\,S^{(1)\dagger},$$

where $S^{(1)}$ is the red-sideband π-swap on ion 1 (part (a)). On the two *internal* qubits (the bus returns to $|0\rangle$),

$$G = \text{diag}\big(e^{-i\gamma},\ 1,\ -1,\ e^{+i\gamma}\big) \quad \text{in the basis } \{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}.$$

Choose single-qubit Z phases

$$Z_c = \text{diag}\big(1,\,e^{-i(\pi+\gamma)}\big), \qquad Z_t = \text{diag}\big(1,\,e^{-i\gamma}\big).$$

Then

$$(Z_c \otimes Z_t)\, G = e^{-i\gamma}\ \mathrm{diag}(1, 1, 1, -1) = e^{-i\gamma}\, \mathrm{CZ},$$

so, up to a global phase,

$$\boxed{\mathrm{CZ}\ =\ (Z_c \otimes Z_t)\ S^{(1)}\, U^{(2)}\, S^{(1)\dagger}}.$$

Now, one can simply perform a basis transform via Hadamard gates to produce the desired CNOT gate $\mathrm{CNOT}_{1\to 2} = (I \otimes H)\, \mathrm{CZ}\, (I \otimes H)$.

□

8 Quantum noise and quantum operations

Exercise 8.1: Unitary evolution as a quantum operation

Pure states evolve under unitary transforms as $|\psi\rangle \mapsto U|\psi\rangle$. Show that, equivalently, we may write $\rho \mapsto \mathcal{E}(\rho) \equiv U\rho U^\dagger$, for $\rho = |\psi\rangle\langle\psi|$.

Solution

Concepts Involved: Unitary Operators, Density Operators

Under unitary evolution, a pure state maps to another pure state via $|\psi\rangle \mapsto U|\psi\rangle$. The density operator corresponding to the final state is then $\rho_{U\psi} = U|\psi\rangle\langle\psi|U^\dagger = U\rho U^\dagger$. We can see that this final state is equivalent to if we started in the density operator formalism $\rho = |\psi\rangle\langle\psi|$ and acted on it with the quantum operation $\mathcal{E}(\rho) = U\rho U^\dagger$. □

Exercise 8.2: Measurement as a quantum operation

Recall from Section 2.2.3 (on page 84) that a quantum measurement with outcomes labeled by m is described by a set of measurement operators M_m such that $\sum_m M_m^\dagger M_m = I$. Let the state of the system immediately before the measurement be ρ. Show that for $\mathcal{E}_m(\rho) \equiv M_m\rho M_m^\dagger$, the state of the system immediately after the measurement is

$$\frac{\mathcal{E}_m(\rho)}{\operatorname{tr}\left(\mathcal{E}_m(\rho)\right)}$$

Also show that the probability of obtaining this measurement result is $p(m) = \operatorname{tr}\left(\mathcal{E}_m(\rho)\right)$.

Solution

Concepts Involved: Quantum Measurement, Density Operators

It suffices to prove the claim for pure states $\rho = |\psi\rangle\langle\psi|$, as mixed states are a linear combination of pure states and $\mathcal{E}_m(\cdot)$ and the trace are both linear in their arguments.
The postulate of quantum measurement tells us that the probability of measuring outcome m given state $|\psi\rangle$ is:

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

For pure $\rho = |\psi\rangle\langle\psi|$ we have:

$$\operatorname{tr}\big(\mathcal{E}_m(\rho)\big) = \operatorname{tr}\Big(M_m\rho M_m^\dagger\Big) = \operatorname{tr}\Big(M_m|\psi\rangle\langle\psi|M_m^\dagger\Big) = \langle\psi| M_m^\dagger M_m|\psi\rangle = p(m)$$

where we have used that $\operatorname{tr}\big(|\alpha\rangle\langle\beta|\big) = \langle\alpha|\beta\rangle$. The postulate of measurement tells us that the post-measurement state (when measuring a pure state) is:

$$|\psi\rangle \to \frac{M_m |\psi\rangle}{\sqrt{\langle\psi| M_m^\dagger M_m |\psi\rangle}}.$$

The density operator corresponding to the post measurement state is:

$$\frac{M_m |\psi\rangle}{\sqrt{\langle\psi| M_m^\dagger M_m |\psi\rangle}} \left(\frac{\langle\psi| M_m^\dagger}{\sqrt{\langle\psi| M_m^\dagger M_m |\psi\rangle}} \right) = \frac{M_m |\psi\rangle\langle\psi| M_m^\dagger}{\langle\psi| M_m^\dagger M_m |\psi\rangle} = \frac{M_m \rho M_m^\dagger}{p(m)} = \frac{\mathcal{E}_m(\rho)}{\mathrm{tr}(\mathcal{E}_m(\rho))}$$

which is what we wished to show. □

Exercise 8.3

Our derivation of the operator-sum representation implicitly assumed that the input and output spaces for the operation were the same. Suppose a composite system AB initially in an unknown quantum state ρ is brought into contact with a composite system CD initially in some standard state $|0\rangle$, and the two systems interact according to a unitary interaction U. After the interaction we discard systems A and D, leaving a state ρ' of system BC. Show that the map $\mathcal{E}(\rho) = \rho'$ satisfies

$$\mathcal{E}(\rho) = \sum_k E_k \rho E_k^\dagger$$

for some set of linear operators E_k from the state space of system AB to the state space of system BC, and such that $\sum_k E_k^\dagger E_k = I$.

Solution

Concepts Involved: Stinespring Representation, Quantum Operations, Density Operators, Partial Trace

---

This problem involves deriving the operator-sum representation for the map $\mathcal{E}(\rho) = \rho'$, where ρ' is the state of subsystems BC after discarding subsystems AD.
Step 1: State evolution via unitary U
Initially, the total state of the system $ABCD$ is

$$\rho_{AB} \otimes |0\rangle_C\langle 0| \otimes |0\rangle_D\langle 0|.$$

After applying the unitary U, the state becomes

$$U(\rho_{AB} \otimes |0\rangle_C\langle 0| \otimes |0\rangle_D\langle 0|)U^\dagger.$$

Step 2: Partial trace over AD
After the interaction, systems A and D are discarded. To describe the remaining state of systems BC, we need to trace out the subsystems AD. The state of BC after tracing out AD is

$$\rho'_{BC} = \mathrm{Tr}_{AD}\left(U(\rho_{AB} \otimes |0\rangle_C\langle 0| \otimes |0\rangle_D\langle 0|)U^\dagger\right).$$

Step 3: Kraus decomposition

Now, we can express this map $\mathcal{E}$ in terms of Kraus operators. To this end, we express

$$\mathcal{E}(\rho_{AB}) = \sum_k E_k \rho_{AB} E_k^\dagger,$$

where the operators E_k are linear operators from the state space of system AB to the state space of system BC.

Specifically, the Kraus operators E_k are given by:

$$E_k = \langle e_k | U | 0 \rangle_D,$$

where $\{|e_k\rangle\}$ is some orthonormal basis for the combined system AD.

Step 4: Trace-preserving condition

Finally, the requirement that $\mathcal{E}$ is trace-preserving implies that the Kraus operators must satisfy

$$\sum_k E_k^\dagger E_k = I.$$

This follows simply from the cyclicity of the trace.

□

Exercise 8.4: Measurement

Suppose we have a single qubit principal system, interacting with a single qubit environment through the transform

$$U = P_0 \otimes I + P_1 \otimes X$$

where X is the usual Pauli matrix (acting on the environment), and $P_0 \equiv |0\rangle\langle 0|$, $P_1 \equiv |1\rangle\langle 1|$ are projectors (acting on the system). Give the quantum operation for this process, in the operator-sum representation, assuming the environment starts in the state $|0\rangle$.

Solution

Concepts Involved: Kraus Representation, Quantum Measurements, Projectors, Partial Trace

Simply using the natural form of the Kraus operators $E_k \equiv \langle e_k | U | e_0 \rangle$, we have:

$$E_0 = \langle 0_E | U | 0_E \rangle = P_0,$$
$$E_1 = \langle 1_E | U | 0_E \rangle = P_1.$$

□

Exercise 8.5: Spin flips

Just as in the previous exercise, but now let

$$U = \frac{X}{\sqrt{2}} \otimes I + \frac{Y}{\sqrt{2}} \otimes X$$

Give the quantum operation for this process, in the operator sum representation.

Solution

Concepts Involved: Kraus Represenations, Partial Trace

Using the same idea,

$$E_0 = \langle 0_E|U|0_E\rangle = \frac{X}{\sqrt{2}},$$
$$E_1 = \langle 1_E|U|0_E\rangle = \frac{Y}{\sqrt{2}}.$$

□

Exercise 8.6: Composition of quantum operations

Suppose $\mathcal{E}$ and $\mathcal{F}$ are quantum operations on the same quantum system. Show that the composition $\mathcal{F} \circ \mathcal{E}$ is a quantum operation, in the sense that is has an operator-sum representation. State and prove an extension of this result to the case where $\mathcal{E}$ and $\mathcal{F}$ do not necessarily have the same input and output spaces.

Solution

Concepts Involved: Krauss Represenation, Linear Maps, Compositions

Step 1: Composition of Two Quantum Operations

For an input state ρ, the action of the composition is:

$$(\mathcal{F} \circ \mathcal{E})(\rho) = \mathcal{F}(\mathcal{E}(\rho)).$$

First, apply $\mathcal{E}(\rho)$ using its Kraus decomposition:

$$\mathcal{E}(\rho) = \sum_k E_k \rho E_k^\dagger.$$

Next, apply $\mathcal{F}$ to the result of $\mathcal{E}(\rho)$:

$$\mathcal{F}(\mathcal{E}(\rho)) = \mathcal{F}\left(\sum_k E_k \rho E_k^\dagger\right) = \sum_k \mathcal{F}(E_k \rho E_k^\dagger).$$

Since $\mathcal{F}$ is linear, we can apply it to each term in the sum separately: $\mathcal{F}(E_k \rho E_k^\dagger) = \sum_j F_j E_k \rho E_k^\dagger F_j^\dagger$.
Thus, the composition $\mathcal{F} \circ \mathcal{E}$ can be written as:

$$(\mathcal{F} \circ \mathcal{E})(\rho) = \sum_{j,k} F_j E_k \rho E_k^\dagger F_j^\dagger.$$

This expression is again in the form of an operator-sum representation, where the Kraus operators for the composed operation $\mathcal{F} \circ \mathcal{E}$ are given by the products $F_j E_k$. Hence, we have shown that the composition of two quantum operations is itself a quantum operation, with Kraus operators $F_j E_k$.

Step 2: Trace-Preserving Condition

To check that the map $\mathcal{F} \circ \mathcal{E}$ is trace-preserving, we compute:

$$\sum_{j,k} (F_j E_k)^\dagger (F_j E_k) = \sum_{j,k} E_k^\dagger F_j^\dagger F_j E_k.$$

Since $\mathcal{F}$ is trace-preserving, we know that $\sum_j F_j^\dagger F_j = I$, so this simplifies to:

$$\sum_k E_k^\dagger I E_k = \sum_k E_k^\dagger E_k = I,$$

Step 3: Extension to Different Input and Output Spaces

Suppose that:

- $\mathcal{E}$ maps states from the Hilbert space $\mathcal{H}_A$ to $\mathcal{H}_B$,
- $\mathcal{F}$ maps states from $\mathcal{H}_B$ to $\mathcal{H}_C$.

We can still write the composition $\mathcal{F} \circ \mathcal{E}$, but now $\mathcal{E}$ will have Kraus operators $E_k : \mathcal{H}_A \to \mathcal{H}_B$ and $\mathcal{F}$ will have Kraus operators $F_j : \mathcal{H}_B \to \mathcal{H}_C$.
The composition will act as:

$$(\mathcal{F} \circ \mathcal{E})(\rho) = \sum_{j,k} F_j E_k \rho E_k^\dagger F_j^\dagger,$$

where ρ is a state on $\mathcal{H}_A$, and the result will be a state on $\mathcal{H}_C$. The Kraus operators for the composed map are still given by the products $F_j E_k$, which now act from $\mathcal{H}_A$ to $\mathcal{H}_C$.

□

Exercise 8.7

Suppose that instead of doing a projective measurement on the combined principal system and environment we had performed a general measurement described by measurement operators $\{M_m\}$. Find operator-sum representations for the corresponding quantum operations $\mathcal{E}_m$ on the principal system, and show that the respective measurement probabilities are $\mathrm{tr}\big[\mathcal{E}(\rho)\big]$.

Solution

Concepts Involved: Kraus Representation, Quantum Measurements

---

Let the principal system S start in ρ and the environment E in $|0\rangle_E$. A joint unitary U acts on $S \otimes E$, followed by a general measurement on SE with measurement operators $\{M_m\}$ satisfying $\sum_m M_m^\dagger M_m = I_{SE}$.
The (unnormalized) conditional post-measurement state of S given outcome m is

$$\mathcal{E}_m(\rho) = \mathrm{Tr}_E\Big[M_m\, U\big(\rho \otimes |0\rangle\langle 0|\big)U^\dagger M_m^\dagger\Big].$$

Fix an orthonormal basis $\{|\alpha\rangle\}$ of E and define Kraus operators on S by

$$K_{m,\alpha} \;:=\; \langle\alpha|\, M_m\, U\, |0\rangle \in \mathcal{L}(S).$$

Then $\mathcal{E}_m$ has the operator-sum form

$$\boxed{\mathcal{E}_m(\rho) = \sum_\alpha K_{m,\alpha}\, \rho\, K_{m,\alpha}^\dagger}.$$

Moreover,

$$\sum_{m,\alpha} K_{m,\alpha}^\dagger K_{m,\alpha} = \langle 0|\, U^\dagger\Big(\sum_m M_m^\dagger M_m\Big)U\, |0\rangle = \langle 0|\, U^\dagger U\, |0\rangle = I_S,$$

so $\sum_m \mathcal{E}_m$ is trace preserving and each $\mathcal{E}_m$ is completely positive.
The probability of outcome m is

$$\boxed{p(m) = \mathrm{Tr}\big[\mathcal{E}_m(\rho)\big]} = \mathrm{Tr}\Big[M_m\, U\big(\rho \otimes |0\rangle\langle 0|\big)U^\dagger M_m^\dagger\Big],$$

and equivalently $p(m) = \mathrm{Tr}(F_m \rho)$ with POVM elements $F_m := \sum_\alpha K_{m,\alpha}^\dagger K_{m,\alpha} \geq 0$ and $\sum_m F_m = I_S$.
(If E starts in a mixed state $\sigma_E = \sum_j \lambda_j |e_j\rangle\langle e_j|$, take $K_{m,\alpha j} := \sqrt{\lambda_j}\, \langle\alpha| M_m U |e_j\rangle$ and sum over α, j.)

□

Exercise 8.8: Non-trace-preserving quantum operations

Explain how to construct a unitary operator for a system–environment model of a non-trace-preserving quantum operation, by introducing an extra operator, E_∞, into the set of operation elements E_k, chosen so that when summing over the complete set of k, including $k = \infty$, one obtains $\sum_k E_k^\dagger E_k = I$.

Solution

Concepts Involved: Kraus Representation, Completeness

Non trace preserving quantum operators are characterized by the relation

$$\sum_k E_k^\dagger E_k \leq I.$$

To embed this operation into a unitary evolution, we introduce an auxiliary operator E_∞ such that

$$\sum_k E_k^\dagger E_k + E_\infty^\dagger E_\infty = I,$$

where E_∞ is defined as

$$E_\infty := \sqrt{I - \sum_k E_k^\dagger E_k}.$$

To construct a unitary U, introduce an environment Hilbert space with an orthonormal basis $\{|e_k\rangle\}$ and define

$$U|\psi\rangle|e_0\rangle = \sum_k (E_k|\psi\rangle) \otimes |e_k\rangle + (E_\infty|\psi\rangle) \otimes |e_\infty\rangle.$$

The unitarity condition $U^\dagger U = I$ follows from the completeness relation of the extended Krauss operators, ensuring that the process remains reversible in the extended system-environment space. □

Exercise 8.9: Measurement model

If we are given a set of quantum operations $\{\mathcal{E}_m\}$ such that $\sum_m \mathcal{E}_m$ is trace-preserving, then it is possible to construct a *measurement model* giving rise to this set of quantum operations. For each m, let E_{mk} be a set of operation elements for $\mathcal{E}_m$. Introduce an environmental system, E, with an orthonormal basis $|m, k\rangle$ in one-to-one correspondence with the set of indices for the operation elements. Analogously to the earlier construction, define an operator U such that

$$U|\psi\rangle|e_0\rangle = \sum_{m,k} E_{mk}|\psi\rangle|m, k\rangle.$$

Next, define projectors $P_m \equiv \sum_k |m, k\rangle\langle m, k|$ on the environmental system, E. Show that performing U on $\rho \otimes |e_0\rangle\langle e_0|$, then measuring P_m gives m with probability $\operatorname{tr}\big(\mathcal{E}_m(\rho)\big)$, and the corresponding post-measurement state of the principal system is $\mathcal{E}_m(\rho)/\operatorname{tr}\big(\mathcal{E}_m(\rho)\big)$.

Solution

Concepts Involved: Kraus Representation, Projectors, Quantum Measurement

Performing U we have:

$$\rho \otimes |e_0\rangle\langle e_0| \mapsto U(\rho \otimes |e_0\rangle\langle e_0|)U^\dagger = \sum_{n,k}\sum_{n',k'}(E_{nk}\rho E^\dagger_{n'k'} \otimes |n,k\rangle\langle n,k'|)$$

Now measuring P_m, the post-measurement state of the entire system is given by:

$$\begin{aligned} P_m U(\rho \otimes |e_0\rangle\langle e_0|)U^\dagger P_m &= (\sum_l |m,l\rangle\langle m,l|)(\sum_{n,k}\sum_{n,k'} E_{mk}\rho E^\dagger_{mk'} \otimes |n,k\rangle\langle n',k'|)(\sum_l |m,l\rangle\langle m,l|) \\ &= \sum_{k,k'} E_{mk}\rho E^\dagger_{mk'} \otimes |m,k\rangle\langle m,k'| \end{aligned}$$

where we have used the orthogonality of the $|m,k\rangle$ states.
Tracing out the environmental system yields the (unnormalized) post-measurement state on the principal system:

$$\begin{aligned} \mathrm{Tr}_E(P_m U(\rho \otimes |e_0\rangle\langle e_0|)U^\dagger P_m) &= \mathrm{Tr}_E(\sum_{k,k'} E_{mk}\rho E^\dagger_{mk'} \otimes |m,k\rangle\langle m,k'|) \\ &= \sum_k E_{mk}\rho E^\dagger_{mk} \\ &= \mathcal{E}_m(\rho) \end{aligned}$$

where the trace only picks out the diagonal elements in the sum. To get the probability of this measurement outcome, we take the trace of the above (noting that $\mathrm{Tr} = \mathrm{Tr}\,\mathrm{Tr}_E$):

$$\begin{aligned} p(m) &= \mathrm{Tr}\Big(P_m U(\rho \otimes |e_0\rangle\langle e_0|)U^\dagger P_m\Big) \\ &= \mathrm{Tr}\Big(\mathrm{Tr}_E(P_m U(\rho \otimes |e_0\rangle\langle e_0|)U^\dagger P_m)\Big) \\ &= \mathrm{Tr}\big(\mathcal{E}_m(\rho)\big) \end{aligned}$$

And hence the normalized post-measurement state of the principal system is:

$$\mathcal{E}_m(\rho)/p(m) = \mathcal{E}_m(\rho)/\,\mathrm{Tr}\big(\mathcal{E}_m(\rho)\big)$$

□

Exercise 8.10

Give a proof of Theorem 8.3 based on the freedom in the operator-sum representation, as follows. Let $\{E_j\}$ be a set of operation elements for $\mathcal{E}$. Define a matrix $W_{jk} \equiv \mathrm{tr}\Big(E_j^\dagger E_k\Big)$. Show that the matrix W is Hermitian and of rank at most d^2, and thus there is a unitary matrix u such that $uWu^\dagger$ is diagonal with at most d^2 non-zero entries. Use u to define a new set of at most d^2 non-zero operation elements $\{F_j\}$ for $\mathcal{E}$.

Solution

Concepts Involved: Kraus Representation, Unitary Operators

Let $\mathcal{E}(\rho) = \sum_j E_j \rho E_j^\dagger$ be a quantum channel with operation elements $\{E_j\}$. Define the Hermitian matrix

$$W_{jk} = \text{Tr}(E_j^\dagger E_k).$$

Since W is constructed from inner products, it satisfies

$$W_{jk}^* = \text{Tr}((E_j^\dagger E_k)^*) = \text{Tr}(E_k^\dagger E_j) = W_{kj},$$

proving that W is Hermitian.

Since each E_j is a $d \times d$ matrix, the space of all such matrices has dimension at most d^2. Thus, the rank of W is at most d^2,

$$\text{rank}(W) \leq d^2.$$

Since W is Hermitian, there exists a unitary matrix u such that

$$uWu^\dagger = \Lambda,$$

where Λ is diagonal with at most d^2 nonzero entries.

Define new operation elements

$$F_j = \sum_k u_{jk} E_k.$$

Since u is unitary,

$$E(\rho) = \sum_j E_j \rho E_j^\dagger = \sum_j F_j \rho F_j^\dagger.$$

Thus, the same quantum channel can be represented using at most d^2 nonzero operators $\{F_j\}$. $\square$

Exercise 8.11

Suppose $\mathcal{E}$ is a quantum operation mapping a d-dimensional input space to a d'-dimensional output space. Show that $\mathcal{E}$ can be described using a set of at most dd' operation elements $\{E_k\}$.

Solution

Concepts Involved: Kraus Representation

Again we let Let $\mathcal{E}(\rho) = \sum_j E_j \rho E_j^\dagger$ be a quantum channel with operation elements $\{E_j\}$, where each E_j is now a $d' \times d$ matrix. Now let $D = \max(d, d')$ and construct the operation elements $\left\{E_j'\right\}$ where each E_j' is a $D \times D$ matrix, having padded the deficient rows/columns with zeros. We can then construct $W_{ij} = \text{Tr}\left(E_j'^\dagger E_k'\right)$ as in the last problem, wherein W now satisfies $\text{rank}(W) \leq dd'$ as each E_j' has only at most dd' nonzero entries. Following the previous exercise, we can construct dd' nonzero operators $\left\{F'_j\right\}$ (which are $D \times D$ matrices) which we can then trim back off the zero row/columns to get back to $d' \times d$ operators $\left\{F_j\right\}$. □

Exercise 8.12

Why can we assume that O has determinant 1 in the decomposition (8.93)?

Solution

Concepts Involved: Affine Maps, Orthogonal Matrices, Polar Decomposition

By construction the O appearing in $M = OS$ is orthogonal, which means $\det(O) = \pm 1$. In the case that $\det(O) = -1$, we may write $M = (-O)(-S) = O'S'$ in which $O' = -O$ is real, orthogonal, and has $\det(O') = 1$ and $S' = -S$ is still symmetric and real. Hence we may assume that O has determinant 1. □

Remark: In the polar decomposition discussed in Chapter 2.1.10, we write $M = UJ$ where U is unitary J is positive, not just real and symmetric. Here we relax the positivity condition so that we may absorb a potential negative sign from U/O.

Exercise 8.13

Show that unitary transformations correspond to rotations of the Bloch sphere.

Solution

Concepts Involved: Unitary Operators, Rotations, Bloch Sphere

Any single qubit density matrix can be expressed as $\rho = \frac{1}{2}(I + \mathbf{r} \cdot \boldsymbol{\sigma})$, with $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ and $\mathbf{r} \in \mathbb{R}^3$. Let the unitary be the axis–angle element

$$U = e^{-i\frac{\theta}{2}\hat{n}\cdot\boldsymbol{\sigma}} = c\,I - is\,(\hat{n}\cdot\boldsymbol{\sigma}), \qquad c = \cos\frac{\theta}{2}, \quad s = \sin\frac{\theta}{2},$$

where $\|\hat{n}\| = 1$. The evolved state is $\rho' = U\rho U^\dagger = \frac{1}{2}\big(I + \mathbf{r}'\cdot\boldsymbol{\sigma}\big)$ with

$$\mathbf{r}'\cdot\boldsymbol{\sigma} = U\,(\mathbf{r}\cdot\boldsymbol{\sigma})\,U^\dagger.$$

Using the Pauli identity $(\boldsymbol{\sigma}\cdot a)(\boldsymbol{\sigma}\cdot b) = (a\cdot b)\,I + i\,\boldsymbol{\sigma}\cdot(a\times b)$, set $A = \boldsymbol{\sigma}\cdot\mathbf{r}$, $B = \hat{n}\cdot\boldsymbol{\sigma}$, we compute

$$UAU^\dagger = (cI - isB)\,A\,(cI + isB) = c^2 A + ics[A,B] + s^2 BAB.$$

From the product rule,

$$[A,B] = 2i\,\boldsymbol{\sigma}\cdot(\mathbf{r}\times\hat{n}) = -2i\,\boldsymbol{\sigma}\cdot(\hat{n}\times\mathbf{r}), \qquad BAB = -\,\boldsymbol{\sigma}\cdot\mathbf{r} + 2(\hat{n}\cdot\mathbf{r})\,\boldsymbol{\sigma}\cdot\hat{n}.$$

Inserting this into the expression of interest and using $c^2 - s^2 = \cos\theta$, $2cs = \sin\theta$, $2s^2 = 1 - \cos\theta$, we have

$$\begin{aligned} U(\boldsymbol{\sigma}\cdot\mathbf{r})U^\dagger &= (c^2 - s^2)\,\boldsymbol{\sigma}\cdot\mathbf{r} + 2cs\,\boldsymbol{\sigma}\cdot(\hat{n}\times\mathbf{r}) + 2s^2(\hat{n}\cdot\mathbf{r})\,\boldsymbol{\sigma}\cdot\hat{n} \\ &= \boldsymbol{\sigma}\cdot\Big(\mathbf{r}\cos\theta + (\hat{n}\times\mathbf{r})\sin\theta + \hat{n}(\hat{n}\cdot\mathbf{r})(1-\cos\theta)\Big). \end{aligned}$$

Therefore the Bloch vector rotates as

$$\boxed{\mathbf{r}' = \mathbf{r}\cos\theta + (\hat{n}\times\mathbf{r})\sin\theta + \hat{n}(\hat{n}\cdot\mathbf{r})\,(1-\cos\theta)}$$

i.e. Rodrigues' formula for a rotation by angle θ about axis $\hat{n}$.
Hence U acts as a rotation on the Bloch sphere. □

Exercise 8.14

Show that $\det(S)$ need not be positive.

Solution

Concepts Involved: Affine Maps, Symmetric Matrices, Polar Decomposition, Bloch Sphere

---

It suffices to provide an example. Consider $M = \text{diag}(1, 1, -1)$, corresponding to a reflection $z \leftrightarrow -z$ about the xy plane of the Bloch sphere. Evidently this maps the Bloch sphere to itself. However, we note that $\det(M) = -1$ and thus $\det(M) = \det(OS) = \det(O)\det(S) = \det(S) = -1$ and hence $\det(S)$ is negative in this case. Generally, any map involving a reflection will result in $\det(S) < 0$ (since we absorb the negative signs arising from reflections that were in O into S as per Ex. **??**). □

Exercise 8.15

Suppose a projective measurement is performed on a single qubit in the basis $|+\rangle, |-\rangle$, where $|\pm\rangle \equiv (|0\rangle \pm |1\rangle)/\sqrt{2}$. In the event that we are ignorant of the result of the measurement, the density matrix evolves according to the equation

$$\rho \mapsto \mathcal{E}(\rho) = |+\rangle\langle+|\,\rho\,|+\rangle\langle+| \;+\; |-\rangle\langle-|\,\rho\,|-\rangle\langle-|$$

Illustrate this transformation on the Bloch sphere.

Solution

Concepts Involved: Density Operators, Quantum Measurement, Bloch Sphere

---

Using the standard pauli projectors, $|\pm\rangle\langle\pm| = \frac{1}{2}(I \pm X)$, we have

$$\mathcal{E}(\rho) = |+\rangle\langle+|\rho|+\rangle\langle+| + |-\rangle\langle-|\rho|-\rangle\langle-| = \tfrac{1}{2}\big(\rho + X\rho X\big).$$

Now, we note that $\rho = \frac{1}{2}(I + \mathbf{r}\cdot\boldsymbol{\sigma})$ with $\mathbf{r} = (r_x, r_y, r_z)$. Since $X\sigma_x X = \sigma_x$, $X\sigma_y X = -\sigma_y$, $X\sigma_z X = -\sigma_z$,

$$\mathcal{E}(\rho) = \tfrac{1}{2}(I + r_x\sigma_x) \quad\Rightarrow\quad \mathbf{r}' = (r_x,\, 0,\, 0).$$

Thus the nonselective measurement in the $\{|+\rangle, |-\rangle\}$ basis acts as *complete dephasing in the X basis*: the Bloch vector is orthogonally projected onto the x-axis (only r_x survives, r_y and r_z vanish). □

Exercise 8.16

The graphical method for understanding single qubit quantum operations was derived for trace-preserving quantum operations. Find an explicit example of a non-trace-preserving quantum operation which cannot be described as a deformation of the Bloch sphere, followed by a rotation and a displacement.

Solution

Concepts Involved: Density Operators, Bloch Sphere

Consider the filter

$$\mathcal{E}(\rho) = K\rho K^\dagger, \qquad K = |0\rangle\langle 0| + \lambda\, |1\rangle\langle 1|\,, \quad 0 < \lambda < 1.$$

Then $K^\dagger K = |0\rangle\langle 0| + \lambda^2\, |1\rangle\langle 1| \leq I$, so $\mathcal{E}$ is completely positive and trace-non-increasing . Write $\rho = \frac{1}{2}(I + x\sigma_x + y\sigma_y + z\sigma_z)$, so

$$\mathcal{E}(\rho) = \frac{1}{2}\begin{pmatrix} 1+z & \lambda(x-iy) \\ \lambda(x+iy) & \lambda^2(1-z) \end{pmatrix}, \qquad p(x,y,z) = \mathrm{tr}\big[\mathcal{E}(\rho)\big] = \frac{(1+\lambda^2) + (1-\lambda^2)z}{2}.$$

Conditioned on the outcome (i.e. after renormalization $\rho' = \mathcal{E}(\rho)/p$), the output Bloch vector $r' = (x', y', z')$ is

$$x' = \frac{\lambda x}{p}, \qquad y' = \frac{\lambda y}{p}, \qquad z' = \frac{(1-\lambda^2) + (1+\lambda^2)z}{(1+\lambda^2) + (1-\lambda^2)z}. \tag{$\star$}$$

This dependence on p (which itself depends on z) makes $r \mapsto r'$ *fractional-linear*, not affine. But any "deformation of the Bloch sphere, followed by a rotation and a displacement" is an *affine* map $r' = Ar + t$. A concrete convexity check shows the failure of affinity. Let $\rho_1 = |0\rangle\langle 0|$ ($r_1 = (0,0,1)$), $\rho_2 = |1\rangle\langle 1|$ ($r_2 = (0,0,-1)$), and $\bar{\rho} = \frac{I}{2}$ ($\bar{r} = 0$). From ($\star$):

$$r_1' = (0,0,1), \qquad r_2' = (0,0,-1), \qquad \bar{r}' = \Big(0, 0, \frac{1-\lambda^2}{1+\lambda^2}\Big).$$

Yet $\frac{1}{2}(r_1' + r_2') = (0,0,0) \neq \bar{r}'$. Therefore this non–trace-preserving operation cannot be represented by the Bloch-sphere graphical recipe (affine transform).

□

Exercise 8.17

Verify (8.101) as follows. Define

$$\mathcal{E}(A) \equiv \frac{A + XAX + YAY + ZAZ}{4}$$

and show that

$$\mathcal{E}(I) = I;\ \mathcal{E}(X) = \mathcal{E}(Y) = \mathcal{E}(Z) = 0$$

Now use the Bloch sphere representation for single qubit density matrices to verify (8.101).

Solution

Concepts Involved: Depolarizing Channel, Bloch Sphere

It is straightforward to check

$$\mathcal{E}(X) \equiv \frac{X + XXX + YXY + ZXZ}{4} = \frac{X + X + YXY + ZXZ}{4} = \frac{X + X - X - X}{4} = 0$$

Similarly, it follows $\mathcal{E}(Y) = \mathcal{E}(Z) = 0$.Now, we simply employ linearity and the single qubit density matrix representation to arrive at

$$\mathcal{E}(\rho) = \mathcal{E}\left(\frac{I + r.\sigma}{2}\right) = \mathcal{E}\left(\frac{I}{2}\right) + 0$$
$$\implies \frac{\rho + X\rho X + Y\rho Y + Z\rho Z}{4} = \frac{I}{2}$$

which we set out to prove. □

Exercise 8.18

For $k \geq 1$ show that $\mathrm{tr}\left(\rho^k\right)$ is never increased by the action of the depolarizing channel.

Solution

Concepts Involved: Trace, Depolarizing Channel, Binomial expansion

$$\mathcal{E}_p(\rho) = (1-p)\rho + p\frac{I}{d}$$

$$[\mathcal{E}_p(\rho)]^k = \left((1-p)\rho + p\frac{I}{d}\right)^k$$

Now taking the trace, we have

$$\begin{aligned}
\mathrm{Tr}\left([\mathcal{E}_p(\rho)]^k\right) &= \sum_{m=0}^{k} \binom{k}{m} (1-p)^{k-m} p^m \,\mathrm{Tr}\left(\rho^{k-m}\right) \left(\frac{I}{d}\right)^m \\
&\leq \mathrm{Tr}\left(\rho^k\right) \sum_{m=0}^{k} \binom{k}{m} (1-p)^{k-m} p^m \left(\frac{1}{d}\right)^m && \text{as } \mathrm{Tr}(\rho^k) \geq \mathrm{Tr}(\rho^{k-m})\ \forall m \geq 0 \\
&= \mathrm{Tr}\left(\rho^k\right) (1-p+\frac{p}{d})^k \\
&\leq \mathrm{Tr}\left(\rho^k\right) (1-p+p)^k && \forall\, d \geq 1 \\
&= \mathrm{Tr}\left(\rho^k\right)
\end{aligned}$$

□

Exercise 8.19

Find an operator-sum representation for a generalized depolarizing channel acting on a d-dimensional Hilbert space.

Solution

Concepts Involved: Kraus Representation, Depolarizing Channel

---

The generalized depolarizing channel $\mathcal{E}$ on a d-dimensional Hilbert space is defined as

$$\mathcal{E}(\rho) = (1-p)\rho + \frac{p}{d}I,$$

where $0 \leq p \leq 1$ is the depolarizing probability and I is the identity operator in the d-dimensional space. The operator-sum (Kraus) representation for this channel is

$$\begin{aligned}
E_0 &= \sqrt{1-p}\, I \\
E_{i,j} &= \sqrt{\frac{p}{d}}\, |i\rangle\langle j| \qquad i,j = 1, \ldots, d.
\end{aligned}$$

□

Exercise 8.20

Show that the circuit in Figure 8.13 models the amplitude damping quantum operation, with $\sin^2(\theta/2) = \gamma$.

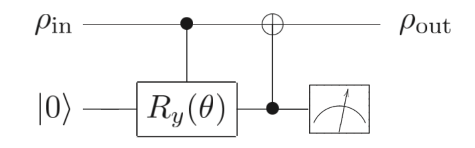


Figure 8.13. Circuit model for amplitude damping

Solution

Concepts Involved: Kraus Represnetation, Amplitude Damping Channel

It is instructive to write the unitary interaction as

$$\begin{aligned} U &= (I \otimes P_0 + X \otimes P_1)(P_0 \otimes I + P_1 \otimes R_y(\theta)) \\ &= P_0 \otimes P_0 + P_1 \otimes P_0 R_y(\theta) + X P_0 \otimes P_1 + X P_1 \otimes P_1 R_y(\theta) \end{aligned}$$

Thus, the Kraus operators can be calculated using

$$\begin{aligned} E_0 &\equiv \langle 0_E|U|0_E\rangle = P_0\langle 0_E|P_0|0_E\rangle + P_1\langle 0_E|P_0 R_y(\theta)|0_E\rangle = P_0 + P_1 . \cos\big(\theta/2\big) \\ E_1 &\equiv \langle 1_E|U|0_E\rangle = XP_1\langle 1_E|P_1 R_y(\theta)|0_E\rangle = XP_1 . \sin\big(\theta/2\big) \end{aligned}$$

To match this with the matrix representation in Eq. (8.108), we should have

$$\begin{aligned} \cos\big(\theta/2\big) &= \sqrt{1-\gamma} \\ \implies \gamma &= \sin^2(\theta/2). \end{aligned}$$

□

Exercise 8.21: Amplitude damping of a harmonic oscillator

Suppose that our principal system, a harmonic oscillator, interacts with an environment, modeled as another harmonic oscillator, through the Hamiltonian

$$H = \chi(a^\dagger b + b^\dagger a)$$

where a and b are the annihilation operators for the respective harmonic oscillators, as defined in Section 7.3.

(1) Using $U = \exp(-iH\Delta t)$, denoting the eigenstates of $b^\dagger b$ at $|k_b\rangle$, and selecting the vacuum state $|0_b\rangle$ as the initial state of the environment, show that the operation elements $E_k = \langle k_b|U|0_b\rangle$ are found to be

$$E_k = \sum_n \sqrt{\binom{n}{k}}\sqrt{(1-\gamma)^{n-k}\gamma^k}\,|n-k\rangle\langle n|$$

where $\gamma = 1 - \cos^2(\chi\Delta t)$ is the probability of losing a single quantum of energy, and states such as $|n\rangle$ are eigenstates of $a^\dagger a$.

(2) Show that the operation elements E_k define a trace-preserving quantum operation.

Solution

Concepts Involved: Creation/Annihilation Operators, Amplitude Daming Channel, Kraus Representation

---

Let $\theta \equiv \chi\,\Delta t$ and $U = e^{-iH\Delta t}$ with $H = \chi(a^\dagger b + b^\dagger a)$. The Heisenberg evolution gives the standard beam-splitter mixing

$$U^\dagger a\,U = a\cos\theta - i\,b\sin\theta, \qquad U^\dagger b\,U = b\cos\theta - i\,a\sin\theta.$$

Since $U\,|0_a, 0_b\rangle = |0_a, 0_b\rangle$, for a system number state $|n\rangle \equiv |n_a\rangle$,

$$U\,|n,0\rangle_{ab} = \frac{(Ua^\dagger U^\dagger)^n}{\sqrt{n!}}\,|0,0\rangle = \frac{(a^\dagger\cos\theta - i\,b^\dagger\sin\theta)^n}{\sqrt{n!}}\,|0,0\rangle\,.$$

Expanding binomially and using $a^{\dagger\,n-k}b^{\dagger\,k}\,|0,0\rangle = \sqrt{(n-k)!\,k!}\,|n-k,k\rangle$ yields

$$U\,|n,0\rangle = \sum_{k=0}^{n}\sqrt{\binom{n}{k}}\,(\cos\theta)^{\,n-k}\,(-i\sin\theta)^{\,k}\,\,|n-k\rangle_a \otimes |k\rangle_b\,.$$

Projecting the environment onto $\langle k_b|$ defines the Kraus action $E_k = \langle k_b|\,U\,|0_b\rangle$:

$$E_k\,|n\rangle = \sqrt{\binom{n}{k}}\,(\cos\theta)^{\,n-k}\,(-i\sin\theta)^{\,k}\,\,|n-k\rangle \qquad (0 \le k \le n),$$

and $E_k\,|n\rangle = 0$ for $k > n$. Absorbing the phase $(-i)^k$ into a redefinition of the environment basis, and writing $\gamma \equiv \sin^2\theta = 1 - \cos^2\theta$, we obtain

$$E_k = \sum_{n=k}^{\infty} \sqrt{\binom{n}{k}} \sqrt{(1-\gamma)^{\,n-k}\,\gamma^{\,k}}\; |n-k\rangle\langle n| \,.$$

To show trace preservation, compute for any number state $|n\rangle$:

$$E_k^\dagger E_k \,|n\rangle = \binom{n}{k}(1-\gamma)^{\,n-k}\gamma^{\,k}\,|n\rangle \quad (0 \le k \le n),$$

hence, by the binomial theorem,

$$\Big(\sum_{k=0}^{\infty} E_k^\dagger E_k\Big)\,|n\rangle = \sum_{k=0}^{n}\binom{n}{k}(1-\gamma)^{\,n-k}\gamma^{\,k}\,|n\rangle = \big[(1-\gamma)+\gamma\big]^{n}\,|n\rangle = |n\rangle \,.$$

Therefore $\sum_k E_k^\dagger E_k = \mathbb{I}$, and the map is trace-preserving. □

Exercise 8.22: Amplitude damping of single qubit density matrix

For the general single qubit state

$$\rho = \begin{bmatrix} a & b \\ b^* & c \end{bmatrix}$$

show that the amplitude damping leads to

$$\mathcal{E}_{AD}(\rho) = \begin{bmatrix} 1-(1-\gamma)(1-a) & b\sqrt{1-\gamma} \\ b^*\sqrt{1-\gamma} & c(1-\gamma) \end{bmatrix}$$

Solution

Concepts Involved: Amplitude Damping Channel, Density Operators

For the general single qubit state

$$\rho = \begin{pmatrix} a & b \\ b^* & c \end{pmatrix}, \quad a + c = 1,$$

the amplitude damping channel is

$$\mathcal{E}_{AD}(\rho) = K_0\rho K_0^\dagger + K_1\rho K_1^\dagger,$$

where the Kraus operators are

$$K_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{pmatrix}, \quad K_1 = \begin{pmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{pmatrix}.$$

Now, applying the Kraus operators

$$K_0 \rho K_0^\dagger = \begin{pmatrix} a & b\sqrt{1-\gamma} \\ b^*\sqrt{1-\gamma} & c(1-\gamma) \end{pmatrix},$$
$$K_1 \rho K_1^\dagger = \begin{pmatrix} \gamma c & 0 \\ 0 & 0 \end{pmatrix}.$$

Summing the two results

$$\mathcal{E}_{AD}(\rho) = \begin{pmatrix} a + \gamma(1-a) & b\sqrt{1-\gamma} \\ b^*\sqrt{1-\gamma} & c(1-\gamma) \end{pmatrix}.$$

Since $a + c = 1$, we arrive at the desired form

$$\mathcal{E}_{AD}(\rho) = \begin{pmatrix} 1 - (1-\gamma)(1-a) & b\sqrt{1-\gamma} \\ b^*\sqrt{1-\gamma} & c(1-\gamma) \end{pmatrix}.$$

□

Exercise 8.23: Amplitude damping of dual-rail qubits

Suppose that a single qubit state is represented by using two qubits, as

$$|\psi\rangle = a|01\rangle + b|10\rangle.$$

Show that $\mathcal{E}_{AD} \otimes \mathcal{E}_{AD}$ applied to this state gives a process which can be described by the operation elements

$$E_0^{\text{dr}} = \sqrt{1-\gamma}\, I$$
$$E_1^{\text{dr}} = \sqrt{\gamma}\left[|00\rangle\langle 01| + |00\rangle\langle 10|\right]$$

that is, either nothing (E_0^{dr}) happens to the qubit, or the qubit is transformed (E_1^{dr}) into the state $|00\rangle$, which is orthogonal to $|\psi\rangle$. This is a simple error-detection code, and is also the basis for the robustness of the 'dual-rail' qubit discussed in Section 7.4.

Solution

Concepts Involved: Amplitude Damping Channel, Dual-Rail Representation

The amplitude damping channel $\mathcal{E}_{AD}$ is described by the Kraus operators

$$K_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{pmatrix}, \qquad K_1 = \begin{pmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{pmatrix}.$$

The action of $\mathcal{E}_{AD} \otimes \mathcal{E}_{AD}$ on $|\psi\rangle$ results in

$$\begin{aligned} K_0 \otimes K_0 |\psi\rangle &= \sqrt{1-\gamma}\, |\psi\rangle, \\ K_1 \otimes K_0 |\psi\rangle &= a\sqrt{\gamma}\, |00\rangle\,, \\ K_0 \otimes K_1 |\psi\rangle &= b\sqrt{\gamma}\, |00\rangle\,, \\ K_1 \otimes K_1 |\psi\rangle &= 0. \end{aligned}$$

Thus, the total channel action is given by

$$\begin{aligned} \mathcal{E}_{AD} \otimes \mathcal{E}_{AD}(|\psi\rangle\langle\psi|) &= (1-\gamma)\, |\psi\rangle\langle\psi| + |a|^2 \gamma\, |00\rangle\langle 00| + |b|^2 \gamma\, |00\rangle\langle 00| \\ &= (1-\gamma)\, |\psi\rangle\langle\psi| + \gamma\, |00\rangle\langle 00| \end{aligned}$$

This process can be described by the operation elements

$$\begin{aligned} E_0^{\text{dr}} &= \sqrt{1-\gamma} I, \\ E_1^{\text{dr}} &= \sqrt{\gamma}\left(|00\rangle\langle 01| + |00\rangle\langle 10|\right). \end{aligned}$$

Hence, $\mathcal{E}_{AD} \otimes \mathcal{E}_{AD}$ either applies no change to the qubits (E_0^{dr}) or transforms the state into $|00\rangle$ (E_1^{dr}).

□

Exercise 8.24: Spontaneous emission is amplitude damping

A single atom coupled to a single mode of electromagnetic radiation undergoes spontaneous emission, as was described in Section 7.6.1. To see that this process is just amplitude damping, take the unitary operation resulting from the Jaynes–Cummings interaction, Equation (7.77), with detuning $\delta = 0$, and give the quantum operation resulting from tracing over the field.

Solution

Concepts Involved: Kraus Representation, Amplitude Damping Channel, JC Model

Eq. (7.77) with $\delta = 0$ is given by:

$$U = |00\rangle\langle 00| + \cos(gt)(|01\rangle\langle 01| + |10\rangle\langle 10|) - i\sin(gt)(|01\rangle\langle 10| + |10\rangle\langle 01|)$$

So the quantum operation on the atom obtained by tracing over the field (the first qubit) is given by:

$$E_0 = \langle 0|_F\, U\, |0\rangle_F = |0\rangle\langle 0| + \cos(gt)\, |1\rangle\langle 1|$$

$$E_1 = \langle 1|_F\, U\, |0\rangle_F = -i\sin(gt)\, |0\rangle\langle 1|\,.$$

When we look at the action of this quantum operation, since $E_1^\dagger = i\sin(gt)\,|1\rangle\langle 0|$ and the operation elements act on ρ via conjugation, we can discard the phase factor (as it drops out) and so:

$$E_0 = |0\rangle\langle 0| + \cos(gt)\,|1\rangle\langle 1| \cong \begin{bmatrix} 1 & 0 \\ 0 & \cos(gt) \end{bmatrix}$$
$$E_1 = \sin(gt)\,|01\rangle\langle 01| \cong \begin{bmatrix} 0 & \sin(gt) \\ 0 & 0 \end{bmatrix}$$

which are the operation elements for amplitude damping with $\gamma = \sin^2(gt)$. □

Exercise 8.25

If we define the temperature T of a qubit by assuming that in equilibrium the probabilities of being in the $|0\rangle$ or $|1\rangle$ states satisfy a Boltzmann distribution, that is $r_0 = e^{-E_0/k_B t}/\mathcal{Z}$ and $p_1 = e^{-E_1/k_B T}/\mathcal{Z}$, where E_0 is the energy of the state $|0\rangle$, E_1 the energy of the state $|1\rangle$, and $\mathcal{Z} = e^{-E_0/k_B t} + e^{-E_1/k_B T}$, what temperature described the state ρ_∞?

Solution

Concepts Involved: Density Operators, Mixed States

To determine the temperature T that describes the state ρ_∞, we assume that in equilibrium, the probability of the qubit being in states $|0\rangle$ and $|1\rangle$ follows the Boltzmann distribution:

$$p_0 = \frac{e^{-E_0/k_B T}}{\mathcal{Z}},$$
$$p_1 = \frac{e^{-E_1/k_B T}}{\mathcal{Z}},$$

where the partition function is:

$$\mathcal{Z} = e^{-E_0/k_B T} + e^{-E_1/k_B T}.$$

By defining the energy difference between the two states as:

$$\Delta E = E_1 - E_0,$$

we rewrite the probability of being in state $|1\rangle$ as:

$$p_1 = \frac{e^{-E_1/k_B T}}{e^{-E_0/k_B T} + e^{-E_1/k_B T}}$$
$$= \frac{e^{-\Delta E/k_B T}}{1 + e^{-\Delta E/k_B T}}.$$

Thus, the ratio of the probabilities satisfies:

$$\frac{p_1}{p_0} = e^{-\Delta E/k_B T}.$$

Given a state ρ_∞ with diagonal elements $(\rho_\infty)_{00} = p$ and $(\rho_\infty)_{11} = 1-p$, we can determine its temperature by solving:

$$\frac{1-p}{p} = e^{-\Delta E/k_B T}.$$

Solving for T:

$$T = \frac{\Delta E}{k_B \ln\left(\frac{p}{1-p}\right)}.$$

□

Exercise 8.26: Circuit model for phase damping

Show that the circuit in Figure 8.15 can be used to model the phase damping quantum operation, provided θ is chosen appropriately.

Solution

Concepts Involved: Kraus Representation, Phase Damping Channel

---

It is instructive to write the unitary interaction as

$$U = (P_0 \otimes I + P_1 \otimes R_y(\theta))$$

Thus, the Kraus operators can be calculated using

$$E_0 \equiv \langle 0_E|U|0_E\rangle = P_0\langle 0_E|I|0_E\rangle + P_1\langle 0_E|R_y(\theta)|0_E\rangle = P_0 + P_1.\cos\big(\theta/2\big)$$
$$E_1 \equiv \langle 1_E|U|1_E\rangle = P_0\langle 1_E|I|0_E\rangle + P_1\langle 1_E|R_y(\theta)|0_E\rangle = P_1.\sin\big(\theta/2\big)$$

This is the phase damping operation, provided θ is chosen as

$$\sin^2(\theta/2) = \lambda.$$

□

Exercise 8.27: Phase damping = phase flip channel

Give the unitary transformation which related the operation elements of (8.127)-(8.128) to those of (8.129)-(8.130); that is, find u such that $\tilde{E}_k = \sum_j u_{kj} E_j$.

Solution

Concepts Involved: Unitary Operators, Kraus Representations, Phase Damping Channel, Phase Flip Channel

---

Let us first choose $\lambda = \sin\theta$, which gives us $\alpha = \cos^2(\theta/2)$.
We can write down the Kraus operators as

$$E_0 = \begin{pmatrix} 1 & 0 \\ 0 & \cos\theta \end{pmatrix}, \quad K_1 = \begin{pmatrix} 0 & 0 \\ 0 & \sin\theta \end{pmatrix}.$$

and

$$\tilde{E}_0 = \begin{pmatrix} \cos(\theta/2) & 0 \\ 0 & \cos(\theta/2) \end{pmatrix}, \quad \tilde{E}_1 = \begin{pmatrix} \sin(\theta/2) & 0 \\ 0 & -\sin(\theta/2) \end{pmatrix}.$$

Using the transform equations,

$$\tilde{E}_0 = u_{00}E_0 + u_{01}E_1$$

yields $u_{00} = \cos(\theta/2)$ and $u_{01} = \sin(\theta/2)$.
Similar equations for $\tilde{E}_1$ gives us, $u_{10} = \sin(\theta/2)$ and $u_{11} = -\cos(\theta/2)$. Thus the unitary transformation relating the two representations is

$$U = \begin{pmatrix} \cos(\theta/2) & \sin(\theta/2) \\ \sin(\theta/2) & -\cos(\theta/2) \end{pmatrix}$$

where $\cos(\theta/2) = \sqrt{\lambda}$.

□

Exercise 8.28: One CNOT phase damping model circuit

Show that a single CNOT gate can be used as a model for phase damping, if we let the initial state of the environment be a mixed state, where the amount of damping is determined by the probability of the states in the mixture.

Solution

Concepts Involved: Controlled Operations, Phase Damping Channel

Let us express the initial state of the environment in the computational basis as

$$\rho_E \cong \begin{pmatrix} p & a \\ a^* & 1-p \end{pmatrix}$$

.
and the pure system state under evolution as

$$\rho_S \cong \begin{pmatrix} |\alpha|^2 & \alpha\beta^* \\ \alpha^*\beta & |\beta|^2 \end{pmatrix}$$

.
The output of the channel is given by,

$$\rho_{\text{out}} = \text{Tr}_E\left(U(\rho_S \otimes \rho_E)U^\dagger\right)$$

$$\rho_{\text{out}} \cong \text{Tr}_E\left(\begin{pmatrix}\mathbf{I} & \mathbf{0}\\ \mathbf{0} & \mathbf{X}\end{pmatrix}\cdot(\rho_S\otimes\rho_E)\cdot\begin{pmatrix}\mathbf{I} & \mathbf{0}\\ \mathbf{0} & \mathbf{X}\end{pmatrix}\right) = \text{Tr}_E\begin{pmatrix}|\alpha|^2\rho_{\mathbf{E}} & \alpha\beta^*\rho_{\mathbf{E}}\cdot\mathbf{X}\\ \alpha^*\beta\mathbf{X}\cdot\rho_{\mathbf{E}} & |\beta|^2\mathbf{X}\cdot\rho_{\mathbf{E}}\cdot\mathbf{X}\end{pmatrix}$$

Thus,

$$\rho_{\text{out}} \cong \begin{pmatrix}|\alpha|^2 & 2\text{Re}(a)\cdot\alpha\beta^*\\ 2\text{Re}(a)\cdot\alpha^*\beta & |\beta|^2\end{pmatrix}$$

This is the phase damping channel with $e^{-\lambda} = 2\text{Re}(a)$ □

Exercise 8.29: Unitality

A quantum process $\mathcal{E}$ is *unital* if $\mathcal{E}(I) = I$. Show that the depolarizing and phase damping channels are unital, while amplitude damping is not.

Solution

Concepts Involved: Unitality, Depolarizing Channel, Phase Damping Channel

We compute the action of the channels on the density matrix. For the depolarizing channel, for the correct normalization we input I/d (a normalized quantum state) rather than I

$$\mathcal{E}_{\text{depol}}(\frac{I}{d}) = p\frac{I}{d} + (1-p)\frac{I}{d} = \frac{I}{d}$$

So it is indeed unital. For phase damping we have

$$\begin{aligned}\mathcal{E}_{PD}(I) &= E_0 I E_0^\dagger + E_1 I E_1^\dagger\\ &= E_0E_0^\dagger + E_1E_1^\dagger\\ &= \begin{bmatrix}1 & 0\\ 0 & \sqrt{1-\lambda}\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & \sqrt{1-\lambda}\end{bmatrix} + \begin{bmatrix}0 & 0\\ 0 & \sqrt{\lambda}\end{bmatrix}\begin{bmatrix}0 & 0\\ 0 & \sqrt{\lambda}\end{bmatrix}\\ &= \begin{bmatrix}1 & 0\\ 0 & 1-\lambda\end{bmatrix} + \begin{bmatrix}0 & 0\\ 0 & \lambda\end{bmatrix}\\ &= I\end{aligned}$$

so it is also unital. For the amplitude damping channel we have

$$\begin{aligned}\mathcal{E}_{\mathsf{AD}}(I) &= E_0 I E_0^\dagger + E_1 I E_1^\dagger \\ &= E_0 E_0^\dagger + E_1 E_1^\dagger \\ &= \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{bmatrix} + \begin{bmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ \sqrt{\gamma} & 0 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1-\gamma \end{bmatrix} + \begin{bmatrix} \gamma & 0 \\ 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 1+\gamma & 0 \\ 0 & 1-\gamma \end{bmatrix} \\ &\neq I\end{aligned}$$

so it is not unital. □

Exercise 8.30: $T_2 \leq T_1/2$

The T_2 phase coherence relaxation rate is just the exponential decay rate of the off-diagonal elements in the qubit density matrix, while T_1 is the decay rate of the diagonal elements (see Equation (7.144)). Amplitude damping has both nonzero T_1 and T_2 rates; show that for amplitude damping $T_2 = T_1/2$. Also show that if amplitude and phase damping are both applied then $T_2 \leq T_1/2$.

Solution

Concepts Involved: Density Operators, Amplitude Damping Channel, Phase Damping Channel

---

Eq. (7.144) describes the T_1/T_2 decay:

$$\begin{bmatrix} a & b \\ b^* & 1-a \end{bmatrix} \mapsto \begin{bmatrix} (a-a_0)e^{-t/T_1} + a_0 & be^{-t/T_2} \\ b^* e^{-t/T_2} & (a_0 - a)e^{-t/T_1} + 1 - a_0 \end{bmatrix}$$

We can compare this to the result of Ex. 8.22:

$$\begin{bmatrix} a & b \\ b^* & 1-a \end{bmatrix} \mapsto \begin{bmatrix} 1-(1-\gamma)(1-a) & b\sqrt{1-\gamma} \\ b^*\sqrt{1-\gamma} & (1-a)(1-\gamma) \end{bmatrix}$$

wherein for amplitude damping we can identify $a_0 = 1$, $(1-\gamma) = e^{-t/T_1}$, and $\sqrt{1-\gamma} = e^{-t/T_2}$. Therein:

$$(e^{-t/T_2})^2 = e^{-2t/T_2} = e^{-t/T_1}$$

and so $T_2 = T_1/2$.
If we have phase damping inb addition to amplitude damping, the off-diagonal/coherence terms decay

with parameter λ (See Eq. (8.125)), so:

$$\begin{bmatrix} a & b \\ b^* & 1-a \end{bmatrix} \mapsto \begin{bmatrix} 1-(1-\gamma)(1-a) & b\sqrt{1-\gamma}e^{-\lambda} \\ b^*\sqrt{1-\gamma}e^{-\lambda} & (1-a)(1-\gamma) \end{bmatrix}$$

wherein we now identify $\sqrt{1-\gamma}e^{-\lambda} = e^{-t/T_2}$. We then can calculate:

$$(e^{-t/T_2})^2 = e^{-2t/T_2} = (\sqrt{1-\gamma}e^{-\lambda})^2 = (1-\gamma)e^{-2\lambda} = e^{-t/T_1 - 2\lambda}$$

so:

$$T_2 = \frac{T_1}{2} + \lambda \leq \frac{T_1}{2}$$

□

Exercise 8.31: Exponential sensitivity to phase damping

Using (8.126), show that the element $\rho_{nm} = \langle n|\rho|m\rangle$ in the density matrix of the harmonic oscillator decays exponentially as $e^{-\lambda(n-m)^2}$ under the effect of phase damping, for some constant λ.

Solution

Concepts Involved: Density Operators, Phase Damping Channel

Consider pure dephasing with Hamiltonian $H = \omega N$ and Lindblad operator $L = N$ (set $\hbar = 1$). The master equation is

$$\dot{\rho} \;=\; -i[H,\rho] + \gamma\Big(L\rho L - \tfrac{1}{2}\{L^2,\rho\}\Big) \;=\; -i\omega[N,\rho] - \tfrac{\gamma}{2}[N,[N,\rho]].$$

In the number basis $N|n\rangle = n|n\rangle$, taking matrix elements gives

$$\dot{\rho}_{nm} \;=\; -i\omega(n-m)\rho_{nm} \;-\; \tfrac{\gamma}{2}(n-m)^2\,\rho_{nm}.$$

Solving,

$$\rho_{nm}(t) \;=\; \rho_{nm}(0)\,e^{-i\omega(n-m)t}\,e^{-\frac{\gamma t}{2}(n-m)^2},$$

so the magnitude decays as

$$|\rho_{nm}(t)| \;=\; |\rho_{nm}(0)|\,e^{-\lambda(n-m)^2}, \qquad \lambda = \tfrac{\gamma t}{2}.$$

□

Exercise 8.32

Explain how to extend quantum process tomography to the case of non-trace-preserving quantum operations, such as arise in the study of measurement.

Solution

Concepts Involved: Quantum Process Tomography, Kraus Representation

In standard quantum process tomography (QPT), one reconstructs a completely positive, trace-preserving (CPTP) map $\mathcal{E}$. In Kraus form,

$$\mathcal{E}(\rho) = \sum_k A_k \rho A_k^\dagger, \qquad \sum_k A_k^\dagger A_k = I.$$

Equivalently, the Choi operator

$$J(\mathcal{E}) = (\mathcal{E} \otimes \mathcal{I})(|\Omega\rangle\langle\Omega|)$$

satisfies the trace-preserving constraint

$$\mathrm{Tr}_{\mathsf{out}}[J(\mathcal{E})] = I.$$

For non–trace-preserving operations, such as measurement outcomes, the map is still completely positive but only trace–*nonincreasing*:

$$\mathcal{E}_m(\rho) = \sum_k A_{m,k}\, \rho\, A_{m,k}^\dagger, \qquad \sum_k A_{m,k}^\dagger A_{m,k} \;\leq\; I.$$

The corresponding Choi operator is positive semidefinite,

$$J(\mathcal{E}_m) \geq 0,$$

but now obeys the weaker condition

$$\mathrm{Tr}_{\mathsf{out}}[J(\mathcal{E}_m)] = P_m \;\leq\; I,$$

where P_m is the *success operator*. The probability of outcome m on input ρ is

$$p_m(\rho) = \mathrm{Tr}[\,P_m \rho\,].$$

Thus the extension of QPT to non–trace-preserving maps is simply: in the reconstruction procedure, do not impose the trace-preserving constraint $\mathrm{Tr}_{\mathsf{out}}\, J = I$, but instead allow the more general $\mathrm{Tr}_{\mathsf{out}}\, J \leq I$, and use tomographic data to estimate both the dynamical map $J(\mathcal{E}_m)$ and its associated success operator P_m. □

Exercise 8.33: Specifying a quantum process

Suppose that one wished to completely specify an arbitrary single qubit operation $\mathcal{E}$ by describing how a set of points on the Bloch sphere $\{\mathbf{r}_k\}$ transform under $\mathcal{E}$. Prove that the set must contain at least four points.

Solution

Concepts Involved: Affine Maps, Bloch Sphere

Any single-qubit state can be written as $\rho = \frac{1}{2}(I + \mathbf{r}\cdot\boldsymbol{\sigma})$ with $\|\mathbf{r}\| \le 1$. Every CPTP map $\mathcal{E}$ acts *affinely* on Bloch vectors

$$\mathbf{r} \;\mapsto\; \mathbf{r}' = T\mathbf{r} + \mathbf{t}, \qquad T \in \mathbb{R}^{3\times 3},\; \mathbf{t} \in \mathbb{R}^3.$$

If we try to determine $\mathcal{E}$ by prescribing the images of m Bloch-sphere points $\{\mathbf{r}_k\}_{k=1}^m$, we obtain the linear system

$$T\mathbf{r}_k + \mathbf{t} = \mathbf{r}'_k \quad (k = 1, \ldots, m),$$

which yields at most $3m$ independent scalar equations for the 12 real unknowns (the 9 entries of T and the 3 entries of $\mathbf{t}$). Thus a *necessary* condition for unique determination of $(T, \mathbf{t})$ in general is $3m \ge 12$, i.e. $m \ge 4$.

Moreover, the CPTP constraints do not rescue the case $m = 3$. Fix three sphere points $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$ and define $\mathbf{d}_1 = \mathbf{r}_1 - \mathbf{r}_3$, $\mathbf{d}_2 = \mathbf{r}_2 - \mathbf{r}_3$. Choose $\mathbf{v} \neq \mathbf{0}$ orthogonal to both $\mathbf{d}_1, \mathbf{d}_2$ and any $\mathbf{u} \neq \mathbf{0}$, and set

$$\Delta T = \mathbf{u}\,\mathbf{v}^\top, \qquad \Delta\mathbf{t} = -\Delta T\,\mathbf{r}_1.$$

Then $\Delta T\,\mathbf{r}_k + \Delta\mathbf{t} = \mathbf{0}$ for $k = 1, 2, 3$. Let $(T_0, \mathbf{t}_0)$ be any qubit CPTP map in the interior of the CPTP set (e.g., the completely depolarizing channel). For sufficiently small $\varepsilon \neq 0$, the perturbed map $(T_0 + \varepsilon\Delta T,\; \mathbf{t}_0 + \varepsilon\Delta\mathbf{t})$ remains CPTP, yet it agrees with $(T_0, \mathbf{t}_0)$ on $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$. Hence three points never suffice to uniquely specify $\mathcal{E}$. Therefore, at least four (affinely independent) Bloch-sphere points are necessary to completely determine an arbitrary single-qubit operation. $\square$

Exercise 8.34: Process tomography for two qubits

Show that the χ_2 describing the black box operations on two qubits can be expressed as

$$\chi_2 = \Lambda_2 \bar{\rho}' \Lambda_2$$

where $\Lambda_2 = \Lambda \otimes \Lambda, \Lambda$ is as defined in Box 8.5, and $\bar{\rho}'$ is a block matrix of 16 measured density matrices,

$$\bar{\rho}' = P^T \begin{bmatrix} \rho'_{11} & \rho'_{12} & \rho'_{13} & \rho'_{14} \\ \rho'_{21} & \rho'_{22} & \rho'_{23} & \rho'_{24} \\ \rho'_{31} & \rho'_{32} & \rho'_{33} & \rho'_{34} \\ \rho'_{41} & \rho'_{42} & \rho'_{43} & \rho'_{44} \end{bmatrix} P$$

where $\rho'_{nm} = \mathcal{E}(\rho_{nm}), \rho_{nm} = T_n\,|00\rangle\langle 00|\,T_m, T_1 = I \otimes I, T_2 = I \otimes X$, $T_3 = X \otimes I, T_4 = X \otimes X$, and $P = I \otimes \left[(\rho_{00} + \rho_{12} + \rho_{21} + \rho_{33}) \otimes I\right]$ is a permutation matrix.

Solution

Concepts Involved: Quantum Process Tomography, Density Operators,

Let $\{|1\rangle, |2\rangle, |3\rangle, |4\rangle\} = \{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$. Define

$$T_1 = I \otimes I, \quad T_2 = I \otimes X, \quad T_3 = X \otimes I, \quad T_4 = X \otimes X.$$

Then

$$\rho_{nm} = T_n|00\rangle\langle 00|T_m = |n\rangle\langle m| \equiv E_{nm}.$$

Passing these through $\mathcal{E}$ gives

$$\rho'_{nm} = \mathcal{E}(\rho_{nm}) = \mathcal{E}(E_{nm}),$$

and assembling them yields the 4×4 block matrix

$$B = [\,\rho'_{nm}\,]_{n,m=1}^{4}.$$

The Choi matrix of $\mathcal{E}$ is

$$J(\mathcal{E}) = \sum_{n,m} \mathcal{E}(E_{nm}) \otimes E_{nm} = \sum_{n,m} \rho'_{nm} \otimes E_{nm}.$$

This is a reshuffling of B, implemented by a fixed permutation matrix P. Define

$$\bar{\rho}' = P^T B P,$$

so that $\bar{\rho}'$ is exactly $J(\mathcal{E})$ written in the computational operator basis.

For a single qubit, Box 8.5 introduces a 4×4 change-of-basis matrix Λ that maps from the matrix-units basis $\{|i\rangle\langle j|\}$ to the normalized Pauli basis $\{I, X, Y, Z\}/\sqrt{2}$. For two qubits this factorizes as

$$\Lambda_2 = \Lambda \otimes \Lambda.$$

By definition, the χ-matrix is the Choi matrix expressed in the Pauli-product basis. Therefore,

$$\chi_2 = \Lambda_2\, \bar{\rho}'\, \Lambda_2,$$

as required. □

Exercise 8.35: Process tomography example

Consider a one qubit black box of unknown dynamics $\mathcal{E}_1$. Suppose that the following four density matrices are obtained from experimental measurements, performed according to Equations (8.173)–(8.176)

$$\rho'_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\rho'_2 = \begin{bmatrix} 0 & \sqrt{1-\gamma} \\ 0 & 0 \end{bmatrix}$$

$$\rho'_3 = \begin{bmatrix} 0 & 0 \\ \sqrt{1-\gamma} & 0 \end{bmatrix}$$

$$\rho'_4 = \begin{bmatrix} \gamma & 0 \\ 0 & 1-\gamma \end{bmatrix}$$

where γ is a numerical parameter. From an independent study of each of these input-output relations, one could make several important observations: the ground state $|0\rangle$ is left invariant by $\mathcal{E}_1$. the excited state $|1\rangle$ partially decays to the ground state, and superposition states are damped. Determine the χ matrix for this process.

Solution

Concepts Involved: Quantum Processing Tomography, Density Operators

The input/output data correspond to the amplitude–damping channel with parameter γ, whose Kraus operators may be chosen as

$$K_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{pmatrix}, \qquad K_1 = \begin{pmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{pmatrix}.$$

Express K_i in the normalized Pauli basis $\{E_0, E_1, E_2, E_3\} = \{I, X, Y, Z\}/\sqrt{2}$. With $s = \sqrt{1-\gamma}$,

$$K_0 = \alpha I + \beta Z \ = \sqrt{2}\alpha\, E_0 + \sqrt{2}\beta\, E_3, \quad \alpha = \tfrac{1+s}{2},\ \beta = \tfrac{1-s}{2},$$

$$K_1 = \tfrac{\sqrt{\gamma}}{2}(X + iY) \ = \ \tfrac{\sqrt{\gamma}}{\sqrt{2}}\, E_1 + \tfrac{i\sqrt{\gamma}}{\sqrt{2}}\, E_2.$$

Let $e_0 = (\sqrt{2}\alpha, 0, 0, \sqrt{2}\beta)^\top$ and $e_1 = (0, \frac{\sqrt{\gamma}}{\sqrt{2}}, \frac{i\sqrt{\gamma}}{\sqrt{2}}, 0)^\top$ be the coefficient vectors of K_0, K_1 in this basis. Then the process matrix is $\chi = \sum_i e_i e_i^\dagger$, i.e.

$$\boxed{\chi = \begin{pmatrix} 1 - \frac{\gamma}{2} + s & 0 & 0 & \frac{\gamma}{2} \\ 0 & \frac{\gamma}{2} & -\frac{i\gamma}{2} & 0 \\ 0 & \frac{i\gamma}{2} & \frac{\gamma}{2} & 0 \\ \frac{\gamma}{2} & 0 & 0 & 1 - \frac{\gamma}{2} - s \end{pmatrix}_{\{I,X,Y,Z\}/\sqrt{2}}, \qquad s = \sqrt{1-\gamma}.}$$

This χ is positive semidefinite and satisfies the trace preserving constraint in the normalized Pauli basis. □

Problem 8.1: Lindblad form to quantum operation

In the notation of Section 8.4.1, explicitly work through the steps to solve the differential equation

$$\dot{\rho} = -\frac{\lambda}{2}(\sigma_+\sigma_-\rho + \rho\sigma_+\sigma_- - 2\sigma_-\rho\sigma_+)$$

for $\rho(t)$. Express the map $\rho(0) \mapsto \rho(t)$ as $\rho(t) = \sum_k E_k(t)\rho(0)E_k^\dagger(t)$.

Solution

Concepts Involved: Differential Equations, Kraus Represenations

Let us use the master equation to obtain the differential equations for the diagonal and off-diagonal elements: - **Population (diagonal) elements**:

$$\dot{\rho}_{11}(t) = -\lambda\rho_{11}(t), \quad \dot{\rho}_{00}(t) = \lambda\rho_{11}(t),$$

with solutions:

$$\rho_{11}(t) = \rho_{11}(0)e^{-\lambda t}, \quad \rho_{00}(t) = \rho_{00}(0) + \rho_{11}(0)(1 - e^{-\lambda t}).$$

- **Coherence (off-diagonal) elements**:

$$\dot{\rho}_{10}(t) = -\frac{\lambda}{2}\rho_{10}(t), \quad \dot{\rho}_{01}(t) = -\frac{\lambda}{2}\rho_{01}(t),$$

with solutions:

$$\rho_{10}(t) = \rho_{10}(0)e^{-\lambda t/2}, \quad \rho_{01}(t) = \rho_{01}(0)e^{-\lambda t/2}.$$

The time evolution of the density matrix can be expressed using Kraus operators for amplitude damping:

$$E_0(t) = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{e^{-\lambda t}} \end{pmatrix}, \quad E_1(t) = \begin{pmatrix} 0 & \sqrt{1 - e^{-\lambda t}} \\ 0 & 0 \end{pmatrix}.$$

These Kraus operators satisfy the completeness relation:

$$E_0^\dagger(t)E_0(t) + E_1^\dagger(t)E_1(t) = I,$$

and the solution to the master equation is:

$$\rho(t) = E_0(t)\rho(0)E_0^\dagger(t) + E_1(t)\rho(0)E_1^\dagger(t).$$

□

Problem 8.2: Teleportation as a quantum operation

Suppose Alice is in possession of a single qubit, denoted as system 1, which she wishes to teleport to Bob. Unfortunately, she and Bob only share an imperfectly entangled pair of qubits. Alice's half of this pair is denoted system 2, and Bob's half is denoted system 3. Suppose Alice performs a measurement described by a set of quantum operations $\mathcal{E}_m$ with result m on systems 1 and 2. Show that this induces an operation $\tilde{E}_m$ relating the initial stat of the system 1 to the final state of system 3, and that teleportation is accomplished if Bob can reverse this operation using a trace-preserving quantum operation $\mathcal{R}_m$, to obtain

$$\mathcal{R}_m\left(\frac{\tilde{\mathcal{E}}_m(\rho)}{\operatorname{tr}\left[\tilde{\mathcal{E}}_m(\rho)\right]}\right) = \rho,$$

where ρ is the initial state of system 1.

Solution

Concepts Involved: Quantum Teleporation, Density Operators, Kraus Representation

Teleportation Protocol

Let us consider a scenario where Alice holds qubit 1 (state ρ) to be teleported, and shares an imperfectly entangled pair of qubits with Bob, where Alice holds qubit 2 and Bob holds qubit 3. The initial state of the combined system (qubits 1, 2, and 3) is:

$$\rho_{123} = \rho_1 \otimes |\Psi\rangle_{23} \langle\Psi| \, ,$$

where $|\Psi\rangle_{23}$ represents the (imperfectly) entangled state between qubits 2 and 3.

Step 1: Alice's Measurement on Systems 1 and 2

Alice performs a measurement on systems 1 and 2, described by the quantum operation $\mathcal{E}_m$, corresponding to Kraus operators $E_m^{(12)}$, and records the measurement outcome m. After this measurement, the state of the combined system becomes:

$$\tilde{\rho}_{123}^{(m)} = \frac{(E_m^{(12)} \otimes I_3)\rho_{123}(E_m^{(12)} \otimes I_3)^\dagger}{\mathrm{Tr}[(E_m^{(12)} \otimes I_3)\rho_{123}(E_m^{(12)} \otimes I_3)^\dagger]}.$$

Step 2: Induced Operation on Bob's Qubit

The measurement on systems 1 and 2 induces an operation $\tilde{\mathcal{E}}_m$ on Bob's qubit (system 3), which can be expressed by tracing out systems 1 and 2:

$$\tilde{\mathcal{E}}_m(\rho_1) = \mathrm{Tr}_{12}\left[(E_m^{(12)} \otimes I_3)(\rho_1 \otimes |\Psi\rangle_{23} \langle\Psi|)(E_m^{(12)} \otimes I_3)^\dagger\right].$$

This describes how Alice's measurement modifies the state of Bob's qubit, based on the initial state ρ_1 of system 1.

Step 3: Bob's Recovery Operation

Once Alice communicates the measurement outcome m to Bob, he applies a trace-preserving quantum operation $\mathcal{R}_m$ to reverse the effect of the induced operation $\tilde{\mathcal{E}}_m$. The condition for successful teleportation is:

$$\mathcal{R}_m\left(\frac{\tilde{\mathcal{E}}_m(\rho_1)}{\mathrm{Tr}[\tilde{\mathcal{E}}_m(\rho_1)]}\right) = \rho_1,$$

where the division by $\mathrm{Tr}[\tilde{\mathcal{E}}_m(\rho_1)]$ ensures normalization.

□

Problem 8.3: Random unitary channels

It is tempting to believe that all unital channels, that is, those for which $\mathcal{E}(I) = I$, result from averaging over random unitary operations, that is, $\mathcal{E}(\rho) = \sum_k p_k \, U_k \, \rho \, U_k^\dagger$, where U_k are unitary operations and the p_k form a probability distribution. Show that while this is true for single qubits, it is untrue for larger systems.

Solution

Concepts Involved: Unitary Operators, Unitality, Kraus Representation

Let $\mathcal{X} = \mathbb{C}^3$ and define the (Werner-Holevo) map

$$\mathcal{E}(X) = \frac{1}{2}\text{Tr}(X)I - \frac{1}{2}X^T$$

for all $X \in L(\mathcal{X})$. It can be confirmed that $\mathcal{E}$ is a channel and unital, but it is not a mixed-unitary channel. To demonstrate that $\mathcal{E}$ is not a mixed-unitary channel, observe the decomposition

$$\mathcal{E}(X) = A_1XA_1^\dagger + A_2XA_2^\dagger + A_3XA_3^\dagger$$

for all $X \in L(\mathcal{X})$, where the matrices are given as

$$A_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} \\ 0 & -\frac{1}{\sqrt{2}} & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 0 & \frac{1}{\sqrt{2}} \\ 0 & 0 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The validity of this expression for all $X \in L(\mathcal{X})$ can be verified through the Choi representation of the map, which matches the right-hand side.
Recalling the Choi theorem,
Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces. Suppose $\mathcal{E} \in C(\mathcal{X}, \mathcal{Y})$ is a quantum channel, and let $\{A_a : a \in \Sigma\} \subset L(\mathcal{X}, \mathcal{Y})$ be a linearly independent set of operators. Then

$$\mathcal{E}(X) = \sum_{a \in \Sigma} A_a X A_a^\dagger$$

for all $X \in L(\mathcal{X})$. The channel $\mathcal{E}$ is an extreme point of the set of channels $C(\mathcal{X}, \mathcal{Y})$ if and only if the collection

$$\{A_b^\dagger A_a : (a, b) \in \Sigma \times \Sigma\} \subset L(\mathcal{X})$$

is linearly independent.
For us, the set $\{A_j^\dagger A_k : 1 \leq j, k \leq 3\}$ contains the following operators:

$$A_1^\dagger A_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}, \quad A_1^\dagger A_2 = \begin{pmatrix} 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A_1^\dagger A_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \end{pmatrix},$$

$$A_2^\dagger A_1 = \begin{pmatrix} 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A_2^\dagger A_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}, \quad A_2^\dagger A_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 \end{pmatrix},$$

$$A_3^\dagger A_1 = \begin{pmatrix} 0 & 0 & -\frac{1}{2} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A_3^\dagger A_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 \end{pmatrix}, \quad A_3^\dagger A_3 = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

This collection forms a linearly independent set and thus $\mathcal{E}$ is an extreme point of the set of channels. Since $\mathcal{E}$ is not a unitary channel, it follows that it cannot be written as a convex combination of unitary channels.

□

9 Distance measures for quantum information

Exercise 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 $(3/4, 1/8, 1/8)$?

Solution

Concepts Involved: Trace Distance

The trace distance between $(1, 0)$ and $\left(\frac{1}{2}, \frac{1}{2}\right)$ is:

$$D = \frac{1}{2}\left(|1 - \tfrac{1}{2}| + |0 - \tfrac{1}{2}|\right) = \frac{1}{2}.$$

The trace distance between $\left(\frac{1}{2}, \frac{1}{3}, \frac{1}{6}\right)$ and $\left(\frac{3}{4}, \frac{1}{8}, \frac{1}{8}\right)$ is:

$$D = \frac{1}{2}\left(\left|\tfrac{1}{2} - \tfrac{3}{4}\right| + \left|\tfrac{1}{3} - \tfrac{1}{8}\right| + \left|\tfrac{1}{6} - \tfrac{1}{8}\right|\right) = \frac{1}{2}\left(\tfrac{1}{4} + \tfrac{5}{24} + \tfrac{1}{24}\right) = \frac{1}{2}.$$

□

Exercise 9.2

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

Solution

Concepts Involved: Trace Distance

For $P = (p, 1 - p)$ and $Q = (q, 1 - q)$,

$$D(P, Q) = \frac{1}{2}\left(|p - q| + |(1 - p) - (1 - q)|\right) = |p - q|.$$

□

Exercise 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 $(3/4, 1/8, 1/8)$?

Solution

Concepts Involved: Fidelity

The fidelity between classical probability distributions $p = (p_1, \ldots, p_n)$ and $q = (q_1, \ldots, q_n)$ is

$$F(p, q) = \sum_i \sqrt{p_i q_i}.$$

For $(1,0)$ and $\left(\frac{1}{2},\frac{1}{2}\right)$,

$$F = \sqrt{1 \cdot \tfrac{1}{2}} + \sqrt{0 \cdot \tfrac{1}{2}} = \sqrt{\tfrac{1}{2}} = \frac{1}{\sqrt{2}}.$$

For $\left(\frac{1}{2},\frac{1}{3},\frac{1}{6}\right)$ and $\left(\frac{3}{4},\frac{1}{8},\frac{1}{8}\right)$,

$$F = \sqrt{\tfrac{1}{2} \cdot \tfrac{3}{4}} + \sqrt{\tfrac{1}{3} \cdot \tfrac{1}{8}} + \sqrt{\tfrac{1}{6} \cdot \tfrac{1}{8}} = \sqrt{\tfrac{3}{8}} + \sqrt{\tfrac{1}{24}} + \sqrt{\tfrac{1}{48}}.$$

□

Exercise 9.4

Prove (9.3)

Solution

Concepts Involved: Trace Distance

We are given two probability distributions p_x and q_x over a finite sample space $\mathcal{X}$, and aim to show:

$$D(p,q) = \frac{1}{2}\sum_x |p_x - q_x| = \max_{S \subseteq \mathcal{X}} \left| \sum_{x \in S} (p_x - q_x) \right|.$$

Let $A := \{x \in \mathcal{X} : p_x \geq q_x\}$. Then:

$$\sum_x |p_x - q_x| = \sum_{x \in A} (p_x - q_x) + \sum_{x \notin A} (q_x - p_x) = 2\sum_{x \in A} (p_x - q_x).$$

Thus,

$$D(p,q) = \frac{1}{2}\sum_x |p_x - q_x| = \sum_{x \in A} (p_x - q_x) = p(A) - q(A).$$

Since this value is achieved by some subset $A \subseteq \mathcal{X}$, it follows that:

$$D(p,q) = \max_{S \subseteq \mathcal{X}} |p(S) - q(S)|.$$

□

Exercise 9.5

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

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

Solution

Concepts Involved: Trace Distance

The trace distance between distributions p and q is

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

Let $\bar{S}$ be the complement of S. Then

$$\sum_{x \in \bar{S}} (p_x - q_x) = \sum_x (p_x - q_x) - \sum_{x \in S} (p_x - q_x) = -\sum_{x \in S} (p_x - q_x),$$

since $\sum_x p_x = \sum_x q_x = 1$. Thus,

$$\left| \sum_{x \in S} (p_x - q_x) \right| = \max \left\{ \sum_{x \in S} (p_x - q_x), \sum_{x \in \bar{S}} (p_x - q_x) \right\}.$$

Therefore,

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

□

Exercise 9.6

What is the trace distance between the density operators

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

Between

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

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

Solution

Concepts Involved: Trace Distance

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

These states are diagonal in the computational basis. Their difference is

$$\rho - \sigma = \left(\frac{3}{4} - \frac{2}{3} \right) |0\rangle \langle 0| + \left(\frac{1}{4} - \frac{1}{3} \right) |1\rangle \langle 1| = \frac{1}{12} |0\rangle \langle 0| - \frac{1}{12} |1\rangle \langle 1|$$

This operator has eigenvalues $\pm\frac{1}{12}$, so its trace norm is:

$$\|\rho - \sigma\|_1 = \left|\frac{1}{12}\right| + \left|-\frac{1}{12}\right| = \frac{1}{6}$$

Thus, the trace distance is

$$D(\rho, \sigma) = \frac{1}{2}\|\rho - \sigma\|_1 = \frac{1}{12}$$

For the second part, we have

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

Recall that

$$|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \quad |-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$$

We express $|\pm\rangle\langle\pm|$ in the Pauli basis

$$|+\rangle\langle +| = \frac{1}{2}(I + X), \quad |-\rangle\langle -| = \frac{1}{2}(I - X)$$

So, we have

$$\rho = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| + \frac{1}{4}|0\rangle\langle 0| - \frac{1}{4}|1\rangle\langle 1| = \frac{I}{2} + \frac{1}{4}Z$$

$$\sigma = \frac{2}{3}\cdot\frac{1}{2}(I + X) + \frac{1}{3}\cdot\frac{1}{2}(I - X) = \frac{1}{2}I + \frac{1}{6}X$$

The Bloch vectors are

$$\vec{r} = (0, 0, \tfrac{1}{2}), \quad \vec{s} = (\tfrac{1}{3}, 0, 0)$$

The trace distance between qubit states is

$$D(\rho, \sigma) = \frac{1}{2}\|\vec{r} - \vec{s}\| = \frac{1}{2}\sqrt{\left(0 - \tfrac{1}{3}\right)^2 + \left(\tfrac{1}{2} - 0\right)^2} = \frac{1}{2}\sqrt{\frac{1}{9} + \frac{1}{4}} = \frac{1}{2}\sqrt{\frac{13}{36}} = \frac{\sqrt{13}}{12}$$

□

Exercise 9.7

Show that for any states ρ and σ, one may write $\rho - \sigma = Q - S$, where Q and S are positive operators with support on orthogonal vector spaces. (*Hint:* use the spectral decomposition $\rho - \sigma = UDU^\dagger$, and split the diagonal matrix D intro positive and negative parts. This fact will continue to be useful later.)

Solution

Concepts Involved: Positive Operators, Spectral Decomposition

---

Let $A := \rho - \sigma$, which is Hermitian. By the spectral theorem, we can write

$$A = UDU^\dagger,$$

where $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ is a real diagonal matrix and U is unitary.
Define the positive and negative parts of D as:

$$D_+ := \mathrm{diag}(\max\{\lambda_i, 0\}), \quad D_- := \mathrm{diag}(\max\{-\lambda_i, 0\}).$$

Then:

$$D = D_+ - D_- \quad \Rightarrow \quad A = UD_+U^\dagger - UD_-U^\dagger.$$

Define:

$$Q := UD_+U^\dagger, \quad S := UD_-U^\dagger.$$

Then $\rho - \sigma = Q - S$, where $Q \geq 0$, $S \geq 0$. Since D_+ and D_- act on disjoint eigenspaces, their supports are orthogonal, and unitaries preserve orthogonality. Therefore:

$$\mathrm{supp}(Q) \perp \mathrm{supp}(S).$$

□

Exercise 9.8: Convexity of the trace distance

Show that the trace distance is convex in its first input,

$$D\left(\sum_i p_i \rho_i, \sigma\right) \leq \sum_i p_i D(p_i, \sigma)$$

By symmetry convexity in the second entry follows from convexity in the first.

Solution

Concepts Involved: Trace distance, Convexity, Triangle inequality

---

Let $D(\rho, \sigma) := \frac{1}{2}\|\rho - \sigma\|_1$. We want to show:

$$D\left(\sum_i p_i \rho_i, \sigma\right) \leq \sum_i p_i D(\rho_i, \sigma).$$

Define $\tau := \sum_i p_i \rho_i$. Then:

$$D(\tau, \sigma) = \frac{1}{2}\left\|\sum_i p_i(\rho_i - \sigma)\right\|_1.$$

Using the triangle inequality and linearity of the trace norm:

$$\left\|\sum_i p_i(\rho_i - \sigma)\right\|_1 \leq \sum_i p_i \|\rho_i - \sigma\|_1.$$

Hence:

$$D\left(\sum_i p_i\rho_i, \sigma\right) \le \sum_i p_i D(\rho_i, \sigma).$$

□

Remark: This proves convexity of the trace distance in the first argument. By symmetry $D(\rho, \sigma) = D(\sigma, \rho)$, convexity in the second argument follows. More generally, the trace distance is jointly convex:

$$D\left(\sum_i p_i\rho_i, \sum_i p_i\sigma_i\right) \le \sum_i p_i D(\rho_i, \sigma_i),$$

a fact useful in quantum information theory and quantum hypothesis testing.

Exercise 9.9: Existence of fixed points

Schauder's fixed point theorem is a classic result from mathematics that implies that any continuous map on a convex, compact subset of a Hilbert space has a fixed point. Use Schauder's fixed point theorem to prove that any trace-preserving quantum operation $\mathcal{E}$ has a fixed point, that is, ρ such that $\mathcal{E}(\rho) = \rho$.

Solution

Concepts Involved: Schauder's Fixed-Point Theorem, CPTP Maps, Convex Compact Sets, Continuity

Let $\mathcal{E}$ be a trace-preserving quantum operation on density matrices. The set $\mathcal{D}$ of density matrices on a finite-dimensional Hilbert space $\mathcal{H}$ is convex, compact, and closed in the trace norm topology.
Since $\mathcal{E}$ is a quantum channel, it is completely positive and trace-preserving (CPTP), hence continuous in the trace norm. Thus, $\mathcal{E} : \mathcal{D} \to \mathcal{D}$ is a continuous map from a convex compact subset of a Banach space (here, a finite-dimensional space suffices) to itself.
By Schauder's fixed point theorem, there exists $\rho \in \mathcal{D}$ such that $\mathcal{E}(\rho) = \rho$. □

Remark: In finite dimensions, the Banach space structure is not essential — the convexity and compactness of the set of density matrices, plus continuity of $\mathcal{E}$, suffice. The fixed point ρ can be interpreted physically as a steady state of the quantum channel.

Exercise 9.10

Suppose $\mathcal{E}$ is a *strictly contractive* trace-preserving quantum operation, that is, for any ρ and σ, $D(\mathcal{E}(\rho), \mathcal{E}(\sigma)) < D(\rho, \sigma)$. Show that $\mathcal{E}$ has a unique fixed point.

Solution

Concepts Involved: Strict Contraction, Trace Distance, Banach Fixed Point Theorem

Let $\mathcal{E}$ be a strictly contractive, trace-preserving quantum operation. That is,

$$D(\mathcal{E}(\rho), \mathcal{E}(\sigma)) < D(\rho, \sigma) \quad \text{for all } \rho \neq \sigma.$$

The trace distance $D(\cdot, \cdot)$ is a metric on the compact convex set $\mathcal{D}$ of density matrices.
This strict contractivity implies that $\mathcal{E}$ is a strict contraction with respect to a complete metric (the trace distance), hence the Banach fixed point theorem applies. Therefore, $\mathcal{E}$ has a unique fixed point $\rho^* \in \mathcal{D}$ such that $\mathcal{E}(\rho^*) = \rho^*$. □

Remark: Strict contractivity guarantees not only existence but also uniqueness of the fixed point. Moreover, iterating the map $\rho_{n+1} := \mathcal{E}(\rho_n)$ leads to exponential convergence to the unique fixed point ρ^*, making it a strong attractor under repeated application of $\mathcal{E}$.

Exercise 9.11

Suppose $\mathcal{E}$ is a trace-preserving quantum operation for which there exists a density operator ρ_0 and a trace-preserving quantum operation $\mathcal{E}'$ such that

$$\mathcal{E}(\rho) = p\rho_0 + (1-p)\mathcal{E}'(\rho),$$

for some p, $0 < p \leq 1$. Physically, this means that with probability p the input state is thrown out and replaced with the fixed state ρ_0, while with probability $1-p$ the operation $\mathcal{E}'$ occurs. Use joint convexity to show that $\mathcal{E}$ is a strictly contractive quantum operation, and thus has a unique fixed point.

Solution

Concepts Involved: Trace Distance, Joint Convexity, Density Operators, Strict Contraction, Fixed Points

Given:

$$\mathcal{E}(\rho) = p\rho_0 + (1-p)\mathcal{E}'(\rho), \quad \text{with } 0 < p \leq 1.$$

Let ρ, σ be two arbitrary density operators. Then:

$$D(\mathcal{E}(\rho), \mathcal{E}(\sigma)) = D(p\rho_0 + (1-p)\mathcal{E}'(\rho),\ p\rho_0 + (1-p)\mathcal{E}'(\sigma)).$$

Apply joint convexity of trace distance:

$$D(p\rho_0 + (1-p)\mathcal{E}'(\rho),\ p\rho_0 + (1-p)\mathcal{E}'(\sigma)) \leq pD(\rho_0, \rho_0) + (1-p)D(\mathcal{E}'(\rho), \mathcal{E}'(\sigma)).$$

Since $D(\rho_0, \rho_0) = 0$ and $\mathcal{E}'$ is trace-preserving (thus contractive):

$$D(\mathcal{E}(\rho), \mathcal{E}(\sigma)) \leq (1-p)D(\mathcal{E}'(\rho), \mathcal{E}'(\sigma)) \leq (1-p)D(\rho, \sigma).$$

Since $0 < p \leq 1$, we have $(1-p) < 1$, so:

$$D(\mathcal{E}(\rho), \mathcal{E}(\sigma)) < D(\rho, \sigma),$$

i.e., $\mathcal{E}$ is strictly contractive. □

Remark: By the previous result, any strictly contractive trace-preserving quantum operation has a unique fixed point. Here, it is easy to see that $\mathcal{E}(\rho_0) = \rho_0$, so ρ_0 is the unique fixed point.

Exercise 9.12

Consider the depolarizing channel introduced in Section 8.3.4 on page 378, $\mathcal{E}(\rho) = pI/2 + (1-p)\rho$. For arbitrary ρ and σ, find $D(\mathcal{E}(\rho), \mathcal{E}(\sigma))$ using the Bloch representation, and prove explicitly that the map $\mathcal{E}$ is strictly contractive, that is, $D(\mathcal{E}(\rho), \mathcal{E}(\sigma)) < D(\rho, \sigma)$.

Solution

Concepts Involved: Depolarizing Channel, Bloch Sphere, Trace Distance, Strinct Contraction

The depolarizing channel acts as

$$\mathcal{E}(\rho) = p\frac{I}{2} + (1-p)\rho, \quad \text{with } 0 < p \leq 1.$$

Represent qubit states ρ and σ in the Bloch form

$$\rho = \frac{1}{2}(I + \vec{r} \cdot \vec{\sigma}), \quad \sigma = \frac{1}{2}(I + \vec{s} \cdot \vec{\sigma}).$$

Then,

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

So the Bloch vector of $\mathcal{E}(\rho)$ is scaled by $1-p$. The trace distance between qubit states is

$$D(\rho, \sigma) = \frac{1}{2}\|\vec{r} - \vec{s}\|, \quad D(\mathcal{E}(\rho), \mathcal{E}(\sigma)) = \frac{1}{2}\|(1-p)(\vec{r} - \vec{s})\| = (1-p)D(\rho, \sigma).$$

Since $0 < p \leq 1$, we have $0 \leq 1 - p < 1$, so

$$D(\mathcal{E}(\rho), \mathcal{E}(\sigma)) < D(\rho, \sigma),$$

for all $\rho \neq \sigma$. Thus, $\mathcal{E}$ is strictly contractive. $\square$

Remark: The depolarizing channel shrinks all Bloch vectors toward the maximally mixed state $I/2$, reducing distinguishability between any pair of states. Hence, it is strictly contractive and has a unique fixed point: $I/2$.

Exercise 9.13

Show that the bit flip channel (Section 8.3.3) is contractive but not strictly contractive. Find the set of fixed points for the bit flip channel.

Solution

Concepts Involved: Bit Flip Channel, Trace Distance, Contraction, Strict Contraction, Fixed Points

The bit flip channel is defined as

$$\mathcal{E}(\rho) = (1-p)\rho + pX\rho X,$$

where X is the Pauli-X operator and $0 \leq p \leq 1$.

To show contractivity, let ρ, σ be density matrices:

$$\mathcal{E}(\rho - \sigma) = (1-p)(\rho - \sigma) + pX(\rho - \sigma)X.$$

Using unitary invariance and convexity of the trace norm

$$\|\mathcal{E}(\rho - \sigma)\|_1 \leq (1-p)\|\rho - \sigma\|_1 + p\|X(\rho - \sigma)X\|_1 = \|\rho - \sigma\|_1.$$

Thus, $D(\mathcal{E}(\rho), \mathcal{E}(\sigma)) \leq D(\rho, \sigma)$, so $\mathcal{E}$ is contractive.
To see it is not strictly contractive, consider ρ, σ diagonal in the $\{|+\rangle, |-\rangle\}$ basis, i.e., eigenstates of X. Then $X\rho X = \rho$, so $\mathcal{E}(\rho) = \rho$, and likewise for σ, hence

$$D(\mathcal{E}(\rho), \mathcal{E}(\sigma)) = D(\rho, \sigma).$$

For fixed points, solve

$$\mathcal{E}(\rho) = \rho \Rightarrow (1-p)\rho + pX\rho X = \rho \Rightarrow X\rho X = \rho.$$

This implies ρ commutes with X, i.e., $[X, \rho] = 0$. The matrices commuting with X are exactly those diagonal in the $\{|+\rangle, |-\rangle\}$ basis. Hence, the fixed points are the set of density operators diagonal in the Hadamard basis. □

Exercise 9.14: Invariance of fidelity under unitary transforms

Prove (9.61) by using the fact that for any positive operator A, $\sqrt{UAU^\dagger} = U\sqrt{A^\dagger}$.

Solution

Concepts Involved: Trace Norm, Positivite Operators, Unitary Operators

Equation (9.61) states that the trace norm is unitarily invariant:

$$\|UAU^\dagger\|_1 = \|A\|_1 \quad \text{for all operators } A \text{ and unitaries } U.$$

Let A be arbitrary. By definition of the trace norm:

$$\|A\|_1 = \mathrm{Tr}\sqrt{A^\dagger A}.$$

Let $B := A^\dagger A$, which is a positive operator. Then:

$$\|UAU^\dagger\|_1 = \mathrm{Tr}\sqrt{(UAU^\dagger)^\dagger(UAU^\dagger)} = \mathrm{Tr}\sqrt{UA^\dagger U^\dagger UAU^\dagger} = \mathrm{Tr}\sqrt{UA^\dagger AU^\dagger}.$$

Now apply the fact that for any positive operator B, $\sqrt{UBU^\dagger} = U\sqrt{B}U^\dagger$. So:

$$\|UAU^\dagger\|_1 = \mathrm{Tr}(U\sqrt{B}U^\dagger) = \mathrm{Tr}(\sqrt{B}) = \|A\|_1,$$

using cyclicity of the trace. Hence, $\|UAU^\dagger\|_1 = \|A\|_1$. □

Exercise 9.15

Show that

$$F(\rho, \sigma) = \max_{|\varphi\rangle} \big|\langle\psi|\varphi\rangle\big|$$

Where $|\psi\rangle$ is any *fixed* purification of ρ, and the maximization is over all purifications of σ.

Solution

Concepts Involved: Fidelity, Purifications, Uhlmann's theorem

Let $|\psi\rangle \in \mathcal{H} \otimes \mathcal{K}$ be a fixed purification of ρ, i.e., $\mathrm{Tr}_{\mathcal{K}} |\psi\rangle\langle\psi| = \rho$.
Any purification $|\varphi\rangle$ of σ in the same extended space can be written as

$$|\varphi\rangle = (I \otimes U)\,|\tilde{\varphi}\rangle\,,$$

where $|\tilde{\varphi}\rangle$ is a fixed purification of σ, and U is a unitary operator on the ancillary Hilbert space $\mathcal{K}$. This follows from the fact that any two purifications of the same state are related by a unitary on the purifying system.
Therefore, the overlap between $|\psi\rangle$ and $|\varphi\rangle$ takes the form

$$\langle\psi|\varphi\rangle = \langle\psi|\,(I \otimes U)\,|\tilde{\varphi}\rangle\,,$$

and the maximal overlap over all purifications of σ becomes

$$\max_{|\varphi\rangle} |\,\langle\psi|\varphi\rangle\,| = \max_{U} \big|\langle\psi|\,(I \otimes U)\,|\tilde{\varphi}\rangle\big|\,.$$

Uhlmann's theorem states that

$$F(\rho, \sigma) = \max_{U} \big|\langle\psi|\,(I \otimes U)\,|\tilde{\varphi}\rangle\big|\,,$$

for any fixed purifications $|\psi\rangle$ of ρ and $|\tilde{\varphi}\rangle$ of σ.
Thus, we conclude

$$F(\rho, \sigma) = \max_{|\varphi\rangle} \big|\langle\psi|\varphi\rangle\big|.$$

□

Remark: This expression shows that fidelity between mixed states equals the best achievable overlap between a fixed purification of ρ and all purifications of σ, emphasizing its interpretation as an optimal transition amplitude between purifications.

Exercise 9.16: The Hilbert–Schmidt inner product and entanglement

Suppose R and Q are two quantum systems with the same Hilbert space. Let $|i_R\rangle$ and $|i_Q\rangle$ be orthonormal basis sets for R and Q. Let A be an operator on R and B be an operator on Q. Define $|m\rangle \equiv \sum_i |i_R\rangle|i_Q\rangle$. Show that

$$\operatorname{tr}\left(A^\dagger B\right) = \langle m|(A \otimes B)|m\rangle$$

where the multiplication on the left hand side is of *matrices*, and it is understood that the matrix elements of A are taken with respect to the basis $|i_R\rangle$ and those for B with respect to the basis $|i_Q\rangle$.

Solution

Concepts Involved: Trace, Tensor Products, Maximally Entangled State, Hilbert-Schmidt Inner Product.

Let $\mathcal{H}_R \cong \mathcal{H}_Q \cong \mathbb{C}^d$ be Hilbert spaces for systems R and Q, with orthonormal bases $\{|i_R\rangle\}$ and $\{\left|i_Q\right\rangle\}$. Define the (unnormalized) maximally entangled state:

$$|m\rangle = \sum_{i=1}^{d} |i_R\rangle \otimes \left|i_Q\right\rangle \in \mathcal{H}_R \otimes \mathcal{H}_Q.$$

Let A be an operator on R, and B an operator on Q. Then:

$$\begin{aligned}
\langle m| (A \otimes B) |m\rangle &= \left(\sum_i \langle i_R| \otimes \left\langle i_Q\right|\right) (A \otimes B) \left(\sum_j |j_R\rangle \otimes \left|j_Q\right\rangle\right) \\
&= \sum_{i,j} \langle i_R| A |j_R\rangle \cdot \left\langle i_Q\right| B \left|j_Q\right\rangle \\
&= \sum_{i,j} A_{ij} B_{ij}^* \\
&= \sum_{i,j} (A^\dagger)_{ji} B_{ij} \\
&= \operatorname{tr}\left(A^\dagger B\right).
\end{aligned}$$

□

Remark: This relation plays a key role in the Choi–Jamiołkowski isomorphism and quantum process tomography.

Exercise 9.17

Show that $0 \leq A(\rho, \sigma) \leq \pi/2$, with equality in the first inequality if and only if $\rho = \sigma$.

Solution

Concepts Involved: Density Operators, Angle Between States, Fidelity

The angle between quantum states ρ and σ is defined as:

$$A(\rho, \sigma) = \arccos F(\rho, \sigma),$$

where $F(\rho, \sigma) = \left(\operatorname{tr} \sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}\right)^2$ is the fidelity.

Step 1: Range of Fidelity

By definition, fidelity satisfies:

$$0 \leq F(\rho, \sigma) \leq 1,$$

with equality $F = 1$ if and only if $\rho = \sigma$, and $F = 0$ if and only if ρ and σ have orthogonal support.

Step 2: Apply arccos to the range

Since $\arccos : [0, 1] \to [0, \pi/2]$ is decreasing, we have:

$$0 \leq A(\rho, \sigma) \leq \frac{\pi}{2}.$$

Step 3: Equality conditions

- $A(\rho, \sigma) = 0 \iff F(\rho, \sigma) = 1 \iff \rho = \sigma,$
- $A(\rho, \sigma) = \frac{\pi}{2} \iff F(\rho, \sigma) = 0 \iff$ the supports of ρ and σ are orthogonal.

Thus, the first inequality is saturated if and only if $\rho = \sigma$. □

Remark: This shows that $A(\rho, \sigma)$ behaves like a proper distance measure (though not a metric), vanishing only when the states are identical and attaining its maximum when they are perfectly distinguishable.

Exercise 9.18: Contractivity of the angle

Let $\mathcal{E}$ be a trace-preserving quantum operation. Show that

$$A(\mathcal{E}(\rho), \mathcal{E}(\sigma)) \leq A(\rho, \sigma)$$

Solution

Concepts Involved: Density Operators, Angle Between States, Fidelity

---

We are given a trace-preserving quantum operation $\mathcal{E}$. Define the quantum angle as

$$A(\rho, \sigma) = \arccos F(\rho, \sigma).$$

Fidelity is monotonic under CPTP maps:

$$F(\mathcal{E}(\rho), \mathcal{E}(\sigma)) \geq F(\rho, \sigma).$$

Since $\arccos(x)$ is decreasing on $[0, 1]$, applying it gives

$$A(\mathcal{E}(\rho), \mathcal{E}(\sigma)) \leq A(\rho, \sigma).$$

□

Exercise 9.19: Joint concavity of fidelity

Prove that fidelity is *jointly concave*,

$$F\left(\sum_i p_i\rho_i, \sum_i p_i\sigma_i\right) \geq \sum_i p_i F(p_i, \sigma_i).$$

Solution

Concepts Involved: Fidelity, Concavity, Joint Concavity, Uhlmann's Theorem, Purifications, Jensen's inequality.

Let $\{p_i\}$ be a probability distribution, and let ρ_i and σ_i be density operators. We aim to prove:

$$F\left(\sum_i p_i\rho_i,\ \sum_i p_i\sigma_i\right) \geq \sum_i p_i F(\rho_i, \sigma_i).$$

By Uhlmann's theorem, the fidelity satisfies:

$$F(\rho, \sigma) = \max_{|\psi\rangle, |\varphi\rangle} |\langle\psi|\varphi\rangle|,$$

where the maximization is over all purifications $|\psi\rangle$ of ρ and $|\varphi\rangle$ of σ.
For each pair (ρ_i, σ_i), let $|\psi_i\rangle$ and $|\varphi_i\rangle$ be purifications such that

$$F(\rho_i, \sigma_i) = |\langle\psi_i|\varphi_i\rangle|.$$

Define the following purifications:

$$|\Psi\rangle = \sum_i \sqrt{p_i}\,|\psi_i\rangle \otimes |i\rangle\,, \quad |\Phi\rangle = \sum_i \sqrt{p_i}\,|\varphi_i\rangle \otimes |i\rangle\,,$$

where $\{|i\rangle\}$ is an orthonormal basis for an auxiliary system.
Then $|\Psi\rangle$ and $|\Phi\rangle$ are purifications of $\sum_i p_i\rho_i$ and $\sum_i p_i\sigma_i$, respectively. Thus,

$$\begin{aligned} F\left(\sum_i p_i\rho_i, \sum_i p_i\sigma_i\right) &\geq |\langle\Psi|\Phi\rangle| \\ &= \left|\sum_i p_i \langle\psi_i|\varphi_i\rangle\right| \\ &\geq \sum_i p_i |\langle\psi_i|\varphi_i\rangle| \\ &= \sum_i p_i F(\rho_i, \sigma_i), \end{aligned}$$

where the second inequality follows from the triangle inequality.

□

Remark: Alternatively, if we use the strong concavity property the result is immediate:

$$F\left(\sum_i p_i\rho_i, \sum_i p_i\sigma_i\right) \geq \sum_i \sqrt{p_ip_i}F(\rho_i,\sigma_i) = \sum_i p_iF(\rho_i,\sigma_i).$$

This inequality shows that averaging quantum states cannot increase their distinguishability, making fidelity a robust tool under mixing and quantum operations.

Exercise 9.20: Concavity of fidelity

Prove that fidelity is concave in the first entry,

$$F\left(\sum_i p_i\rho_i, \sigma\right) \geq \sum_i p_iF(\rho_i,\sigma).$$

By symmetry the fidelity is also concave in the second entry.

Solution

Concepts Involved: Fidelity, Concavity, Uhlmann's theorem, Purifications, Triangle Inequality

Let $\{p_i\}$ be a probability distribution and fix a purification $|\varphi\rangle$ of σ. For each ρ_i, choose a purification $|\psi_i\rangle$ such that

$$F(\rho_i,\sigma) = |\langle\psi_i|\varphi\rangle|.$$

Define

$$|\Psi\rangle = \sum_i \sqrt{p_i}\,|\psi_i\rangle \otimes |i\rangle\,, \quad |\Phi\rangle = |\varphi\rangle \otimes \sum_i \sqrt{p_i}\,|i\rangle\,,$$

where $\{|i\rangle\}$ is an orthonormal basis of an auxiliary system. Then $|\Psi\rangle$ purifies $\sum_i p_i\rho_i$ and $|\Phi\rangle$ purifies σ, so by Uhlmann's theorem,

$$\begin{aligned} F\left(\sum_i p_i\rho_i,\ \sigma\right) &\geq |\langle\Psi|\Phi\rangle| \\ &= \left|\sum_i p_i\,\langle\psi_i|\varphi\rangle\right| \\ &\geq \sum_i p_i|\langle\psi_i|\varphi\rangle| \\ &= \sum_i p_iF(\rho_i,\sigma). \end{aligned}$$

□

Remark: Alternatively, the result immediately follows from the joint concavity:

$$F\left(\sum_i p_i\rho_i, \sigma\right) = F\left(\sum_i p_i\rho_i, \sum_i p_i\sigma\right) \geq \sum_i p_i F(\rho_i, \sigma).$$

Exercise 9.21

When comparing pure states and mixed states it is possible to make a stronger statement than (9.110) about the relationship between trace distance and fidelity. Prove that

$$1 - F(|\psi\rangle, \sigma)^2 \leq D(|\psi\rangle, \sigma).$$

Solution

Concepts Involved: Pure States, Mixed States, Trace Distance, Fidelity

---

Let $|\psi\rangle$ be a pure state and σ a density operator. The fidelity is given by

$$F(|\psi\rangle, \sigma) = \sqrt{\langle\psi|\sigma|\psi\rangle}.$$

From the classical fidelity bound (Eq. 9.108),

$$1 - F(|\psi\rangle, \sigma) \leq D(|\psi\rangle, \sigma),$$

where D is the trace distance. Squaring both sides gives

$$\big(1 - F(|\psi\rangle, \sigma)\big)^2 \leq D(|\psi\rangle, \sigma)^2.$$

Expanding the left-hand side:

$$1 - 2F + F^2 \leq D^2,$$

which implies

$$1 - F^2 \leq D^2 + 2F(1 - F) \leq D,$$

since $F \in [0, 1]$ and $D \leq 1$. Therefore,

$$1 - F(|\psi\rangle, \sigma)^2 \leq D(|\psi\rangle, \sigma).$$

$\square$

Exercise 9.22: Chaining property for fidelity measures

Suppose U and V are unitary operators, and $\mathcal{E}$ and $\mathcal{F}$ are trace-preserving quantum operations meant to approximate U and V. Letting $d(\cdot,\cdot)$ be any metric on the space of density matrices satisfying $d(U\rho U^\dagger, U\sigma U^\dagger) = d(\rho,\sigma)$ for all density matrices ρ and σ and unitary U (such as the angle $\arccos\big(F(\rho,\sigma)\big)$), define the corresponding *error* $E(U,\mathcal{E})$ by

$$E(U,\mathcal{E}) \equiv \max_\rho d(U\rho U^\dagger, \mathcal{E}(\rho))$$

and show that $E(VU, \mathcal{F}\circ\mathcal{E}) \leq E(U,\mathcal{E}) + E(V,\mathcal{F})$. Thus, to perform a quantum computation with high fidelity it suffices to complete each step of the computation with high fidelity.

Solution

Concepts Involved: Unitary Operators, Fidelity, Triangle Inequality

Let $d(\cdot,\cdot)$ be a unitary-invariant metric on quantum states, meaning

$$d(U\rho U^\dagger, U\sigma U^\dagger) = d(\rho,\sigma)$$

for all unitaries U and density operators ρ, σ. Define the approximation error as

$$E(U,\mathcal{E}) := \max_\rho d(U\rho U^\dagger, \mathcal{E}(\rho)).$$

We aim to prove that

$$E(VU, \mathcal{F}\circ\mathcal{E}) \leq E(U,\mathcal{E}) + E(V,\mathcal{F}).$$

Let ρ be any density matrix. Then

$$\begin{aligned} d(VU\rho U^\dagger V^\dagger, (\mathcal{F}\circ\mathcal{E})(\rho)) &= d(V(U\rho U^\dagger)V^\dagger, \mathcal{F}(\mathcal{E}(\rho))) \\ &= d(V\sigma V^\dagger, \mathcal{F}(\tau)), \end{aligned}$$

where we set $\sigma := U\rho U^\dagger$, $\tau := \mathcal{E}(\rho)$.
By the triangle inequality and unitary invariance,

$$\begin{aligned} d(V\sigma V^\dagger, \mathcal{F}(\tau)) &\leq d(V\sigma V^\dagger, \mathcal{F}(\sigma)) + d(\mathcal{F}(\sigma), \mathcal{F}(\tau)) \\ &= d(\sigma, \mathcal{F}^\dagger\mathcal{F}(\sigma)) + d(\sigma,\tau) \\ &\leq \max_\sigma d(V\sigma V^\dagger, \mathcal{F}(\sigma)) + \max_\rho d(U\rho U^\dagger, \mathcal{E}(\rho)) \\ &= E(V,\mathcal{F}) + E(U,\mathcal{E}). \end{aligned}$$

Maximizing over ρ, we conclude

$$E(VU, \mathcal{F}\circ\mathcal{E}) \leq E(U,\mathcal{E}) + E(V,\mathcal{F}).$$

$\square$

Exercise 9.23

show that $\bar{F} = 1$ if and only if $\mathcal{E}(\rho_j) = \rho_j$ for all j such that $p_j > 0$.

Solution

Concepts Involved: Average Fidelity, Density Operators

Let $\bar{F} := \sum_j p_j F(\rho_j, \mathcal{E}(\rho_j))$ denote the average fidelity of a quantum operation $\mathcal{E}$ on an ensemble $\{p_j, \rho_j\}$. We will prove

$$\bar{F} = 1 \quad \iff \quad \mathcal{E}(\rho_j) = \rho_j \quad \text{for all } j \text{ such that } p_j > 0.$$

(⇒) Suppose $\bar{F} = 1$.
Since fidelity satisfies $0 \leq F(\rho_j, \mathcal{E}(\rho_j)) \leq 1$, and all $p_j \geq 0$, the only way the weighted sum $\bar{F}$ can equal 1 is if:

$$F(\rho_j, \mathcal{E}(\rho_j)) = 1 \quad \text{whenever } p_j > 0.$$

But fidelity is 1 if and only if the two states are equal. Hence,

$$\mathcal{E}(\rho_j) = \rho_j \quad \text{for all } j \text{ with } p_j > 0.$$

(⇐) Suppose $\mathcal{E}(\rho_j) = \rho_j$ for all j such that $p_j > 0$.
Then for each such j, we have

$$F(\rho_j, \mathcal{E}(\rho_j)) = F(\rho_j, \rho_j) = 1,$$

so the average fidelity is

$$\bar{F} = \sum_j p_j \cdot 1 = 1.$$

□

Problem 9.1: Alternate characterization of the fidelity

Show that

$$F(\rho, \sigma) = \inf_P \operatorname{tr}(\rho P) \operatorname{tr}\left(\sigma P^{-1}\right),$$

where the infimum is taken over all invertible positive matrices P.

Solution

Concepts Involved: Fidelity, Density Operators, Positive Operators

Recall that

$$F(\rho, \sigma) \;=\; \|\sqrt{\rho}\sqrt{\sigma}\|_1^2 = \left(\operatorname{Tr}\sqrt{\sqrt{\rho}\,\sigma\,\sqrt{\rho}}\right)^2.$$

For any $P > 0$ and any unitary U,

$$|\operatorname{Tr}(\sqrt{\rho}\sqrt{\sigma}U)| = |\operatorname{Tr}\left(\sqrt{P}\sqrt{\rho}\,P^{-1/2}\sqrt{\sigma}U\right)| \leq \|\sqrt{P}\sqrt{\rho}\|_2\,\|\sqrt{P^{-1}}\sqrt{\sigma}U\|_2,$$

where $\|\cdot\|_2$ is the Hilbert–Schmidt norm. This gives

$$|\operatorname{Tr}(\sqrt{\rho}\sqrt{\sigma}U)| \leq \sqrt{\operatorname{Tr}(\rho P)}\sqrt{\operatorname{Tr}(\sigma P^{-1})}.$$

Taking the supremum over U yields

$$\operatorname{Tr}\sqrt{\sqrt{\rho}\,\sigma\,\sqrt{\rho}} \;\leq\; \sqrt{\operatorname{Tr}(\rho P)\,\operatorname{Tr}(\sigma P^{-1})}.$$

Squaring and infimizing over $P > 0$ gives

$$F(\rho,\sigma) \;\leq\; \inf_{P>0} \operatorname{Tr}(\rho P)\operatorname{Tr}\Big(\sigma P^{-1}\Big).$$

For the reverse inequality, define

$$X := \sqrt{\sqrt{\rho}\,\sigma\,\sqrt{\rho}}, \qquad P_\star := \rho^{-1/2}\,X\,\rho^{-1/2}.$$

Then

$$\operatorname{Tr}(\rho P_\star) = \operatorname{Tr}(X), \qquad \operatorname{Tr}\Big(\sigma P_\star^{-1}\Big) = \operatorname{Tr}(X).$$

Thus

$$\operatorname{Tr}(\rho P_\star)\operatorname{Tr}\Big(\sigma P_\star^{-1}\Big) = (\operatorname{Tr}X)^2 = F(\rho,\sigma).$$

Hence the infimum is attained at $P_\star$, and we obtain the equality

$$F(\rho,\sigma) = \inf_{P>0} \operatorname{Tr}(\rho P)\operatorname{Tr}\Big(\sigma P^{-1}\Big).$$

□

Problem 9.2

Let $\mathcal{E}$ be a trace-preserving quantum operation. Show that for each ρ there is a set of operation elements $\{E_i\}$ for $\mathcal{E}$ such that

$$F(\rho,\mathcal{E}) = \big|\operatorname{tr}(\rho E_1)\big|^2.$$

Solution

Concepts Involved: Kraus Representation, Fidelity, Density Operators

---

Let $\mathcal{E}$ be a quantum operation with Kraus decomposition $\mathcal{E}(\rho) = \sum_i E_i \rho E_i^\dagger$. The fidelity between ρ and $\mathcal{E}(\rho)$ is given by

$$F(\rho,\mathcal{E}) = \left(\max_{|\psi\rangle,|\varphi\rangle} |\langle\psi|\varphi\rangle|\right)^2,$$

where $|\psi\rangle$ and $|\varphi\rangle$ are purifications of ρ and $\mathcal{E}(\rho)$, respectively. By Uhlmann's theorem,

$$F(\rho,\mathcal{E}) = \max_U \big|\langle\psi|(I\otimes U)|\Phi\rangle\big|^2,$$

where U ranges over unitaries on the environment, and $|\Phi\rangle$ is a purification of $\mathcal{E}(\rho)$.
Using the Stinespring dilation $\mathcal{E}(\rho) = \mathrm{Tr}_E\left[U(\rho \otimes |0\rangle\langle 0|)U^\dagger\right]$, fix a purification $|\psi\rangle = \sum_j \sqrt{p_j}|j\rangle \otimes |j\rangle$ of ρ. Then a purification of $\mathcal{E}(\rho)$ is

$$|\varphi\rangle = (I \otimes U)(|\psi\rangle \otimes |0\rangle).$$

Therefore,

$$F(\rho, \mathcal{E}) = \left|\langle\psi| \otimes \langle 0|U|\psi\rangle \otimes |0\rangle\right|^2.$$

Choose the Kraus representation $E_i = \langle i|U|0\rangle$ by fixing an orthonormal basis $\{|i\rangle\}$ for the environment. Define $E_1 = \langle 0|U|0\rangle$, then we have

$$F(\rho, \mathcal{E}) = |\mathrm{tr}(\rho E_1)|^2.$$

□

Problem 9.3

Prove fact (5) on this page: Suppose that $\langle\psi|\mathcal{E}(|\psi\rangle\langle\psi|)|\psi\rangle \geq 1 - \eta$ for all $|\psi\rangle$ in the support of ρ for some η. Then $F(\rho, \mathcal{E}) \geq 1 - (3\eta/2)$.

Solution

Concepts Involved: Density Operators, Fidelity, Purification, Spectral Decomposition

---

Let $\rho = \sum_k \lambda_k |k\rangle\langle k|$ be the spectral decomposition and

$$|\Psi\rangle = \sum_k \sqrt{\lambda_k}\, |k\rangle_R \otimes |k\rangle_Q$$

a purification. Then

$$F(\rho, \mathcal{E}) = \sum_{k,\ell} \lambda_k \lambda_\ell\, \langle \ell|\mathcal{E}(|k\rangle\langle \ell|)|\ell\rangle.$$

To control the cross terms, consider

$$|\psi(\boldsymbol{\varphi})\rangle = \sum_k \sqrt{\lambda_k}\, e^{i\varphi_k} |k\rangle,$$

with arbitrary phases $\varphi_k \in [0, 2\pi)$. By hypothesis,

$$\langle\psi(\boldsymbol{\varphi})|\, \mathcal{E}(|\psi(\boldsymbol{\varphi})\rangle\langle\psi(\boldsymbol{\varphi})|)\, |\psi(\boldsymbol{\varphi})\rangle \;\geq\; 1 - \eta.$$

Expanding in the $\{|k\rangle\}$ basis and averaging uniformly over all phases eliminates all oscillatory terms, leaving

$$\overline{f} = F(\rho, \mathcal{E}) + \sum_k \sum_{m \neq k} \lambda_k \lambda_m\, \langle m|\mathcal{E}(|k\rangle\langle k|)|m\rangle \;\geq\; 1 - \eta.$$

Thus,

$$F(\rho, \mathcal{E}) \;\geq\; 1 - \eta - \sum_{k} \sum_{m \neq k} \lambda_k \lambda_m \, \langle m|\mathcal{E}(|k\rangle\langle k|)|m\rangle. \tag{7}$$

Now we use trace preservation. For each k,

$$\sum_{m} \langle m|\mathcal{E}(|k\rangle\langle k|)|m\rangle = 1,$$

and by hypothesis $\langle k|\mathcal{E}(|k\rangle\langle k|)|k\rangle \geq 1 - \eta$, so

$$\sum_{m \neq k} \langle m|\mathcal{E}(|k\rangle\langle k|)|m\rangle \leq \eta.$$

Next, we order the eigenvalues so that $\lambda_1 \geq \lambda_2 \geq \cdots$. Split the double sum in (9) into two pieces
- For $k = 1$,

$$\sum_{m \neq 1} \lambda_1 \lambda_m \langle m|\mathcal{E}(|1\rangle\langle 1|)|m\rangle \;\leq\; \lambda_1 \lambda_2 \, \eta.$$

- For $k \neq 1$,

$$\sum_{m \neq k} \lambda_k \lambda_m \langle m|\mathcal{E}(|k\rangle\langle k|)|m\rangle \;\leq\; \lambda_k \lambda_1 \, \eta. \tag{8}$$

Summing over $k \neq 1$ gives

$$\sum_{k \neq 1} \sum_{m \neq k} \lambda_k \lambda_m \langle m|\mathcal{E}(|k\rangle\langle k|)|m\rangle \;\leq\; (1 - \lambda_1)\lambda_1 \, \eta.$$

Therefore,

$$F(\rho, \mathcal{E}) \;\geq\; 1 - \Big(1 + \lambda_1 \lambda_2 + (1 - \lambda_1)\lambda_1\Big)\eta.$$

For fixed λ_1, the RHS is minimized at $\lambda_2 = 1 - \lambda_1$, yielding

$$F(\rho, \mathcal{E}) \;\geq\; 1 - \big(1 + 2\lambda_1(1 - \lambda_1)\big)\eta.$$

The factor $1 + 2\lambda_1(1 - \lambda_1)$ attains its maximum at $\lambda_1 = \frac{1}{2}$, where it equals $\frac{3}{2}$. Thus

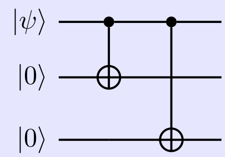
$$F(\rho, \mathcal{E}) \;\geq\; 1 - \tfrac{3}{2} \, \eta.$$

$\square$

10 Quantum error-correction

Exercise 10.1

Verify that the encoding circuit in Figure 10.2 works as claimed.



that is, it encodes $|\psi\rangle = a\,|0\rangle + b\,|1\rangle$ to $a\,|000\rangle + b\,|111\rangle$.

Solution

Concepts Involved: Bit Flip Code, Controlled Operations

We calculate:

$$\begin{aligned}
|\psi\rangle \otimes |0\rangle \otimes |0\rangle &\mapsto CX_{1,3}CX_{1,2}\,|\psi\rangle \otimes |0\rangle \otimes |0\rangle \\
&= aCX_{1,3}CX_{1,2}\,|000\rangle + bCX_{1,3}CX_{1,2}\,|100\rangle \\
&= a\,|000\rangle + b\,|111\rangle
\end{aligned}$$

where we use the definition of the controlled-X gate $CX_{1,2}\,|0\rangle \otimes |\varphi\rangle = |0\rangle \otimes |\varphi\rangle$ and $CX_{1,2}\,|1\rangle \otimes |\varphi\rangle = |1\rangle \otimes X_2\,|\varphi\rangle$ in the last line. $\square$

Exercise 10.2

The action of the bit flip channel can be described by the quantum operation $\mathcal{E}(\rho) = (1-p)\rho + pX\rho X$. Show that this may be given an alternative operator-sum representation, as $\mathcal{E}(\rho) = (1-2p)\rho + 2pP_+\rho P_+ + 2pP_-\rho P_-$ where P_+ and P_- are projectors onto the $+1$ and -1 eigenstates of X, $(|0\rangle + |1\rangle)/\sqrt{2}$ and $(|0\rangle - |1\rangle)/\sqrt{2}$, respectively. This latter representation can be understood as a model in which the qubit is left alone with probability $1-2p$, and is 'measured' by the environment in the $|+\rangle, |-\rangle$ basis with probability $2p$.

Solution

Concepts Involved: Bit Flip Channel, Projectors, Spectral Decomposition

First, note that we may write:

$$P_+ = |+\rangle\langle+| = \frac{|+\rangle\langle+| + |-\rangle\langle-| + |+\rangle\langle+| - |-\rangle\langle-|}{2} = \frac{I + X}{2}$$

where we have used $X = |+\rangle\langle+| - |-\rangle\langle-|$ as the spectral decomposition of X and $I = |+\rangle\langle+| + |-\rangle\langle-|$ the resolution of the identity. We can obtain the analogous relation for P_-:

$$P_- = \frac{I - X}{2}$$

Therefore we can expand:

$$\begin{aligned}\mathcal{E}(\rho) &= (1-2p)\rho + 2pP_+\rho P_+ + 2pP_-\rho P_- \\ &= (1-2p)\rho + 2p\frac{I+X}{2}\rho\frac{I+X}{2} + 2p\frac{I-X}{2}\rho\frac{I-X}{2} \\ &= (1-2p)\rho + \frac{p}{2}(\rho + X\rho + \rho X + X\rho X) + \frac{p}{2}(\rho - X\rho - \rho X + X\rho X) \\ &= (1-p)\rho + pX\rho X\end{aligned}$$

which proves the claim. □

Exercise 10.3

Show by explicit calculation that measuring Z_1Z_2 followed by Z_2Z_3 is equivalent, up to labeling of the measurement outcomes, to measuring the four projectors defined by (10.5)–(10.8), in the sense that both procedures result in the same measurement statistics and post-measurement states.

Solution

Concepts Involved: Bit Flip Code, Quantum Measurement, Projectors, Spectral Decomposition.
The four projection operators corresponding to the four error syndromes of the three-qubit repetition code are given by:

$$\begin{aligned}P_0 &\equiv |000\rangle\langle 000| + |111\rangle\langle 111| \\ P_1 &\equiv |100\rangle\langle 100| + |011\rangle\langle 011| \\ P_2 &\equiv |010\rangle\langle 010| + |101\rangle\langle 101| \\ P_3 &\equiv |001\rangle\langle 001| + |110\rangle\langle 110|\end{aligned}$$

It suffices to show that the composition of projectors corresponding to measurements of Z_1Z_2 and Z_2Z_3 yield the same four projectors as the above. Z_1Z_2 has spectral decomposition:

$$Z_1Z_2 = (|00\rangle\langle 00| + |11\rangle\langle 11|)\otimes I - (|01\rangle\langle 01| + |10\rangle\langle 10|)\otimes I$$

which corresponds to a projective measurement with projectors:

$$\begin{aligned}P_{Z_1Z_2=+1} &= (|00\rangle\langle 00| + |11\rangle\langle 11|)\otimes I \\ P_{Z_1Z_2=-1} &= (|01\rangle\langle 01| + |10\rangle\langle 10|)\otimes I\end{aligned}$$

analogously, Z_2Z_3 has spectral decomposition:

$$Z_2Z_3 = I\otimes(|00\rangle\langle 00| + |11\rangle\langle 11|) - I\otimes(|01\rangle\langle 01| + |10\rangle\langle 10|)$$

which corresponds to a projective measurement with projectors:

$$\begin{aligned}P_{Z_2Z_3=+1} &= I\otimes(|00\rangle\langle 00| + |11\rangle\langle 11|) \\ P_{Z_2Z_3=-1} &= I\otimes(|01\rangle\langle 01| + |10\rangle\langle 10|).\end{aligned}$$

Now, we observe that:

$$P_{Z_1Z_2=+1}P_{Z_2Z_3=+1} = [(|00\rangle\langle 00| + |11\rangle\langle 11|) \otimes I][I \otimes (|00\rangle\langle 00| + |11\rangle\langle 11|)] = |000\rangle\langle 000| + |111\rangle\langle 111| = P_0$$

$$P_{Z_1Z_2=-1}P_{Z_2Z_3=+1} = [(|01\rangle\langle 01| + |10\rangle\langle 10|) \otimes I][I \otimes (|00\rangle\langle 00| + |11\rangle\langle 11|)] = |011\rangle\langle 011| + |100\rangle\langle 100| = P_1$$

$$P_{Z_1Z_2=-1}P_{Z_2Z_3=-1} = [(|01\rangle\langle 01| + |10\rangle\langle 10|) \otimes I][I \otimes (|01\rangle\langle 01| + |10\rangle\langle 10|)] = |010\rangle\langle 010| + |101\rangle\langle 101| = P_2$$

$$P_{Z_1Z_2=+1}P_{Z_2Z_3=-1} = [(|00\rangle\langle 00| + |11\rangle\langle 11|) \otimes I][I \otimes (|01\rangle\langle 01| + |10\rangle\langle 10|)] = |001\rangle\langle 001| + |110\rangle\langle 110| = P_3$$

so the claim is proven. □

Exercise 10.4

Consider the three qubit bit flip code. Suppose we had performed the error syndrome measurement by measuring the eight orthogonal projectors corresponding to projections onto the eight computational basis states.

(1) Write out the projectors corresponding to this measurement, and explain how the measurement result can be used to diagnose the error syndrome: either *no bits flipped* or *bit number j flipped*, where j is in the range one to three.

(2) Show that the recovery procedure works only for computational basis states.

(3) What is the minimum fidelity for the error-correction procedure?

Solution

Concepts Involved: Bit Flip Code, Projectors, Error Syndrome and Recovery, Fidelity

---

(1) The projectors are:

$$|000\rangle\langle 000|, |001\rangle\langle 001|, |010\rangle\langle 010|, |100\rangle\langle 100|, |011\rangle\langle 011|, |101\rangle\langle 101|, |110\rangle\langle 110|, |111\rangle\langle 111|.$$

Since the encoded state is $|\psi\rangle = a|000\rangle + b|111\rangle$, if we measure one of $|000\rangle, |111\rangle$ we would conclude no bits have been flipped. If we measure $|001\rangle$ or $|110\rangle$, we would conclude that bit 1 was flipped (and so the state had become $a|001\rangle + b|110\rangle$. If we measure $|010\rangle$ or $|101\rangle$, then we would conclude that bit 2 was flipped (and so the state had become $a|010\rangle + b|101\rangle$). Finally, if we measure $|100\rangle$ or $|011\rangle$, we would conclude that bit 3 was flipped (and so the state had become $a|100\rangle + b|011\rangle$).

(2) The recovery procedure involves doing nothing if the error syndrome tells us that no bits are flipped, or flipping the jth bit if the jth bit was flipped. In the no bit flip case, after recovery we have $|000\rangle \to |000\rangle, |111\rangle \to |111\rangle$. In the case we conclude the first bit was flipped, after recovery we have $|001\rangle \to |000\rangle, |110\rangle \to |111\rangle$. In the case we conclude the second bit was flipped, after recovery we have $|010\rangle \to |000\rangle, |101\rangle \to |111\rangle$. Finally if we conclude that the third bit was flipped,

after recovery we have $|100\rangle \to |000\rangle, |011\rangle \to |111\rangle$. In all cases, the post-recovery state is one of the computational basis states $|000\rangle$ / $|111\rangle$, so the recovery only succeeds if the initial state was one of the computational basis states.

(3) Supposing we use the three qubit error-correcting code to protect $|\psi\rangle = a|0\rangle + b|1\rangle$. The encoded state is $|\Psi\rangle = a|000\rangle + b|111\rangle$. After applying the noise channel (i.e. each bit flips with probability p), the state we have is:

$$\begin{aligned}\mathcal{E}(\rho_\Psi) = &\left(|a|^2(1-p)^3 + |b|^2p^3\right)|000\rangle\langle 000| \\ &+ \left(|a|^2p(1-p)^2 + |b|^2p^2(1-p)\right)\left(|001\rangle\langle 001| + |010\rangle\langle 010| + |100\rangle\langle 100|\right) \\ &+ \left(|b|^2p(1-p)^2 + |a|^2p^2(1-p)\right)\left(|110\rangle\langle 110| + |101\rangle\langle 101| + |011\rangle\langle 011|\right) \\ &+ \left(|b|^2(1-p)^3 + |a|^2p^3\right)|111\rangle\langle 111|\end{aligned}$$

The recovery procedure maps any state with ≥ 2 zeros to $|000\rangle\langle 000|$ (and hence back to $|0\rangle\langle 0|$ after decoding) and any state with ≥ 2 ones to $|111\rangle\langle 111|$ (and hence back to $|1\rangle\langle 1|$ after decoding), so the final state is:

$$\begin{aligned}\rho \coloneqq \mathcal{R}(\mathcal{E}(\rho_\Psi)) = &\left(|a|^2\left[(1-p)^3 + 3p(1-p)^2\right] + |b|^2\left[p^3 + 3p^2(1-p)\right]\right)|0\rangle\langle 0| \\ &+ \left(|b|^2\left[(1-p)^3 + 3p(1-p)^2\right] + |a|^2\left[p^3 + 3p^2(1-p)\right]\right)|1\rangle\langle 1|\end{aligned}$$

To find the minimum fidelity, it suffices to minimize $\langle\psi|\rho|\psi\rangle$, which we compute to be:

$$\begin{aligned}\langle\psi|\rho|\psi\rangle = &|a|^2\left(|a|^2\left[(1-p)^3 + 3p(1-p)^2\right] + |b|^2\left[p^3 + 3p^2(1-p)\right]\right) \\ &+ |b|^2\left(|b|^2\left[(1-p)^3 + 3p(1-p)^2\right] + |a|^2\left[p^3 + 3p^2(1-p)\right]\right)\end{aligned}$$

With the normalization constraint on the initial state of $|a|^2 + |b|^2 = 1$, we can rewrite the above in terms of of $|a|^2$ alone:

$$\langle\psi|\rho|\psi\rangle = \left(2(|a|^2)^2 - 2|a|^2 + 1\right)\left[(1-p)^3 + 3p(1-p)^2\right] + \left(2|a|^2 - 2(|a|^2)^2\right)\left[p^3 + 3p^2(1-p)\right]$$

We take the derivative of the above w.r.t. $|a|^2$ and set it to zero to find the minimzing value of $|a|^2$:

$$\frac{\partial}{\partial|a|^2}\langle\psi|\rho|\psi\rangle = (4|a|^2 - 2)\left[(1-p)^3 + 3p(1-p)^2 - p^3 + 3p^2(1-p)\right] = 0$$

Which for any value of p is satisfied when $|a|^2 = \frac{1}{2}$, i.e. $|a| = \frac{1}{\sqrt{2}}$ (and so $|b| = \frac{1}{\sqrt{2}}$ as well). So, the fidelity is minimized for $|\psi\rangle = \frac{1}{\sqrt{2}}(e^{i\varphi_0}|0\rangle + e^{i\varphi_1}|1\rangle)$ (which is what we might have expected - given that the error correction succeeds only for the computational basis states, we should have the worst fidelity for states which are the furthest from both). For these states, the fidelity is (plugging into our former general expression):

$$F_{\text{min}} = \sqrt{\langle\psi|\rho|\psi\rangle} = \sqrt{\frac{(1-p)^3 + 3p(1-p)^2 + 3p^2(1-p) + p^3}{2}} = \frac{1}{\sqrt{2}}$$

□

Exercise 10.5

Show that the syndrome measurement for detecting phase flip errors in the Shor code corresponds to measuring the observables $X_1X_2X_3X_4X_5X_6$ and $X_4X_5X_6X_7X_8X_9$.

Solution

Concepts Involved: Shor Code, Eigenvalues, Eigenvectors, Error Syndrome

---

We have the codewords:

$$|0_L\rangle = \frac{(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)}{2\sqrt{2}}$$
$$|1_L\rangle = \frac{(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)}{2\sqrt{2}}$$

A phase flip error on a given block amounts to flipping the phase of the block:

$$Z_i(|000\rangle \pm |111\rangle) = |000\rangle \mp |111\rangle$$

Measuring $X_{1-2} = X_1X_2X_3X_4X_5X_6$ compares the phase of the first and second blocks (with $+1$ if they are the same, -1 if they are different) - we can see this from the eigenvalue relations:

$$\begin{aligned}
X_1X_2X_3X_4X_5X_6(|000\rangle + |111\rangle)(|000\rangle + |111\rangle) &= +1(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)\\
X_1X_2X_3X_4X_5X_6(|000\rangle - |111\rangle)(|000\rangle + |111\rangle) &= +1(|111\rangle - |000\rangle)(|000\rangle + |111\rangle)\\
&= -1(|000\rangle - |111\rangle)(|000\rangle + |111\rangle)\\
X_1X_2X_3X_4X_5X_6(|000\rangle + |111\rangle)(|000\rangle - |111\rangle) &= +1(|000\rangle + |111\rangle)(|111\rangle - |000\rangle)\\
&= -1(|000\rangle + |111\rangle)(|000\rangle - |111\rangle)\\
X_1X_2X_3X_4X_5X_6(|000\rangle - |111\rangle)(|000\rangle - |111\rangle) &= +1(|111\rangle - |000\rangle)(|111\rangle - |000\rangle)\\
&= +1(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)
\end{aligned}$$

analogously, measuring $X_{2-3} = X_4X_5X_6X_7X_8X_9$ compares the phase of the second and third blocks. The codewords have all three blocks with the same phase, so if we measure $X_{1-2} = X_{2-3} = +1$ then we conclude no phase flip error occured. If we measure $X_{1-2} = +1$ and $X_{2-3} = -1$ then we conclude that a phase flip must have occured in the third block (one of qubits 7/8/9). If we measure $X_{1-2} = -1$ and $X_{2-3} = +1$ then we conclude that a phase flip occured on the first block (one of qubits 1/2/3). If we measure $X_{1-2} = X_{2-3} = -1$ then we conclude that a phase flip error occured on the second block (one of qubits 4/5/6). We thus conclude that measuring these two operators yields the syndrome for detecting phase flip errors. □

Exercise 10.6

Show that recovery from a phase flip on any of the first three qubits may be accomplished by applying the operator $Z_1Z_2Z_3$.

Solution

Concepts Involved: Shor Code, Error Recovery

We have the encoded state:

$$\begin{aligned}
|\psi_L\rangle &= \alpha|0_L\rangle + \beta|1_L\rangle \\
&= \alpha\frac{(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)}{2\sqrt{2}} + \beta\frac{(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)}{2\sqrt{2}}
\end{aligned}$$

We saw that $Z_i(|000\rangle \pm |111\rangle) = |000\rangle \mp |111\rangle$ for $i \in \{1,2,3\}$, so if a phase flip error occurs on the first qubit, we have:

$$|\psi_L\rangle \xrightarrow{\mathcal{E}} \alpha\frac{(|000\rangle - |111\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)}{2\sqrt{2}} + \beta\frac{(|000\rangle + |111\rangle)(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)}{2\sqrt{2}}$$

Now since $Z_1Z_2Z_3(|000\rangle \pm |111\rangle) = |000\rangle \pm (-1)^3|111\rangle = |000\rangle \mp |111\rangle$, we have:

$$\begin{aligned}
&Z_1Z_2Z_3\left(\alpha\frac{(|000\rangle - |111\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)}{2\sqrt{2}} + \beta\frac{(|000\rangle + |111\rangle)(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)}{2\sqrt{2}}\right) \\
&= \alpha\frac{(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)}{2\sqrt{2}} + \beta\frac{(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)}{2\sqrt{2}} \\
&= |\psi_L\rangle
\end{aligned}$$

so the error correction is accomplished. □

Exercise 10.7

Consider the three qubit bit flip code of Section 10.1.1, with corresponding projector $P = |000\rangle\langle 000| + |111\rangle\langle 111|$. The noise process this code protects against has operation elements $\left\{\sqrt{(1-p)^3}I, \sqrt{p(1-p)^2}X_1, \sqrt{p(1-p)^2}X_2, \sqrt{p(1-p)^2}X_3\right\}$ where p is the probability that the bit flips. Note that this quantum operation is not trace-preserving, since we have omitted operation elements corresponding to bit flips on two and three qubits. Verify the quantum error-correction conditions for this code and noise process.

Solution

Concepts Involved: Bit Flip Code, Projectors, Error Correction Conditions

We calculate $PE_i^\dagger E_jP$ for each of the errors E_i. Note that in this case the errors are all Hermitian, so this reduces to calculation of PE_iE_jP for all combinations of errors. Furthermore, note that the set of errors $\left\{\sqrt{(1-p)^3}I, \sqrt{p(1-p)^2}X_1, \sqrt{p(1-p)^2}X_2, \sqrt{p(1-p)^2}X_3\right\}$ is mutually commuting, so

$PE_iE_jP = PE_jE_iP$ so we only need to check all combinations (and the order does not effect the result).

$$P\sqrt{(1-p)^3}I\sqrt{(1-p)^3}IP = (1-p)^3P^2 = (1-p)^3P$$

$$P\sqrt{p(1-p)^2}X_1\sqrt{p(1-p)^2}X_1P = p(1-p)^2PIP = p(1-p)^2P^2 = p(1-p)^2P$$

$$P\sqrt{p(1-p)^2}X_2\sqrt{p(1-p)^2}X_2P = p(1-p)^2PIP = p(1-p)^2P^2 = p(1-p)^2P$$

$$P\sqrt{p(1-p)^2}X_3\sqrt{p(1-p)^2}X_3P = p(1-p)^2PIP = p(1-p)^2P^2 = p(1-p)^2P$$

where we have used that Paulis square to the identity and projectors are idempotent. For the other combinations we find:

$$P\sqrt{(1-p)^3}I\sqrt{p(1-p)^2}X_1P = \sqrt{p(1-p)^5}(|000\rangle\langle 000| + |111\rangle\langle 111|)(|100\rangle\langle 000| + |011\rangle\langle 111|) = 0 = 0P$$

$$P\sqrt{(1-p)^3}I\sqrt{p(1-p)^2}X_2P = \sqrt{p(1-p)^5}(|000\rangle\langle 000| + |111\rangle\langle 111|)(|010\rangle\langle 000| + |101\rangle\langle 111|) = 0 = 0P$$

$$P\sqrt{(1-p)^3}I\sqrt{p(1-p)^2}X_3P = \sqrt{p(1-p)^5}(|000\rangle\langle 000| + |111\rangle\langle 111|)(|001\rangle\langle 000| + |110\rangle\langle 111|) = 0 = 0P$$

$$P\sqrt{p(1-p)^2}X_1\sqrt{p(1-p)^2}X_2P = p(1-p)^2(|000\rangle\langle 100| + |111\rangle\langle 011|)(|010\rangle\langle 000| + |101\rangle\langle 111|) = 0 = 0P$$

$$P\sqrt{p(1-p)^2}X_1\sqrt{p(1-p)^2}X_3P = p(1-p)^2(|000\rangle\langle 100| + |111\rangle\langle 011|)(|001\rangle\langle 000| + |110\rangle\langle 111|) = 0 = 0P$$

$$P\sqrt{p(1-p)^2}X_2\sqrt{p(1-p)^2}X_3P = p(1-p)^2(|000\rangle\langle 010| + |111\rangle\langle 101|)(|001\rangle\langle 000| + |110\rangle\langle 111|) = 0 = 0P$$

So we therefore find:

$$PE_i^\dagger E_jP = \alpha_{ij}P$$

where:

$$\alpha = \begin{bmatrix} (1-p)^3 & 0 & 0 & 0 \\ 0 & p(1-p)^2 & 0 & 0 \\ 0 & 0 & p(1-p)^2 & 0 \\ 0 & 0 & 0 & p(1-p)^2 \end{bmatrix}$$

is Hermitian. The error-correction conditions are therefore verified. □

Exercise 10.8

Verify that the three qubit phase flip code $|0_L\rangle = |+++\rangle, |1_L\rangle = |---\rangle$ satisfies the quantum error-correction conditions for the set of error operators $\{I, Z_1, Z_2, Z_3\}$.

Solution

Concepts Involved: Phase Flip Code, Projectors, Error Correction Conditions

First, note the projector onto the code C in this case is:

$$P = |+++\rangle\langle+++| + |---\rangle\langle---|$$

We calculate $PE_i^\dagger E_j P$ for each of the errors $E_i \in \{I, Z_1, Z_2, Z_3\}$. The calculation is completely analogous to that in Ex. 10.7, simply replacing $X \to Z$, $|0\rangle \to |+\rangle, |1\rangle \to |-\rangle$, and setting all of the pre-factors $\sqrt{(1-p)^3}$ and $\sqrt{p(1-p)^2}$ to one. The result we find is:

$$PE_i^\dagger E_j P = \alpha_{ij} P$$

where:

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

which is of course Hermitian, and the QEC conditions are thus satisfied. □

Exercise 10.9

Again, consider the three qubit phase flip code. Let P_i and Q_i be the projectors into the $|0\rangle$ and $|1\rangle$ states, respectively, of the ith qubit. Prove that the three qubit phase flip code protects against the error set $\{I, P_1, Q_1, P_2, Q_2, P_3, Q_3\}$.

Solution

Concepts Involved: Phase Flip Code, Error Correction Conditions, Discretization of Errors

By Theorem 10.2 in the text, we know that if C is a quantum code and $\mathcal{R}$ is the error-correction procedure constructed via the error correction conditions that corrects for a noise process $\mathcal{E}$ with operation elements $\{E_i\}$, then $\mathcal{R}$ also corrects for arbitrary complex linear combinations of the E_i. We then note that all errors in $\{I, P_1, Q_1, P_2, Q_2, P_3, Q_3\}$ can be written as linear combinations of errors in $\{I, Z_1, Z_2, Z_3\}$:

$$I = I, \quad P_i = \frac{I + Z_i}{2}, \quad Q_i = \frac{I - Z_i}{2}$$

therefore by Theorem 10.2 and the result of Ex. 10.8, the phase flip code can protect against $\{IP_1, Q_1, P_2, Q_2, P_3, Q_3\}$. □

Exercise 10.10

Explicitly verify the quantum error-correction conditions for the Shor code, for the error set containing I and the error operators X_j, Y_j, Z_j for $j = 1$ through 9.

Solution

Concepts Involved: Shor Code, Projectors, Error Correction Conditions

We have

$$P = |0_L\rangle\langle 0_L| + |1_L\rangle\langle 1_L|$$

where

$$|0_L\rangle = \frac{(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)}{2\sqrt{2}}$$
$$|1_L\rangle = \frac{(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)}{2\sqrt{2}}$$

All E_i we consider in the set of errors are Hermitian, so $E_i^\dagger = E_i$. Firstly, since all Pauli operators square to the identity, we find for all E_i that

$$PE_i^\dagger E_i P = PE_i^2 P = PIP = P^2 = P$$

Next, we observe

$$PIX_kP = PIY_kP = PIZ_kP = PX_kZ_kP = PZ_kX_kP = PX_kY_KP = PY_kX_kP = PY_kZ_kP = PZ_kY_kP = 0 = 0P$$

as all possible bit/phase flips on a single qubit map the codewords to states orthogonal to both codewords. Next, we find that for $k \neq l$ that

$$PX_kX_lP = PY_kY_lP = PX_kY_lP = PY_kX_lP = PX_kZ_lP = PZ_kX_lP = PY_kZ_lP = PZ_kY_lP = 0 = 0P$$

as in each case we have a bit flip on one or two qubits which maps the codewords to state orthogonal to both codwords.
Finally, we find that

$$PZ_kZ_lP = \begin{cases} P & \text{if } k,l \text{ are in the same block} \\ 0 & \text{if } k,l \text{ are in different blocks} \end{cases}$$

as in the former case the two phase flips cancel out (and thus P is preserved), and in the latter case we have phase flips on two different blocks and the codewords are mapped to states orthogonal to both codewords.

We thus conclude that

$$PE_i^\dagger E_j P = \alpha_{ij} P$$

where α_{ij} is Hermitian (as it has 1s on the diagonal, 1s on entries for $Z_i Z_j$ with i, j in the same block (thus symmetric across the diagonal), and zero elsewhere). We have thus verified the quantum error correction conditions. □

Exercise 10.11

Construct operation elements for a single qubit quantum operation $\mathcal{E}$ that upon input of any state ρ replaces it with the completely randomized state $I/2$. It is amazing that even such noise models as this may be corrected by codes such as the Shor code!

Solution

Concepts Involved: Kraus Representation, Density Operators

We wish to find operation elements $\{E_k\}$ such that:

$$\mathcal{E}(\rho) = \sum_k E_k \rho E_k^\dagger = \frac{I}{2}$$

for any input single-qubit state ρ. We claim that $\left\{\frac{1}{2}I, \frac{1}{2}X, \frac{1}{2}Y, \frac{1}{2}Z\right\}$ are the operation elements with this desired property. First, we verify that they satisfy the completeness relation:

$$\sum_k E_k^\dagger E_k = \frac{1}{4}(I^2 + X^2 + Y^2 + Z^2) = \frac{1}{4}(4I) = I$$

Next, we verify that they have the claimed property of sending every initial qubit state to the maximally mixed state. From Ex. 2.72, we can write a single-qubit density operator as:

$$\rho = \frac{I + r_x X + r_y Y + r_z Z}{2}$$

where $\mathbf{r} = (r_x, r_y, r_z) \in \mathbb{R}^3$ and $\|\mathbf{r}\| = 1$. Now calculating $\mathcal{E}(\rho)$, we have:

$$\begin{aligned}
\mathcal{E}(\rho) &= \frac{1}{4}(\rho + X\rho X + Y\rho Y + Z\rho Z) \\
&= \frac{1}{8}\Big((I + X^2 + Y^2 + Z^2) + r_x(X + X^3 + YXY + ZXZ) \\
&+ r_y(Y + XYX + Y^3 + ZYZ) + r_z(Z + XZX + YZY + Z^3)\Big) \\
&= \frac{1}{8}\left(4I + r_x(2X + 2iYZ) + r_y(2Y + 2iZX) + r_z(2Z + 2iXY)\right) \\
&= \frac{1}{8}\left(4I + r_x(2X + 2i(iX) + r_y(2Y + 2i(iY)) + r_z(2Z + 2i(iZ))\right) \\
&= \frac{1}{8}(4I) \\
&= \frac{I}{2}
\end{aligned}$$

where we use that $XY = iZ, YZ = iX$, and $XZ = iY$. The claim is thus proven. ☐

Remark:
The described channel is of course the single-qubit depolarizing channel of full strength.

Exercise 10.12

Show that the fidelity between the state $|0\rangle$ and $\mathcal{E}(|0\rangle\langle 0|)$ is $\sqrt{1 - 2p/3}$, and use this to argue that the minimum fidelity for the depolarizing channel is $\sqrt{1 - 2p/3}$.

Solution

Concepts Involved: Fidelity, Depolarizing Channel

The density operator corresponding to $|0\rangle$ is $|0\rangle\langle 0|$, and sending this through the depolarizing channel we have:

$$\begin{aligned}
\mathcal{E}(|0\rangle\langle 0|) &= (1 - p)\,|0\rangle\langle 0| + \frac{p}{3}(X\,|0\rangle\langle 0|\,X + Y\,|0\rangle\langle 0|\,Y + Z\,|0\rangle\langle 0|\,Z) \\
&= (1 - p)\,|0\rangle\langle 0| + \frac{p}{3}(|1\rangle\langle 1| + |1\rangle\langle 1| + |0\rangle\langle 0|) \\
&= (1 - \frac{2p}{3})\,|0\rangle\langle 0| + \frac{2p}{3}\,|1\rangle\langle 1|
\end{aligned}$$

and so:

$$F(|0\rangle, \mathcal{E}(|0\rangle\langle 0|)) = \sqrt{\langle 0|\left((1 - \frac{2p}{3})\,|0\rangle\langle 0| + \frac{2p}{3}\,|1\rangle\langle 1|\right)|0\rangle} = \sqrt{1 - \frac{2p}{3}} \tag{9}$$

as claimed. Because the depolarizing channel is symmetric in $X/Y/Z$, it is therefore symmetric in possible input states and so for any input state $|\psi\rangle$ we would find that $F(|0\rangle, \mathcal{E}(|0\rangle\langle 0|)) = \sqrt{1 - 2p/3}$. As such, this is the minimum fidelity.

For a more rigorous argument; by from Ex. 2.72, we can write a single-qubit density operator as

$$\rho = \frac{I + r_x X + r_y Y + r_z Z}{2}$$

where $\mathbf{r} = (r_x, r_y, r_z) \in \mathbb{R}^3$ and $\|\mathbf{r}\| = 1$ when $\rho = |\psi\rangle\langle\psi|$ a pure state. We then have that:

$$\langle\psi| X |\psi\rangle = \mathrm{Tr}\big(\rho_\psi X\big) = \mathrm{Tr}\left(\frac{X + r_x I + r_y YX + r_z ZX}{2}\right) = r_x$$

and analogously for Y/Z. We then have:

$$\begin{aligned}
F(|\psi\rangle, \mathcal{E}(|\psi\rangle\langle\psi|)) &= \sqrt{\langle\psi| \left((1-p)|\psi\rangle\langle\psi| + \frac{p}{3}(X|\psi\rangle\langle\psi| X + Y|\psi\rangle\langle\psi| Y + Z|\psi\rangle\langle\psi| Z)\right)} \\
&= \sqrt{(1-p)|\psi\rangle\langle\psi|^2 + \frac{p}{3}(\langle\psi| X |\psi\rangle^2 + \langle\psi| Y |\psi\rangle^2 + \langle\psi| Z |\psi\rangle^2} \\
&= \sqrt{(1-p) + \frac{p}{3}(r_x^2 + r_y^2 + r_z^2)} \\
&= \sqrt{(1-p) + \frac{p}{3}} \\
&= \sqrt{1 - \frac{2p}{3}}
\end{aligned}$$

where we use that $\|\mathbf{r}\| = 1$ in the fourth equality. Since this is true for all pure states, it must be the minimum fidelity. □

Exercise 10.13

Show that the minimum fidelity $F(|\psi\rangle, \mathcal{E}(|\psi\rangle\langle\psi|))$ when $\mathcal{E}$ is the amplitude damping channel when $\mathcal{E}$ is the amplitude damping channel with parameter γ is $\sqrt{1-\gamma}$.

Solution

Concepts Involved: Fidelity, Amplitude Damping Channel

Let $|\psi\rangle = a|0\rangle + b|1\rangle$ with $|a|^2 + |b|^2 = 1$. Then we have:

$$|\psi\rangle\langle\psi| = \begin{bmatrix} |a|^2 & ab^* \\ a^*b & |b|^2 \end{bmatrix}$$

and after applying the amplitude damping channel we have (from Ex. 8.22):

$$\mathcal{E}(|\psi\rangle\langle\psi|) = \begin{bmatrix} 1 - (1-\gamma)(1-|a|^2) & ab^*\sqrt{1-\gamma} \\ a^*b\sqrt{1-\gamma} & |b|^2(1-\gamma) \end{bmatrix}$$

Calculating the fidelity, we then have:

$$\begin{aligned}
F(|\psi\rangle, \mathcal{E}(|\psi\rangle\langle\psi|)) &= \sqrt{\langle\psi| \mathcal{E}(|\psi\rangle\langle\psi|) |\psi\rangle} \\
&= \sqrt{\begin{bmatrix} a^* & b^* \end{bmatrix} \begin{bmatrix} 1-(1-\gamma)(1-|a|^2) & ab^*\sqrt{1-\gamma} \\ a^*b\sqrt{1-\gamma} & |b|^2(1-\gamma) \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}} \\
&= \sqrt{|a|^2(1-(1-\gamma)(1-|a|^2)) + 2|a|^2|b|^2\sqrt{1-\gamma} + |b|^4(1-\gamma)}
\end{aligned}$$

Using the normalization condition, $|b|^2 = 1 - |a|^2$ so we can write the above in terms of $|a|^2$ alone:

$$\begin{aligned}
F(|\psi\rangle, \mathcal{E}(|\psi\rangle\langle\psi|)) &= \sqrt{|a|^2(1-(1-\gamma)(1-|a|^2)) + 2|a|^2(1-|a|^2)\sqrt{1-\gamma} + (1-|a|^2)^2(1-\gamma)} \\
&= \sqrt{2(1-\sqrt{1-\gamma}-\gamma)(|a|^2)^2 + (-2+2\sqrt{1-\gamma}+3\gamma)|a|^2 + 1 - \gamma}
\end{aligned}$$

This is minimized when the expression under the square root is minimized. Taking the derivative w.r.t $|a|^2$, we find:

$$\frac{\partial}{\partial |a|^2}(F(|\psi\rangle, \mathcal{E}(|\psi\rangle\langle\psi|)))^2 = 4(1-\sqrt{1-\gamma}-\gamma)|a|^2 - 2 + 2\sqrt{1-\gamma} + 3\gamma$$

which is non-negative for $|a|^2 \in [0,1]$ (which is the domain over which it is defined). Therefore $F(|\psi\rangle, \mathcal{E}(|\psi\rangle\langle\psi|))$ is an increasing function of $|a|^2$ over $[0,1]$, and hence $F(|\psi\rangle, \mathcal{E}(|\psi\rangle\langle\psi|))$ is minimized when $a = 0$ and therefore $|\psi\rangle = |1\rangle$. In this case, the fidelity is:

$$F(|\psi\rangle, \mathcal{E}(|\psi\rangle\langle\psi|))_{\text{min}} = F(|1\rangle, \mathcal{E}(|1\rangle\langle 1|)) = \sqrt{1-\gamma}$$

as claimed. □

Exercise 10.14

Write an expression for a generator matrix encoding k bits using r repetitions for each bit. This is an $[rk, k]$ linear code, and should have an $rk \times k$ generator matrix.

Solution

Concepts Involved: Repetition Code, Generator Matrices

The claimed generator matrix G is an $rk \times k$ matrix such that:

$$G_{ij} = \begin{cases} 1 & r(j-1) < i \leq rj \\ 0 & \text{otherwise.} \end{cases}$$

By matrix multiplication, we can see that:

$$G(x_1, x_2, \ldots, x_k) = (\overbrace{x_1, \ldots, x_1}^{r \text{ times}}, \overbrace{x_2, \ldots, x_2}^{r \text{ times}}, \ldots, \overbrace{x_k, \ldots, x_k}^{r \text{ times}})$$

□

Exercise 10.15

Show that adding one column of G to another results in a generator matrix generating the same code.

Solution

Concepts Involved: Generator Matrices

The set of possible codewords of a code corresponds to the vector space spanned by the columns of G. So if $G = [\mathbf{v}_1, \mathbf{v}_2, \dots \mathbf{v}_k]$ then the possible codewords are:

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_k\mathbf{v}_k$$

where $c_i \in \mathbb{Z}_2$ and addition is done modulo 2. WLOG suppose we add column 2 to colummn 1 (permuting the columns of G clearly preserves the codespace, as it just amounts to permuting labels in the equation above), so we have $G' = [\mathbf{v}_1 + \mathbf{v}_2, \mathbf{v}_2, \dots, \mathbf{v}_k]$, so the possible codewords are:

$$\mathbf{v}' = c_1'(\mathbf{v}_1 + \mathbf{v}_2) + c_2'\mathbf{v}_2 + \ldots + c_k'\mathbf{v}_k$$

where $c_i' \in \mathbb{Z}_2$. By defining $c_1' = c_1, c_2' = c_1 + c_2$, and $c_j' = c_j$ for $j \geq 2$ we can see that the codewords $\mathbf{v}'$ are the same as the codewords $\mathbf{v}$, and thus G and G' generate the same code. □

Exercise 10.16

Show that adding one row of the parity check matrix to another does not change the code. Using Gaussian elimination and swapping of bits it is therefore possible to assume that the parity check matrix has the *standard form* $[A|I_{n-k}]$ where A is an $(n-k) \times k$ matrix.

Solution

Concepts Involved: Parity Check Matrices, Gaussian Elimination

In the parity check matrix formulation, an $[n, k]$ code is all $\mathbf{x} \in \mathbb{Z}_2^n$ such that $Hx = 0$ where $H \in \mathbb{Z}_2^{(n-k)\times n}$ is the parity check matrix. We can write:

$$H = \begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \\ \vdots \\ \mathbf{v}_n^T \end{bmatrix}$$

with $\mathbf{v}_i^T$ the rows of H. By the definition of matrix multiplication, it follows from $Hx = 0$ that:

$$\mathbf{v}_i \cdot \mathbf{x} = 0$$

for each i and for all codewords $\mathbf{x}$. WLOG suppose we add row 2 to row 1 (permuting the rows of H

clearly preserves the codespace, as the above condition is unchanged). We then have:

$$H' = \begin{bmatrix} \mathbf{v}_1^T + \mathbf{v}_2^T \\ \mathbf{v}_2^T \\ \vdots \\ \mathbf{v}_n^T \end{bmatrix}$$

It then follows that if $Hx = 0$, then $H'x = 0$ as:

$$(\mathbf{v}_1 + \mathbf{v}_2) \cdot \mathbf{x} = \mathbf{v}_1 \cdot \mathbf{x} + \mathbf{v}_2 \cdot \mathbf{x} = 0 + 0 = 0$$

and $\mathbf{v}_i \cdot \mathbf{x} = 0$ for $i \geq 2$. Furthermore, if $H'x = 0$ then $Hx = 0$ as:

$$\mathbf{v}_1 \cdot \mathbf{x} = (\mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_2) \cdot \mathbf{x} = (\mathbf{v}_1 \cdot \mathbf{v}_2) \cdot \mathbf{x} + \mathbf{v}_2 \cdot \mathbf{x} = 0 + 0 = 0$$

and $\mathbf{v}_i \cdot \mathbf{x} = 0$ for $i \geq 2$. Therefore, H, H' correspond to the same code.
Since Gaussian elimination only involves swapping rows (does nothing to H), swapping columns (changes the labels of the qubits) and adding rows to each other (does nothing as shown above), we can thus always assume that the parity check matrix can be brought to standard form. □

Exercise 10.17

Find a parity check matrix for the $[6, 2]$ repetition code defined by the generator matrix in (10.54).

Solution

Concepts Involved: Repetition Code, Generator Matrices, Parity Check Matrices
The [6, 2] repetition code has generator matrix:

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

To construct H, we pick out $6 - 2 = 4$ linearly independent vectors orthogonal to the columns of G. Four such vectors are:

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

Therefore one such parity check matrix is:

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

□

Exercise 10.18

Show that the parity check matrix H and generator matrix G for the same linear code satisfy $HG = 0$

Solution

Concepts Involved: Generator Matrices, Parity Check Matrices.

For a given $[n, k]$ code with parity check matrix H and generator matrix G, we can write:

$$H = \begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \\ \vdots \\ \mathbf{v}_{n-k}^T \end{bmatrix}, \quad G = \begin{bmatrix} \mathbf{y}_1 & \mathbf{y}_2 & \cdots & \mathbf{y}_k \end{bmatrix}$$

where $\mathbf{v}_i, \mathbf{y}_j$ are each n-dimensional vectors which are orthogonal to one another. By the definition of matrix multiplication, HG is a $n - k \times k$ matrix with entries:

$$(HG)_{ij} = \mathbf{v}_i \cdot \mathbf{y}_j = 0$$

where the last equality follows by orthogonality, thus proving the claim. □

Exercise 10.19

Suppose an $[n, k]$ linear code C has a parity check matrix of the form $H = [A|I_{n-k}]$, for some $(n - k) \times k$ matrix A. Show that the corresponding generator matrix is

$$G = \left[\frac{I_k}{-A}\right].$$

Solution

Concepts Involved: Generator Matrices, Parity Check Matrices

To get a generator matrix from a parity check matrix, we pick k linearly independent vectors $\mathbf{y}_1, \ldots \mathbf{y}_k$ spanning the kernel of H, and set G to have columns $\mathbf{y}_1$ through $\mathbf{y}y_k$. In this case, H has standard form

$H = [A|I_{n-k}]$ and so we may write:

$$H = \begin{bmatrix} \mathbf{v}_1^T & \mathbf{e}_{1,n-k}^T \\ \mathbf{v}_2^T & \mathbf{e}_{2,n-k}^T \\ \vdots & \\ \mathbf{v}_{n-k}^T & \mathbf{e}_{n-k,n-k}^T \end{bmatrix}$$

where $\mathbf{v}_i^T$ are the rows of A and $\mathbf{e}_{i,n-k}$ is a $n-k$ length vector with 1 in the ith position and 0s elsewhere. We want to find vectors in the kernel of H, i.e. vectors $\mathbf{y}$ that satisfy:

$$\mathbf{v}_i \cdot \mathbf{y}_{1,\ldots k} + \mathbf{e}_{i,n-k} \cdot \mathbf{y}_{k+1,\ldots n}$$

for every $i \in \{1, \ldots, n-k\}$. A clear choice that satisfies this relation is $\mathbf{y}_{1,\ldots,n-k} = \mathbf{e}_{n,k}$ and $\mathbf{y}_{k+1,\ldots,n} = -\mathbf{w}_n$ where $\mathbf{w}_n$ is the nth column of A; this choice satisfies the above as:

$$\mathbf{v}_i \cdot \mathbf{e}_{n,k} + \mathbf{e}_{i,n-k} \cdot (-\mathbf{w}_n) = a_{in} - a_{in} = 0$$

This yields $\mathbf{y}_i, \ldots \mathbf{y}_k$, one for each column of A. We therefore construct G as:

$$G = \begin{bmatrix} \mathbf{e}_{1,k} & \mathbf{e}_{2,k} & \ldots & \mathbf{e}_{k,k} \\ -\mathbf{w}_1 & -\mathbf{w}_2 & \cdots & -\mathbf{w}_n \end{bmatrix} = \begin{bmatrix} I_k \\ -A \end{bmatrix}$$

which was what we wished to show. □

Exercise 10.20

Let H be a parity check matrix such that any $d-1$ columns are linearly independent, but there exists a set of d linearly dependent columns. Show that the code defined by H has distance d.

Solution

Concepts Involved: Parity Check Matrices, Code Distance

Given a parity check matrix H, a codeword $\mathbf{x} = (x_1, x_2, \ldots x_n)$ satisfies $H\mathbf{x} = 0$, which implies:

$$x_1\mathbf{h}_1 + x_2\mathbf{h}_2 + \ldots + x_n\mathbf{h}_n = 0$$

where $\mathbf{h}_i$ are the columns of H. Since any $d-1$ columns are linearly independent, it follows that any sum of $d-1$ (or less) columns of H is nonzero, and therefore there are no codewords of weight $d-1$ or lower. However, there exists a set of d linearly dependent columns, and therefore there exists a codeword that is 1 for each of those columns and 0 elsewhere, and is therefore of weight d. There are no codewords of smaller weight, and therefore we conclude that the code has distance d. □

Exercise 10.21: Singleton bound

Show than an $[n, k, d]$ code must satisfy $n - k \geq d - 1$.

Solution

Concepts Involved: Parity Check Matrices, Code Distance

An $[n, k, d]$ code has an $n-k \times n$ parity check matrix H. It therefore has at most $n-k$ linearly independent columns. From the solution to the previous exercise, if a code has distance d then the parity check matrix must have all sets of $d-1$ columns be linearly independent (else there would exist a codeword of weight $d-1$ and hence the code would have distance $< d$, a contradiction). Combining these two facts, it follows that:

$$n - k \geq d - 1$$

as claimed. ☐

Exercise 10.22

Show that all Hamming codes have distance 3, and thus can correct an error on a single bit. The Hamming codes are therefore $[2^r - 1, 2^r - r - 1, 3]$ codes.

Solution

Concepts Involved: Hamming Code, Parity Check Matrices, Code Distance
Recall that the $[2^r - 1, 2^r - r - 1]$ Hamming code (for $r \geq 2$) is a linear code having parity check matrix whose columns are all $2^r - 1$ bit strings of length r which are not zero.

Any two columns of H are different and therefore linearly independent. Furthermore, H has the 3 columns:

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

which are linearly dependent as they add together to make zero. By Ex. 10.20, any Hamming code has distance 3. Since a distance $d \geq 2t + 1$ code can correct errors on t bits, all Hamming codes can correct errors on one bit. ☐

Exercise 10.23

($*$) Prove the Gilbert-Varshamov bound.

Solution

Concepts Involved: Binary Entropy, Hamming Weight, Entropy Bound
We want to show that for every n and t, there exists a binary linear code of length n and dimension k correcting t errors such that

$$\frac{k}{n} \geq 1 - H_2\left(\frac{2t}{n}\right),$$

where $H_2(x) = -x \log_2 x - (1-x)\log_2(1-x)$ is the binary entropy function.

Proof. Let

$$B(n,r) \;=\; \sum_{i=0}^{r} \binom{n}{i}$$

denote the volume of a Hamming ball of radius r in $\{0,1\}^n$.

Step 1: Greedy code construction. Start with the full set $\{0,1\}^n$. Pick an arbitrary string as a codeword and delete its Hamming ball of radius $2t$. Repeat until no strings remain. Since balls of radius $2t$ around distinct codewords are disjoint if the minimum distance is $\geq 2t+1$, the number of codewords M chosen satisfies

$$M \;\geq\; \frac{2^n}{B(n,2t)}.$$

Thus there exists a code with dimension

$$k \;=\; \log_2 M \;\geq\; n - \log_2 B(n,2t).$$

Step 2: Volume estimate. For $\alpha = r/n$ with $0 \leq \alpha \leq \frac{1}{2}$,

$$B(n,\alpha n) \;\leq\; 2^{\,nH_2(\alpha)}.$$

This follows from the entropy bound $\binom{n}{\alpha n} \leq 2^{nH_2(\alpha)}$ and the inequality $B(n,\alpha n) \leq (\alpha n+1)\binom{n}{\alpha n}$, whose polynomial prefactor is negligible compared to the exponential term.

Step 3: Conclusion. Substituting $\alpha = 2t/n$ gives

$$k \;\geq\; n - nH_2\!\left(\tfrac{2t}{n}\right),$$

or equivalently

$$\frac{k}{n} \;\geq\; 1 - H_2\!\left(\tfrac{2t}{n}\right).$$

□

Exercise 10.24

Show that a code with generator matrix G is weakly self-dual if and only if $G^T G = 0$.

Solution

Concepts Involved: Generator Matrices, Dual Codes

Recall that if we have a $[n,k]$ code C with generator matrix G and parity check matrix H, then the dual code $C^\perp$ is the code consisting of all codewords y such that y is orthogonal to all codewords in C. Furthermore, recall that a code is weakly self-dual if $C \subseteq C^\perp$.

$\boxed{\implies}$: Suppose that G is weakly self dual. Then, it follows that every codeword $y = Gx \in C$ is contained in $C^\perp$. Since the parity check matrix applied to a codeword of a code gives zero, for $C^\perp$ we have that $G^T y = 0$ and thus $G^T y = G^T G x = 0$ for all length k binary vectors x, but this is only possible if $G^T G = 0$.

$\boxed{\Longleftarrow}$: Suppose that $G^T G = 0$. Suppose that $y = Gx, y' = Gx'$ are codewords of C. We then see that $y' \cdot y = x'^T G^T G x = x'^T 0 x = 0$. Thus, all codewords of C are orthogonal to each other, and thus all codewords y of C are codewords of $C^\perp$ by definition. Thus $C \subseteq C^\perp$. □

Exercise 10.25

Let C be a linear code. Show that if $x \in C^\perp$ then $\sum_{y \in C} (-1)^{x \cdot y} = |C|$, while if $x \notin C^\perp$ then $\sum_{y \in C} (-1)^{x \cdot y} = 0$.

Solution

Concepts Involved: Dual Codes

Suppose $x \in C^\perp$. Then it follows that $y \cdot x = 0$ for every codeword $y \in C$. Thus, $\sum_{y \in C}(-1)^{x \cdot y} = \sum_{y \in C}(-1)^0 = \sum_{y \in C} 1 = |C|$. Suppose instead that $x \notin C^\perp$. Then we claim that half the codewords of C are orthogonal to x with $x \cdot y = 0$, and the other half are not orthogonal with $x \cdot y = 1$. As proof of this fact, suppose there are n basis codewords $y_1, \ldots, y_n$ of C; since $x \notin C^\perp$, it follows that at least one of these basis codewords is not orthogonal to x. Let $y_1, \ldots, y_k$ be the non-orthogonal basis codewords and $y_{k+1}, \ldots, y_n$ be the orthogonal codewords. To form an arbitrary codeword of C, we take a linear combination of the basis codewords. If an even number fo $y_1, \ldots, y_k$ are included in the sum then the codeword is orthogonal to x and otherwise the codeword is not orthogonal to x. There are 2^{k-1} to choose an even number of elements out of a set of k elements and 2^{k-1} ways to choose an odd number, and therefore there are the same number of codewords in C that are orthogonal to x as there are those that are not orthogonal. Hence, the contributions of these two evenly weighted parts cancel in the sum to yield $\sum_{y \in C}(-1)^{x \cdot y} = 0$. □

Exercise 10.26

Suppose H is a parity check matrix. Explain how to compute the transformation $|x\rangle|0\rangle \mapsto |x\rangle|Hx\rangle$ using a circuit composed entirely of controlled-`NOT`s.

Solution

Concepts Involved: Parity Check Matrices, Controlled Operations

By the definition of matrix multiplication:

$$(Hx)_i = \sum_j H_{ij} x_j \tag{10}$$

So, if we want the ith bit in the second register to become $\left|(Hx)_i\right\rangle$, this can be realized by applying a CNOT gate with control on the jth qubit of the first register (in $|x\rangle$) and acting on the ith qubit of the second register if $H_{ij} = 1$ (and doing nothing if $H_{ij} = 0$); each gate (or identity) realizing one term in the above sum. □

Exercise 10.27

Show that the codes defined by

$$|x + C_2\rangle \equiv \frac{1}{\sqrt{|C_2|}} \sum_{y \in C_2} (-1)^{u \cdot y} |x + y + v\rangle$$

as parameterized by u and v are equivalent to **CSS**(C_1, C_2) in the sense that they have the same error-correcting properties. These codes, which we'll refer to as **CSS**$_{u,v}(C_1, C_2)$ will be useful later in our study of quantum key distribution, in Section 12.6.5.

Solution

Concepts Involved: CSS Codes

Let us start by "corrupting" the states by applying a bit flip (X) where ever v is nonzero and a phase flip (Z) where ever u is nonzero. First applying the bit flips:

$$X^v |x + C_2\rangle = \frac{1}{\sqrt{|C_2|}} \sum_{y \in C_2} (-1)^{u \cdot y} |x + y + v + v\rangle = \frac{1}{\sqrt{|C_2|}} \sum_{y \in C_2} (-1)^{u \cdot y} |x + y\rangle$$

where we use that $v + v = 0$ under mod 2 addition. Next applying the phase flips:

$$Z^u X^v |x + C_2\rangle = \frac{1}{\sqrt{|C_2|}} \sum_{y \in C_2} (-1)^{u \cdot y} (-1)^{u \cdot (x+y)} |x + y\rangle = \frac{1}{\sqrt{|C_2|}} \sum_{y \in C_2} (-1)^{u \cdot x} |x + y\rangle = (-1)^{u \cdot x} \sum_{y \in C_2} |x + y\rangle$$

Up to the phase factor $(-1)^{u \cdot x}$, the states appearing on the RHS are identical to the states defining the code **CSS**(C_1, C_2) (Eq. (10.64) in the text). This phase factor can be neglected in the bit/phase error detection and correction analysis carried out in Section 10.4.2, and thus we conclude that **CSS**$_{u,v}(C_1, C_2)$ has the same error correcting properties as **CSS**(C_1, C_2). □

Exercise 10.28

Verify that the transpose of the matrix in (10.77) is the generator of the $[7, 4, 3]$ Hamming code.

Solution

Concepts Involved: Hamming Code, Parity Check Matrices, Generator Matrices

The matrix in (10.77) is:

$$H[C_2] = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}$$

Taking the transpose, we have:

$$H[C_2]^T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \end{bmatrix}$$

Looking at the parity check matrix for the $[7, 4, 3]$ Hamming code, we have:

$$H[C_1] = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}$$

In order to confirm $H[C_2]^T$ as the generator matrix of the code, we require that the columns are linearly independent and lie in the kernel of $H[C_1]$. The linearly independence is immediately clear by inspection (as columns $i = 1, 2, 3, 4$ uniquely have a non-zero ith entry). Let us then just verify that they lie in the kernel of H, which can equivalently be done by checking that $H[C_1]H[C_2]^T = 0$:

$$\begin{aligned} H[C_1]H[C_2]^T &= \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1+1 & 1+1 & 1+1 & 1+1+1+1 \\ 1+1 & 1+1 & 1+1 & 1+1 \\ 1+1 & 1+1 & 1+1 & 1+1 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

□

Exercise 10.29

Show that an arbitrary linear combination of any two elements of V_S is also in V_S. Therefore, V_S is a subspace of the n qubit state space. Show that V_S is the intersection of the subspaces fixed by each operator in S (that is, the eigenvalue one eigenspaces of elements of S).

Solution

Concepts Involved: Group Theory, Pauli Group, Stabilizer Formalism, Vector Subspace

Let $S \subseteq \mathcal{P}_n$ be a subgroup of the n-qubit Pauli group. Define the stabilized subspace:

$$V_S = \left\{ |\psi\rangle \in \mathbb{C}^{2^n} \;\middle|\; g|\psi\rangle = |\psi\rangle \text{ for all } g \in S \right\}.$$

To show that V_S is a subspace, take $|\psi\rangle, |\varphi\rangle \in V_S$, and $\alpha, \beta \in \mathbb{C}$. Then for any $g \in S$,

$$\begin{aligned} g(\alpha|\psi\rangle + \beta|\varphi\rangle) &= \alpha g|\psi\rangle + \beta g|\varphi\rangle \\ &= \alpha|\psi\rangle + \beta|\varphi\rangle, \end{aligned}$$

so $\alpha|\psi\rangle + \beta|\varphi\rangle \in V_S$. Hence, V_S is a subspace.
Now define for each $g \in S$ the stabilized subspace:

$$\text{Stab}(g) := \left\{ |\psi\rangle \in \mathbb{C}^{2^n} \;\middle|\; g|\psi\rangle = |\psi\rangle \right\}.$$

Then by definition,

$$V_S = \bigcap_{g \in S} \text{Stab}(g),$$

as a state is in V_S if and only if it is stabilized (i.e., fixed) by all elements of S. $\square$

Remark: For an abelian $S \subseteq \mathcal{P}_n$ with $|S| = 2^r$ and $-I \notin S$, the dimension of V_S is 2^{n-r}, characterizing the number of logical qubits.

Exercise 10.30

Show that $-I \notin S$ implies $\pm iI \notin S$.

Solution

Concepts Involved: Group Theory, Pauli Group

Subgroups are groups and therefore closed under multiplication. If $iI \in S$, then $(iI)^2 = (i)^2 I = -I \in S$. The claim is thus shown via the contrapositive. $\square$

Exercise 10.31

Suppose that S is a subgroup of G_n generated by elements $g_1, \ldots, g_l$. Show that all the elements of S commute if and only if g_i and g_j commute for each pair i, j.

Solution

Concepts Involved: Group Theory, Pauli Group, Group Generators

$\boxed{\implies}$: Suppose all elements of S commute. Since each generator is itself an element of S, any pair of generators will commute.

$\boxed{\Longleftarrow}$: Suppose any pair of generators g_i, g_j of S commute. Any two elements $s_1, s_2 \in S$ can be written as a product of generators:

$$s_1 = g_1^{n_1} g_2^{n_2} \dots g_k^{n_k}, \quad s_2 = g_1^{m_1} g_2^{m_2} \dots g_k^{m_k}$$

Using the pairwise commutation of the generators, we can move the g_is together to find that $s_1 s_2 = s_2 s_1 = g_1^{n_1+m_1} g_2^{n_2+m_2} \dots g_k^{n_k+m_k}$ and so all elements of S commute. □

Remark: The above argument holds for any group, not just subgroups of G_n.

Exercise 10.32

Verify that the generators in Figure 10.6 (reproduced below) stabilize the codewords for the Steane code, as described in Section 10.4.2.

Name	Operator
g_1	$IIIXXXX$
g_2	$IXXIIXX$
g_3	$XIXIXIX$
g_4	$IIIZZZZ$
g_5	$IZZIIZZ$
g_6	$ZIZIZIZ$

Solution

Concepts Involved: Steane Code, Stabilizer Formalism

The logical codewords for the Steane code are given by:

$$\begin{aligned} |0_L\rangle = \frac{1}{\sqrt{8}} \big[&|0000000\rangle + |1010101\rangle + |0110011\rangle + |1100110\rangle \\ &+ |0001111\rangle + |1011010\rangle + |0111100\rangle + |1101001\rangle \big] \\ |1_L\rangle = \frac{1}{\sqrt{8}} \big[&|1111111\rangle + |0101010\rangle + |1001100\rangle + |0011001\rangle \\ &+ |1110000\rangle + |0100101\rangle + |1000011\rangle + |0010110\rangle \big] \end{aligned}$$

Let us first check that $|0_L\rangle$ is stabilized by the 6 generators:

$$\begin{aligned}
g_1 |0_L\rangle &= \frac{1}{\sqrt{8}}\big[\,|0001111\rangle + |1011010\rangle + |0111100\rangle + |1101001\rangle \\
&\quad + |0000000\rangle + |1010101\rangle + |0110011\rangle + |1100110\rangle\big] \\
&= |0_L\rangle \\
g_2 |0_L\rangle &= \frac{1}{\sqrt{8}}\big[\,|0110011\rangle + |1100110\rangle + |0000000\rangle + |1010101\rangle \\
&\quad + |0111100\rangle + |1101001\rangle + |0001111\rangle + |1011010\rangle\big] \\
&= |0_L\rangle \\
g_3 |0_L\rangle &= \frac{1}{\sqrt{8}}\big[\,|1010101\rangle + |0000000\rangle + |1100110\rangle + |0110011\rangle \\
&\quad + |1011010\rangle + |0001111\rangle + |1101001\rangle + |0111100\rangle\big] \\
&= |0_L\rangle \\
g_4 |0_L\rangle &= \frac{1}{\sqrt{8}}\Big[(-1)^0|0000000\rangle + (-1)^2|1010101\rangle + (-1)^2|0110011\rangle + (-1)^2|1100110\rangle \\
&\quad + (-1)^4|0001111\rangle + (-1)^2|1011010\rangle + (-1)^2|0111100\rangle + (-1)^2|1101001\rangle\Big] \\
&= |0_L\rangle \\
g_5 |0_L\rangle &= \frac{1}{\sqrt{8}}\Big[(-1)^0|0000000\rangle + (-1)^2|1010101\rangle + (-1)^4|0110011\rangle + (-1)^2|1100110\rangle \\
&\quad + (-1)^2|0001111\rangle + (-1)^2|1011010\rangle + (-1)^2|0111100\rangle + (-1)^2|1101001\rangle\Big] \\
&= |0_L\rangle \\
g_6 |0_L\rangle &= \frac{1}{\sqrt{8}}\Big[(-1)^0|0000000\rangle + (-1)^4|1010101\rangle + (-1)^2|0110011\rangle + (-1)^2|1100110\rangle \\
&\quad + (-1)^2|0001111\rangle + (-1)^2|1011010\rangle + (-1)^2|0111100\rangle + (-1)^2|1101001\rangle\Big] \\
&= |0_L\rangle
\end{aligned}$$

For $|1_L\rangle$, first observe that $|1_L\rangle = XXXXXXX|0_L\rangle = \bar{X}|0_L\rangle$, and further that $[\bar{X}, g_i] = 0$ for each i (the commutation with the X generators is trivial, and the commutation with the Z generators is seen by observing that there are an even number of Zs). Hence, for each g_i we have $g_i|1_L\rangle = g_i\bar{X}|0_L\rangle = \bar{X}g_i|0_L\rangle = \bar{X}|0_L\rangle = |1_L\rangle$ as we have already shown that $|0_L\rangle$ is stabilized. We conclude that both codewords are stabilized by all generators. □

Exercise 10.33

Show that g and g' commute if and only if $r(g)\Lambda r(g')^T = 0$. (In the check matrix representation, arithmetic is done modulo two.)

Solution

Concepts Involved: Stabilizer Formalism, Check Matrix Representation, Symplectic Inner Product

Each n-qubit Pauli operator g is represented as a binary vector

$$r(g) = (a_1, \ldots, a_n \mid b_1, \ldots, b_n) \in \mathbb{F}_2^{2n} \tag{11}$$

where $a_j = 1$ if the jth component is X or Y, and $b_j = 1$ if it is Z or Y.
Let g, g' have binary vectors $r(g) = (a \mid b)$, $r(g') = (a' \mid b')$. Define the symplectic matrix

$$\Lambda = \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix} \tag{12}$$

Then the symplectic inner product is given by

$$\begin{aligned} r(g)\Lambda r(g')^T &= a \cdot b' + b \cdot a' \\ &= \sum_{j=1}^{n} (a_j b'_j + b_j a'_j) \mod 2 \end{aligned}$$

This value is 0 if and only if g and g' commute. Therefore,

$$g \text{ and } g' \text{ commute} \iff r(g)\Lambda r(g')^T = 0 \tag{13}$$

□

Exercise 10.34

Let $S = \langle g_1, \ldots, g_l \rangle$. Show that $-I$ is not an element of S if and only if $g_j^2 = I$ for all j, and $g_j \neq -I$ for all j.

Solution

Concepts Involved: Group Theory, Pauli Group, Order

Let $S = \langle g_1, \ldots, g_l \rangle$ be a subgroup of the n-qubit Pauli group generated by elements g_j. We claim:

$$-I \notin S \quad \text{if and only if} \quad g_j^2 = I \text{ and } g_j \neq -I \text{ for all } j.$$

$\boxed{\Longrightarrow}$: Assume $-I \notin S$. Then for each generator g_j, we must have $g_j^2 = I$. If instead $g_j^2 = -I$, then $-I \in S$, contradicting the assumption. Also, if any $g_j = -I$, then clearly $-I \in S$. Therefore, we must have both $g_j^2 = I$ and $g_j \neq -I$ for all j.
$\boxed{\Longleftarrow}$: Now assume that for all j, $g_j^2 = I$ and $g_j \neq -I$. Then each generator is of order 2 and Hermitian. Products of such generators can only produce other Hermitian Pauli operators (up to sign), and since none of them is $-I$, and none squares to $-I$, no combination of them can yield $-I$. Therefore, $-I \notin S$.
Conclusion: $-I \notin S$ if and only if each g_j is of order 2 and not equal to $-I$. □

Exercise 10.35

Let S be a subgroup of G_n such that $-I$ is not an element of S. Show that $g^2 = I$ for all $g \in S$, and thus $g^\dagger = g$.

Solution

Concepts Involved: Group Theory, Pauli Group

Let $S \subseteq G_n$ be a subgroup of the Pauli group such that $-I \notin S$. Every element $g \in G_n$ has the form $g = \alpha P$, where $\alpha \in \{\pm 1, \pm i\}$ and P is a tensor product of Pauli matrices. Since each Pauli matrix squares to I, we have

$$g^2 = \alpha^2 I \in \{I, -I\}$$

But $-I \notin S$, so $g^2 = I$ for all $g \in S$. Now, since each Pauli matrix is Hermitian and $\alpha = \pm 1$, it follows that

$$g^\dagger = \bar{\alpha} P = \alpha P = g$$

So every $g \in S$ is Hermitian.

$\square$

Exercise 10.36

Explicitly verify that $UX_1U^\dagger = X_1X_2$, $UX_2U^\dagger = X_2$, $UZ_1U^\dagger = Z_1$, and $UZ_2U^\dagger = Z_1Z_2$. These and other useful conjugation relations for the Hadamard, phase, and Pauli gates are summarized in Figure 10.7.

Solution

Concepts Involved: Unitary Operators, Controlled Operations

First note our ability to write the controlled-NOT operation in block diagonal form

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

By explicit (block) matrix multiplication we then have:

$$UX_1U^\dagger = \begin{bmatrix} I & 0 \\ 0 & X \end{bmatrix} \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & X \end{bmatrix} = \begin{bmatrix} 0 & I \\ X & 0 \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & X \end{bmatrix} = \begin{bmatrix} 0 & X \\ X & 0 \end{bmatrix} = X_1X_2$$

$$UX_2U^\dagger = \begin{bmatrix} I & 0 \\ 0 & X \end{bmatrix} \begin{bmatrix} X & 0 \\ 0 & X \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & X \end{bmatrix} = \begin{bmatrix} X & 0 \\ 0 & X^2 \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & X \end{bmatrix} = \begin{bmatrix} X & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & X \end{bmatrix} = \begin{bmatrix} X & 0 \\ 0 & X \end{bmatrix} = X_2$$

$$UZ_1U^\dagger = \begin{bmatrix} I & 0 \\ 0 & X \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & X \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & -X \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & X \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & -X^2 \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} = Z_1$$

$$UZ_2U^\dagger = \begin{bmatrix} I & 0 \\ 0 & X \end{bmatrix}\begin{bmatrix} Z & 0 \\ 0 & Z \end{bmatrix}\begin{bmatrix} I & 0 \\ 0 & X \end{bmatrix} = \begin{bmatrix} Z & 0 \\ 0 & XZ \end{bmatrix}\begin{bmatrix} I & 0 \\ 0 & X \end{bmatrix} = \begin{bmatrix} Z & 0 \\ 0 & XZX \end{bmatrix} = \begin{bmatrix} Z & 0 \\ 0 & -Z \end{bmatrix} = Z_1Z_2$$

□

Exercise 10.37

What is $UY_1U^\dagger$?

Solution

Concepts Involved: Unitary Operators, Controlled Operations

Note from the previous Exercise that $UX_1U^\dagger = X_1X_2$ and $UZ_1U^\dagger = Z_1$. Hence writing $Y = iZX$ and using that $U^\dagger U = I$:

$$UY_1U^\dagger = iUZXU^\dagger = iUZU^\dagger UXU^\dagger = iZ_1X_1X_2 = iY_1X_2.$$

□

Exercise 10.38

Suppose U and V are unitary operators on two qubits which transform Z_1, Z_2, X_1 and X_2 by conjugation in the same way. Show this implies that $U = V$.

Solution

Concepts Involved: Unitary Operators, Pauli Group

We observe that $UAU^\dagger = VAV^\dagger$ for $A \in \{Z_1, Z_2, X_1, X_2\}$ implies that $UAU^\dagger = VAV^\dagger$ for any two qubit Pauli as $UAU^\dagger UBU^\dagger = VAV^\dagger VBV^\dagger \implies UABU^\dagger = VABV^\dagger$ for any two operators A, B, and $\{Z_1, Z_2, X_1, X_2\}$ generates all two qubit Paulis via multiplication (which is seen from $Z^2 = I$ and $ZX = iY$). Then, observing that the two-qubit Pauli operators form a basis for operators acting on the Hilbert space $\mathbb{C}^4$ (formally - they are 16 linearly independent vectors in $\mathbb{C}^{4\times 4}$ over field $\mathbb{C}$, and since $\dim \mathbb{C}^{4\times 4} = 16$ they form a basis) we can write any operator $O \in \mathbb{C}^{4\times 4}$ as a linear combination of the two qubit Paulis, $O = \sum_i c_iA_i$. We then have:

$$UOU^\dagger = U\left(\sum_i c_iA_i\right)U^\dagger = \sum_i c_iUA_iU^\dagger = \sum_i c_iVA_iV^\dagger = V\left(\sum_i c_iA_i\right)V^\dagger = VOV^\dagger$$

So $UOU^\dagger = VOV^\dagger$ for all $O \in \mathbb{C}^{4\times 4}$, which is only possible if $U = V$. □

Exercise 10.39

Verify (10.91).

Solution

Concepts Involved: Unitary Operators

We verify by matrix multiplication:

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

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

□

Exercise 10.40

Provide an inductive proof of Theorem 10.6 as follows.

1. Prove that the Hadamard and phase gates can be used to perform any normalizer operation on a single qubit.
2. Suppose U is an $n+1$ qubit gate in $N(G_{n+1})$ such that $UZ_1U^\dagger = X_1 \otimes g$ and $UX_1U^\dagger = Z_1 \otimes g'$ for some $g, g' \in G_n$. Define U' on n qubits by $U'|\psi\rangle \equiv \sqrt{2}\langle 0|U(|0\rangle \otimes |\psi\rangle)$. Use the inductive hypothesis to show that the construction for U in Figure 10.9 may be implemented using $O(n^2)$ Hadamard, phase, and controlled-NOT gates.
3. Show that any gate $U \in N(G_{n+1})$ may be implemented using $O(n^2)$ Hadamard, phase, and controlled-NOT gates.

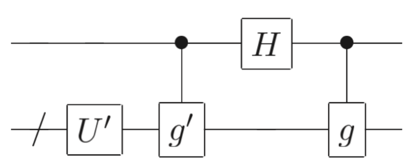


Figure 10.9. Construction used to prove that the Hadamard, phase and controlled-NOT gates generate the normalizer $N(G_n)$.

Solution

Concepts Involved: Group theory, Pauli Group, Stabilizer Formalism, Normalizers, Controlled Operations, Induction

1. H **and** S **generate all single-qubit normalizer operations.** Let G_1 be the single-qubit Pauli group and $N(G_1)$ its normalizer (single-qubit Clifford). Conjugation by any $U \in N(G_1)$ permutes $\{X, Y, Z\}$ up to a sign. The maps induced by

$$HXH = Z,\ HZH = X,\ HYH = -Y, \qquad SXS^\dagger = Y,\ SYS^\dagger = -X,\ SZS^\dagger = Z$$

show H implements the transposition $(X\,Z)$ and S the 3-cycle $(X\,Y\,Z)$ (up to signs). These generate all signed permutations of $\{X, Y, Z\}$ modulo global phase, hence $\langle H, S\rangle = N(G_1)$ (up to phase).

2. Canonical $(n{+}1)$**-qubit case.** Assume the inductive hypothesis: every $V \in N(G_n)$ is implementable with $O(n^2)$ gates from $\{H, S, \mathrm{CNOT}\}$. Let $U \in N(G_{n+1})$ obey

$$UZ_1U^\dagger = X_1 \otimes g, \qquad UX_1U^\dagger = Z_1 \otimes g',$$

for some $g, g' \in G_n$ on qubits $2, \dots, n{+}1$. Define U' on n qubits by

$$U' \,|\psi\rangle \;:=\; \sqrt{2}\,\langle 0|\, U\big(|0\rangle \otimes |\psi\rangle\big)\,.$$

Claim: $U' \in N(G_n)$. For any $h \in G_n$, since U normalizes G_{n+1} there exist a phase ω, a single-qubit Pauli P_1, and $p_h \in G_n$ s.t. $U\,(I \otimes h)\,U^\dagger = \omega\,(P_1 \otimes p_h)$. Using the given images of X_1 and Z_1, one checks that the $\langle 0|\,(\cdot)\,|0\rangle$ matrix element vanishes unless $P_1 = I$; in that case

$$U'\,h\,U'^\dagger \;\propto\; p_h \in G_n,$$

so U' maps Paulis to Paulis and hence $U' \in N(G_n)$ (global phase irrelevant).
Realizing the g, g' factors with $O(n)$ gates. We can wire the first qubit to imprint the tensor factors of g and g' using $\{H, S, \mathrm{CNOT}\}$, leaving the remaining n-qubit block otherwise untouched:

- To multiply Z_1 by a Z_j factor (in g): apply $\mathrm{CNOT}(j \to 1)$, which conjugates $Z_1 \mapsto Z_1Z_j$ and leaves X_1 invariant.
- To multiply Z_1 by an X_j factor: apply H_j, then $\mathrm{CNOT}(j \to 1)$, then H_j.
- To multiply X_1 by an X_j factor (in g'): apply $\mathrm{CNOT}(1 \to j)$, which conjugates $X_1 \mapsto X_1X_j$ and leaves Z_1 invariant.
- To multiply X_1 by a Z_j factor: apply H_j, then $\mathrm{CNOT}(1 \to j)$, then H_j.

Processing all non-identity factors in g, g' uses $O(n)$ one- and two-qubit Clifford gates.
Complete the construction. By the claim, $U' \in N(G_n)$ and by the inductive hypothesis can be implemented on the last n qubits with $O(n^2)$ gates. Conjugating $I \otimes U'$ by the $O(n)$-gate wiring above, and composing with a single-qubit H on qubit 1 (to swap $X_1 \leftrightarrow Z_1$ as specified) yields a circuit whose conjugation matches U on both the first-qubit generators and on $I \otimes G_n$. Total gate count: $O(n) + O(n^2) = O(n^2)$.

3. General $(n{+}1)$**-qubit case.** For any $U \in N(G_{n+1})$ there exist $P_1, P_1' \in \{X_1, Y_1, Z_1\}$ and $g, g' \in G_n$ with

$$UZ_1U^\dagger = \pm P_1 \otimes g, \qquad UX_1U^\dagger = \pm P_1' \otimes g',$$

obeying the Pauli commutation/anticommutation relations. Using only H and S on the first qubit (constant cost), map the pair (P_1, P_1') to the canonical pair (X_1, Z_1) or (Z_1, X_1) (signs are irrelevant up to global phase). This reduces to the canonical case handled in Part 2, yielding an $O(n^2)$ implementation over $\{H, S, \mathrm{CNOT}\}$. $\square$

Exercise 10.41

Verify Equations (10.92) through (10.95).

Solution

Concepts Involved: Controlled Operations

Toffoli projector form. Let $U = \mathrm{CCNOT}_{1,2\to 3}$ and $\Pi = \frac{1}{4}(I - Z_1)(I - Z_2)$. Set $A := \Pi \otimes (X_3 - I_3)$. Then $U = I + A$ and

$$UOU^\dagger = (I + A)O(I + A) = O + \{A, O\} + AOA$$

for any operator O.

(1) $UZ_1U^\dagger = Z_1$ and $UZ_2U^\dagger = Z_2$. Since Z_1 (and Z_2) commute with Π and with $X_3 - I_3$, we have $[A, Z_j] = 0$ and $AZ_jA = 0$. Hence $UZ_jU^\dagger = Z_j$ for $j = 1, 2$.

(2) $UX_3U^\dagger = X_3$. Here X_3 commutes with Π. Compute $\{A, X_3\} = \Pi\otimes\{X_3-I_3, X_3\} = \Pi\otimes 2(I_3-X_3)$, and $AX_3A = \Pi\otimes(X_3-I_3)X_3(X_3-I_3) = \Pi\otimes 2(X_3-I_3)$. These two terms cancel, giving $UX_3U^\dagger = X_3$.

(3) $UX_1U^\dagger = \dfrac{1}{2}X_1\,(I + Z_2 + X_3 - Z_2X_3)$. Use

$$\{\Pi, X_1\} = \tfrac{1}{4}\,\{I - Z_1, X_1\}(I - Z_2) = \tfrac{1}{2}\,X_1\,(I - Z_2), \qquad \Pi X_1\Pi = 0,$$

(the last because $(I - Z_1)X_1(I - Z_1) = 0$). Thus

$$\begin{aligned} UX_1U^\dagger &= X_1 + \{A, X_1\} + AX_1A = X_1 + \{\Pi, X_1\} \otimes (X_3 - I_3) + 0 \\ &= X_1 + \tfrac{1}{2}X_1(I - Z_2)(X_3 - I_3) = \tfrac{1}{2}\,X_1\,(I + Z_2 + X_3 - Z_2X_3). \end{aligned}$$

(4) $UX_2U^\dagger = \dfrac{1}{2}X_2\,(I + Z_1 + X_3 - Z_1X_3)$. This is identical to (3) by the $1 \leftrightarrow 2$ symmetry of U.

(5) $UZ_3U^\dagger = \dfrac{1}{2}Z_3\,(I+Z_1+Z_2-Z_1Z_2)$. We have $\{X_3-I_3, Z_3\} = -2Z_3$ and $(X_3-I_3)Z_3(X_3-I_3) = 0$, and Z_3 commutes with Π. Therefore

$$UZ_3U^\dagger = Z_3 + \{A, Z_3\} + AZ_3A = Z_3 + \Pi \otimes (-2Z_3) + 0 = Z_3\,(I - 2\Pi).$$

Since $I - 2\Pi = \frac{1}{2}\,(I + Z_1 + Z_2 - Z_1Z_2)$, the claim follows. $\square$

Exercise 10.42

Use the stabilizer formalism to verify that the circuit of Figure 1.13 on page 27 teleports qubits, as claimed. Note that the stabilizer formalism restricts the class of states being teleported, so in some sense this is not a complete description of teleportation, nevertheless it does allow an understanding of the dynamics of teleportation.

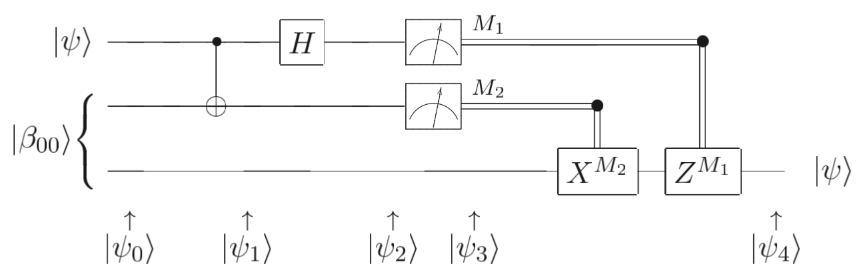


Figure 1.13. Quantum circuit for teleporting a qubit. The two top lines represent Alice's system, while the bottom line is Bob's system. The meters represent measurement, and the double lines coming out of them carry classical bits (recall that single lines denote qubits).

Solution

Concepts Involved: Stabilizer Formalism, Quantum Teleporation

---

Let us label the qubits as follows: qubit 1 is Alice's input (A), qubit 2 is the resource qubit (R), and qubit 3 is Bob's qubit (B). The Bell state $|\beta_{00}\rangle_{RB}$ has stabilizers $Z_R Z_B$ and $X_R X_B$. Alice performs CNOT $A \to R$, then H_A, then measures Z_A with outcome m_1 and Z_R with outcome m_2. Bob applies $X^{m_2} Z^{m_1}$ to qubit B.

Case I: input $|0\rangle_A$ (stabilizer Z_A). We start with a state with the stabilizer group

$$S = \langle Z_A,\ Z_R Z_B,\ X_R X_B \rangle.$$

After CNOT $A \to R$,

$$S = \langle Z_A,\ Z_A Z_R Z_B,\ X_R X_B \rangle.$$

After H_A,

$$S = \langle X_A,\ X_A Z_R Z_B,\ X_R X_B \rangle \equiv \langle X_A,\ Z_R Z_B,\ X_R X_B \rangle$$

Measure Z_A (outcome m_1): replace X_A by $(-1)^{m_1} Z_A$:

$$S = \langle (-1)^{m_1} Z_A,\ Z_R Z_B,\ X_R X_B \rangle.$$

Measure Z_R (outcome m_2): replace $X_R X_B$ by $(-1)^{m_2} Z_R$ and $Z_R Z_B$ by $(-1)^{m_2} Z_B$

$$S = \langle (-1)^{m_1} Z_A,\ (-1)^{m_2} Z_B,\ (-1)^{m_2} Z_R \rangle.$$

Now, looking at Bob's stabilizer, we have $\langle (-1)^{m_2} Z_B \rangle$. Thus Bob's state is $Z^{m_1} X^{m_2} |0\rangle$.

Case II: input $|+\rangle_A$ (stabilizer X_A). Initial stabilizer group to start with

$$S = \langle X_A,\ Z_R Z_B,\ X_R X_B \rangle.$$

After CNOT $A \to R$:

$$S = \langle X_A X_R,\ Z_A Z_R Z_B,\ X_R X_B \rangle.$$

After H_A,

$$S = \langle Z_A X_R,\ X_A Z_R Z_B,\ X_R X_B \rangle.$$

Measure Z_A (outcome m_1): replace $X_A Z_R Z_B$ by $(-1)^{m_1} Z_A$ and reduce $Z_A X_R$ to $(-1)^{m_1} X_R$:

$$S = \langle (-1)^{m_1} X_R,\ X_R X_B,\ (-1)^{m_1} Z_A \rangle.$$

Measure Z_R (outcome m_2): replace X_R by $(-1)^{m_2} Z_R$ and update $X_R X_B \mapsto (-1)^{m_1} X_B$:

$$S = \langle (-1)^{m_2} Z_R,\ (-1)^{m_1} X_B,\ (-1)^{m_1} Z_A \rangle.$$

Eliminating A, R, Bob's stabilizer is $\langle (-1)^{m_1} X_B \rangle$. Thus Bob's state is $Z^{m_1} X^{m_2} |+\rangle$.

□

Remark: Technically speaking, one should also consider starting with $|1\rangle$ or $|-\rangle$ as input; following the sign changes through the stabilizer updates leads to the same conclusion as in Cases I and II.

Exercise 10.43

Show that $S \subseteq N(S)$ for any subgroup S of G_n.

Solution

Concepts Involved: Group Theory, Normalizer

By definition,

$$N(S) = \{g \in G_n \mid gsg^{-1} \in S \text{ for all } s \in S\}$$

$$\text{For any } s \in S, \text{ and all } s' \in S, \text{ we have } ss's^{-1} \in S \text{ since } S \text{ is a group.}$$

$$\Rightarrow s \in N(S) \text{ for all } s \in S \Rightarrow S \subseteq N(S)$$

□

Remark: $S = N(S)$ iff S is self-normalizing. For stabilizer codes, $N(S) \setminus S$ contains logical operators.

Exercise 10.44

Show that $N(S) = Z(S)$ for any subgroup S of G_n not containing $-I$.

Solution

Concepts Involved: Group Theory, Pauli Group, Normalizer, Centralizer Pauli group G_n, Group theory.

$$N(S) = \{g \in G_n \mid gsg^{-1} \in S \text{ for all } s \in S\}, \quad Z(S) = \{g \in G_n \mid gs = sg \text{ for all } s \in S\}$$

Assume $-I \notin S$. Let $g \in N(S)$. Then for all $s \in S$,

$$gsg^{-1} = \omega_s s \quad \text{for some } \omega_s \in \{\pm 1, \pm i\} \Rightarrow gs = \omega_s sg$$

If any $\omega_s \neq 1$, then $gs \neq sg \Rightarrow g \notin Z(S)$, but $\omega_s g \in S \Rightarrow \pm i \in S$, implying $-I \in S$, contradiction. So all $\omega_s = 1$, and thus $g \in Z(S)$. Therefore, $N(S) \subseteq Z(S)$. Since always $Z(S) \subseteq N(S)$, we conclude:

$$N(S) = Z(S)$$

□

Remark: If $-I \in S$, then elements may normalize S while anti-commute with some $s \in S$, breaking the equality. This makes the exclusion of $-I$ essential in stabilizer codes.

Exercise 10.45: Correcting located errors

Suppose $C(S)$ is an $[n, k, d]$ stabilizer code. Suppose k qubits are encoded in n qubits using this code, which is then subjected to noise. Fortunately, however, we are told that only $d-1$ of the qubits are affected by the noise, and moreover, we are told precisely which $d-1$ qubits have been affected. Show that it is possible to correct the effects of such *located* errors.

Solution

Concepts Involved: Stabilizer Codes, Code Distance, Error Correction Conditions

Let $C(S)$ be an $[n, k, d]$ stabilizer code with code projector P. Suppose noise acts on a known set $L \subseteq [n]$ of ℓ qubits with $\ell \leq d-1$ ("located" errors). To show these errors are correctable, it suffices (by linearity) to verify the Knill–Laflamme conditions for a Pauli basis of errors supported on L.
Let $\{E_\alpha\}$ be all n-qubit Pauli operators supported on L. For any α, β we have $E_\alpha^\dagger E_\beta$ supported on L, hence

$$\mathrm{wt}(E_\alpha^\dagger E_\beta) \;\leq\; \ell \;\leq\; d-1.$$

By distance d, any nontrivial Pauli of weight $< d$ either (i) maps the code space to an orthogonal subspace (i.e. is detectable), in which case $PE_\alpha^\dagger E_\beta P = 0$, or (ii) lies in the stabilizer S, in which case $PE_\alpha^\dagger E_\beta P = P$. Thus there exist scalars $c_{\alpha\beta}$ with

$$PE_\alpha^\dagger E_\beta P \;=\; c_{\alpha\beta}\, P,$$

namely $c_{\alpha\beta} = 1$ if $E_\alpha^\dagger E_\beta \in S$ and $c_{\alpha\beta} = 0$ otherwise.
These are exactly the Knill–Laflamme error-correction conditions, so the set of all errors supported on L is correctable. Equivalently, any $[n, k, d]$ code can correct up to $d-1$ located errors on known positions. □

Exercise 10.46

Show that the stabilizer for the three qubit phase flip code is generated by X_1X_2 and X_2X_3.

Solution

Concepts Involved: Phase Flip Codes, Stabilizer codes, Generators

The three-qubit phase-flip code encodes

$$|0_L\rangle = |+++\rangle\,, \qquad |1_L\rangle = |---\rangle\,,$$

with $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$. Since $X\,|+\rangle = |+\rangle$ and $X\,|-\rangle = -\,|-\rangle$, we can check the stabilizers

$$g_1 = X_1 X_2, \qquad g_2 = X_2 X_3.$$

On $|0_L\rangle = |+++\rangle$ both g_1 and g_2 act trivially, giving eigenvalue $+1$. On $|1_L\rangle = |---\rangle$, two X operators act, so the minus signs cancel, again giving eigenvalue $+1$. Thus both logical states are stabilized by g_1 and g_2. These generate the stabilizer group

$$S = \{I,\, X_1X_2,\, X_2X_3,\, X_1X_3\}.$$

Since there are $n = 3$ qubits and $r = 2$ independent generators, the stabilized subspace has dimension $2^{n-r} = 2$, which is precisely the logical qubit of the code. Therefore, the stabilizer of the three-qubit phase-flip code is generated by X_1X_2 and X_2X_3. □

Remark: This stabilizer structure is directly analogous to that of the three-qubit bit-flip code, whose stabilizer is generated by Z_1Z_2 and Z_2Z_3. The two codes are related by applying Hadamard gates to each qubit, $H^{\otimes 3}$, which interchange $Z \leftrightarrow X$. Thus, the phase-flip code is simply the Hadamard-transformed version of the bit-flip code.

Exercise 10.47

Verify that the generators of Figure 10.11 generate the two codewords of Equation (10.13).

Name	Operator
g_1	$Z\,Z\,I\,I\,I\,I\,I\,I\,I$
g_2	$I\,Z\,Z\,I\,I\,I\,I\,I\,I$
g_3	$I\,I\,I\,Z\,Z\,I\,I\,I\,I$
g_4	$I\,I\,I\,I\,Z\,Z\,I\,I\,I$
g_5	$I\,I\,I\,I\,I\,I\,Z\,Z\,I$
g_6	$I\,I\,I\,I\,I\,I\,I\,Z\,Z$
g_7	$X\,X\,X\,X\,X\,X\,I\,I\,I$
g_8	$I\,I\,I\,X\,X\,X\,X\,X\,X$
$\bar{Z}$	$X\,X\,X\,X\,X\,X\,X\,X\,X$
$\bar{X}$	$Z\,Z\,Z\,Z\,Z\,Z\,Z\,Z\,Z$

Figure 10.11. The eight generators for the Shor nine qubit code, and the logical Z and logical X operations. (Yes, they really are the reverse of what one might naively expect!)

Solution

Concepts Involved: Shor Code, Stabilizer Code, Generators

Letting $S = \langle g_1, g_2, g_3, g_4, g_5, g_6, g_7, g_8 \rangle$, we have $n - k = 8$ independent and mutually commuting generators from G_7, and hence V_S is $2^1 = 2$-dimensional (there are only 2 codewords) from Proposition 10.5. We could generate the two codewords via brute force by defining the projector:

$$P_S^{(0,0,0,0,0,0,0,0)} = \frac{\prod_{i=1}^{8}(I + g_i)}{2^7}$$

and finding the two eigenvectors corresponding to non-zero eigenvalues. However, since we are given the codewords:

$$|0_L\rangle = \frac{(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)}{2\sqrt{2}}$$
$$|1_L\rangle = \frac{(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)}{2\sqrt{2}}$$

to verify them we only need verify that they are both $+1$ eigenvalues of $g_1, \ldots, g_8$.
First, note that for the generators $g_1, \ldots g_6$ that they are all pairs of Zs that act on the same 3-qubit blocks, and hence leave the codewords invariant (the sign of $|111\rangle$ in a given block is flipped twice). For g_7, g_8, note that this involves swapping $|000\rangle \leftrightarrow |111\rangle$ in two blocks; for $|0_L\rangle$ this does nothing (every block is left invariant by XXX) and for $|1_L\rangle$ we a two minus signs (one per block) which cancels. Hence the two codewords are eigenstates of the generators. □

Exercise 10.48

Show that the operations $\bar{Z} = X_1X_2X_3X_4X_5X_6X_7X_8X_9$ and $\bar{X} = Z_1Z_2Z_3Z_4Z_5Z_6Z_7Z_8Z_9$ act as logical Z and X operations on a Shor-code encoded qubit. That is, show that this $\bar{Z}$ is independent of and commutes with the generators of the Shor code, and that $\bar{X}$ is independent of and commutes with the generators of the Shor code, and anti-commutes with $\bar{Z}$.

Solution

Concepts Involved: Shor Code, Stabilizer Code, Logicals

Let S be the stabilizer of the Shor $[9, 1, 3]$ code with eight generators: six Z-type pair checks within each 3-qubit block and two X-type weight-6 checks that couple blocks. Consider

$$\overline{Z} = X_1X_2 \cdots X_9 = X^{\otimes 9}, \qquad \overline{X} = Z_1Z_2 \cdots Z_9 = Z^{\otimes 9}.$$

(1) $\overline{Z}$ commutes with S. For any Z-pair generator Z_iZ_j (within a block), we have $X^{\otimes 9}(Z_iZ_j) = (-Z_i)(-Z_j)X^{\otimes 9} = Z_iZ_jX^{\otimes 9}$, since there are two overlaps with Z, giving a net phase $(-1)^2 = +1$. For each weight-6 X-type generator, all overlapping factors are X, hence commute trivially with $X^{\otimes 9}$. Therefore $\overline{Z} \in N(S)$.
(2) $\overline{X}$ commutes with S. It clearly commutes with every Z-pair generator. For a weight-6 X-type generator, the overlap with $Z^{\otimes 9}$ is on 6 qubits, and $ZX = -XZ$ per overlap; thus the total phase is $(-1)^6 = +1$, so they commute. Hence $\overline{X} \in N(S)$.
(3) Independence: $\overline{Z}, \overline{X} \notin S$. Every element of S has X-support on either zero or exactly two of the 3-qubit blocks, since products of the two weight-6 X-generators toggle blocks in pairs; thus no stabilizer element has X on all three blocks simultaneously, so $X^{\otimes 9} \notin S$. Likewise, within any single block the

Z-pair generators toggle Z's two at a time, so every stabilizer element has an even number of Z's per block; therefore $Z^{\otimes 3}$ on a block (odd parity) cannot arise, and $Z^{\otimes 9} \notin S$.

(4) $\overline{X}$ *anti-commutes with* $\overline{Z}$*.* They overlap on 9 qubits, so

$$\overline{X}\,\overline{Z} = (-1)^9\,\overline{Z}\,\overline{X} = -\,\overline{Z}\,\overline{X}.$$

Combining (1)–(4), $\overline{Z}, \overline{X} \in N(S) \setminus S$ and satisfy the Pauli relation $\overline{X}\,\overline{Z} = -\overline{Z}\,\overline{X}$, hence they act as the logical Z and X operators on the Shor-code encoded qubit. □

Exercise 10.49

Use Theorem 10.8 to verify that the five qubit code can protect against an arbitrary single-qubit error.

Solution

Concepts Involved: Five Qubit Code, Stabilizer Code, Error Correction Conditions

---

Let the five-qubit code have stabilizer

$$S = \langle g_1, g_2, g_3, g_4 \rangle, \qquad g_1 = XZZXI,\;\; g_2 = IXZZX,\;\; g_3 = XIXZZ,\;\; g_4 = ZXIXZ,$$

and logical operators $\bar{Z} = ZZZZZ$, $\bar{X} = XXXXX$. By Theorem 10.8, a set of errors $\{E_j\}$ is correctable iff for all j, k, either $E_j^\dagger E_k \in S$ or $E_j^\dagger E_k$ anticommutes with at least one generator of S (i.e. yields a nonzero syndrome).

Consider $\mathcal{E} = \{I, X_i, Y_i, Z_i \;:\; i = 1, \ldots, 5\}$. For a single-qubit Pauli on site i, its 4-bit syndrome $s(E) = (s_1, \ldots, s_4) \in \{\pm 1\}^4$ is determined by commutation with the generators:

$$s_\ell(E) = -1 \iff Eg_\ell = -g_\ell E, \qquad s_\ell(E) = +1 \iff Eg_\ell = g_\ell E.$$

Equivalently, writing a binary check-matrix where rows index g_ℓ and columns index qubits, the syndrome of X_i is the "Z-column" at i (which g_ℓ have Z on qubit i); the syndrome of Z_i is the "X-column" at i; and the syndrome of Y_i is the bitwise XOR of those two (since Y anticommutes with both X and Z).

Reading off from $(g_1, \ldots, g_4)$ gives (rows ordered g_1, g_2, g_3, g_4):

	$i = 1$	2	3	4	5
X-column	$[1,0,1,0]$	$[0,1,0,1]$	$[0,0,1,0]$	$[1,0,0,1]$	$[0,1,0,0]$
Z-column	$[0,0,0,1]$	$[1,0,0,0]$	$[1,1,0,0]$	$[0,1,1,0]$	$[0,0,1,1]$

Hence the five X-columns are pairwise distinct and nonzero; the five Z-columns are pairwise distinct and nonzero; and each Y-column (XOR of its X- and Z-columns) is also distinct and different from any X- or Z-column. Therefore all 15 nontrivial errors in $\mathcal{E}$ have pairwise distinct nonzero syndromes. Consequently, for any distinct single-qubit errors $E_j \neq E_k$, the product $E_j^\dagger E_k$ anticommutes with at least one generator of S, satisfying Theorem 10.8. It follows that the five-qubit code corrects an arbitrary single-qubit error. (Equivalently, $d = 3$ so $t = 1$ is correctable.) □

Exercise 10.50

Show that the five qubit code saturates the quantum Hamming bound, that is, it satisfies the inequality of (10.51) with equality.

Solution

Concepts Involved: Five Qubit Code, Quantum Hamming Bound

The quantum Hamming bound for a non-degenerate $[n,k]$ code correcting up to t errors is

$$\sum_{j=0}^{t} \binom{n}{j} 3^j\, 2^k \;\leq\; 2^n.$$

This expresses that each possible error (identity plus all weight-$\leq\, t$ Pauli errors) produces a distinct 2^k-dimensional subspace, all of which must fit inside the 2^n-dimensional Hilbert space.
For the five-qubit code, $n=5$, $k=1$, and $t=1$. Thus,

$$\left[\binom{5}{0}3^0 + \binom{5}{1}3^1\right]2^1 = (1+15)\cdot 2 = 32 = 2^5.$$

So the inequality holds with equality. Therefore the five-qubit $[5,1,3]$ code exactly saturates the quantum Hamming bound, making it a *perfect code*. □

Exercise 10.51

Verify that the check matrix defined in (10.106) corresponds to the stabilizer of the CSS code $CSS(C_1, C_2)$, and use Theorem 10.8 to show that arbitrary errors on up to t qubits may be corrected by this code.

Solution

Concepts Involved: Stabilizer Codes, CSS Codes, Check Matrix Representation, Code Distance

The check matrix

$$[Z|X] = \begin{bmatrix} H(C_2^\perp) & 0 \\ 0 & H(C_1) \end{bmatrix}$$

defines Z-type stabilizers Z^z for $z \in C_2$ (rows of $H(C_2^\perp)$) and X-type stabilizers X^x for $x \in C_1^\perp$ (rows of $H(C_1)$). Since $C_2 \subseteq C_1$ we have $C_1^\perp \subseteq C_2^\perp$, hence $H(C_2^\perp)H(C_1)^T = 0$ and all stabilizers commute. The code space stabilized is spanned by

$$\{|x + C_2\rangle = \tfrac{1}{\sqrt{|C_2|}} \sum_{y \in C_2} |x+y\rangle : x \in C_1\},$$

which is exactly $CSS(C_1, C_2)$ of dimension $2^{k_1-k_2}$. The normalizer is

$$N(S) = \{X^u Z^v : u \in C_1,\ v \in C_2^\perp\}, \quad S = \{X^u Z^v : u \in C_2,\ v \in C_1^\perp\}.$$

Thus if $F = X^u Z^v \in N(S) \setminus S$, then either $u \in C_1 \setminus C_2$ or $v \in C_2^\perp \setminus C_1^\perp$. In the first case, u is a nonzero codeword of C_1 not contained in C_2, so by definition $\mathrm{wt}(u) \geq d(C_1)$. In the second case, v is a nonzero codeword of $C_2^\perp$ not in $C_1^\perp$, so $\mathrm{wt}(v) \geq d(C_2^\perp)$. Hence any nontrivial logical operator has weight at least $\min\{d(C_1), d(C_2^\perp)\}$.
Let $\mathcal{E}$ be all Pauli errors of weight $\leq t$. For $E_j, E_k \in \mathcal{E}$ we have $F = E_j^\dagger E_k = X^u Z^v$ of weight $\leq 2t$. If $F \in N(S) \setminus S$, then either $\mathrm{wt}(u) \geq d(C_1)$ or $\mathrm{wt}(v) \geq d(C_2^\perp)$. Under the assumption $d(C_1), d(C_2^\perp) \geq 2t+1$ this is impossible since $\mathrm{wt}(F) \leq 2t$. Thus $E_j^\dagger E_k \notin N(S) \setminus S$ for all j, k, and by Theorem 10.8 the code corrects all errors on up to t qubits. □

Exercise 10.52

Verify by direct operation on the codewords that the operators of (10.107) act appropriately, as logical Z and X.

Solution

Concepts Involved: Steane Code, Stabilizer Codes, CSS Codes, Logicals

---

The Steane code is a CSS code based on the $[7, 4, 3]$ Hamming code C and its dual $C^\perp$. The logical codewords can be written as

$$|0_L\rangle = \frac{1}{\sqrt{|C^\perp|}} \sum_{x \in C^\perp} |x\rangle, \qquad |1_L\rangle = \frac{1}{\sqrt{|C^\perp|}} \sum_{x \in C^\perp} |x \oplus u\rangle,$$

where $u = 1111111$.

Action of $Z^{\otimes 7}$. For any string y, $Z^{\otimes 7}|y\rangle = (-1)^{\mathrm{wt}(y)}|y\rangle$. Every $x \in C^\perp$ has even weight (0 or 4), hence

$$Z^{\otimes 7}|0_L\rangle = |0_L\rangle.$$

For $x \oplus u$, the weight is odd since $\mathrm{wt}(u) = 7$ and $\mathrm{wt}(x)$ is even, so

$$Z^{\otimes 7}|1_L\rangle = -|1_L\rangle.$$

Thus $Z^{\otimes 7}$ acts as the logical $\bar{Z}$ operator.

Action of $X^{\otimes 7}$. Bitwise X flips bits: $X^{\otimes 7}|y\rangle = |y \oplus u\rangle$. Therefore

$$X^{\otimes 7}|0_L\rangle = \frac{1}{\sqrt{|C^\perp|}} \sum_{x \in C^\perp} |x \oplus u\rangle = |1_L\rangle,$$

and similarly $X^{\otimes 7}|1_L\rangle = |0_L\rangle$. Thus $X^{\otimes 7}$ acts as the logical $\bar{X}$ operator.
Thus, the transversal operators of (10.107),

$$\bar{Z} = Z^{\otimes 7}, \qquad \bar{X} = X^{\otimes 7},$$

indeed act on the Steane codewords as the encoded Pauli Z and X. □

Exercise 10.53

Prove that the encoded Z operators are independent of one another.

Solution

Concepts Involved: Stabilizer Codes, Check Matrix Representation, Logicals

We put the stabilizer check matrix into the standard form of Eq. (10.111). In this form, we choose the k encoded Z operators with check matrix

$$G_Z^{(\mathrm{enc})} = \left[0 \; 0 \; 0 \; \middle| \; A_2^\mathsf{T} \; 0 \; I_k \right].$$

Each $\bar{Z}_j$ has no X part and a Z part that contains a 1 in the j-th of the last k columns, together with additional entries given by A_2^T.

Now suppose that a nontrivial product of the encoded Z's equals a stabilizer element. Equivalently, let $v \in \mathbb{F}_2^k \setminus \{0\}$ and consider

$$r = v^\mathsf{T} G_Z^{(\mathrm{enc})}.$$

The vector r has X-part equal to 0, and its Z-part has last k entries equal to $v \neq 0$.

Next, we examine all possible stabilizer vectors. Every stabilizer is a binary sum of rows of the stabilizer check matrix G. In standard form, the stabilizer rows place an identity matrix in the *earlier* blocks of the Z-part, but they contain no terms in the last k Z-columns reserved for $G_Z^{(\mathrm{enc})}$. Therefore, no linear combination of stabilizer rows can yield a vector r with X-part zero and last-k Z columns equal to a nonzero v.

This contradiction shows that the only solution is $v = 0$. Hence, no nontrivial product of the encoded Z's lies in the stabilizer, and we conclude that the encoded Z operators are independent of one another. □

Exercise 10.54

Prove that with the check matrix for the encoded X operators defined as above, the encoded X operators are independent of one another and of the generators, commute with the generators of the stabilizer, with each other, and $\bar{X}_j$ commutes with all the $\bar{Z}_k$ except with $\bar{Z}_j$, with which it anti-commutes.

Solution

Concepts Involved: Stabilizer Codes, Check Matrix Representation, Logicals

We work in the binary symplectic representation and put the stabilizer check matrix into the standard form of Eq. (10.111). In this form we *choose* the k encoded Z and X operators to have $k \times 2n$ check matrices

$$G_Z = \left[0 \; 0 \; 0 \; \middle| \; A_2^\mathsf{T} \; 0 \; I_k \right], \qquad G_X = \left[0 \; E^\mathsf{T} \; I_k \; \middle| \; C^\mathsf{T} \; 0 \; 0 \right].$$

Thus a row of G_X (the jth $\bar{X}_j$) has X-part supported only in the middle block E^T and in the *last* k *X-columns* (with a 1 at column $n - k + j$), and Z-part supported only in the left block C^T. A row of G_Z (the jth $\bar{Z}_j$) has no X-part and a Z-part whose *last* k *Z-columns* contain the unit vector e_j.

Independence of the $\bar{X}_j$ and independence from the stabilizer. We take a binary combination $v \in \mathbb{F}_2^k$ of the rows of G_X. By construction, the X-part of $\sum_j v_j \bar{X}_j$ has its last k columns equal to v, thanks to the terminal I_k. In standard form, stabilizer rows have no support in those last k X-columns, so no nonzero v can be realized by a stabilizer combination. Therefore the $\bar{X}_j$ are independent of one another and of the stabilizer.

Commutation with the stabilizer. Write a stabilizer row as $[x_s|z_s]$ and a logical-X row as $[x_X|z_X]$. The commutation condition for Paulis is the vanishing symplectic product

$$\langle [x_X|z_X], [x_s|z_s] \rangle \;=\; x_X z_s^\mathsf{T} + z_X x_s^\mathsf{T} \;=\; 0 \quad \text{(over } \mathbb{F}_2\text{)}.$$

Because x_X has support only in (E^T, I_k) and z_X only in C^T, while $[x_s|z_s]$ has support confined to the stabilizer blocks of Eq. (10.111), the two cross-terms cancel blockwise, so each $\bar{X}_j$ commutes with every generator of S.

Mutual commutation of the encoded X's. For $i \neq j$, take two rows $[x_X^{(i)}|z_X^{(i)}]$ and $[x_X^{(j)}|z_X^{(j)}]$ of G_X. The symplectic product $x_X^{(i)} z_X^{(j)\mathsf{T}} + z_X^{(i)} x_X^{(j)\mathsf{T}}$ vanishes: the terminal I_k support of $x_X^{(i)}$ has no matching Z-support in $z_X^{(j)}$, and the shared $(E^\mathsf{T}, C^\mathsf{T})$ blocks are arranged to be orthogonal in the standard form. Hence $[\bar{X}_i, \bar{X}_j] = 0$.

Commutation/anticommutation with the encoded Z's. A row of G_Z has $x_Z = 0$ and z_Z with last-k Z-columns equal to e_j. A row of G_X has x_X with last-k X-columns equal to e_i and z_X with no support in the last-k Z-columns. Therefore

$$\langle \bar{X}_i, \bar{Z}_j \rangle = x_X z_Z^\mathsf{T} + z_X x_Z^\mathsf{T} = e_i e_j^\mathsf{T} = \delta_{ij}.$$

Thus $\bar{X}_j$ commutes with all $\bar{Z}_k$ for $k \neq j$, and anti-commutes with $\bar{Z}_j$.

Conclusion. With $G_X = [0\, E^\mathsf{T}\, I_k \mid C^\mathsf{T}\, 0\, 0]$ in standard form, the encoded X operators are independent of one another and of the stabilizer, commute with the stabilizer and with each other, and satisfy the logical Pauli relations $\bar{X}_j \bar{Z}_k = (-1)^{\delta_{jk}} \bar{Z}_k \bar{X}_j$.

□

Exercise 10.55

Find the $\bar{X}$ operator for the standard form of the Steane code.

Solution

Concepts Involved: Shor Code, Stabilizer Codes, Check Matrix Representation, Logicals, Standard Form

From the standard-form check matrix of the Steane code (Eq. 10.112 in Nielsen & Chuang), we can read off

$$A_2 = (1, 1, 0),$$

so the encoded logical operator is

$$\bar{Z} = Z_1 Z_2 Z_7 \quad \text{(standard-form labeling)}.$$

To determine $\bar{X}$, we require an operator that (i) commutes with all stabilizer generators, and (ii) anti-commutes with $\bar{Z}$. Because the code is CSS, we can restrict attention to X-type operators. To anti-commute with $Z_1 Z_2 Z_7$, the operator must overlap an odd number of times with $\{1, 2, 7\}$. A minimal choice is

$$\bar{X} = X_1 X_2 X_7.$$

Undoing the swaps used to bring the check matrix into standard form relabels the qubits, yielding

$$\bar{X} = X_2 X_4 X_6 \quad \text{(original Steane labeling)}.$$

This logical operator is equivalent (up to multiplication by stabilizers) to the transversal form

$$\bar{X} = X^{\otimes 7}.$$

□

Exercise 10.56

Show that replacing an encoded X or Z operator by g times that operator, where g is an element of the stabilizer, does not change the action of the operator on the code.

Solution

Concepts Involved: Stabilizer Codes, Logicals

The code space is defined by the stabilizer group S as

$$\mathcal{C} = \{\, |\psi\rangle : g|\psi\rangle = |\psi\rangle \;\; \forall g \in S \,\}.$$

Let $\bar{O}$ be a logical operator (for example $\bar{X}$ or $\bar{Z}$). If we replace it by $g\bar{O}$ with $g \in S$, then for any code state $|\psi\rangle \in \mathcal{C}$ we have

$$(g\bar{O})|\psi\rangle = g(\bar{O}|\psi\rangle).$$

Because $\bar{O}|\psi\rangle$ is again a code state, and g acts as the identity on all code states, it follows that

$$g(\bar{O}|\psi\rangle) = \bar{O}|\psi\rangle.$$

Therefore

$$(g\bar{O})|\psi\rangle = \bar{O}|\psi\rangle \quad \forall\, |\psi\rangle \in \mathcal{C},$$

showing that $g\bar{O}$ and $\bar{O}$ have identical action on the code. Thus multiplying a logical operator by any stabilizer element does not change its encoded action. □

Exercise 10.57

Give the check matrices for the five and nine qubit codes in standard form.

Solution

Concepts Involved: Five Qubit Code, Shor Code, Check Matrix Representation, Standard Form

For an $[n, k]$ stabilizer with $(n-k)$ generators, the check matrix is $G = [\, G_X \mid G_Z \,] \in \mathbb{F}_2^{(n-k)\times 2n}$. Row operations and simultaneous column swaps (qubit relabeling) bring G to

$$G \sim \left[\begin{array}{cc|cc} I_r & A & B & 0 \\ 0 & 0 & D & I_{n-k-r} \end{array}\right], \quad r = \operatorname{rank}(G_X).$$

Five-qubit code $[\![5, 1, 3]\!]$. Stabilizers:

$$g_1 = XZZXI, \quad g_2 = IXZZX, \quad g_3 = XIXZZ, \quad g_4 = ZXIXZ.$$

Standard-form matrices:

$$G_X = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}, \qquad G_Z = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 \end{bmatrix},$$

with $r = 4 = n - k$.

Nine-qubit Shor code $[\![9, 1, 3]\!]$. Stabilizers:

$$\begin{aligned} &Z_1Z_2,\ Z_2Z_3,\ Z_4Z_5,\ Z_5Z_6,\ Z_7Z_8,\ Z_8Z_9, \\ &X_1 \cdots X_6, \quad X_4 \cdots X_9. \end{aligned}$$

After row/column operations, one explicit standard form is

$$G_X = \begin{bmatrix} 1&0&0&0&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0 \end{bmatrix}, \quad G_Z = \begin{bmatrix} 0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0 \\ 0&0&1&0&0&0&0&0&0 \\ 0&0&0&1&0&0&0&0&0 \\ 0&0&0&0&1&0&0&0&0 \\ 0&0&0&0&0&1&0&0&0 \\ 0&0&0&0&0&0&0&1&0 \\ 0&0&0&0&0&0&0&0&1 \end{bmatrix},$$

with $r = 2$ and $n - k - r = 6$.

Thus the five-qubit code has full X-rank, while the nine-qubit code splits into two X-checks and six Z-checks, reflected by the I_2 and I_6 blocks. □

Exercise 10.58

Verify that the circuits in Figures 10.13–10.15 work as described, and check the claimed circuit equivalences.

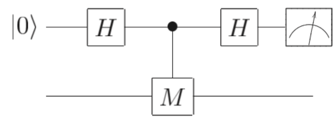


Figure 10.13. Quantum circuit for measuring a single qubit operator M with eigenvalues ± 1. The top qubit is the ancilla used for the measurement, and the bottom qubit is being measured.

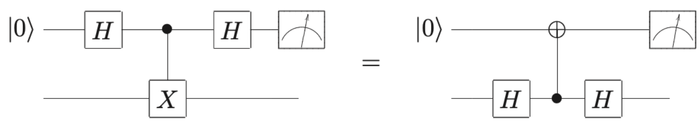


Figure 10.14. Quantum circuit for measuring the X operator. Two equivalent circuits are given; the one on the left is the usual construction (as in Figure 10.13), and the one on the right is a useful equivalent circuit.

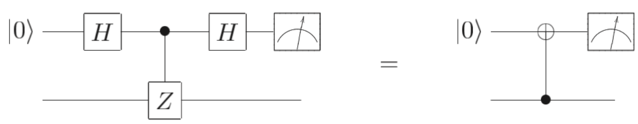


Figure 10.15. Quantum circuit for measuring the Z operator. Two equivalent circuits are given; the one on the left is the usual construction (as in Figure 10.13), and the one on the right is a useful simplification.

Solution

Concepts Involved: Quantum Measurement, Controlled Operations

Measure X_s (Fig. 10.14)	**Measure Z_s (Fig. 10.15)**
Ancilla initial: $\langle Z_a \rangle$	Ancilla initial: $\langle Z_a \rangle$
H_s: $Z_s \leftrightarrow X_s$	—
CNOT $s \to a$: $Z_a \mapsto Z_s Z_a$	CNOT $s \to a$: $Z_a \mapsto Z_s Z_a$
Now stabilizer: $\langle Z_s Z_a \rangle$	Now stabilizer: $\langle Z_s Z_a \rangle$
Final H_s: $Z_s \mapsto X_s \;\Rightarrow\; \langle X_s Z_a \rangle$	—
Measure Z_a: outcome $+1:\ \langle X_s, Z_a \rangle$ outcome $-1:\ \langle -X_s, Z_a \rangle$	Measure Z_a: outcome $+1:\ \langle Z_s, Z_a \rangle$ outcome $-1:\ \langle -Z_s, Z_a \rangle$

In the Fig. 10.14 (15) circuit, the ancilla measurement outcome projects the system into the corresponding ± 1 eigenspace of $X_s(Z_s)$.

□

Exercise 10.59

Show that by using the identities of Figures 10.14 and 10.15, the syndrome circuit of Figure 10.16 can be replaced with the circuit of Figure 10.17.

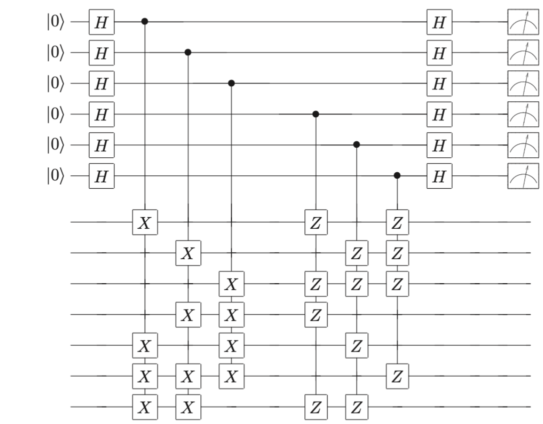


Figure 10.16. Quantum circuit for measuring the generators of the Steane code, to give the error syndrome. The top six qubits are the ancilla used for the measurement, and the bottom seven are the code qubits.

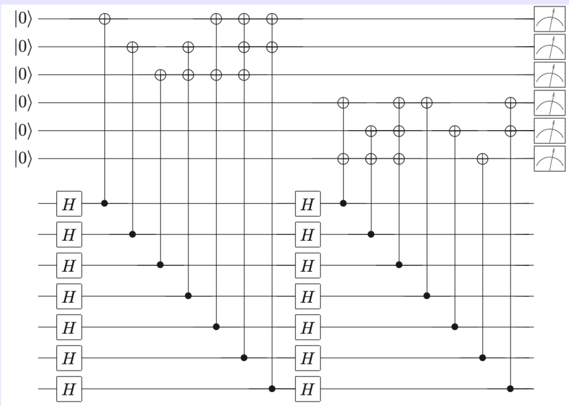


Figure 10.17. Quantum circuit equivalent to the one in Figure 10.16.

Solution

Concepts Involved: Error Syndrome, Controlled Operations

The circuit of Fig. 10.16 extracts both Z- and X-type Steane stabilizers using ancillas with Hadamards. We show it is equivalent to the Hadamard-free circuit of Fig. 10.17 by applying the identities of Figs. 10.14–10.15.

Step 1 (basis change). For the X-type checks, Fig. 10.16 prepares an ancilla in $|0\rangle$, applies H, couples it to data qubits via CNOTs, applies H, then measures in the Z basis. Using the identities

$$HXH = Z, \quad HZH = X, \quad (I \otimes H)\,\mathrm{CNOT}\,(I \otimes H) = \mathrm{CZ},$$

this procedure is equivalent to preparing $|0\rangle$, using CNOTs with *swapped control/target*, and finally measuring in Z. Thus all ancillas can be measured in Z, and the X-check ancillas simply use reversed CNOT orientation.

Step 2 (uniform fan-outs). For Z-type checks, the data act as controls and the ancilla as target. For X-type checks, after Step 1, the ancilla acts as control and the data as targets. In both cases the syndrome is copied into the ancilla by a fan-out of CNOT gates, with the qubits involved determined by the rows of the Steane parity-check matrices.

Step 3 (commutation). CNOTs with common control or disjoint wires commute, so each ancilla's set of CNOTs may be regrouped into a left-to-right "fan-out" block. This yields exactly the structure of Fig. 10.17: three ancillas coupled with data-as-control (for Z stabilizers) and three ancillas coupled with ancilla-as-control (for X stabilizers), all measured in Z.

Hence, by eliminating the explicit Hadamards and commuting CNOTs, the syndrome-extraction circuit of Fig. 10.16 is transformed into the simpler and equivalent circuit of Fig. 10.17. □

Exercise 10.60

Construct a syndrome measuring circuit analogous to that in Figure 10.16, but for the nine and five qubit codes.

Solution

Concepts Involved: Five Qubit Code, Shor Code, Error Syndrome, Controlled Operations

For a stabilizer $g = \bigotimes_i P_i$ on data qubits:

- *Z-type* ($P_i \in \{Z, I\}$): prepare ancilla $|0\rangle$; for each i with $P_i = Z$ apply $\mathrm{CNOT}_{i\to a}$; measure Z_a.
- *X-type* ($P_i \in \{X, I\}$): prepare ancilla $|+\rangle$; for each i with $P_i = X$ apply $\mathrm{CNOT}_{a\to i}$; measure X_a (i.e., H_a, then Z_a).
- *Mixed Pauli* ($P_i \in \{X, Z, I\}$): either
(A) **ancilla-operations:** prepare $|+\rangle_a$; for each X-site do $\mathrm{CNOT}_{a\to i}$, for each Z-site do $\mathrm{CNOT}_{i\to a}$; measure X_a; or
(B) **data-conjugation:** apply H on data qubits where $P_i = X$ (so $X \to Z$), run the Z-type pattern, undo those H's.

In Heisenberg form, these implement the projective measurement $\{(I \pm g)/2\}$.

Nine-qubit (Shor) code $[[9, 1, 3]]$**.** Choose generators

$$\underbrace{Z_1Z_2,\ Z_2Z_3,\ Z_4Z_5,\ Z_5Z_6,\ Z_7Z_8,\ Z_8Z_9}_{6\ Z\text{-checks}}, \qquad \underbrace{X_1X_2X_3X_4X_5X_6,\ \ X_4X_5X_6X_7X_8X_9}_{2\ X\text{-checks}}.$$

Syndrome circuits (use one ancilla per row, or reuse sequentially):

- For each pair $(i, j) \in \{(1,2),(2,3),(4,5),(5,6),(7,8),(8,9)\}$ $|0\rangle_a \xrightarrow{\mathrm{CNOT}_{i\to a},\,\mathrm{CNOT}_{j\to a}}$ measure Z_a.
- For $S_1 = \{1,2,3,4,5,6\}$ and $S_2 = \{4,5,6,7,8,9\}$: $|+\rangle_a \xrightarrow{\prod_{i\in S_r}\mathrm{CNOT}_{a\to i}}$ measure X_a.

Counts: $6 \times 2 + 6 + 6 = 24$ CNOTs, 6 Z-basis and 2 X-basis ancilla measurements; $n - k = 8$ syndrome bits.

Five-qubit code $[[5, 1, 3]]$**.** Use the standard generators (cyclic shifts of $XZZXI$)

$$g_1 = X_1Z_2Z_3X_4I_5, \quad g_2 = I_1X_2Z_3Z_4X_5,$$
$$g_3 = X_1I_2X_3Z_4Z_5, \quad g_4 = Z_1X_2I_3X_4Z_5.$$

Measure each with one ancilla the following protocol

$$\begin{aligned}
g_1 &: \ |+\rangle_a,\ \mathrm{CNOT}_{a\to1}, \mathrm{CNOT}_{2\to a}, \mathrm{CNOT}_{3\to a}, \mathrm{CNOT}_{a\to4}, \quad \text{measure } X_a,\\
g_2 &: \ |+\rangle_a,\ \mathrm{CNOT}_{a\to2}, \mathrm{CNOT}_{3\to a}, \mathrm{CNOT}_{4\to a}, \mathrm{CNOT}_{a\to5}, \quad \text{measure } X_a,\\
g_3 &: \ |+\rangle_a,\ \mathrm{CNOT}_{a\to1}, \mathrm{CNOT}_{a\to3}, \mathrm{CNOT}_{4\to a}, \mathrm{CNOT}_{5\to a}, \quad \text{measure } X_a,\\
g_4 &: \ |+\rangle_a,\ \mathrm{CNOT}_{1\to a}, \mathrm{CNOT}_{a\to2}, \mathrm{CNOT}_{a\to4}, \mathrm{CNOT}_{5\to a}, \quad \text{measure } X_a.
\end{aligned}$$

□

Exercise 10.61

Describe explicit recovery operations $E_j^\dagger$ corresponding to the different possible error syndromes that may be measured using the circuit in Figure 10.16.

Solution

Concepts Involved: Steane Code, Error Syndrome and Recovery

The Steane $[[7,1,3]]$ code is a CSS code built from the classical $[7,4,3]$ Hamming code. Its stabilizer check matrix is derived from the Hamming parity-check matrix

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

Each column of H labels one of the 7 physical qubits, and each 3-bit vector is the *syndrome* produced when that qubit alone suffers an error.

In the Steane syndrome circuit (Fig. 10.16):

- The three Z-type stabilizers give a 3-bit outcome s_X that detects whether an X or Y error has occurred, and on which qubit.
- The three X-type stabilizers give a 3-bit outcome s_Z that detects whether a Z or Y error has occurred, and on which qubit.

Error correction rule (for weight-1 errors):

$$E^\dagger = \begin{cases} I, & \text{if } s_X = 000,\ s_Z = 000, \\ X_j, & \text{if } s_X = \text{col}_j(H),\ s_Z = 000, \\ Z_j, & \text{if } s_X = 000,\ s_Z = \text{col}_j(H), \\ Y_j, & \text{if } s_X = \text{col}_j(H) = s_Z. \end{cases}$$

That is, a nonzero s_X means "apply an X on qubit j," a nonzero s_Z means "apply a Z on qubit j," and if both syndromes are nonzero and identical, the error was Y_j.

Column → qubit map. The 3-bit syndromes correspond to the columns of H as follows

j	1	2	3	4	5	6	7
$\text{col}_j(H)$	$\begin{bmatrix}1\\0\\0\end{bmatrix}$	$\begin{bmatrix}0\\1\\0\end{bmatrix}$	$\begin{bmatrix}1\\1\\0\end{bmatrix}$	$\begin{bmatrix}0\\0\\1\end{bmatrix}$	$\begin{bmatrix}1\\0\\1\end{bmatrix}$	$\begin{bmatrix}0\\1\\1\end{bmatrix}$	$\begin{bmatrix}1\\1\\1\end{bmatrix}$

The Steane $[[7,1,3]]$ code is a CSS code built from the classical $[7,4,3]$ Hamming code. Its stabilizer

check matrix is derived from the Hamming parity-check matrix

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

Each column of H labels one of the 7 physical qubits, and each 3-bit vector is the *syndrome* produced when that qubit alone suffers an error.

In the Steane syndrome circuit (Fig. 10.16):

- The three Z-type stabilizers give a 3-bit outcome s_X that detects whether an X or Y error has occurred, and on which qubit.
- The three X-type stabilizers give a 3-bit outcome s_Z that detects whether a Z or Y error has occurred, and on which qubit.

Error correction rule (for weight-1 errors):

$$E^\dagger = \begin{cases} I, & \text{if } s_X = 000,\ s_Z = 000, \\ X_j, & \text{if } s_X = \text{col}_j(H),\ s_Z = 000, \\ Z_j, & \text{if } s_X = 000,\ s_Z = \text{col}_j(H), \\ Y_j, & \text{if } s_X = \text{col}_j(H) = s_Z. \end{cases}$$

That is, a nonzero s_X means "apply an X on qubit j," a nonzero s_Z means "apply a Z on qubit j," and if both syndromes are nonzero and identical, the error was Y_j.

Column → qubit map. The 3-bit syndromes correspond to the columns of H as follows:

j	1	2	3	4	5	6	7
$\text{col}_j(H)$	$\begin{bmatrix}1\\0\\0\end{bmatrix}$	$\begin{bmatrix}0\\1\\0\end{bmatrix}$	$\begin{bmatrix}1\\1\\0\end{bmatrix}$	$\begin{bmatrix}0\\0\\1\end{bmatrix}$	$\begin{bmatrix}1\\0\\1\end{bmatrix}$	$\begin{bmatrix}0\\1\\1\end{bmatrix}$	$\begin{bmatrix}1\\1\\1\end{bmatrix}$

Summary. Each nontrivial syndrome corresponds to exactly one qubit, identified by its column of H. The type of error (X, Z, or Y) is determined by whether s_X, s_Z, or both are nonzero. If the two syndromes point to different qubits, this indicates an error of weight ≥ 2, which the Steane code cannot guarantee to correct.

□

Exercise 10.62

Show by explicit construction of generators for the stabilizer that concatenating an $[n_1, 1]$ stabilizer code with an $[n_2, 1]$ stabilizer code gives an $[n_1 n_2, 1]$ stabilizer code.

Solution

Concepts Involved: Stabilizer Codes, Generators, Code Concatenation

Let C_{out} be an $[n_1, 1]$ stabilizer code with stabilizer $S_{\text{out}} \subset \mathcal{P}_{n_1}$ and chosen logical Paulis $\overline{X}_{\text{out}}, \overline{Z}_{\text{out}} \in N(S_{\text{out}}) \setminus S_{\text{out}}$ and C_{in} be an $[n_2, 1]$ stabilizer code with stabilizer $S_{\text{in}} \subset \mathcal{P}_{n_2}$ and logical Paulis $\overline{X}_{\text{in}}, \overline{Z}_{\text{in}} \in N(S_{\text{in}}) \setminus S_{\text{in}}$.

Concatenation replaces each physical qubit of C_{out} by an n_2-qubit block encoded by C_{in}, giving n_1 blocks of size n_2 (total $n_1 n_2$ qubits). Define the blockwise "lift"

$$\varphi : \; \mathcal{P}_{n_1} \longrightarrow \mathcal{P}_{n_1 n_2}, \qquad \varphi\left(\bigotimes_{j=1}^{n_1} P_j\right) = \bigotimes_{j=1}^{n_1} \overline{P_j}_{\text{in}}^{(j)},$$

where $\overline{I}_{\text{in}} := I^{\otimes n_2}$, $\overline{X}_{\text{in}}, \overline{Z}_{\text{in}}$ are the inner logicals, $\overline{Y}_{\text{in}} := i\,\overline{X}_{\text{in}}\overline{Z}_{\text{in}}$, and the superscript $^{(j)}$ indicates action on block j only. Because $\overline{X}_{\text{in}}, \overline{Z}_{\text{in}}$ reproduce Pauli commutation relations modulo S_{in}, φ preserves commutation.

Concatenated stabilizer. Define $S_{\text{con}} \subset \mathcal{P}_{n_1 n_2}$ to be generated by

$$\bigcup_{j=1}^{n_1} S_{\text{in}}^{(j)} \quad \cup \quad \{\varphi(g) : g \in \mathcal{G}_{\text{out}}\},$$

where $S_{\text{in}}^{(j)} := \{s^{(j)} : s \in S_{\text{in}}\}$ and $\mathcal{G}_{\text{out}}$ is any independent generating set of S_{out} (of size $n_1 - 1$).

Abelianness and stabilization. Elements within each $S_{\text{in}}^{(j)}$ commute, and different blocks have disjoint support, so $\bigcup_j S_{\text{in}}^{(j)}$ is abelian. Since $\overline{X}_{\text{in}}, \overline{Z}_{\text{in}} \in N(S_{\text{in}})$, every $\varphi(g)$ commutes with every $S_{\text{in}}^{(j)}$. If $g, h \in S_{\text{out}}$ commute, then, by symplectic preservation, $[\varphi(g), \varphi(h)] = 0$. Hence S_{con} is abelian. Writing the encoder as

$$\mathcal{E}_{\text{con}} = (\mathcal{E}_{\text{in}})^{\otimes n_1} \circ \mathcal{E}_{\text{out}},$$

each $S_{\text{in}}^{(j)}$ fixes $\text{Im}(\mathcal{E}_{\text{in}})$ in block j, and for $g \in S_{\text{out}}$, $\varphi(g)$ acts as the outer stabilizer on the encoded logicals of the blocks, so it also fixes $\text{Im}(\mathcal{E}_{\text{con}})$. Thus S_{con} stabilizes the concatenated code space.

Independence and parameters. Choose independent generators for each block: $|S_{\text{in}}| = n_2 - 1$ for an $[n_2, 1]$ code, so $\sum_j |S_{\text{in}}^{(j)}| = n_1(n_2 - 1)$. The lifted set $\{\varphi(g) : g \in \mathcal{G}_{\text{out}}\}$ contributes $n_1 - 1$ *additional* independent generators: reduce modulo the subgroup generated by $\bigcup_j S_{\text{in}}^{(j)}$ (i.e., project each block to the logical coset space $N(S_{\text{in}})/S_{\text{in}}$). Under this projection, every $S_{\text{in}}^{(j)}$ becomes trivial, while $\varphi(g)$ maps to g on the n_1 outer qubits; since the g are independent in S_{out}, the $\varphi(g)$ cannot be products of inner stabilizers. Therefore

$$\text{rank}(S_{\text{con}}) = n_1(n_2 - 1) + (n_1 - 1) = n_1 n_2 - 1,$$

so on $n_1 n_2$ physical qubits the concatenated code encodes $k = n_1 n_2 - (n_1 n_2 - 1) = 1$ qubit; i.e., it is an $[n_1 n_2, 1]$ stabilizer code.

A compatible choice of logical operators is

$$\overline{X}_{\text{con}} = \varphi(\overline{X}_{\text{out}}), \qquad \overline{Z}_{\text{con}} = \varphi(\overline{Z}_{\text{out}}),$$

which commute with S_{con}, anticommute with each other, and are independent modulo S_{con}. ☐

Exercise 10.63

Suppose U is any unitary operation mapping the Steane code into itself, and such that $U\bar{Z}U^\dagger = \bar{X}$ and $U\bar{X}U^\dagger = \bar{Z}$. Prove that up to a global phase the action of U on the encoded states $|0_L\rangle$ and $|1_L\rangle$ is $|0_L\rangle \mapsto (|0_L\rangle + |1_L\rangle)/\sqrt{2}$ and $|1_L\rangle \mapsto (|0_L\rangle - |1_L\rangle)/\sqrt{2}$.

Solution

Concepts Involved: Logicals

Using that $\bar{Z}\left|0_L/1_L\right\rangle = \pm\left|0_L/1_L\right\rangle$ and that $U^\dagger U = I$ we find:

$$U|0_L\rangle = U(+1)|0_L\rangle = U\bar{Z}|0_L\rangle = U\bar{Z}U^\dagger U|0_L\rangle = \bar{X}U|0\rangle_L$$

$$U|1_L\rangle = U(-1)^2|1_L\rangle = -U(-1)|1_L\rangle = -U\bar{Z}|1_L\rangle = -U\bar{Z}U^\dagger U|1_L\rangle = -\bar{X}U|1_L\rangle$$

The above algebra shows that $U\left|0_L/1_L\right\rangle$ are eigenstates of $\bar{X}$ with eigenvalue ± 1, hence:

$$U|0_L\rangle \mapsto \frac{|0_L\rangle + |1_L\rangle}{\sqrt{2}}, \quad U|1_L\rangle \mapsto \frac{|0_L\rangle - |1_L\rangle}{\sqrt{2}}$$

up to a possible global phase. □

Exercise 10.64: Back propogation of errors

It is clear that an X error on the control qubit of a `CNOT` gate propagates to the target qubit. In addition, it turns out that a Z error on the target propagates back to the control! Show this using the stabilizer formalism, and also directly using quantum circuit identities. You may find Exercise 4.20 on page 179 useful.

Solution

Concepts Involved: Stabilizer Formalism, Errors

Let $U = \mathrm{CNOT}_{c\to t}$, i.e. $U|x,y\rangle = |x,\, y\oplus x\rangle$.

(A) Stabilizer / Heisenberg-picture propagation. We simply compute Pauli conjugations

$$\begin{aligned} UX_cU^\dagger|x,y\rangle &= UX_c|x,\, y\oplus x\rangle = U|x\oplus 1,\, y\oplus x\rangle \\ &= \big|x\oplus 1,\, (y\oplus x)\oplus(x\oplus 1)\big\rangle = |x\oplus 1,\, y\oplus 1\rangle = (X_cX_t)|x,y\rangle, \end{aligned}$$

so $UX_cU^\dagger = X_cX_t$ (a control-X propagates to the target).

$$\begin{aligned} UZ_tU^\dagger|x,y\rangle &= U\big((-1)^y|x,\, y\oplus x\rangle\big) = (-1)^{y\oplus x}|x,y\rangle \\ &= (-1)^x(-1)^y|x,y\rangle = (Z_cZ_t)|x,y\rangle, \end{aligned}$$

so $UZ_tU^\dagger = Z_cZ_t$ (a target-Z kicks back to the control).

Thus the CNOT conjugation table is

$$X_c \mapsto X_c X_t, \quad Z_c \mapsto Z_c, \quad X_t \mapsto X_t, \quad Z_t \mapsto Z_c Z_t.$$

(B) Direct circuit identities (projector form). Use

$$\mathrm{CNOT}_{c\to t} = |0\rangle\langle 0|_c \otimes I_t + |1\rangle\langle 1|_c \otimes X_t.$$

Control-X propagation.

$$\begin{aligned}(X_c\otimes I_t)\,\mathrm{CNOT} &= (X\otimes I)\big(\,|0\rangle\langle 0|\otimes I + |1\rangle\langle 1|\otimes X\big)\\ &= |1\rangle\langle 0|\otimes I + |0\rangle\langle 1|\otimes X\\ &= \big(\,|0\rangle\langle 0|\otimes I + |1\rangle\langle 1|\otimes X\big)(X\otimes X)\\ &= \mathrm{CNOT}\,(X_c\otimes X_t).\end{aligned}$$

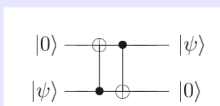
Hence X_c pushed through CNOT picks up an X_t.
Target-Z back-propagation.

$$\begin{aligned}(I_c\otimes Z_t)\,\mathrm{CNOT} &= (I\otimes Z)\big(\,|0\rangle\langle 0|\otimes I + |1\rangle\langle 1|\otimes X\big)\\ &= |0\rangle\langle 0|\otimes Z + |1\rangle\langle 1|\otimes ZX\\ &= |0\rangle\langle 0|\otimes Z - |1\rangle\langle 1|\otimes XZ\\ &= \big(\,|0\rangle\langle 0|\otimes I + |1\rangle\langle 1|\otimes X\big)\,(Z_c\otimes Z_t)\\ &= \mathrm{CNOT}\,(Z_c\otimes Z_t),\end{aligned}$$

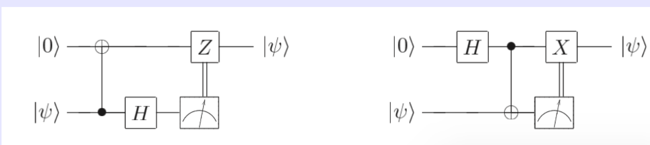
where we used $ZX = -XZ$ and $Z_c\,|0\rangle\langle 0| = +\,|0\rangle\langle 0|$, $Z_c\,|1\rangle\langle 1| = -\,|1\rangle\langle 1|$. Thus pushing Z_t through CNOT produces an extra Z_c on the control.
Both derivations establish X on the control spreads to the target, and Z on the target kicks back to the control. □

Exercise 10.65

An unknown qubit in the state $|\psi\rangle$ can be swapped with a second qubit which is prepared in the state $|0\rangle$ using only two controlled-`NOT` gates, with the circuit



Show that the two circuits below, which use only a single `CNOT` gate, with measurement and a classically controlled single qubit operation, also accomplish the same task



Solution

Concepts Involved: Controlled Operations, Quantum Measurement

Let the top wire be A (initially $|0\rangle$) and the bottom wire be B (initially $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$).
Circuit A (CNOT $B\to A$, measure B in X, apply Z on A if $m=1$).

$$|0\rangle_A |\psi\rangle_B \xrightarrow{\mathrm{CNOT}_{B\to A}} \alpha|00\rangle + \beta|11\rangle .$$

Write B in the X basis, $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$:

$$\alpha|00\rangle + \beta|11\rangle = \frac{1}{\sqrt{2}}\Big(|+\rangle_B (\alpha|0\rangle_A + \beta|1\rangle_A) + |-\rangle_B (\alpha|0\rangle_A - \beta|1\rangle_A)\Big).$$

Measuring B in the X basis gives outcome $m \in \{0,1\}$ with post-measurement state on A

$$m=0: \; |\psi\rangle, \qquad m=1: \; Z|\psi\rangle .$$

Applying the classical correction Z^m on A yields $|\psi\rangle_A$ deterministically. The measured qubit B is known and can be reset to $|0\rangle$.
Circuit B (H on A, CNOT $A\to B$, measure B in Z, apply X on A if $m=1$).

$$|0\rangle_A |\psi\rangle_B \xrightarrow{H_A} |+\rangle_A |\psi\rangle_B = \frac{1}{\sqrt{2}}\Big(|0\rangle_A (\alpha|0\rangle_B + \beta|1\rangle_B) + |1\rangle_A (\alpha|1\rangle_B + \beta|0\rangle_B)\Big).$$

Measure B in the computational basis with outcome $m \in \{0,1\}$:

$$m=0: \; \alpha|0\rangle_A + \beta|1\rangle_A = |\psi\rangle_A, \qquad m=1: \; \beta|0\rangle_A + \alpha|1\rangle_A = X|\psi\rangle_A .$$

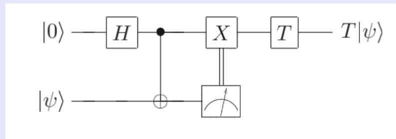
Applying X^m on A gives $|\psi\rangle_A$ deterministically. Again, B is known (in $|m\rangle$) and can be reset to $|0\rangle$.
Hence, in both circuits the net effect is

$$|0\rangle_A |\psi\rangle_B \longmapsto |\psi\rangle_A |0\rangle_B ,$$

i.e., the unknown state is transferred to the top wire using one CNOT plus measurement and a classically controlled Pauli. □

Exercise 10.66

One way to implement a $\pi/8$ gate is to first swap the qubit state $|\psi\rangle$ you wish to transform with some unknown state $|0\rangle$, then to apply a $\pi/8$ gate to the resulting qubit. Here is a quantum circuit which does that:



Doing this does not seem particularly useful, but actually it leads to something which is! Show that by using the relations $TX = \exp(-i\pi/4)SX$ and $TU = UT$ (U is the controlled-NOT gate, and T acts on the control qubit) we may obtain the circuit of Figure 10.25.

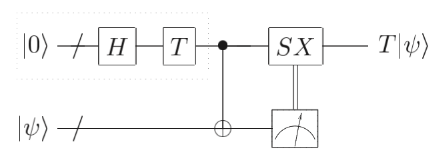


Figure 10.25. Quantum circuit which fault-tolerantly implements a $\pi/8$ gate. The dashed box represents a (non-fault-tolerant) preparation procedure for the ancilla state $(|0\rangle + \exp(i\pi/4)|1\rangle)/\sqrt{2}$; how to do this preparation fault-tolerantly is explained in the text. The slash on the wire denotes a bundle of seven qubits, and the double-line wire represents the classical bit resulting from the measurement. Note that the final SX operation is controlled by the measurement result.

There is an error in the problem statement, and the corrected first relation is that $TX = e^{-i\pi/4}SXT$.

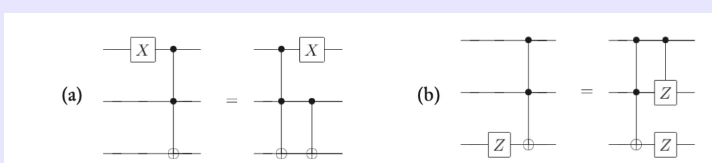
Solution

Concepts Involved: Quantum Measurement, Controlled Operations

Starting wihh the given circuit, we commute the T gate across the measurement-controlled X gate. Using the $TX = e^{-i\pi/4}SXT$ relation (discarding the irrelevant global phase), we get a measurement-controlled SX operation preceded by a T. We then commute the T across the controlled-NOT gate. This yields the circuit of Figure 10.25. □

Exercise 10.67

Show that the following circuit identities hold:



Solution

Concepts Involved: Controlled Operations

Let $U = \mathrm{CCX}_{12\to 3}$ and let $a, b, c \in \{0, 1\}$.

(a) X **on a control.** We compare both sides on $|a, b, c\rangle$:

$$\text{LHS}: \quad U X_1 |a, b, c\rangle = U |a \oplus 1, b, c\rangle = \big|a \oplus 1, b,\ c \oplus (a \oplus 1)b\big\rangle .$$

$$\text{RHS}: \quad X_1\, \mathrm{CX}_{2\to 3}\, U |a, b, c\rangle = X_1\, \mathrm{CX}_{2\to 3} |a, b, c \oplus ab\rangle = X_1 |a, b,\ c \oplus ab \oplus b\rangle = |a \oplus 1, b,\ c \oplus ab \oplus b\rangle .$$

Since $(a \oplus 1)b = ab \oplus b$, the third components coincide; hence LHS = RHS for all a, b, c, so

$$U X_1 \;=\; X_1\, \mathrm{CX}_{2\to 3}\, U.$$

(b) Z **on the target.**

$$\text{LHS}: \quad U Z_3 |a, b, c\rangle = U\big((-1)^c |a, b, c\rangle\big) = (-1)^c |a, b,\ c \oplus ab\rangle .$$

RHS:

$$\begin{aligned} Z_3\, \mathrm{CZ}_{1,2}\, U |a, b, c\rangle = Z_3\, \mathrm{CZ}_{1,2} |a, b,\ c \oplus ab\rangle &= Z_3\big((-1)^{ab} |a, b,\ c \oplus ab\rangle\big) \\ &= (-1)^{ab}(-1)^{c\oplus ab} |a, b,\ c \oplus ab\rangle . \end{aligned}$$

Because $(-1)^{ab}(-1)^{c\oplus ab} = (-1)^c$, the amplitudes match; thus LHS = RHS for all a, b, c, and

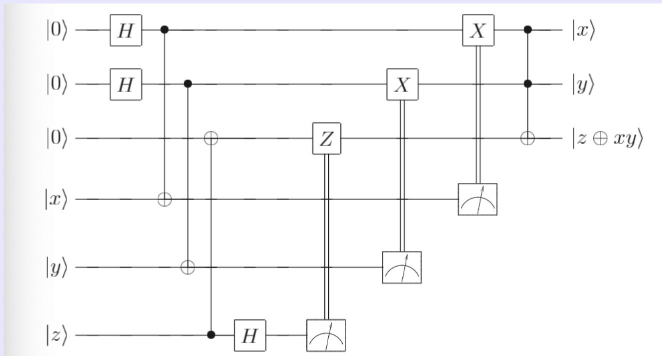
$$U Z_3 \;=\; Z_3\, \mathrm{CZ}_{1,2}\, U.$$

□

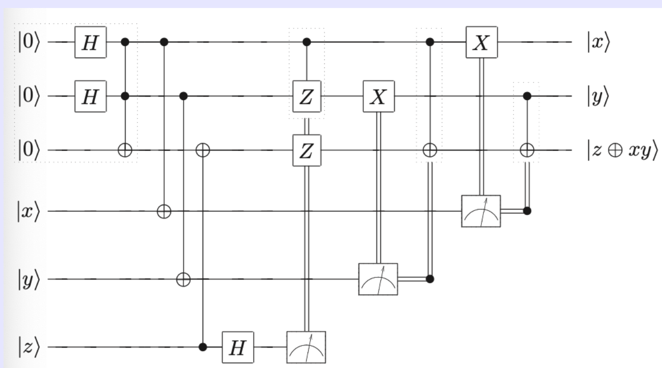
Exercise 10.68: Fault-tolerant Toffoli gate construction

A procedure similar to the above sequence of exercises for the $\pi/8$ gate gives a fault-tolerant Toffoli gate.

(1) First, swap the three qubit state $|xyz\rangle$ you wish to transform with some known state $|000\rangle$, then apply a Toffoli gate to the resulting qubits¿ Show that the following circuit accomplishes this task:



(2) Using the commutation rules from Exercise 10.67, show that moving the final Toffoli gate all the way back to the left side gives the circuit



(3) Assuming the ancilla preparation shown in the leftmost dotted box can be done fault-tolerantly, show that this circuit can be used to give a fault-tolerant implementation of the Toffoli gate using the Steane code.

Solution

Concepts Involved: Fault Tolerance, Controlled Operations

(2) Sliding the Toffoli to the left. Use Heisenberg conjugation relations for CCX (controls $1, 2$; target 3):

$$\mathrm{CCX}\, X_1\, \mathrm{CCX} = X_1, \quad \mathrm{CCX}\, X_2\, \mathrm{CCX} = X_2, \quad \mathrm{CCX}\, X_3\, \mathrm{CCX} = X_3,$$
$$\mathrm{CCX}\, Z_1\, \mathrm{CCX} = Z_1\, \mathrm{CNOT}_{2\to 3}, \qquad \mathrm{CCX}\, Z_2\, \mathrm{CCX} = Z_2\, \mathrm{CNOT}_{1\to 3}, \qquad \mathrm{CCX}\, Z_3\, \mathrm{CCX} = Z_3,$$
$$\mathrm{CCX}\,(\mathrm{CNOT}_{i\to j})\,\mathrm{CCX} \text{ is a CNOT on the same pair, possibly accompanied by Paulis per the above.}$$

Pushing CCX leftward across each teleportation gadget converts that gadget's classically controlled X/Z corrections into the extra single-qubit Z/X boxes and dotted CNOTs drawn in the second figure. After commuting all the way left, the Toffoli resides entirely in the *leftmost* dotted box, and everything to its right is Clifford + measurements. Thus the second circuit is equivalent to the first.

(3) Fault tolerance with the Steane code. In the second circuit the non-Clifford operation (Toffoli) is isolated inside the leftmost dotted box as an *ancilla factory* producing the three-qubit "Toffoli state." Prepare this ancilla *fault-tolerantly* (verify; discard on failure). The remainder of the circuit consists only of transversal Clifford gates (H, CNOT, S), Pauli-basis measurements, and Pauli-frame updates applied between the encoded data blocks and the verified ancilla. For the Steane $[[7, 1, 3]]$ CSS code, these operations are transversal and satisfy the FT criterion (one fault creates at most one data error per block, which is correctable by distance 3). Since the only non-Clifford gate is performed off-line during verified ancilla preparation, the on-line interaction is fault tolerant. Therefore the construction implements a FT logical Toffoli.

□

Exercise 10.69

Show that a single failure anywhere in the ancilla preparation and verification can lead to at most one X or Y error in the ancilla output.

Solution

Concepts Involved: Fault Tolerance

Let us represent possible X/Y errors on the three output qubits by a vector $e = (e_1, e_2, e_3) \in \{0, 1\}^3$, where $e_i = 1$ means that qubit i carries an X or Y component. The verification circuit measures the parities of qubits $(1, 2)$ and $(2, 3)$. In matrix form,

$$H = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}, \qquad \text{accept if } He = 0 \pmod 2.$$

A single *preparation fault* can introduce at most two X/Y errors, so the dangerous cases are $e = 110, 011, 101$. In each case one finds $He \neq 0$, so the verifier detects the error and the run is rejected. Thus, when a preparation fault slips through verification, the error pattern must be either $e = 000$ (no error) or a single flip $e = e_i$.

A single *verification fault* also cannot produce multiple X/Y errors on the outputs: a faulty CNOT touches at most one ancilla output; a single-qubit fault affects only one line; and faults in verifier preparation or measurement flip the syndrome but add no data errors.

Therefore, under the single-fault assumption and the acceptance condition $He = 0$, the only possible error patterns on the ancilla outputs are

$$e \in \{000, 100, 010, 001\},$$

i.e. at most one X or Y error. □

Exercise 10.70

Show that Z errors in the ancilla do not propagate to affect the encoded data, but result in an incorrect measurement result being observed.

Solution

Concepts Involved: Fault Tolerance

Consider the measurement of an X-type stabilizer using an ancilla prepared in $|+\rangle$, with CNOT gates $\mathrm{CNOT}_{a\to i}$ (ancilla as control, data as target) for each qubit i in the stabilizer, followed by a measurement of X_a.

Suppose a Z error occurs on the ancilla. Using the CNOT propagation rules in the Heisenberg picture,

$$\mathrm{CNOT}_{c\to t}: \quad Z_c \mapsto Z_c, \qquad Z_t \mapsto Z_c Z_t,$$

we see that for $\mathrm{CNOT}_{a\to i}$,

$$\mathrm{CNOT}_{a\to i}\, Z_a\, \mathrm{CNOT}^\dagger_{a\to i} = Z_a.$$

Thus the operator Z_a remains confined to the ancilla throughout the fanout, and no factor on any data qubit is created. Therefore a Z error on the ancilla *never propagates to the encoded data*.

At the end of the circuit the ancilla is measured in the X basis. Since Z_a anticommutes with X_a, the presence of a Z_a error flips the measurement outcome:

$$Z_a X_a = -X_a Z_a.$$

Hence the encoded data remain untouched, but the syndrome bit extracted from the ancilla is flipped.

In summary, a Z error on the ancilla does not affect the encoded data, but it results in an incorrect measurement result being observed. □

Exercise 10.71

Verify that when $M = e^{-i\pi/4} SX$ the procedure we have described gives a fault-tolerant method for measuring M.

Solution

Concepts Involved: Fault Tolerance

We wish to measure

$$M = e^{-i\pi/4} SX \;=\; e^{-i(\pi/4)Z} X \;=\; \tfrac{1}{\sqrt{2}}(X + Y). \tag{14}$$

Thus $M^\dagger = M$ and $M^2 = I$, so the eigenvalues are ± 1.

Controlled-M. Using projector decompositions,

$$\mathrm{CNOT}_{a\to d} = \big(|0\rangle\langle 0|_a \otimes I_d\big) + \big(|1\rangle\langle 1|_a \otimes X_d\big), \tag{15}$$
$$\mathrm{CS}_{a\to d} = \big(|0\rangle\langle 0|_a \otimes I_d\big) + \big(|1\rangle\langle 1|_a \otimes S_d\big), \tag{16}$$
$$\mathrm{CZ}_{a\to d} = \big(|0\rangle\langle 0|_a \otimes I_d\big) + \big(|1\rangle\langle 1|_a \otimes Z_d\big), \tag{17}$$
$$T_a = |0\rangle\langle 0|_a + e^{i\pi/4}\,|1\rangle\langle 1|_a. \tag{18}$$

Hence

$$T_a\,\mathrm{CZ}_{a\to d}\,\mathrm{CS}_{a\to d}\,\mathrm{CNOT}_{a\to d} = |0\rangle\langle 0|_a \otimes I_d \;+\; |1\rangle\langle 1|_a \otimes \big(e^{i\pi/4} ZSX\big)_d \tag{19}$$

We must check then that $\bar{M} = (e^{i\pi/4}ZSX)^{\otimes 7}$ on the physical qubits of the Steane code has the logical action of M. First, we calculate:

$$MZM^\dagger = e^{i\pi/4}SXZe^{-i\pi/4}XS^\dagger = S(-ZX)XS^\dagger = -Z$$

$$MXM^\dagger = SXXXS^\dagger = Y$$

and then in the codespace we calculate:

$$\bar{M}\bar{Z}\bar{M}^\dagger = (e^{i\pi/4}ZSX)^{\otimes 7}Z^{\otimes 7}(e^{-i\pi/4}XS^\dagger Z)^{\otimes 7} = (ZSXZXS^\dagger Z)^{\otimes 7} = (Z(-Z)Z)^{\otimes 7} = -(Z)^{\otimes 7} = -\bar{Z}$$

$$\bar{M}\bar{X}\bar{M}^\dagger = (e^{i\pi/4}ZSX)^{\otimes 7}X^{\otimes 7}(e^{-i\pi/4}XS^\dagger Z)^{\otimes 7} = (ZSXXXS^\dagger Z)^{\otimes 7} = (ZYZ)^{\otimes 7} = (-Y)^{\otimes 7} = \bar{Y}$$

where in the last equality we note that $\bar{Y} = i\bar{X}\bar{Z} = iX^{\otimes 7}Z^{\otimes 7} = i(-iY)^{\otimes 7} = i(-i)^7 Y^{\otimes 7} = -Y^{\otimes 7}$. So indeed this is the correct action.

Measurement scheme. Prepare a verified m-qubit cat ancilla

$$|\mathrm{cat}_m\rangle \;=\; \tfrac{1}{\sqrt{2}}\big(|0^m\rangle + |1^m\rangle\big), \tag{20}$$

apply $\mathrm{C}M_{a_j\to d_j}$ transversally between each ancilla qubit a_j and its paired data qubit d_j, then measure all ancillas in the X basis and take the parity. The joint unitary is

$$U \;=\; \prod_j \mathrm{C}M_{a_j\to d_j} \;=\; |0^m\rangle\langle 0^m| \otimes I \;+\; |1^m\rangle\langle 1^m| \otimes M_L, \tag{21}$$

so phase kickback effects a projective measurement of the encoded observable M_L.

Fault tolerance. Transversality ensures any single faulty two-qubit gate touches at most one data qubit. Moreover,

$$[Z_a, \mathrm{CNOT}_{a\to d}] = 0, \qquad [Z_a, \mathrm{CS}_{a\to d}] = 0, \tag{22}$$

so an ancilla Z never propagates to data and only flips the X-basis readout. Ancilla X/Y faults can propagate to the *paired* data qubit but to no others. Verified cat preparation bounds non-Z ancilla faults to at most one, and repeating the measurement with majority vote suppresses readout flips. Hence the procedure both measures M correctly and satisfies the single-fault–single-data-error criterion. □

Exercise 10.72: Fault-tolerant Toffoli ancilla state construction

Show how to fault-tolerantly prepare the state created by the circuit in the dotted box of exercise 10.68, that is,

$$\frac{|000\rangle + |010\rangle + |100\rangle + |111\rangle}{2}$$

You may find it helpful to first give the stabilizer generators for this state.

Solution

Concepts Involved: Fault Tolerance, Stabilizer Formalism

TODO

□

Exercise 10.73: Fault-tolerant encoded state construction

Show that the Steane code encoded $|0\rangle$ state can be constructed fault-tolerantly in the following manner.

(1) Begin with the circuit of Figure 10.16, and replace the measurement of each generator, as shown in Figure 10.30, with each ancilla qubit becoming a cat state $|00\ldots0\rangle + |11\ldots1\rangle$, and the operations rearranged to have their controls on different qubits, so that errors do not propagate within the code block.

(2) Add a stage to fault-tolerantly measure Z.

(3) Calculate the error probability of this circuit, and of the circuit when the generator measurements are repeated three times and majority voting is done.

(4) Enumerate the operations which should be performed conditioned on the measurement results and show that they can be done fault-tolerantly.

Solution

Concepts Involved:

TODO □

Exercise 10.74

Construct a quantum circuit to fault-tolerantly generate the encoded $|0\rangle$ state for the five qubit code (Section 10.5.6).

Solution

Concepts Involved:

TODO □

Problem 10.1

Channels $\mathcal{E}_1$ and $\mathcal{E}_2$ are said to be *equivalent* if there exist unitary channels $\mathcal{U}$ and $\mathcal{V}$ such that $\mathcal{E}_2 = \mathcal{U} \circ \mathcal{E}_1 \circ \mathcal{V}$.

(1) Show that the relation of channel equivalence is an equivalence relation.

(2) Show how to turn an error-correcting code for $\mathcal{E}_1$ into an error-correcting code for $\mathcal{E}_2$. Assume that the error-correction procedure for $\mathcal{E}_1$ is performed as a projective measurement followed by a conditional unitary operation, and explain how the error-correction procedure for $\mathcal{E}_2$ can be performed in the same fashion.

Solution

Concepts Involved: Unitary Operators, Quantum Channels, Quantum Measurement.
Recall that a unitary channel is a quantum channel such that $\mathcal{U}(\rho) = U\rho U^\dagger$ for a unitary operator U.

(1) *Reflexivity.* Take $\mathcal{U} = \mathcal{V} = \mathcal{I}$ (the identity channel, which is unitary). Then $\mathcal{E} = \mathcal{I} \circ \mathcal{E} \circ \mathcal{I}$ so $\mathcal{E} \sim \mathcal{E}$ as desired.

Symmetry. Suppose $\mathcal{E}_1 \sim \mathcal{E}_2$. Then there exists $\mathcal{U}, \mathcal{V}$ such that $\mathcal{E}_2 = \mathcal{U} \circ \mathcal{E}_1 \circ \mathcal{V}$. It then follows that $\mathcal{E}_1 = \mathcal{U}^\dagger \circ \mathcal{E}_2 \circ \mathcal{V}^\dagger$, where $\dagger$ denotes the dual channel, i.e. if $\mathcal{U}(\rho) = U\rho U^\dagger$ then $\mathcal{U}^\dagger(\rho) = U^\dagger \rho U$. The dual channel of a unitary channel is also a unitary channel as $U^\dagger$ is unitary for unitary U. Hence $\mathcal{E}_2 \sim \mathcal{E}_1$ as desired.

Transitivity. Suppose $\mathcal{E}_1 \sim \mathcal{E}_2$ and $\mathcal{E}_2 \sim \mathcal{E}_3$. Then there exist $\mathcal{U}, \mathcal{V}, \mathcal{U}', \mathcal{V}'$ such that $\mathcal{E}_2 = \mathcal{U} \circ \mathcal{E}_1 \circ \mathcal{V}$ and $\mathcal{E}_3 = \mathcal{U}' \circ \mathcal{E}_2 \circ \mathcal{V}'$. It then follows that $\mathcal{E}_3 = \mathcal{U}' \circ \mathcal{U} \circ \mathcal{E}_1 \circ \mathcal{V} \circ \mathcal{V}'$ and since the composition of two unitary channels is another unitary channel (as the composition of two unitary operators is again unitary), it follows that $\mathcal{E}_1 \sim \mathcal{E}_3$ as desired.

We have therefore shown that channel equivalence is an equivalence relation.

(2) Suppose that $\mathcal{E}_2 = \mathcal{U} \circ \mathcal{E}_1 \circ \mathcal{V}$. If the error correcting code for $\mathcal{E}_1$ consists of a projective measurement P_m followed by a conditional unitary C_m (where the subscript denotes a conditioning on the measurement outcome m), this means that for a given state ρ which lies in the support of the code that $C_m P_m \mathcal{E}_1(\rho) P_m C_m^\dagger \propto \rho$.

If we instead have $\mathcal{E}_2$ as our noise channel, let us rotate the measurement basis and carry out the measurement $P'_m = UP_mU^\dagger$ (note that a unitary conjugation of a projector remains a projector, as $P'^2_m = (UP_mU^\dagger)^2 = UP_mU^\dagger UP_mU^\dagger = UP_mU^\dagger = P'_m$). Let us also modify the conditional unitary to be $C'_m = V^\dagger C_m U^\dagger$. We then have:

$$\begin{aligned} C'_m P'_m \mathcal{E}_2(\rho) P'_m C'^\dagger_m &= (V^\dagger C_m U^\dagger)(U P_m U^\dagger) U \mathcal{E}_1(V \rho V^\dagger) U^\dagger (U P_m U^\dagger)(V^\dagger C_m U^\dagger)^\dagger \\ &= V^\dagger C_m U^\dagger U P_m U^\dagger U \mathcal{E}_1(V \rho V^\dagger) U^\dagger U P_m U^\dagger U C_m^\dagger V \\ &= V^\dagger C_m P_m \mathcal{E}_1(V \rho V^\dagger) P_m C_m^\dagger V \\ &\propto V^\dagger V \rho V^\dagger V \\ &= \rho \end{aligned}$$

which shows that the error-correction procedure for $\mathcal{E}_2$ can be performed in the same fashion (with the unitary-modified measurements/unitaries).

□

Problem 10.2: Gilbert–Varshamov bound

Prove the Gilbert-Varshamov bound for CSS codes, namely, that an $[n, k]$ CSS code correcting t errors exists for some k such that

$$\frac{k}{n} \geq 1 - 2H\left(\frac{2t}{n}\right).$$

As a challenge, you may like to try proving the Gilbert–Varshamov bound for a general stabilizer code, namely, that there exists an $[n, k]$ stabilizer code correcting errors on t qubits, with

$$\frac{k}{n} \geq 1 - \frac{2\log(3)t}{n} - H\left(\frac{2t}{n}\right)$$

Solution

Concepts Involved: Entropy, CSS Codes, Stabilizer Codes

CSS Gilbert–Varshamov. A CSS code is specified by linear binary codes $C_2 \subseteq C_1 \subseteq \mathbb{F}_2^n$ and encodes $k = \dim C_1 - \dim C_2$ qubits. It corrects t errors (i.e. has $d \geq 2t+1$) if

$$\mathrm{dist}(C_1) \geq 2t+1 \quad \text{and} \quad \mathrm{dist}(C_2^\perp) \geq 2t+1,$$

which rules out all nontrivial Z-type and X-type logicals of weight $\leq 2t$, respectively.
Let $\mathcal{W} = \{w \in \mathbb{F}_2^n \setminus \{0\} : \mathrm{wt}(w) \leq 2t\}$ with $|\mathcal{W}| = V(n, 2t) - 1$, where $V(n,r) = \sum_{i=0}^{r} \binom{n}{i}$. Choose C_1 uniformly at random among k_1-dimensional subspaces of $\mathbb{F}_2^n$. For any fixed nonzero w,

$$\Pr[w \in C_1] = 2^{k_1 - n}, \qquad \Pr\left[w \in C_1^\perp\right] = 2^{-k_1}.$$

By a union bound over $\mathcal{W}$,

$$\Pr[\exists\, w \in \mathcal{W} \cap C_1] \leq V(n, 2t)\, 2^{k_1 - n}, \qquad \Pr\left[\exists\, w \in \mathcal{W} \cap C_1^\perp\right] \leq V(n, 2t)\, 2^{-k_1}.$$

Thus there exists a particular C_1 with both $\mathrm{dist}(C_1) \geq 2t+1$ and $\mathrm{dist}(C_1^{\perp}) \geq 2t+1$ whenever

$$\log_2 V(n,2t) \;\leq\; k_1 \;\leq\; n - \log_2 V(n,2t). \tag{23}$$

Now condition on such a C_1 and pick C_2 uniformly at random among the k_2-dimensional subspaces of C_1. For any fixed $w \notin C_1^{\perp}$, the event $w \in C_2^{\perp}$ is equivalent to w being orthogonal to C_2, which happens with probability 2^{-k_2}. Another union bound gives

$$\Pr\Big[\exists\, w \in \mathcal{W} \cap C_2^{\perp}\Big] \;\leq\; V(n,2t)\, 2^{-k_2},$$

so there exists $C_2 \subseteq C_1$ with $\mathrm{dist}(C_2^{\perp}) \geq 2t+1$ whenever

$$k_2 \;\geq\; \log_2 V(n,2t). \tag{24}$$

Combining (23)–(24), there exist nested $C_2 \subseteq C_1$ with

$$k = \dim C_1 - \dim C_2 \;\geq\; n - 2\log_2 V(n,2t).$$

Using the entropy bound $\log_2 V(n,2t) \leq n\, H(2t/n)$, we obtain

$$\frac{k}{n} \;\geq\; 1 - 2\, H\Big(\frac{2t}{n}\Big),$$

which is the CSS Gilbert–Varshamov bound. (For finite n, replace the entropy bound by the exact $\log_2 V(n,2t)$ and handle floors/ceilings.)

Stabilizer Gilbert–Varshamov (challenge). A stabilizer code corresponds to an $(n-k)$-dimensional totally isotropic subspace $S \leq \mathbb{F}_2^{2n}$ with respect to the standard symplectic form. It corrects t errors iff no nontrivial Pauli of weight $\leq t$ lies in $N(S) \setminus S$, equivalently no nonzero symplectic vector of weight $\leq 2t$ lies in $S^{\perp} \setminus S$.
The number of nonidentity Pauli errors supported on at most $2t$ qubits is

$$B(n,2t) = \sum_{i=1}^{2t} 3^i \binom{n}{i}.$$

Choose S uniformly at random among $(n-k)$-dimensional isotropic subspaces. A fixed nonidentity Pauli has a uniformly random syndrome; hence

$$\Pr\big[E \in N(S)\big] = 2^{-(n-k)}, \qquad \Pr\big[E \in N(S) \setminus S\big] \leq 2^{-(n-k)}.$$

By a union bound over $B(n,2t)$ errors, a code with distance $\geq 2t+1$ exists whenever $2^{n-k} > B(n,2t)$. Using the bounds $\sum_{i \leq 2t} \binom{n}{i} \leq 2^{\,nH(2t/n)}$ and $3^i \leq 2^{i \log_2 3}$, we get

$$\log_2 B(n,2t) \;\leq\; n\, H\Big(\frac{2t}{n}\Big) + 2t \log_2 3,$$

so it suffices that $n - k > nH(2t/n) + 2t\log_2 3$, i.e.

$$\frac{k}{n} \;\geq\; 1 - H\Big(\frac{2t}{n}\Big) - \frac{2\log 3}{n}\, t,$$

which is the stabilizer Gilbert–Varshamov bound. □

Problem 10.3: Encoding stabilizer codes

Suppose we assume that the generators for the code are in standard form, and that the encoded Z and X operators have been constructed in standard form. Find a circuit taking the $n \times 2n$ check matrix corresponding to a listing of all the generators for the code together with the encoded Z operations from

$$G = \left[\begin{array}{ccc|ccc} 0 & 0 & 0 & I & 0 & 0 \\ 0 & 0 & 0 & 0 & I & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

to the standard form

$$\left[\begin{array}{ccc|ccc} I & A_1 & A_2 & B & 0 & C_2 \\ 0 & 0 & 0 & D & I & E \\ 0 & 0 & 0 & A_2^T & 0 & I \end{array}\right]$$

Solution

Concepts Involved:

TODO □

Problem 10.4: Encoding by teleportation

Suppose you are given a qubit $|\psi\rangle$ to encode in a stabilizer code, but you are not told anything about how $|\psi\rangle$ was constructed: it is an unknown state¿ Construct a circuit to perform the encoding in the following manner:

(1) Explain how to fault-tolerantly construct the partially encoded state

$$\frac{|0\rangle|0_L\rangle + |1\rangle|1_L\rangle}{\sqrt{2}},$$

by writing this as a stabilizer state, so it can be prepared by measuring stabilizer generators.

(2) Show how to fault-tolerantly perform a Bell basis measurement with $|\psi\rangle$ and the unencoded qubit from this state.

(3) Give the Pauli operations which you need to fix up the remaining encoded qubit after this measurement, so that it becomes $|\psi\rangle$, as in the usual quantum teleportation scheme.

Compute the probability of error of this circuit. Also show how to modify the circuit to perform fault-tolerant decoding.

Solution

Concepts Involved: Fault Tolerance

Let B be an $[[n, 1, d]]$ stabilizer code with stabilizer $\langle S_1, \ldots, S_{n-1}\rangle$ and logical Paulis $\bar{X}, \bar{Z}$ acting on the block B. Let q_0 be a single physical qubit, and let $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ be the unknown input qubit to be encoded.

(1) Fault-tolerant preparation of the physical–logical Bell state. Define

$$|\Phi^+_{1\text{-}L}\rangle \;=\; \frac{|0\rangle|0_L\rangle + |1\rangle|1_L\rangle}{\sqrt{2}}. \tag{25}$$

On $q_0 \cup B$, this is the unique $+1$ common eigenstate of the commuting set

$$\mathcal{S}_\Phi \;=\; \big\langle\, S_1, \ldots, S_{n-1},\, Z_{q_0}\bar{Z}_B,\, X_{q_0}\bar{X}_B \,\big\rangle. \tag{26}$$

To prepare it fault-tolerantly:

(1) Project B into the codespace by FT measurements of $\{S_j\}$ (e.g. Shor/Steane extraction with verified ancillas), applying Pauli corrections as needed.

(2) FT-measure $Z_{q_0}\bar{Z}_B$ and $X_{q_0}\bar{X}_B$ using small *verified cat* ancillas that couple disjointly to q_0 and to a transversal implementation of $\bar{Z}, \bar{X}$. Flip by Z on q_0 or $\bar{Z}$ on B (and similarly for X) to enforce the $+1$ outcomes.

Optionally repeat each parity measurement and majority vote to suppress readout error. The result is $|\Phi^+_{1\text{-}L}\rangle$.

(2) Fault-tolerant Bell measurement on (ψ, q_0). A Bell measurement is obtained by jointly measuring the commuting observables $Z_\psi Z_{q_0}$ and $X_\psi X_{q_0}$.

$$\text{Record } m_Z, m_X \in \{0,1\} \text{ such that } Z_\psi Z_{q_0} = (-1)^{m_Z}, \quad X_\psi X_{q_0} = (-1)^{m_X}. \tag{27}$$

Implement each parity measurement fault-tolerantly via verified cat ancillas with *disjoint* data couplings:

- ZZ parity: prepare and verify $|\mathrm{cat}_Z\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$; apply $\mathrm{CNOT}(\psi \to a_1)$ and $\mathrm{CNOT}(q_0 \to a_2)$; measure a_1, a_2 in Z and set $m_Z = z_1 \oplus z_2$.
- XX parity: either Hadamard both data qubits and reuse the ZZ gadget, or prepare/verify $|\mathrm{cat}_X\rangle = (|{+}{+}\rangle + |{-}{-}\rangle)/\sqrt{2}$, couple via $\mathrm{CZ}(\psi, a_1)$ and $\mathrm{CZ}(q_0, a_2)$, measure a_1, a_2 in X and set $m_X = x_1 \oplus x_2$.

Repetition with majority vote further suppresses measurement errors.

(3) Logical Pauli fix-ups (teleportation). Using the teleportation identity with a physical–logical Bell pair,

$$|\psi\rangle_\psi \otimes |\Phi^+_{1\text{-}L}\rangle_{q_0,B} \xrightarrow{\text{measure } ZZ, XX \text{ on } (\psi, q_0)} \bar{X}^{\,m_X}\bar{Z}^{\,m_Z}\,|\psi\rangle_L \;\text{ on } B. \tag{28}$$

Thus apply the logical correction on B

$$(m_Z, m_X) = (0,0) \Rightarrow I, \quad (1,0) \Rightarrow \bar{Z}, \quad (0,1) \Rightarrow \bar{X}, \quad (1,1) \Rightarrow \bar{Z}\bar{X}, \tag{29}$$

yielding the encoded state $|\psi\rangle_L$.

Probability of error. Under an independent local stochastic noise model with physical error rate p per location, a distance-$d \geq 3$ code, verified ancillas, transversal/disjoint data couplings, and (optional) readout repetition, any single physical fault during preparation, Bell parity measurements, or correction produces at most one correctable data error (or a tracked Pauli-frame flip). Hence logical failure requires at least two faults:

$$P_{\text{fail}} \;=\; C\,p^2 + O(p^3), \tag{30}$$

where C counts malignant fault pairs (gadget- and code-dependent). For general distance d with gadgets respecting distance, letting $t = \lfloor (d-1)/2 \rfloor$,

$$P_{\text{fail}} \;=\; \tilde{C}\,p^{\,t+1} + O\Big(p^{\,t+2}\Big). \tag{31}$$

Fault-tolerant decoding (teleport out). To decode $|\psi\rangle_L$ onto a fresh physical qubit q_{out}, prepare $|\Phi^+_{1\text{-}L}\rangle_{q_{\text{out}},B'}$ as above, then FT-measure $\bar{Z}_B\bar{Z}_{B'}$ and $\bar{X}_B\bar{X}_{B'}$. Apply $Z^{m_Z}X^{m_X}$ (or frame update) to q_{out} to obtain $|\psi\rangle$. The same fault-tolerance order and scaling apply. □

Problem 10.5

Suppose $C(S)$ is an $[n, 1]$ stabilizer code capable of correcting errors on a single qubit. Explain how a fault-tolerant implementation of the controlled-`NOT` gate may be implemented between two logical qubits encoded using this code, using only fault-tolerant stabilizer state preparation, fault-tolerant measurement of elements of the stabilizer, and normalizer gates applied transversally.

Solution

Concepts Involved:

TODO □